The Folk Theorem for the Prisoner's Dilemma with Endogenous Private Monitoring*

Guilherme Carmona† Krittanai Laohakunakorn‡

University of Surrey University of Surrey

June 12, 2023

Abstract

We study the repeated prisoner's dilemma with private monitoring under the assumption that the monitoring structure is endogenously chosen by the players in each period. We allow the players to choose from all possible monitoring structures. If the players disagree on the monitoring structure they would like, the realized monitoring structure is determined by a function that aggregates their choices. When one player can dictate the monitoring structure, then the repetition of the stage Nash is the only sequential equilibrium outcome. In contrast, when no player can dictate the monitoring structure, we provide conditions on the aggregation function under which any strictly individually rational and feasible payoff vector can be supported in sequential equilibrium.

*We wish to thank the editor (Tilman Börgers), an associate editor, three anonymous referees and seminar participants at the 2022 PET conference (Marseille, France) for very detailed and helpful comments. Any remaining errors are, of course, ours.

†Address: University of Surrey, School of Economics, Guildford, GU2 7XH, UK; email: g.carmona@surrey.ac.uk.

‡Address: University of Surrey, School of Economics, Guildford, GU2 7XH, UK; email: k.laohakunakorn@surrey.ac.uk.

1 Introduction

The prisoner's dilemma is a prototypical example of a game whose infinite repetition can greatly expand the set of equilibrium outcomes. This conclusion has been established by several folk theorems (see Section 2 for a literature review) for various monitoring structures which are exogenously given and independent of the decisions of the players. This feature can be critical for the construction of equilibria which often rely on strategies (such as belief-free strategies) that depend on the exact details of the fixed monitoring structure.

Thus, typical arguments that establish the folk theorem in the repeated prisoner's dilemma may not apply if players can manipulate the monitoring structure. Such manipulations are a realistic feature of long term relationships, and in fact it seems that the purpose of such manipulations is exactly to solve the coordination problems that arise with private information. For example, people actively design monitoring by asking for reports, having meetings, etc, and often this is done to help those interacting repeatedly to coordinate their activities and obtain good outcomes. Hence, we ask whether cooperation can be sustained — and more generally whether the folk theorem holds — in the repeated prisoner's dilemma when monitoring can be designed.

To take a concrete example that will motivate some of the features we wish to capture in our model, consider a cartel with two members. In order to collude successfully, secret price-cutting behavior must be deterred; this relies on some degree of monitoring. It is natural that the choice of monitoring technology is made by the firms themselves, i.e. monitoring should be endogenous. For example, as described in Marshall and Marx (2012), cartels often hire consulting firms to assist with monitoring.[1] But the cartel members may have conflicting incentives over the choice of monitoring technology; for example, each firm may wish to know the sales figures of

[1]One such consulting firm is Fides, later known as AC-Treuhand. According to the European Court ruling in *Organic Peroxides*, AC-Treuhand, among other things, "collected data on [*Organic Peroxides*] sales and provided the participants with the relevant statistics," "acted as a moderator in case of tensions between members of the agreement and encouraged the parties to find compromises," and "organised the auditing of the data submitted by the parties."

its competitor while not wishing to reveal its own. Moreover, even when a consulting firm is facilitating the monitoring, cartel members can submit misleading data or otherwise try to manipulate the information that the consulting firm will provide. Since each cartel member does not directly observe their competitor's interaction with the consulting firm, it cannot always be sure whether the data provided by the consulting firm is reliable. Thus, each firm may have an incentive to learn about the monitoring choices and the information of its competitor, as well as whether there has been a deviation in the stage game (i.e. secret price-cutting). Our model of endogenous monitoring aims to capture, in a reduced form way, the above features.

We consider a model where the players can choose not only their stage game actions, but also the monitoring structure in each period. Specifically, each player observes a private signal that is possibly informative about the other player's stage game action (and signal). In the standard model, the distribution of the signals is an exogenous function of the stage game actions. From the perspective of one player, the joint distribution of the signals depends on the stage game action chosen by the other player. One can imagine, as in the previous paragraph, that the players can take several actions that influence this dependence; as a reduced form representation, we simply allow each player to choose directly a distribution over the set of signal profiles for each stage game action of his opponent. For example, letting each player's stage game action set be $\{C, D\}$ and signal set be $\{c, d\}$, player 1's chosen monitoring structure could put probability 1 on signal profile (c, c) if player 2's stage game action is C and on signal profile (d, c) if player 2's stage game action is D, in which case player 2 always observes signal c (i.e. the second coordinate of the signal profile is always c) and player 1 observes c if and only if player 2 chooses C.

To capture the idea that there may be several actions available to the players that can affect the monitoring structure, and because we do not wish to rule out any such action by assumption, we allow the players to choose *any* monitoring structure. Thus, it is natural for players to disagree; indeed, as in the above example, each player may want his own private signal to perfectly reveal his opponent's stage game action but his opponent's signal not to reveal anything about his own stage game action. Due to this possibility of disagreement, we will need to specify how the monitoring structure that actually determines the players' signals is obtained from the monitoring structures chosen by the players. It is important to emphasize that in our formalization, each player observes only his own stage game action, choice of monitoring structure and private signal; in particular, neither the monitoring structure chosen by his opponent nor the realized monitoring structure that actually determines the players' signals is observed.[2]

How the monitoring structure that actually determines the players' signals is obtained from the monitoring structures chosen by the players is crucial for determining the equilibrium outcomes of the game. Indeed, if the monitoring structure is always the one chosen by a given player, then repetition of stage Nash is the only equilibrium outcome. This happens because the stage game action of such player's opponent does not depend on his own stage game action, hence his stage game action must be the strictly dominant D; given this, his opponent must play D as well.

It then follows that any degree of cooperation can be an equilibrium outcome only if the monitoring structure reflects the monitoring structures chosen by all the players. We allow for this by specifying that the monitoring structure is determined by an aggregation function that depends on the monitoring structures chosen by the players. In the context of the cartel example, one can interpret the data each cartel member submits to the consulting firm plus any other covert interactions as its choice of monitoring structure; the aggregation function then represents the audit and the subsequent recommendation of the consulting firm (as well as the result of the covert interactions).[3]

[2]Our formalization aims to capture some intuitive features of real-world designs. Indeed, the meetings that people have or reports that they write correspond to the choice of monitoring structures by the players. While some aspects of such choices may be known by everyone (e.g. that a meeting has been arranged or a report commissioned), the exact monitoring choices of each player (e.g. how much to reveal or conceal in a report, how much information to acquire before a meeting) and the monitoring structure that results from these interactions is often not known by those involved in it. As we have argued, all of these features are natural possibilities in cartels.

[3]According to the European Court decision in *Low density polyethylene* (as quoted in Marshall

Our main result is that the folk theorem holds whenever the aggregation function is *responsive*, which means, roughly, that whenever each player proposes for a signal profile to occur with probability 1, any signal profile other than the two proposed occurs with probability 0, and whenever a player proposes that a signal profile should occur with probability 1, it occurs with strictly positive probability. In the context of the cartel example, if conditional on firm 2 having deviated, firm 1 would like the consulting firm to reveal this information while firm 2 requests that it does not, property 1 requires that the consulting firm does not incorrectly announce that it was in fact firm 1 who deviated and property 2 requires that the consulting firm does in fact reveal firm 2's deviation with strictly positive probability.

The intuition behind this result is as follows. We specify a sequence of stage game actions to achieve the desired payoff and a monitoring strategy such that on the equilibrium path, each player observes the signal c with probability 1. On the other hand, if a player deviates from the specified stage game action, his opponent will receive the signal d with strictly positive probability. Thus, having received a d signal, a player is sure that his opponent has deviated. A responsive aggregation function ensures that there exists a strategy with such properties.

According to our strategy, when a player is sure that his opponent is not on the equilibrium path, he plays D forever. In general, this may not be optimal since, for example, even when player 1 is sure that player 2 has deviated, if player 2 assigns low probability to being caught, then player 2, and hence player 1, may prefer to continue cooperating. On the other hand, we exploit the fact that there exist deviations where the deviating player is sure that his deviation is detected. By making such deviations

and Marx (2012)): "Fides is an industry-wide statistical service run by a Zurich-based accounting firm. Subscribing producers supply each month individual data on their production, sales and stock movements to the central office which collates the information from the different producers and draws up global anonymized statistics for the Western European market. From these each producer can determine its own market share but not those of competitors. The system contains confidentiality safeguards but there is nothing to prevent competitors exchanging detailed information themselves in some other forum. The official Fides totals could then be used, as was envisaged, to check the accuracy of the figures exchanged by the producers."

the most likely tremble, whenever a player assigns probability zero to his opponent being on the equilibrium path, he also believes that his opponent assigns probability zero to him being on the equilibrium path.

The above implies that it is optimal for each player to play D forever following any history where there is zero probability that the other player is still on the equilibrium path. In particular, it is optimal to play D forever after receiving a d signal, and this is then sufficient to deter deviations from the equilibrium path. For histories other than the ones on the equilibrium path and the ones where there is zero probability of the other player being on the equilibrium path, we rely on a fixed point argument to show that there exists optimal continuation play at all such histories.[4]

The paper is organized as follows. Section 2 reviews the literature. In Section 3 we present a two-period repeated game to illustrate our main result in a simpler setting. The infinitely repeated prisoner's dilemma that we consider, as well as our results, are in Section 4. Section 5 contains some concluding remarks. The proof of our results are in the Appendix (Section 6). Supplementary material to this paper containing a stronger result for the case where cooperation is to be sustained in every period is available online.

2 Literature review

Various folk theorems have been established under progressively weaker assumptions on the monitoring structure, e.g. Fudenberg and Maskin (1986) with perfect monitoring, Fudenberg, Levine, and Maskin (1994) with imperfect monitoring, Hörner and Olszewski (2006) with private almost-perfect monitoring and Sugaya (2022) with private monitoring; for the particular case of private monitoring in the prisoner's dilemma, see also Sekiguchi (1997), Bhaskar and Obara (2002), Piccione (2002) and

[4]We thank an anonymous referee for suggesting the use of a fixed point argument to us. For the special case of (C, C) in every period and under additional assumptions on the aggregation function, it is possible to dispense with such fixed point argument and obtain an explicitly specified strategy in which players do not mix. Such strategy has some similarities to the ones used by Sekiguchi (1997) and Bhaskar and Obara (2002); see the supplementary material to this paper for details.

Ely and Valimaki (2002) for the case of almost-perfect monitoring, and Matsushima (2004) and Yamamoto (2012) for the case of conditionally independent but not necessarily almost-perfect monitoring. Unlike these papers, we do not take the monitoring structure as given, but instead we allow the players to choose the monitoring structure in each period.

Our model is a special case of a general repeated game with private monitoring.[5] Thus, one could try to obtain our main result by applying Sugaya's (2022) folk theorem to such repeated game. However, the assumptions of the latter do not hold when there are many stage game actions relative to signals. Such situations arise naturally in our model, where the players can choose from a large set of monitoring structures.

Our paper is also related to an important paper by Miyagawa, Miyahara, and Sekiguchi (2008) who prove a folk theorem for repeated games when players have to option to pay for accurate information about their opponents' actions. Thus, they also endogenize the monitoring structure in a repeated game. Differently to us, they assume that each player can decide whether to pay a cost in order to perfectly observe the other players' stage game actions, but observational decisions cannot be observed at any cost. Thus, their model captures the idea that perfect information can be obtained if enough resources are spent acquiring that information. The assumption that observational decisions cannot be observed makes it difficult to motivate players to pay for information; nevertheless, they show that a folk theorem holds generally in their setting. However, it leaves open the question why players are not able to learn about observational decisions. More generally, just as perfect information can be acquired given sufficient effort, players may also exert effort to hide information from their opponents, and the resulting monitoring structure will presumably depend

[5]A general repeated game with private monitoring consists of, for each player $i \in \{1,2\}$, a stage-game action set $\hat{S}_i$, a set of private signals $\hat{Y}_i$, a distribution function $\hat{\gamma} : \hat{S}_1 \times \hat{S}_2 \to \Delta(\hat{Y}_1 \times \hat{Y}_2)$ and a utility function $\hat{u}_i : \hat{S}_1 \times \hat{S}_2 \to \mathbb{R}$. Our model can be embedded by letting $\hat{S}_i = R_i \times S_i$, where R_i and S_i are the set of monitoring and stage game actions from our model, letting $\hat{Y}_i = Y_i$ where Y_i is our set of signals, letting the distribution $\hat{\gamma} : (R_1 \times R_2) \times (S_1 \times S_2) \to \Delta(Y_1 \times Y_2)$ be given by $\hat{\gamma}(r,s) = \alpha(r_1^{s_2}, r_2^{s_1})$, where α is our aggregation function, and by letting the utility function $\hat{u}_i : (R_1 \times R_2) \times (S_1 \times S_2) \to \mathbb{R}$ be given by $\hat{u}_i(r,s) = u_i(s)$ where u_i is our utility function.

on the interaction between these decisions. Our model attempts to capture some of these features omitted from Miyagawa, Miyahara, and Sekiguchi (2008) and thus complements their work.

In contrast to Miyagawa, Miyahara, and Sekiguchi (2008), in our model each player can choose any joint distribution of the signal profile for each stage game action of his opponent; thus, each private signal provides information about the stage game action, the monitoring action and the private signal of the other player. We view this as a reduced form model of *unrestricted* information acquisition, where players can learn about and manipulate each other's information as well as learning about the stage game action. On the other hand, we assume that monitoring is *costless*; indeed, what distinguishes stage game actions from monitoring actions in our model is that monitoring actions do not affect payoffs. Our justification is that in many situations of interest, the costs associated with monitoring are small compared to the benefits from successful cooperation in the stage game; thus, it makes sense to consider the idealized situation where these costs are zero. In addition, if the monitoring costs are large, then our view is that the monitoring actions should then be modelled as part of the stage game.

More generally, for a given stage game, it is important to understand how the freedom to choose different information structures affects the ability to provide long run incentives. For example, we might hope to gain an insight into which information structures are conducive to long run cooperation, abstracting away from any costs associated with the information choices. Two features of the strategy that we use to establish our folk theorem seem likely to play a role in more general constructions: First, on the equilibrium path, each player wishes to receive a signal that is perfectly informative about the stage game action of her opponent. Second, when in the punishment phase, each player proposes that her opponent receives a signal that perfectly reveals the information that she is in the punishment phase.

3 Motivating example

We illustrate our setting and results in a two-period repeated game (based on a similar example in Mailath and Samuelson (2006)) in which the first period is a prisoner's dilemma and the second period is a coordination game.

1\2	C	D
C	$2, 2$	$-1, 3$
D	$3, -1$	$0, 0$

1\2	A	B
A	$3, 3$	$0, 0$
B	$0, 0$	$1, 1$

Figure 1: Stage games

In period 1, each player chooses a monitoring action $r_i \in R_i$ and a stage game action from the prisoner's dilemma on the left of Figure 1. Then each player observes a private signal $y_i \in Y_i = \{c, d\}$. The joint distribution of the private signals (y_1, y_2) is given by $\gamma : (R_1 \times R_2) \times \{C, D\}^2 \to \Delta(Y_1 \times Y_2)$ and depends on the monitoring and stage game actions in period 1 in a way that we will specify in the next paragraph. In period 2, the coordination game on the right of Figure 1 is played. The repeated game payoff is the discounted sum of the payoffs from the two stage games, i.e. the players' payoff is $u = (1 - \delta)u^1 + \delta u^2$, where u^t is the payoff at period t, $t = 1, 2$.

The monitoring action of each player corresponds to a pair of conditional distributions $r_i = (r_i^C, r_i^D)$, where $r_i^C \in \Delta(Y_1 \times Y_2)$ is the distribution that player i wants when $s_{-i} = C$ and $r_i^D \in \Delta(Y_1 \times Y_2)$ is the distribution that player i wants when $s_{-i} = D$.[6] Thus, $R_i \subseteq \Delta(Y_1 \times Y_2) \times \Delta(Y_1 \times Y_2)$.[7] These choices are combined as

[6]The choice of the monitoring action r_i can be interpreted as choosing a statistical test that provides information to player i about the action of player $-i$ through differences in the distributions r_i^C and r_i^D, in the same way that the distribution of outcomes of a medical test depends on whether or not the person being tested is healthy or not. Note, however, that player i's signal also provides information about player $-i$'s signal and vice versa.

[7]The reason we do not let $R_i = \Delta(Y_1 \times Y_2) \times \Delta(Y_1 \times Y_2)$ is to avoid technical difficulties arising when R_i is infinite. As Myerson and Reny (2020) show, defining sequential equilibrium in infinite games is complicated and we get around this problem by focusing on any finite subset $X \subseteq \Delta(Y_1 \times Y_2)$ containing the degenerate distributions and letting $R_i = X^2$. This makes the games we consider finite

follows:

$$\gamma(y_1, y_2|r_1, r_2, s_1, s_2) = \beta r_1^{s_2}(y_1, y_2) + (1-\beta) r_2^{s_1}(y_1, y_2)$$

for each $y_i \in Y_i$, $r_i \in R_i$, $s_i \in \{C, D\}$, and $i \in \{1, 2\}$, where β represents who controls the information. An interpretation of γ is as follows: Each player chooses a monitoring structure and one of the two realizes, the one of player 1 with probability β and that of player 2 with probability $1-\beta$. If the monitoring structure of player i realizes, then a signal profile y is drawn according to $r_i^{s_{-i}}$.

If $\beta = 1$, then player 1 plays D with probability 1 in any sequential equilibrium of the repeated game. Indeed, if player 1 plays C with a strictly positive probability, then his continuation payoff after choosing C must be greater than his continuation payoff after choosing D. But the two are actually the same and equal to

$$\sum_{y_1} \sum_{r_2, s_2, y_2} \sigma_2(r_2, s_2) r_1^{s_2}(y_1, y_2) u_1(a_1(y_1), a_2(r_2, s_2, y_2))$$

where $\sigma_2(r_2, s_2)$ is the probability that player 2 chooses (r_2, s_2) in period 1, $a_2(r_2, s_2, y_2)$ is player 2's (mixed) strategy in period 2 at player 2's private history $h_2 = (r_2, s_2, y_2)$ and $a_1(y_1)$ solves

$$\max_{a_1 \in \{A, B\}} \sum_{r_2, s_2, y_2} \frac{\sigma_2(r_2, s_2) r_1^{s_2}(y_1, y_2)}{\sum_{r_2', s_2', y_2'} \sigma_2(r_2', s_2') r_1^{s_2'}(y_1, y_2')} u_1(a_1, a_2(r_2, s_2, y_2)).$$

In words, nothing in period 2 depends on the choice of s_1, hence, s_1 must be D since it is strictly dominant in period 1's stage game.

Similarly, if $\beta = 0$, then player 2 plays D with probability 1 in any sequential equilibrium of the repeated game. In contrast, if $\beta \in (0, 1)$, then cooperation in the first period is an equilibrium outcome. Formally, *for each* $\beta \in (0, 1)$*, there exists* $\delta^* \in (0, 1)$ *such that, for each* $\delta > \delta^*$ *and each finite subset* X *of* $\Delta(Y_1 \times Y_2)$ *containing* 1_y *for each* $y \in Y_1 \times Y_2$*, there exists a sequential equilibrium where* (C, C) *is played in the first period when* $R_i = X^2$ *for each* $i \in \{1, 2\}$*.*[8]

and we show in Section 4 that our main result holds uniformly, namely for any strictly individually rational and feasible payoff v, if players are sufficiently patient, then for each such finite subset of monitoring actions, there exists a sequential equilibrium with payoff v.

[8] For $y \in Y_1 \times Y_2$, 1_y denotes the probability measure degenerate on y.

We establish the above claim in what follows. Let $\beta \in (0,1)$. For convenience, let $\beta_1 = \beta$ and $\beta_2 = 1-\beta$. Define $\delta^* = \max\left\{\frac{1}{4(1-\beta)+1}, \frac{1}{4\beta+1}\right\}$. Let $\delta > \delta^*$, let X be a finite subset of $\Delta(Y_1 \times Y_2)$ containing 1_y for each $y \in Y_1 \times Y_2$, and let $R_i = X^2$ for each $i \in \{1,2\}$.

We use the following notation: for each $\rho \in \Delta(Y_1 \times Y_2)$, let $\rho_{Y_i} = \sum_{y_{-i}} \rho(\cdot, y_{-i})$ denote the *marginal of ρ on Y_i*; in particular, $r_{i,Y_i}^{s_{-i}}$ is the marginal of $r_i^{s_{-i}}$ on Y_i and $1_{(y_i,y_{-i}),Y_i}$ is the marginal of $1_{(y_i,y_{-i})}$ on Y_i.

An assessment specifies an action and a belief for each (private) history of each player. For each $i \in \{1,2\}$, player i's history h_i is either empty (in period 1) or of the form (r_i, s_i, y_i) (in period 2). Thus, we just need to specify beliefs at the latter histories and these are of the form $\mu(r_{-i}, s_{-i}, y_{-i}|h_i)$.

An intuitive description of the strategy is as follows. Each player cooperates in period 1 and chooses a monitoring structure r_i^* that perfectly reveals the stage game action of his opponent and yields signal c to the opponent. In period 2, each player plays A if and only if he believes that the opponent is still on the equilibrium path with sufficiently high probability.

Formally, the strategy σ_i is as follows: If h_i is the empty history, then $\sigma_i(h_i) = (r_i^*, s_i^*)$ with $s_i^* = C$ and $r_i^{*,C} = 1_{(c,c)}$ and $r_i^{*,D} = 1_{(d,c)}$. Otherwise, i.e. in period 2,

$$\sigma_i(h_i) = \begin{cases} A & \text{if } y_i = c, s_i = C \text{ and } 3\beta_i r_i^C(c,c) + 3(1-\beta_i) \geq \beta_i r_i^C(c,d), \\ A & \text{if } y_i = c, s_i = D \text{ and } 3\beta_i r_i^C(c,c) \geq \beta_i r_i^C(c,d) + 1 - \beta_i, \\ A & \text{if } y_i = d, r_{i,Y_i}^C(d) > 0 \text{ and } 3r_i^C(d,c) \geq r_i^C(d,d), \\ B & \text{otherwise.} \end{cases}$$

Given the beliefs (to be defined), each player will play A at exactly those histories where $\mu(r_{-i}^*, C, c|h_i) \geq 1/4$.

For histories $h_i = (r_i, s_i, y_i)$ that are reached with strictly positive probability if player i plays (r_i, s_i) and player $-i$ plays according to the strategy (i.e. r_{-i}^* and C), beliefs are determined via Bayes' rule. This corresponds to (1) below. In contrast, if player i chooses r_i such that $r_{i,Y_i}^C(d) = 0$ (e.g. r_i^*), then $y_i = d$ happens with zero probability when player $-i$ follows his strategy. In this case, if $y_i = d$ does happen, we specify that player i believes that player $-i$ has deviated to $s_{-i} = D$ and $r_{-i} = 1_{(c,d)}$,[9] i.e. player i believes that player $-i$ has played D and chosen signal c for himself and d for player i. The usefulness of this deviation is that if player $-i$ plays $(D, 1_{(c,d)})$, then regardless of the signal he receives, he will assign zero probability to player i being on the equilibrium path and will therefore play B in the second period according to the strategy.

Formally, the beliefs are as follows. If either $y_i = c$ or $y_i = d$ and $r^C_{i,Y_i}(d) > 0$,[10]

$$\mu(r_{-i}, s_{-i}, y_{-i}|h_i) = \begin{cases} \frac{\beta_i r_i^C(y_i,c)+(1-\beta_i)1_{(c,s_i)}(y_i,c)}{\beta_i r_{i,Y_i}^C(y_i)+(1-\beta_i)1_{(c,s_i),Y_i}(y_i)} & \text{if } r_{-i} = r^*_{-i}, s_{-i} = C, y_{-i} = c, \\ \frac{\beta_i r_i^C(y_i,d)+(1-\beta_i)1_{(c,s_i)}(y_i,d)}{\beta_i r_{i,Y_i}^C(y_i)+(1-\beta_i)1_{(c,s_i),Y_i}(y_i)} & \text{if } r_{-i} = r^*_{-i}, s_{-i} = C, y_{-i} = d, \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

Otherwise:

$$\mu(r_{-i}, s_{-i}, y_{-i}|h_i) = \begin{cases} \frac{\beta_i r_i^D(d,c)+1-\beta_i}{\beta_i r_{i,Y_i}^D(d)+1-\beta_i} & \text{if } r_{-i} = 1_{(c,d)}, s_{-i} = D, y_{-i} = c, \\ \frac{\beta_i r_i^D(d,d)}{\beta_i r_{i,Y_i}^D(d)+1-\beta_i} & \text{if } r_{-i} = 1_{(c,d)}, s_{-i} = D, y_{-i} = d, \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

The beliefs satisfy $\text{supp}(\mu(\cdot|h_i)) \subseteq \{(r^*_{-i}, C, c), (r^*_{-i}, C, d), (1_{(c,d)}, D, c), (1_{(c,d)}, D, d)\}$ and the strategy in period 2 specifies A at exactly those histories where $\mu(r^*_{-i}, C, c|h_i) \geq 1/4$. Moreover, $\sigma_{-i}(r^*_{-i}, C, c) = A$, $\sigma_{-i}(r^*_{-i}, C, d) = B$, $\sigma_{-i}(1_{(c,d)}, D, c) = B$ and $\sigma_{-i}(1_{(c,d)}, D, d) = B$, so in fact player i will play A if and only if he believes that the player $-i$ will play A with probability at least 1/4, as required by sequential rationality in period 2. In period 1, any deviation where D is played results in the opponent receiving a d signal — and thus playing B in period 2 — with strictly positive probability. For sufficiently high δ such deviation is not profitable.

In the remainder of this section we show formally that the above assessment is a sequential equilibrium.

[9] We write $r_{-i} = 1_{(y_{-i},y_i)}$ to mean that player $-i$ chooses $r^C_{-i} = r^D_{-i} = 1_{(y_{-i},y_i)}$, i.e. regardless of the stage game action of player i, player $-i$ would like the signal profile (y_{-i}, y_i) to be realized with probability 1.

[10] With an abuse of notation, we write $1_{(c,s_i)}$ to mean $1_{(c,c)}$ when $s_i = C$ and $1_{(c,d)}$ when $s_i = D$.

Consistency of beliefs: For each $i \in \{1,2\}$, let $\{\sigma_i^k\}_k$ be a sequence of player i's first period strategy such that $\sigma_i^k(r_i^*, C) \to 1$, $\sigma_i^k(1_{(c,d)}, D) = 1/k$, and $\sigma_i^k(r_i, s_i) = 1/k^2$ for each $(r_i, s_i) \notin \{(r_i^*, C), (1_{(c,d)}, D)\}$.

Fix $i \in \{1,2\}$. If $y_i = c$, then $\sum_{y_{-i}} \gamma(c, y_{-i}|r_i, r_{-i}^*, s_i, C) \geq 1-\beta_i > 0$; if $y_i = d$ and $r_{i,Y_i}^C(d) > 0$, then $\sum_{y_{-i}} \gamma(d, y_{-i}|r_i, r_{-i}^*, s_i, C) \geq \beta_i r_{i,Y_i}^C(d) > 0$. Thus, in these cases, i's beliefs are determined by Bayes' rule and $\mu(r_{-i}, s_{-i}, y_{-i}|h_i)$ is given by (1).

If $y_i = d$, $s_i \in \{C, D\}$ and r_i is such that $r_{i,Y_i}^C(d) = 0$, then

$$\gamma(d, y_{-i}|r, s) = \begin{cases} \beta_i r_i^C(d, y_{-i}) = 0 & \text{if } y_{-i} \in \{c,d\}, s_{-i} = C, r_{-i} = r_{-i}^*, \\ \beta_i r_i^D(d, c) + 1 - \beta_i & \text{if } y_{-i} = c, s_{-i} = D, r_{-i} = 1_{(c,d)}, \\ \beta_i r_i^D(d, d) & \text{if } y_{-i} = d, s_{-i} = D, r_{-i} = 1_{(c,d)}. \end{cases}$$

Hence, $\mu(r_{-i}, s_{-i}, y_{-i}|h_i)$ is given by (2).

Sequential rationality in period 2: First, we show that

$$\sigma_i(h_i) = \begin{cases} A & \text{if } h_i = (r_i^*, C, c), \\ B & \text{if } h_i = (r_i^*, C, d), \\ B & \text{if } h_i = (1_{(c,d)}, D, c), \\ B & \text{if } h_i = (1_{(c,d)}, D, d). \end{cases}$$

Indeed, if $h_i = (r_i^*, C, c)$, then $3\beta_i r_i^{*,C}(c,c) + 3(1-\beta_i) = 3\beta_i 1_{(c,c)}(c,c) + 3(1-\beta_i) = 3 > 0 = \beta_i 1_{(c,c)}(c,d) = \beta_i r_i^{*,C}(c,d)$ and, hence, $\sigma_i(r_i^*, C, c) = A$. If $h_i = (r_i^*, C, d)$, then $r_{i,Y_i}^{*,C}(d) = 0$ and, hence, $\sigma_i(r_i^*, C, d) = B$. If $h_i = (1_{(c,d)}, D, c)$, then $3\beta_i 1_{(c,d)}(c,c) = 0 < 1 = \beta_i 1_{(c,d)}(c,d) + 1 - \beta_i$ and, hence, $\sigma_i(1_{(c,d)}, D, c) = B$. Finally, if $h_i = (1_{(c,d)}, D, d)$, then $1_{(c,d),Y_i}(d) = 0$ and, hence, $\sigma_i(1_{(c,d)}, D, d) = B$.

For each h_i such that $y_i = c$ or $y_i = d$ and $r_{i,Y_i}^C(d) > 0$, $\text{supp}(\mu(\cdot|h_i)) \subseteq \{r_{-i}^*\} \times \{C\} \times Y_{-i}$. Given that when $r_{-i} = r_{-i}^*$ and $s_{-i} = C$, player $-i$ will play A in period 2 if and only if $y_{-i} = c$, it follows that $\sigma_i(h_i)$ is optimal because player i plays A if and only if $3\mu(r_{-i}^*, C, c|h_i) \geq \mu(r_{-i}^*, C, d|h_i)$ and this condition is equivalent to $3\beta_i r_i^C(c,c) + 3(1-\beta_i) \geq \beta_i r_i^C(c,d)$ when $y_i = c$ and $s_i = C$, to $3\beta_i r_i^C(c,c) \geq \beta_i r_i^C(c,d) + 1 - \beta_i$ when $y_i = c$ and $s_i = D$, and to $3r_i^C(d,c) \geq r_i^C(d,d)$ when $y_i = d$ and $r_{i,Y_i}^C(d) > 0$.

For each h_i such that $y_i = d$ and $r^C_{i,Y_i}(d) = 0$, $\text{supp}(\mu(\cdot|h_i)) \subseteq \{1_{(c,d)}\} \times \{D\} \times Y_{-i}$ and, thus, player i believes that player $-i$ will play B. Since $\sigma_i(h_i) = B$, it follows that $\sigma_i(h_i)$ is optimal.

Sequential rationality in period 1: Following the strategy yields a payoff of $2(1-\delta) + 3\delta$ since, in period 2, $h_i = (r^*_i, C, c)$ for each i and, thus, both players play A.

Suppose that player i deviates to $(r_i, s_i) \neq (r^*_i, s^*_i)$. If $s_i = s^*_i$, then total payoff is at most $2(1-\delta) + 3\delta$ since the first period payoff is 2 and the second period payoff cannot exceed 3. If $s_i \neq s^*_i$, then $s_i = D$ and either $h_{-i} = (r^*_{-i}, C, c)$ or $h_{-i} = (r^*_{-i}, C, d)$. Thus, player $-i$ plays A in the former case and B in the latter case. The probability of the latter case is at least $1 - \beta_i$ since, given $s_i = D$ and r^*_{-i}, player $-i$ will receive a d signal with probability at least $1 - \beta_i$. Thus, player i's second period payoff is at most $3\beta_i + 1 - \beta_i < 3$ since $\beta_i < 1$. The payoff from the deviation is then at most $3(1-\delta) + (3\beta_i + 1 - \beta_i)\delta$. It follows from $\delta > \delta^*$ that $2(1-\delta) + 3\delta > 3(1-\delta) + (3\beta_i + 1 - \beta_i)\delta$ and, hence, the assessment is sequentially rational.

4 Repeated prisoner's dilemma with endogenous private monitoring

We extend the logic of the motivating example by providing conditions under which the folk theorem holds for the repeated prisoner's dilemma with endogenous private monitoring.

In general, a *repeated game with endogenous monitoring* consists of a set of players I, a stage game $G = (S_i, u_i)_{i \in I}$, a set of private signals Y_i for each player, a set of monitoring actions R_i for each player, a monitoring technology $\gamma : R \times S \to \Delta Y$, and a discount factor δ. We focus on the case where the stage game is the prisoner's dilemma in Figure 2, where $g, l > 0$, and $Y_i = \{c, d\}$ for each $i \in I$.[11]

In each period $t = 1, 2 \ldots$, each player chooses $r_{i,t} \in R_i$ and $s_{i,t} \in S_i$, and then

[11]Note that (C, C) is efficient only if $g \leq l$, although we do not need to assume this.

$1\backslash 2$	C	D
C	$1,1$	$-l, 1+g$
D	$1+g, -l$	$0,0$

Figure 2: Prisoner's dilemma

observes $y_{i,t} \in Y_i$. The joint distribution of $y_t \equiv \times_{i\in I} y_{i,t}$ is given by $\gamma(r_t, s_t)$, where $r_t \equiv \times_{i\in I} r_{i,t}$ and $s_t \equiv \times_{i\in I} s_{i,t}$. That is, the joint distribution of signals is determined by the monitoring and stage game actions through the monitoring technology. Besides his own stage game and monitoring actions, the private signal $y_{i,t}$ is the only variable that player i observes in period t; in particular, payoffs are not observed. Thus, a history for player i at the end of period t takes the form $h_i^t = (r_{i,\tau}, s_{i,\tau}, y_{i,\tau})_{\tau=1}^t$.

Let H_i^t be the set of all histories for player i in period t, with H_i^0 being the singleton set containing the empty history. Let $H_i = \cup_{t=0}^{\infty} H_i^t$ denote the set of player i's histories. A strategy for player i is $\sigma_i : H_i \to \Delta(R_i \times S_i)$. Repeated game payoffs are given by the discounted sum of stage game payoffs, i.e. the monitoring actions have no direct impact on payoffs.

As in the motivating example, each player's monitoring action is a pair of conditional distributions $r_i = (r_i^C, r_i^D)$ with $r_i^{s_{-i}} \in \Delta Y$ for each $i \in \{1,2\}$ and $s_{-i} \in S_{-i}$. Thus, $R_i \subseteq (\Delta Y)^2$ for each $i \in \{1,2\}$. Furthermore, the monitoring actions of the players are aggregated to determine the joint distribution of private signals in the sense that there is an aggregation function $\alpha : (\Delta Y)^2 \to \Delta Y$ such that, for each $r \in R$ and $s \in S$,

$$\gamma(r, s) = \alpha(r_1^{s_2}, r_2^{s_1}).$$

The above formalization allows, of course, for the possibility that *a player controls the monitoring structure*, in the sense that either $\alpha(\rho, \rho') = \rho$ for each $(\rho, \rho') \in (\Delta Y)^2$ or $\alpha(\rho, \rho') = \rho'$ for each $(\rho, \rho') \in (\Delta Y)^2$. This case is easy to analyze since then the stage game Nash equilibrium must be played in every period. Indeed, as in the motivating example, the stage game action of the player who controls the monitoring structure has no impact (holding his monitoring action fixed) on the continuation play of his opponent; hence such player must choose D in every period. Thus, the other player must choose D in every period as well.

Remark 1 *If a player controls the monitoring structure, then, in every sequential equilibrium,* (D, D) *is played in each period.*

When no player controls the monitoring structure, the stage game action s_i of each player may affect the distribution of signals since it determines the probability distribution $r_{-i}^{s_i}$ of his opponent and the aggregation function α may depend on the latter. In what follows we will focus on the case where this dependence of α on each of its coordinates is strong enough in the sense of properties 1 and 2 below.

These properties require the following notation:

(i) For each $(\rho, \rho') \in (\Delta Y)^2$ and $y \in Y$, $\alpha(\rho, \rho')[y]$ is the probability of signal profile y according to $\alpha(\rho, \rho')$.

(ii) Let α_i refer to α when $r_i^{s_{-i}}$ is the first argument, i.e. $\alpha_i(r_i^{s_{-i}}, r_{-i}^{s_i})[(y_i, y_{-i})] = \alpha(r_1^{s_2}, r_2^{s_1})[(y_1, y_2)]$.

(iii) If $r_i^{s_{-i}} = 1_{(y_i, y_{-i})}$ and $r_{-i}^{s_i} = 1_{(y'_{-i}, y'_i)}$, we write $\alpha_i(r_i^{s_{-i}}, r_{-i}^{s_i}) = \alpha_i(1_{(y_i, y_{-i})}, 1_{(y'_i, y'_{-i})})$. The reason for this notation is as follows. When player i chooses a degenerate distribution 1_y, we write (as we did in the motivating example) y by listing player i's signal first, i.e. $y = (y_i, y_{-i})$. Hence, for example, if player 1 chooses $1_{(d,c)}$ (i.e. signal d for himself and c for player 2) and player 2 chooses $1_{(c,d)}$ (i.e. c for himself and d for player 1), then both players are choosing the same distribution over Y. In the above notation, it is clear that both players are choosing the same distribution because the resulting distribution over Y is $\alpha_1(1_{(d,c)}, 1_{(d,c)})$ (or, equivalently, $\alpha_2(1_{(c,d)}, 1_{(c,d)})$).

We say that α is *responsive* if it satisfies:

1. $\alpha(1_y, 1_{y'})[\tilde{y}] = 0$ for each $y, y' \in Y$ and $\tilde{y} \notin \{y, y'\}$.

2. $\min_{\rho \in \Delta Y} \alpha_i(1_y, \rho)[y] > 0$ for each $y \in Y$ and $i \in \{1, 2\}$.

Note that property 2 implies that $\min_{\rho \in \Delta Y} \alpha_i(\rho, 1_y)[y] > 0$ for each $y \in Y$ and $i \in \{1, 2\}$.[12] In words, these properties require the following:

1. If each player proposes that a signal should occur with probability 1, then all signals other than the two proposed happen with probability 0; this property requires in a weak form that α respects the choice of the players on what they agree. In particular, when both players propose the same signal y, this property implies that y happens with probability 1.

2. If one player proposes that a signal y should occur with probability 1, then y happens with strictly positive probability no matter the choice of the other player; this property requires that no player can fully control the monitoring structure.

Properties 1 and 2 are natural and mild conditions to make sure that the aggregation function reflects the choices of both players.

An example of an aggregation function satisfying the above properties is the one considered in the motivating example: $\alpha_\beta(\rho, \rho') = \beta\rho + (1 - \beta)\rho'$. The aggregation function α_β also belongs to a class of aggregation functions that can be viewed as mixed extensions of their restriction to pure monitoring strategies. In general, we can obtain such aggregation function α as the mixed extension of its restriction to pairs of degenerate signal distributions as follows. A pure monitoring strategy is simply $y \in Y$ which is identified with 1_y. Thus, writing $\rho(y)$ for the probability assigned to 1_y for each $\rho \in \Delta Y$ and $y \in Y$, α is a *mixed extension* if, for each $(\rho, \rho') \in (\Delta Y)^2$,

$$\alpha(\rho, \rho') = \sum_{y} \sum_{y'} \rho(y)\rho'(y')\alpha(1_y, 1_{y'}).^{13}$$

For aggregation functions belonging to the class of mixed extensions, α is responsive if the following conditions hold:

---

[12]Indeed, $\alpha_i(\rho, 1_y)[y] = \alpha_{-i}(1_{(y_{-i}, y_i)}, \rho)[(y_{-i}, y_i)] \geq \min_{\rho' \in \Delta Y} \alpha_{-i}(1_{(y_{-i}, y_i)}, \rho')[(y_{-i}, y_i)] > 0$.

[13]This equality holds in the case of α_β and, hence, α_β is the mixed extension of its restrition to pure monitoring strategies. However, there are responsive aggregation functions that are not the mixed extension of its restrition to pure monitoring strategies. An example is as follows: Let $\succ$ be the linear order on Y defined by $(c, c) \succ (c, d) \succ (d, c) \succ (d, d)$ and define $\alpha(\rho, \rho') = \frac{1}{2} 1_{y^1(\rho)} + \frac{1}{2} 1_{y^1(\rho')}$ where for each $\rho \in \Delta Y$, $y^1(\rho) \in Y$ is such that $\rho(y^1) = \max_y \rho(y)$ and there is no $\tilde{y} \in Y$ such that

(a) $\alpha(1_y, 1_{y'})[\tilde{y}] = 0$ for each $y, y' \in Y$ and $\tilde{y} \notin \{y, y'\}$.

(b) $\alpha_i(1_y, 1_{y'})[y] > 0$ for each $y, y' \in Y$ and $i \in \{1, 2\}$.[14]

Whenever monitoring is responsive, we obtain a folk theorem for the repeated prisoner's dilemma with endogenous private monitoring. Let

$$V^* = \{u \in \text{co}(u(S)) : u_i > 0 \text{ for each } i = 1, 2\}$$

be the set of feasible and strictly individually rational payoffs and

$$\mathcal{X} = \{X : X \subseteq \Delta Y, X \text{ is finite and } 1_y \in X \text{ for each } y \in Y\}$$

be the collection of the finite subsets X of ΔY containing all the degenerate probability measures 1_y on Y.

Theorem 1 *If α is responsive, then, for each $v \in V^*$, there exists $\delta^* \in (0, 1)$ such that, for each $\delta \geq \delta^*$ and $X \in \mathcal{X}$, when $R_i = X^2$ for each $i \in \{1, 2\}$, there exists a sequential equilibrium σ and a sequence of stage game actions profiles $\{s_t^*\}_{t=1}^{\infty}$ such that $(1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} u(s_t^*) = v$ and, for each $t \in \mathbb{N}$, s_t^* is played in period t with probability 1 according to σ.*

The proof of Theorem 1 is constructive. Given a feasible and strictly individually rational payoff v, we now describe a strategy profile σ and beliefs μ such that (σ, μ) is a sequential equilibrium (for any X in the statement of theorem) whose payoff is v.

By Fudenberg and Maskin (1991), there exists a sequence $\{s_t^*\}_{t=1}^{\infty}$ such that $v_i = \sum_{t=1}^{\infty} \delta^{t-1} u_i(s_t^*)$ for each $i \in \{1, 2\}$ and the continuation payoff for each player from any time onwards is bounded away from zero. Our strategy profile will specify that each player i chooses the stage game action $s_{i,t}^*$ in period t and the monitoring action

$\rho(\tilde{y}) = \max_y \rho(y)$ and $\tilde{y} \succ y^1(\rho)$. Then α is responsive but it is not the mixed extension of its restriction to pure monitoring strategies. Indeed, if $\rho = \frac{1}{2} 1_{(c,c)} + \frac{1}{2} 1_{(d,d)}$, then $\alpha(\rho, \rho) = 1_{(c,c)}$ and $\sum_y \sum_{y'} \rho(y) \rho'(y') \alpha(1_y, 1_{y'}) = \frac{1}{2} 1_{(c,c)} + \frac{1}{2} 1_{(d,d)}$.

[14] Property (a) is exactly the same as property 1. Property 2 holds since it follows from (b) that $\alpha_i(1_y, \rho)[y] = \sum_{y'} \rho(y') \alpha_i(1_y, 1_{y'})[y] \geq \min_{y' \in Y} \alpha_i(1_y, 1_{y'})[y] > 0$; hence $\min_{\rho \in \Delta Y} \alpha_i(1_y, \rho)[y] > 0$.

$r^*_{i,t} = (r^{*,C}_{i,t}, r^{*,D}_{i,t})$ that gives, with probability one, signal c to i's opponent and, for player i, signal c if and only if $s_{-i} = s^*_{-i,t}$, i.e. $r^{*,s_{-i}}_{i,t} = 1_{(c,c)}$ if $s_{-i} = s^*_{-i,t}$ and $r^{*,s_{-i}}_{i,t} = 1_{(d,c)}$ if $s_{-i} \neq s^*_{-i,t}$; in this monitoring structure, a signal c for player i means that his opponent has chosen the correct action. By property 1, this implies that the signal profile will be (c, c) in every period on the equilibrium path. For each $t \geq 0$, let

$$h^{*,t}_i = \big((r^*_{i,1}, s^*_{i,1}, c), \ldots, (r^*_{i,t}, s^*_{i,t}, c)\big) \text{ and}$$
$$H^*_i = \{h^{*,t}_i : t \in \mathbb{N}_0\}.$$

Under our strategy, $h^{*,t}_i$ is a history where i has neither deviated nor detected a deviation; in particular, $h^{*,0}_i$ is player i's empty history. Thus, our strategy will recommend $r^*_{i,t+1}$ and $s^*_{i,t+1}$ at such histories.

The usual difficulty in private monitoring games is coordination failure, namely, in our context, that a player i may still assign strictly positive probability to $h^{*,t}_{-i}$ despite having deviated from $r^*_{i,k}$ or $s^*_{i,k}$ for some $1 \leq k \leq t$. Thus, even when player i observes a deviation and assigns zero probability to $h^{*,t}_{-i}$, he may be unsure about the continuation play of his opponent who may still assign strictly positive probability to $h^{*,t}_i$. In our construction, such coordination failure can be avoided in the set H^{*0}_i of histories where there is zero probability that player $-i$'s history is in H^*_{-i} as follows. First, we specify that each player i plays D following any $h_i \in H^{*0}_i$. Second, we specify consistent beliefs such that whenever player i's history is in H^{*0}_i, he believes that player $-i$'s history is in H^{*0}_{-i}. Thus, player i plays D and expects player $-i$ to play D at $h_i \in H^{*0}_i$. The set H^{*0}_i is easily characterized in terms of primitive elements of the game, namely

$$H^{*0}_i = \left\{ h_i \in H_i : \prod_{t=1}^{\ell(h_i)} \alpha_i(r^{s^*_{-i,t}}_{i,t}, r^{*,s_{i,t}}_{-i,t})[(y_{i,t}, c)] = 0 \right\},$$

where $\ell(h_i)$ is the length of history h_i; this characterization makes it clear that H^{*0}_i is absorbing in the sense that if a history h_i belongs to it, then so will any of its continuation histories. Thus, once player i observes $h_i \in H^{*0}_i$, he will play D forever and expects player $-i$ to play D forever too. Regarding his monitoring action, we specify that player i plays $1_{(c,d)}$ regardless of player $-i$'s stage game action.

The sequential equilibrium we use to establish Theorem 1 is belief-based and, in fact, beliefs plays an important role in our argument. A key property is that player i's beliefs at any history are concentrated on $H^*_{-i} \cup H^{*0}_{-i}$. In particular, this allows us to specify player i's strategy σ_i only at histories $H^*_i \cup H^{*0}_i$, as we have done, and use a fixed point argument to obtain the strategies at the remaining histories. Specifically, for each specification of both players' strategies ν at those remaining histories, we obtain a fully specified strategy σ^ν by combining ν with the specification of σ_i at $H^*_i \cup H^{*0}_i$ for each i. Moreover, we show that there is a belief system μ^ν such that

(a) (σ^ν, μ^ν) is consistent,

(b) $\operatorname{supp}(\mu^\nu(\cdot|h_i)) \subseteq H^{*0}_{-i}$ for each $i \in \{1,2\}$ and $h_i \in H^{*0}_i$,[15]

(c) $h^{*,\ell(h_i)}_{-i} \in \operatorname{supp}(\mu^\nu(\cdot|h_i)) \subseteq \{h^{*,\ell(h_i)}_{-i}\} \cup H^{*0}_{-i}$ for each $i \in \{1,2\}$ and $h_i \in H_i \setminus H^{*0}_i$, and

(d) the function $\nu \mapsto \mu^\nu$ is continuous.

Property (d) then implies that the correspondence that consists of the mixed actions that at histories h_i maximize i's payoff given beliefs $\mu^\nu(\cdot|h_i)$ and continuation strategy determined by σ^ν is continuous and, thus, has a fixed point ν^*.

The sequential equilibrium we use to establish Theorem 1 is then $(\sigma, \mu) = (\sigma^{\nu^*}, \mu^{\nu^*})$. Sequential rationally follows by construction at histories $h_i \in H_i \setminus (H^*_i \cup H^{*0}_i)$. At histories $h_i \in H^{*0}_i$, player i's beliefs are concentrated on H^{*0}_{-i}, which is absorbing and, thus, implies that player $-i$ will play D forever; thus sequential rationality also holds at $h_i \in H^{*0}_i$. The argument is more involved for histories $h_i \in H^*_i$ but sequential rationality also holds essentially because deviations are caught with strictly positive probability, in which case the continuation payoff will be equal to zero.

Properties (a)–(d) play an important role in our argument. The reason why they hold is roughly as follows. If the strategy σ_i (or σ^ν_i more generally) is followed, then

[15] For any probability measure π, $\operatorname{supp}(\pi)$ denotes the support of π.

the t-period histories of player i that are played with strictly positive probability are $\{h_i^{*,t}\} \cup H_i^D$, where

$$H_i^D = \{h_i \in H_i : h_i = (h_i^{*,t} \cdot (r^*_{i,t+1}, s^*_{i,t+1}, d) \cdot h'_i) \text{ for some } t \geq 0 \text{ and } h'_i \in H_i^C\} \text{ and}$$
$$H_i^C = \{h_i \in H_i : r_{i,t} = 1_{(c,d)} \text{ and } s_{i,t} = D \text{ for all } 1 \leq t \leq \ell(h_i)\}.$$

If $h_i \in H_i^t \setminus H_i^{*0}$, then $h_{-i}^{*,t}$ occurs with strictly positive probability and, thus, $\mu(\cdot|h_i)$ is determined by σ_{-i} and $h_{-i}^{*,t} \in \text{supp}(\mu(\cdot|h_i))$. Furthermore, since the histories h_{-i} that are played with strictly positive probability are in $\{h_{-i}^{*,t}\} \cup H_{-i}^D \subseteq \{h_{-i}^{*,t}\} \cup H_{-i}^{*0}$, it follows that $\text{supp}(\mu(\cdot|h_i)) \subseteq \{h_{-i}^{*,t}\} \cup H_{-i}^{*0}$.

At histories $h_i \in H_i^t \cap H_i^{*0}$, $h_{-i}^{*,t}$ occurs with probability zero. Hence, the only possibility for $\mu(\cdot|h_i)$ to be determined by σ_{-i} is for some history $h_{-i} \in H_{-i}^D$ to occur with strictly positive probability. The set of player i's histories where there is zero probability that player $-i$'s history belongs to H_{-i}^D is

$$H_i^{D0} = \Big\{h_i \in H_i : \text{ for all } 1 \leq n \leq \ell(h_i) \text{ and } (y_{-i,n+1}, \ldots, y_{-i,\ell(h_i)}) \in Y_{-i}^{\ell(h_i)-n},$$
$$\left(\prod_{k=1}^{n-1} \alpha_i(r_{i,k}^{s^*_{-i,k}}, r_{-i,k}^{*,s_{i,k}})[(y_{i,k}, c)]\right) \alpha_i(r_{i,n}^{s^*_{-i,n}}, r_{-i,n}^{*,s_{i,n}})[(y_{i,n}, d)] \left(\prod_{k=n+1}^{\ell(h_i)} \alpha_i(r_{i,k}^D, 1_{(d,c)})[y_k]\right)$$
$$= 0\Big\}.$$

Hence, if $h_i \in H_i^{*0} \setminus H_i^{D0}$, then $\text{supp}(\mu(\cdot|h_i)) \subseteq H_{-i}^{*0}$. Finally, for histories $h_i \in H_i^{*0} \cap H_i^{D0}$, beliefs are not determined by σ. Thus, we choose a strategy $\bar{\sigma}$ that players use to form beliefs in the case $h_i \in H_i^{*0} \cap H_i^{D0}$ to obtain that $\text{supp}(\mu(\cdot|h_i)) \subseteq H_{-i}^{*0}$. In particular, if player i observes an unexpected d signal in period t having been on the equilibrium path, we specify that the most likely tremble is that his opponent has chosen $\bar{\sigma}_{-i}(h_{-i}^{*,t-1}) = (1_{(c,d)}, \bar{s}_{-i})$, where $\bar{s}_{-i} \neq s^*_{-i,t}$. The usefulness of this particular deviation is that $h_{-i}^{*,t-1} \cdot (1_{(c,d)}, \bar{s}_{-i}, y_{-i,t}) \in H_{-i}^{*0}$ for each $y_{-i,t}$; thus $\text{supp}(\mu(\cdot|h_i)) \subseteq H_{-i}^{*0}$.

The proof in Section 6.1 establishes all these claims as well as the omitted elements in detail. In the supplementary material to this paper we show that, under a stronger notion of responsiveness for the aggregation function α, the outcome consisting of (C, C) in every period, and thus the payoff $(1, 1)$, can be sustained with a sequential equilibrium (σ, μ) such that σ is pure and explicitly specified, i.e. we dispense with the "black box" aspect of the fixed point argument we used above. The strategy is the same as above for histories $h_i \in H_i^* \cup H_i^{*0}$ and specifies, for histories $h_i \in H_i^t \setminus (H_i^* \cup H_i^{*0})$, (r_i^*, D) if $0 < \mu(h_{-i}^{*,t}|h_i) < \mu_i^*$, and (r_i^*, C) if $\mu(h_{-i}^{*,t}|h_i) \geq \mu_i^*$, where $\mu_i^* \in (0,1)$.[16]

We conclude this outline of the proof of our main result with a discussion of the role played by properties 1 and 2 in the definition of a responsive aggregation function. For simplicity, consider the case where mutual cooperation is to be sustained. Property 1 implies that on the equilibrium path, where each player i chooses stage game action C and monitoring action r_i^*, the signal profile (c, c) happens with probability 1. This, in turn, implies that player i assigns probability 1 to $h_{-i}^{*,t}$ when he observes $h_i^{*,t}$. In addition, property 1 also implies that if player 1 is in H_1^{*0} (and, thus, plays D and $1_{(c,d)}$) and player 2 is in $H_2 \setminus H_2^{*0}$ (and, thus, plays monitoring action r_2^*), then the signal profile $(y_1, y_2) = (c, d)$ happens with probability 1 since $\alpha_1(1_{(c,d)}, 1_{(c,d)}) = 1_{(c,d)}$.

The usefulness of property 2 can be seen by noting that deviations to D by some player on the equilibrium path will be detected with strictly positive probability by his opponent. Thus, if player 2 unilaterally deviates to stage game action D in period 1, with probability at least $\min_{\rho \in \Delta Y} \alpha_1(1_{(d,c)}, \rho)[(d, c)] > 0$, player 1 will observe signal d in period 1 and play D from period 2 onwards.

5 Concluding remarks

In this paper we have shown how the ability of players to design the monitoring structure of the game affects its equilibrium outcomes. This was shown in the context of the infinitely repeated prisoner's dilemma by contrasting the cases where one player controls the monitoring structure with the case where the monitoring structure is responsive to both players' monitoring choices. Indeed, in the former case, the stage game Nash equilibrium is played in every period in every equilibrium whereas, in the latter case, the folk theorem holds.

The extension of the above results to stage games other than the prisoner's

[16] Note that when $s_{-i,t}^* = C$ for all $t \in \mathbb{N}$, $r_{i,t}^*$ does not depend on t; hence we write r_i^* in this case.

dilemma is outside the scope of this paper. Nevertheless, the extension of our framework is straightforward as we now illustrate.

As already noted, a repeated game with endogenous monitoring consists of a set of players I, a stage game $G = (S_i, u_i)_{i \in I}$, a set of private signals Y_i for each player, a set of monitoring actions $R_i \subseteq (\Delta Y)^{|S_{-i}|}$ for each player, a monitoring technology $\gamma : R \times S \to \Delta Y$, and a discount factor δ. The interpretation is that each player chooses a profile of conditional distributions $r_i = (r_i^{s_{-i}^1}, \ldots, r_i^{s_{-i}^{|S_{-i}|}})$, where $r_i^{s_{-i}} \in \Delta Y$ is the distribution over signals that player i wants when the stage game actions of his opponents is s_{-i}.

In each period $t = 1, 2 \ldots$, each player chooses $r_{i,t} \in R_i$ and $s_{i,t} \in S_i$, and then observes $y_{i,t} \in Y_i$. The joint distribution of $y_t \equiv \times_{i \in I} y_{i,t}$ is given by $\gamma(r_t, s_t)$, where $r_t \equiv \times_{i \in I} r_{i,t}$ and $s_t \equiv \times_{i \in I} s_{i,t}$. The joint distribution of signals is determined by the monitoring and stage game actions through an aggregation function $\alpha : (\Delta Y)^I \to \Delta Y$ that determines the joint distribution of signals from the choices of the players:

$$\gamma(r_t, s_t) = \alpha(r_{1,t}^{s_{-1,t}}, \ldots, r_{|I|,t}^{s_{-|I|,t}})$$

Each player observes his own previous monitoring and stage game actions and private signals but nothing else. Thus, a history for player i at the end of period t takes the form $h_i^t = (r_{i,\tau}, s_{i,\tau}, y_{i,\tau})_{\tau=1}^t$. Letting H_i^t be the set of all histories for player i in period t, a strategy for player i is $\sigma_i : \cup_{t=0}^{\infty} H_i^t \to \Delta(R_i \times S_i)$. Repeated game payoffs are given by the discounted sum of stage game payoffs, i.e. the monitoring actions have no direct impact on payoffs.

We can then ask whether our results extend to such general repeated games with endogenous monitoring. The proof of Remark 1 easily extends to show that, for any stage game G, the player who controls the monitoring structure plays a myopic best-reply to his expectation of his opponents' strategy in every period and every equilibrium. When such player has a dominant strategy in a two-player game, such as the prisoner's dilemma, this implies that a stage game Nash equilibrium is played in every period and in every equilibrium. In general, we can ask: Is there a sharp characterization of equilibria in general repeated games with endogenous monitoring when one player controls the monitoring structure?

While the extension of Theorem 1 to a Nash-threat folk theorem for general games seems feasible, obtaining a folk theorem in full generality seems more difficult but we can ask: Is it the case that the folk theorem holds for any stage game and any responsive monitoring structure? The difficulty in answering this question lies in the extension of our method of proof of Theorem 1 to general repeated games with endogenous monitoring.

6 Appendix

6.1 Proof of Theorem 1

6.1.1 Parametrization

Let $v \in V^*$ and $0 < \varepsilon < \frac{\min_i v_i}{2}$. By Fudenberg and Maskin (1991), there exists $\bar{\delta} \in (0,1)$ such that, for each $\delta \geq \bar{\delta}$, there exists a sequence $\{s_t^*\}_{t=1}^{\infty} = \{(s_{1,t}^*, s_{2,t}^*)\}_{t=1}^{\infty}$ such that $(1-\delta)\sum_{t=1}^{\infty} \delta^{t-1} u(s_t^*) = v$ and for each $i \in \{1,2\}$ and $t \in \mathbb{N}$, $|(1-\delta)\sum_{j=1}^{\infty} \delta^{j-1} u_i(s_{t+j}^*) - v_i| < \varepsilon$.

For each $i \in \{1,2\}$, let $\underline{\alpha}_i = \min_{\rho \in \Delta Y, y \in Y} \alpha_i(\rho, 1_y)[y] > 0$. Then set

$$\delta^* = \max\left\{\bar{\delta}, \max_i \frac{\max\{g,l\}}{\max\{g,l\} + (v_i - \varepsilon)\underline{\alpha}_i}\right\}.$$

Let $\delta \geq \delta^*$, $X \in \mathcal{X}$ and $R_i = X^2$ for each $i \in \{1,2\}$ be fixed for the remainder of this proof.

6.1.2 The assessment

We specify only part of the assessment, the remaining part being obtained via a fixed point argument as detailed below.

For each $t \in \mathbb{N}$, $i \in \{1,2\}$ and $s_{-i} \in \{C,D\}$, let $r_{i,t}^{*,s_{-i}} = 1_{(c,c)}$ if $s_{-i} = s_{-i,t}^*$ and

$r^{*,s_{-i}}_{i,t} = 1_{(d,c)}$ if $s_{-i} \neq s^*_{-i,t}$. Define

$$h_i^{*,t} = \big((r^*_{i,1}, s^*_{i,1}, c), \ldots, (r^*_{i,t}, s^*_{i,t}, c)\big) \text{ for each } t \geq 0,$$
$$H_i^* = \{h_i^{*,t} : t \in \mathbb{N}_0\} \text{ and}$$
$$H_i^{*0} = \{h_i \in H_i : \prod_{t=1}^{\ell(h_i)} \alpha_i(r^{s^*_{-i,t}}_{i,t}, r^{*,s_{i,t}}_{-i,t})[(y_{i,t}, c)] = 0\}.$$

Note that $h_i^{*,0}$ is player i's empty history.

The **strategy** is partly specified as follows. Define, for each $h_i \in H_i$,

$$\sigma_i(h_i) = \begin{cases} (1_{(c,d)}, D) & \text{if } h_i \in H_i^{*0} \\ (r^*_{i,\ell(h_i)+1}, s^*_{i,\ell(h_i)+1}) & \text{if } h_i = h_i^{*,\ell(h_i)}, \end{cases}$$

where $\ell(h_i)$ is the length of history h_i. We also write $\sigma_i(r_i, s_i | h_i)$ for the probability that $\sigma_i(h_i)$ assigns to (r_i, s_i).

Let

$$\Theta = \{(i, h_i) : i \in \{1, 2\} \text{ and } h_i \in H_i \setminus (H_i^* \cup H_i^{*0}),$$
$$\Sigma_{i,h_i} = \Delta(X^2 \times \{C, D\}) \text{ for each } (i, h_i) \in \Theta,$$
$$\Sigma = \prod_{(i,h_i)\in\Theta} \Sigma_{i,h_i}$$

and, for each $\nu \in \Sigma$, σ^ν be the strategy defined by setting, for each $i \in \{1, 2\}$ and $h_i \in H_i$,

$$\sigma_i^\nu(h_i) = \begin{cases} \nu_i(h_i) & \text{if } h_i \in H_i \setminus (H_i^* \cup H_i^{*0}), \\ \sigma_i(h_i) & \text{otherwise.} \end{cases}$$

We will specify beliefs μ^ν for each $\nu \in \Sigma$ as follows.

Lemma 1 *For each $\nu \in \Sigma$, there exists a beliefs system μ^ν such that*

(a) (σ^ν, μ^ν) is consistent,

(b) $\text{supp}(\mu^\nu(\cdot|h_i)) \subseteq H^{*0}_{-i}$ for each $i \in \{1, 2\}$ and $h_i \in H_i^{*0}$,

(c) $h^{*,\ell(h_i)}_{-i} \in \text{supp}(\mu^\nu(\cdot|h_i)) \subseteq \{h^{*,\ell(h_i)}_{-i}\} \cup H^{*0}_{-i}$ for each $i \in \{1, 2\}$ and $h_i \in H_i \setminus H_i^{*0}$, and

(d) the function $\nu \mapsto \mu^\nu$ is continuous.

6.1.3 Proof of Lemma 1

Define a strategy $\bar{\sigma}$ by setting, for each $i \in \{1,2\}$ and $h_i \in H_i$,

$$\bar{\sigma}_i(h_i) = \begin{cases} (1_{(c,d)}, D) & \text{if } h_i \in H_i \setminus H_i^{*0} \text{ and } s^*_{i,\ell(h_i)+1} = C, \\ (1_{(c,d)}, C) & \text{if } h_i \in H_i \setminus H_i^{*0} \text{ and } s^*_{i,\ell(h_i)+1} = D, \\ (1_{(d,c)}, D) & \text{if } h_i \in H_i^{*0}. \end{cases}$$

Let $\hat{\sigma}$ be such that, for each $i \in \{1,2\}$ and $h_i \in H_i$, $\hat{\sigma}_i(h_i)$ is totally mixed. For each $j \in \mathbb{N}$, let σ^j be defined by setting, for each $i \in \{1,2\}$ and $h_i \in H_i$,

$$\sigma^j_i(h_i) = \left(1 - \frac{1}{j} - \frac{1}{j^j}\right) \sigma^\nu_i(h_i) + \frac{1}{j}\bar{\sigma}_i(h_i) + \frac{1}{j^j}\hat{\sigma}_i(h_i).$$

Then $\{\sigma^j\}_{j=1}^\infty$ is a sequence of totally mixed strategies converging to σ^ν.

Let $i \in \{1,2\}$, $t \in \mathbb{N}$ and $h_i = (r_{i,k}, s_{i,k}, y_{i,k})_{k=1}^t \in H_i^t$. Then, for each $h_{-i} = (r_{-i,k}, s_{-i,k}, y_{-i,k})_{k=1}^t \in H_{-i}^t$ and $j \in \mathbb{N}$,

$$\mu^j(h_{-i}|h_i) = \frac{\prod_{k=1}^t \alpha_i(r_{i,k}^{s_{-i,k}}, r_{-i,k}^{s_{i,k}})[y_k]\sigma^j_{-i}(r_{-i,k}, s_{-i,k}|h_{-i}^{k-1})}{\sum_{(\hat{r}_{-i,k},\hat{s}_{-i,k},\hat{y}_{-i,k})_{k=1}^t \in H_{-i}^t} \prod_{k=1}^t \alpha_i(r_{i,k}^{\hat{s}_{-i,k}}, \hat{r}_{-i,k}^{s_{i,k}})[(y_{i,k}, \hat{y}_{-i,k})]\sigma^j_{-i}(\hat{r}_{-i,k}, \hat{s}_{-i,k}|\hat{h}_{-i}^{k-1})}$$

where $h_{-i}^k = (r_{-i,n}, s_{-i,n}, y_{-i,n})_{n=1}^k$ and $\hat{h}_{-i}^k = (\hat{r}_{-i,n}, \hat{s}_{-i,n}, \hat{y}_{-i,n})_{n=1}^k$ for each $k \geq 0$. In Claims 2 – 4 below we show that $\{\mu^j(h_{-i}|h_i)\}_{j=1}^\infty$ converges. Thus, the **beliefs** μ^ν are defined by setting, for each $i \in \{1,2\}$, $h_i \in H_i$ and $h_{-i} \in H_{-i}^{\ell(h_i)}$,

$$\mu^\nu(h_{-i}|h_i) = \lim_j \mu^j(h_{-i}|h_i).$$

It then follows that (σ^ν, μ^ν) is consistent. Thus, part (a) of Lemma 1 follows.

The following notation is useful to describe the expression for $\mu^\nu(h_{-i}|h_i)$. For each strategy σ', $i \in \{1,2\}$, $h_i \in H_i$ and $h_{-i} \in H_{-i}^{\ell(h_i)}$, let

$$\pi_i^k(h_i, h_{-i}, \sigma') = \alpha_i(r_{i,k}^{s_{-i,k}}, r_{-i,k}^{s_{i,k}})[y_k]\sigma'_{-i}(r_{-i,k}, s_{-i,k}|h_{-i}^{k-1}) \text{ for each } 1 \leq k \leq \ell(h_i), \text{ and}$$

$$\pi_i(h_i, h_{-i}, \sigma') = \prod_{k=1}^{\ell(h_i)} \pi_i^k(h_i, h_{-i}, \sigma').$$

Then, with this notation, $\mu^j(h_{-i}|h_i) = \frac{\pi_i(h_i, h_{-i}, \sigma^j)}{\sum_{\hat{h}_{-i} \in H_{-i}^t} \pi_i(h_i, \hat{h}_{-i}, \sigma^j)}$ with $t = \ell(h_i)$. Furthermore, for each $h_i \in H_i^t$, $h_{-i} \in H_{-i}^t$ and $j \in \mathbb{N}$,

$$\mu^j(h_{-i}|h_i) = \frac{\mu^j(h_{-i}^{t-1}|h_i^{t-1})\pi_i^t(h_i, h_{-i}, \sigma^j)}{\sum_{\hat{h}_{-i} \in H_{-i}^{t-1}} \sum_{(\hat{r}_{-i,t},\hat{s}_{-i,t},\hat{y}_{-i,t}) \in H_{-i}^1} \mu^j(\hat{h}_{-i}|h_i^{t-1})\pi_i^t(h_i, \hat{h}_{-i} \cdot (\hat{r}_{-i,t}, \hat{s}_{-i,t}, \hat{y}_{-i,t}), \sigma^j)}.$$

Let

$$H_i^C = \{h_i \in H_i : r_{i,t} = 1_{(c,d)} \text{ and } s_{i,t} = D \text{ for all } 1 \leq t \leq \ell(h_i)\},$$
$$H_i^D = \{h_i \in H_i : h_i = (h_i^{*,t} \cdot (r^*_{i,t+1}, s^*_{i,t+1}, d) \cdot h'_i) \text{ for some } t \geq 0 \text{ and } h'_i \in H_i^C\}.$$

The set H_i^D contains player i's histories that can occur with strictly positive probability under σ_i^ν other than those in H_i^*. Note that $H_i^D \subseteq H_i^{*0}$. Indeed, $r_{i,t+1}^{s^*_{-i,t+1}} = 1_{(c,c)}$, $s_{i,t+1} = s^*_{i,t+1}$ and $y_{i,t+1} = d$ imply that

$$\alpha_i(r_{i,t+1}^{s^*_{-i,t+1}}, r_{-i,t+1}^{*,s_{i,t+1}})[(y_{i,t+1}, c)] = \alpha_i(1_{(c,c)}, 1_{(c,c)})[(d, c)] = 1_{(c,c)}(d, c) = 0$$

by property 1.

Claim 1 *For each $\nu \in \Sigma$, $i \in \{1, 2\}$ and $t \in \mathbb{N}$,*

$$\left\{ h_i \in H_i^t : \prod_{k=1}^{t} \sigma_i^\nu(r_{i,k}, s_{i,k} | h_i^{k-1}) > 0 \right\} \subseteq \{h_i^{*,t}\} \cup H_i^D.$$

Proof. Let h_i be such that $\prod_{k=1}^t \sigma_i^\nu(r_{i,k}, s_{i,k}|h_i^{k-1}) > 0$ and consider first the case where $y_{i,k} = c$ for each $1 \leq k \leq t$. Since $\sigma_i^\nu(h_i^{*,k}) = (r^*_{i,k+1}, s^*_{i,k+1})$ for each $0 \leq k \leq t-1$, it follows that $h_i = h_i^{*,t}$.

Hence, consider next the case where $y_{i,k} \neq c$ for some $1 \leq k \leq t$ and let $\hat{t} = \min\{1 \leq k \leq t : y_{i,k} = d\}$. Since $\sigma_i^\nu(h_i^{*,k}) = (r^*_{i,k+1}, s^*_{i,k+1})$ for each $0 \leq k \leq \hat{t} - 1$, it follows that $(r_{i,k}, s_{i,k}, y_{i,k}) = (r^*_{i,k}, s^*_{i,k}, c)$ for each $1 \leq k \leq \hat{t} - 1$ and $(r_{i,\hat{t}}, s_{i,\hat{t}}, y_{i,\hat{t}}) = (r^*_{i,\hat{t}}, s^*_{i,\hat{t}}, d)$. Then, for each $\hat{t} + 1 \leq k \leq t$, $h_i^{k-1} \in H_i^{*0}$ and $(r_{i,k}, s_{i,k}) = \sigma_i^\nu(h_i^{k-1}) = (1_{(c,d)}, D)$. Thus, $h_i \in H_i^D$. ■

Claim 2 *For each $i \in \{1, 2\}$, $t \in \mathbb{N}$ and $h_i \in H_i^t \setminus H_i^{*0}$:*

1. $\lim_j \mu^j(h_{-i}|h_i) = \frac{\pi_i(h_i, h_{-i}, \sigma^\nu)}{\sum_{\hat{h}_{-i} \in H^t_{-i}} \pi_i(h_i, \hat{h}_{-i}, \sigma^\nu)}$ *for each $h_{-i} \in H^t_{-i}$,*
2. $h^*_{-i} \in \text{supp}(\mu^\nu(\cdot|h_i))$,
3. $\text{supp}(\mu^\nu(\cdot|h_i)) \subseteq \{h^*_{-i}\} \cup H^{*0}_{-i}$ *and*
4. $\lim_j(j^{j-1-t}\mu^j(h_{-i}|h_i)) = 0$ *for each $h_{-i} \in H^t_{-i} \setminus (H^{*0}_{-i} \cup \{h^{*,t}_{-i}\})$.*

Proof. We have that $\pi_i(h_i, h^{*,t}_{-i}, \sigma^\nu) > 0$ since $h_i \in H^t_i \setminus H^{*0}_i$. This implies parts 1 and 2. For part 3, note that if $h_{-i} \in H^t_{-i} \setminus \{h^{*,t}_{-i}\}$ is such that $\mu^\nu(h_{-i}|h_i) > 0$, then $\prod_{k=1}^t \sigma^\nu_{-i}(r_{-i,k}, s_{-i,k}|h^{k-1}_{-i}) > 0$. Thus, by Claim 1, $h_{-i} \in H^D_{-i} \subseteq H^{*0}_{-i}$.

We establish part 4 by induction on t. Let $t = 1$ and consider $h_i = (r_i, s_i, y_i) \in H^1_i \setminus H^{*0}_i$. Since $h_i \notin H^{*0}_i$, we have that $\alpha_i(r_i^{s^*_{-i,1}}, r^{*,s_i}_{-i,1})[(y_i, c)] > 0$. In addition,

$$\sum_{\hat{h}_{-i} \in H^1_{-i}} \pi_i(h_i, \hat{h}_{-i}, \sigma^j) \to \sum_{y_{-i}} \alpha_i(r_i^{s^*_{-i,1}}, r^{*,s_i}_{-i,1})[y_i, y_{-i}].$$

Thus, $(r^*_{-i,1}, s^*_{-i,1}, c) \in \text{supp}(\mu^\nu(\cdot|h_i)) \subseteq \{(r^*_{-i,1}, s^*_{-i,1}, c), (r^*_{-i,1}, s^*_{-i,1}, d)\}$ and note that $(r^*_{-i,1}, s^*_{-i,1}, d) \in H^{*0}_{-i}$. Letting $\bar{s}_{-i} \neq s^*_{-i,1}$, we have that $\{(1_{(c,d)}, \bar{s}_{-i}, c), (1_{(c,d)}, \bar{s}_{-i}, d)\} \subseteq H^{*0}_{-i}$ since $\alpha_{-i}(1_{(c,d)}, 1_{(c,d)})[(y_{-i}, c)] = 0$ by property 1. Hence, for each $h_{-i} = (r_{-i}, s_{-i}, y_{-i}) \in H^1_{-i} \setminus (H^{*0}_{-i} \cup \{h^{*,1}_{-i}\})$ and $\sigma'_{-i} \in \{\sigma^\nu_{-i}, \bar{\sigma}_{-i}\}$, $\sigma'_{-i}(r_{-i}, s_{-i}|h^{*,0}_{-i}) = 0$ and

$$\lim_j (j^{j-2}\mu^j(h_{-i}|h_i)) = \lim_j \frac{j^{j-2}\pi_i(h_i, h_{-i}, \hat{\sigma})j^{-j}}{\sum_{\hat{h}_{-i} \in H^1_{-i}} \pi_i(h_i, \hat{h}_{-i}, \sigma^j)} = 0.$$

Let $t > 1$ and assume that we have established that, for each $k = 1, \ldots, t-1$ and $h_i \in H^k_i \setminus H^{*0}_i$, $\lim_j(j^{j-1-k}\mu^j(h_{-i}|h_i)) = 0$ for each $h_{-i} \in H^k_{-i} \setminus (H^{*0}_{-i} \cup \{h^{*,k}_{-i}\})$. Let $h_i \in H^t_i \setminus H^{*0}_i$ and $h_{-i} \in H^t_{-i} \setminus (H^{*0}_{-i} \cup \{h^{*,t}_{-i}\})$. We have that

$$\lim_j \sum_{\hat{h}_{-i} \in H^{t-1}_{-i}} \sum_{(\hat{r}_{-i,t}, \hat{s}_{-i,t}, \hat{y}_{-i,t}) \in H^1_{-i}} \mu^j(\hat{h}_{-i}|h^{t-1}_i)\pi^t_i(h_i, \hat{h}_{-i} \cdot (\hat{r}_{-i,t}, \hat{s}_{-i,t}, \hat{y}_{-i,t}), \sigma^j) > 0$$

because $h_i \in H^t_i \setminus H^{*0}_i$ and, thus, $\mu^\nu(h^{*,t-1}_{-i}|h^{t-1}_i) > 0$ and $\pi^t_i(h_i, h^{*,t}_i, \sigma^\nu) > 0$. Hence, if $h^{t-1}_{-i} \neq h^{*,t-1}_{-i}$, then $h^{t-1}_{-i} \notin H^{*0}_{-i}$ (since, otherwise, $h_{-i} \in H^{*0}_{-i}$) and $\lim_j(j^{j-1-(t-1)}\mu^j(h^{t-1}_{-i}|h^{t-1}_i)) = 0$. Thus, $\lim_j(j^{j-1-t}\mu^j(h_{-i}|h_i)) = 0$.

If, instead, $h^{t-1}_{-i} = h^{*,t-1}_{-i}$, note that $h^{*,t-1}_{-i} \cdot (r^*_{-i,t}, s^*_{-i,t}, c) = h^{*,t}_{-i}$, $h^{*,t-1}_{-i} \cdot (r^*_{-i,t}, s^*_{-i,t}, d) \in H^{*0}_{-i}$ and that for $\bar{s}_{-i} \neq s^*_{-i,t}$, $h^{*,t-1}_{-i} \cdot (1_{(c,d)}, \bar{s}_{-i}, y_{-i}) \in H^{*0}_{-i}$ for each $y_{-i} \in Y_{-i}$. Thus, in this case,

$$(r_{-i,t}, s_{-i,t}, y_{-i,t}) \notin \{(r^*_{-i,t}, s^*_{-i,t}, c), (r^*_{-i,t}, s^*_{-i,t}, d), (1_{(c,d)}, \bar{s}_{-i}, c), (1_{(c,d)}, \bar{s}_{-i}, d)\}$$

and $\sigma'_{-i}(r_{-i,t}, s_{-i,t}|h^{*,t-1}_{-i}) = 0$ for each $\sigma'_{-i} \in \{\sigma^\nu_{-i}, \bar{\sigma}_{-i}\}$. Hence, the numerator of $(j^{j-1-t}\mu^j(h_{-i}|h_i))$ is

$$j^{j-1-t}\mu^j(h^{*,t-1}_{-i}|h^{t-1}_i)\pi^t_i(h_i, h^{*,t-1}_{-i} \cdot (r_{-i,t}, s_{-i,t}, y_{-i,t}), \hat{\sigma})j^{-j}$$

and, hence, $\lim_j (j^{j-1-t}\mu^j(h_{-i}|h_i)) = 0$. ■

Define

$$H_i^{D0} = \Big\{h_i \in H_i : \text{ for all } 1 \le n \le \ell(h_i) \text{ and } (y_{-1,n+1}, \ldots, y_{-i,\ell(h_i)}) \in Y_{-i}^{\ell(h_i)-n},$$
$$\left(\prod_{k=1}^{n-1} \alpha_i(r_{i,k}^{s^*_{-i,k}}, r_{-i,k}^{*,s_{i,k}})[(y_{i,k}, c)]\right) \alpha_i(r_{i,n}^{s^*_{-i,n}}, r_{-i,n}^{*,s_{i,n}})[(y_{i,n}, d)] \left(\prod_{k=n+1}^{\ell(h_i)} \alpha_i(r_{i,k}^D, 1_{(d,c)})[y_k]\right) = 0\Big\}.$$

This is the set of player i's histories that happen with probability zero when $h_{-i} \in H_{-i}^D$, i.e. player $-i$ follows σ_{-i} and, for some $1 \le n \le \ell(h_i)$, $h_{-i}^k \in H_{-i}^*$ for all $k < n$ and $h_{-i}^k \in H_{-i}^{*0}$ for all $k \ge n$.

Claim 3 *For each $i \in \{1,2\}$, $t \in \mathbb{N}$ and $h_i \in H_i^t \cap (H_i^{*0} \setminus H_i^{D0})$:*

1. $\lim_j \mu^j(h_{-i}|h_i) = \frac{\pi_i(h_i,h_{-i},\sigma^\nu)}{\sum_{\hat{h}_{-i} \in H_{-i}^t} \pi_i(h_i,\hat{h}_{-i},\sigma^\nu)}$ *for each $h_{-i} \in H_{-i}^t$,*
2. $\mathrm{supp}(\mu^\nu(\cdot|h_i)) \subseteq H_{-i}^{*0}$, *and*
3. $\lim_j (j^{j-1-t}\mu^j(h_{-i}|h_i)) = 0$ *for each $h_{-i} \in H_{-i}^t \setminus H_{-i}^{*0}$.*

Proof. We have that $\pi_i(h_i, h_{-i}^{*,t}, \sigma^\nu) = 0$ since $h_i \in H_i^{*0}$. Since $h_i \notin H_i^{D0}$, there exist $1 \le n \le t$ and $(y_{-1,n+1}, \ldots, y_{-i,t}) \in Y_{-i}^{t-n}$ such that

$$\left(\prod_{k=1}^{n-1} \alpha_i(r_{i,k}^{s^*_{-i,k}}, r_{-i,k}^{*,s_{i,k}})[(y_{i,k}, c)]\right) \alpha_i(r_{i,n}^{s^*_{-i,n}}, r_{-i,n}^{*,s_{i,n}})[(y_{i,n}, d)] \left(\prod_{k=n+1}^{t} \alpha_i(r_{i,k}^D, 1_{(d,c)})[y_k]\right) > 0.$$

Hence, letting for each $1 \le k \le t$,

$$(r_{-i,k}, s_{-i,k}, y_{-i,k}) = \begin{cases} (r^*_{-i,k}, s^*_{-i,k}, c) & \text{if } k \le n-1, \\ (r^*_{-i,k}, s^*_{-i,k}, d) & \text{if } k = n, \\ (1_{(c,d)}, D, y_{-i,k}) & \text{if } k \ge n+1, \end{cases} \tag{3}$$

it follows that $\pi_i(h_i, h_{-i}, \sigma^\nu) > 0$. This then implies part 1 and that $\mu^\nu(h_{-i}^{*,t}|h_i) = 0$.

For part 2, note that if $h_{-i} \in H_{-i}^t$ is such that $\mu^\nu(h_{-i}|h_i) > 0$, then $h_{-i} \ne h_{-i}^{*,t}$ and $\prod_{k=1}^t \sigma_{-i}^\nu(r_{-i,k}, s_{-i,k}|h_{-i}^{k-1}) > 0$. Thus, by Claim 1, $h_{-i} \in H_{-i}^D \subseteq H_{-i}^{*0}$.

We establish part 3 by induction on t. Let $t = 1$ and consider $h_i = (r_i, s_i, y_i) \in H_i^1 \cap (H_i^{*0} \setminus H_i^{D0})$. Then $\alpha_i(r_i^{s^*_{-i,1}}, r_{-i,1}^{*,s_i})[(y_i, d)] > 0$ and $\alpha_i(r_i^{s^*_{-i,1}}, r_{-i,1}^{*,s_i})[(y_i, c)] = 0$

since, respectively, $h_i \in H_i^1 \setminus H_i^{D0}$ and $h_i \in H_i^1 \cap H_i^{*0}$. In addition, for each j, $\mu^j(h_{-i}^{*,1}|h_i) = 0$ and

$$\sum_{\hat{h}_{-i} \in H_{-i}^1} \pi_i(h_i, \hat{h}_{-i}, \sigma^j) \to \alpha_i(r_i^{s_{-i,1}^*}, r_{-i,1}^{*,s_i})[(y_i, d)].$$

Thus, $\text{supp}(\mu^\nu(\cdot|h_i)) = \{(r_{-i,1}^*, s_{-i,1}^*, d)\} \subseteq H_{-i}^{*0}$. We have that $\{(1_{(c,d)}, \bar{s}_{-i}, c), (1_{(c,d)}, \bar{s}_{-i}, d)\} \subseteq H_{-i}^{*0}$ for $\bar{s}_{-i} \neq s_{-i,1}^*$ since $\alpha_{-i}(1_{(c,d)}, 1_{(c,d)})[(y_{-i}, c)] = 0$ by property 1. Hence, for each $h_{-i} \in H_{-i} \setminus H_{-i}^{*0}$,

$$\lim_j (j^{j-2} \mu^j(h_{-i}|h_i)) = \lim_j \frac{j^{j-2} \pi_i(h_i, h_{-i}, \hat{\sigma}) j^{-j}}{\sum_{\hat{h}_{-i} \in H_{-i}^1} \pi_i(h_i, \hat{h}_{-i}, \sigma^j)} = 0.$$

Let $t > 1$ and assume that we have established that, for each $k = 1, \ldots, t-1$ and $h_i \in H_i^k \cap (H_i^{*0} \setminus H_i^{D0})$, $\lim_j (j^{j-1-k} \mu^j(h_{-i}|h_i)) = 0$ for each $h_{-i} \in H_{-i}^k \setminus H_{-i}^{*0}$. Let $h_i \in H_i^t \cap (H_i^{*0} \setminus H_i^{D0})$. We have that

$$\lim_j \sum_{\hat{h}_{-i} \in H_{-i}^{t-1}} \sum_{(\hat{r}_{-i,t}, \hat{s}_{-i,t}, \hat{y}_{-i,t}) \in H_{-i}^1} \mu^j(\hat{h}_{-i}|h_i^{t-1}) \pi_i^t(h_i, \hat{h}_{-i} \cdot (\hat{r}_{-i,t}, \hat{s}_{-i,t}, \hat{y}_{-i,t}), \sigma^j) > 0$$

because $h_i \notin H_i^{D0}$; thus, letting $\tilde{h}_{-i}$ be defined by (3), it follows that $\mu^\nu(\tilde{h}^{t-1}|h_i^{t-1}) > 0$ and $\pi_i^t(h_i, \tilde{h}_{-i}, \sigma^\nu) > 0$. Hence, for each $h_{-i} \in H_{-i}^t \setminus H_{-i}^{*0}$, $\lim_j (j^{j-1-t} \mu^j(h_{-i}|h_i)) = 0$ since $\lim_j (j^{j-1-(t-1)} \mu^j(h_{-i}^{t-1}|h_i^{t-1})) = 0$. ■

For each $i \in \{1, 2\}$, let

$$\bar{H}_i = \left\{ h_i \in H_i : \pi_i(h_i, h_{-i}, \bar{\sigma}) > 0 \text{ for some } h_{-i} \in H_{-i}^{\ell(h_i)} \right\}.$$

Claim 4 *For each $i \in \{1, 2\}$, $t \in \mathbb{N}$ and $h_i \in H_i^t \cap H_i^{*0} \cap H_i^{D0}$:*

1. *for each $h_{-i} \in H_{-i}^t$,*

$$\lim_j \mu^j(h_{-i}|h_i) = \begin{cases} \frac{\pi_i(h_i, h_{-i}, \bar{\sigma})}{\sum_{\hat{h}_{-i} \in H_{-i}^t} \pi_i(h_i, \hat{h}_{-i}, \bar{\sigma})} & \text{if } h_i \in \bar{H}_i \\ \frac{\pi_i(h_i, h_{-i}, \hat{\sigma})}{\sum_{\hat{h}_{-i} \in H_{-i}^t} \pi_i(h_i, \hat{h}_{-i}, \hat{\sigma})} & \text{otherwise,} \end{cases}$$

and

2. $\text{supp}(\mu^\nu(\cdot|h_i)) \subseteq H_{-i}^{*0}$.

Proof. Part 1 follows because $\pi_i(h_i, h_{-i}, \sigma^\nu) = 0$ for each $h_{-i} \in H^t_{-i}$ since $h_i \in H_i^{*0} \cap H_i^{D0}$. Indeed, if $h_i \in \bar{H}_i$, let $h_{-i} \in H^t_{-i}$ be such that $\pi_i(h_i, h_{-i}, \bar{\sigma}) > 0$; then $\sum_{\hat{h}_{-i} \in H^t_{-i}} \pi_i(h_i, \hat{h}_{-i}, \bar{\sigma}) > 0$ and

$$\begin{aligned}\lim_j \mu^j(h_{-i}|h_i) &= \lim_j \frac{j^{-1}\pi_i(h_i, h_{-i}, \bar{\sigma}) + j^{-j}\pi_i(h_i, h_{-i}, \hat{\sigma})}{\sum_{\hat{h}_{-i} \in H^t_{-i}} \left(j^{-1}\pi_i(h_i, \hat{h}_{-i}, \bar{\sigma}) + j^{-j}\pi_i(h_i, \hat{h}_{-i}, \hat{\sigma})\right)} \\ &= \lim_j \frac{\pi_i(h_i, h_{-i}, \bar{\sigma}) + j^{-j+1}\pi_i(h_i, h_{-i}, \hat{\sigma})}{\sum_{\hat{h}_{-i} \in H^t_{-i}} \left(\pi_i(h_i, \hat{h}_{-i}, \bar{\sigma}) + j^{-j+1}\pi_i(h_i, \hat{h}_{-i}, \hat{\sigma})\right)} \\ &= \frac{\pi_i(h_i, h_{-i}, \bar{\sigma})}{\sum_{\hat{h}_{-i} \in H^t_{-i}} \pi_i(h_i, \hat{h}_{-i}, \bar{\sigma})}.\end{aligned}$$

If instead $h_i \notin \bar{H}_i$, then $\pi_i(h_i, h_{-i}, \bar{\sigma}) = 0$ for each $h_{-i} \in H^t_{-i}$. Since $\hat{\sigma}$ is totally mixed, let $h_{-i} \in H^t_{-i}$ be such that $\pi_i(h_i, h_{-i}, \hat{\sigma}) > 0$; for instance, let $h_{-i} = (r_{-i,k}, s_{-i,k}, y_{-i,k})_{k=1}^t$ with $(r_{-i,k}, s_{-i,k}, y_{-i,k}) = (1_{(c,y_{i,k})}, C, c)$ for each $1 \le k \le t$ and note that $\pi_i(h_i, h_{-i}, \hat{\sigma}) > 0$ since $\alpha_i(r^C_{i,k}, 1_{(y_{i,k},c)})[(y_{i,k}, c)]\hat{\sigma}_{-i}(1_{(c,y_{i,k})}, C|h^{k-1}_{-i}) > 0$ for each $1 \le k \le t$ by property 2 and because $\hat{\sigma}$ is totally mixed. Then $\sum_{\hat{h}_{-i} \in H^t_{-i}} \pi_i(h_i, \hat{h}_{-i}, \hat{\sigma}) > 0$ and

$$\lim_j \mu^j(h_{-i}|h_i) = \frac{\pi_i(h_i, h_{-i}, \hat{\sigma})}{\sum_{\hat{h}_{-i} \in H^t_{-i}} \pi_i(h_i, \hat{h}_{-i}, \hat{\sigma})}.$$

We establish part 2 by induction on t. Let $t = 1$ and consider $h_i = (r_i, s_i, y_i) \in H_i^1 \cap H_i^{*0} \cap H_i^{D0}$. Since $h_i \in H_i^{*0} \cap H_i^{D0}$, we have that $\alpha_i(r_i^{s^*_{-i,1}}, r^{*,s_i}_{-i,1})[(y_i, d)] = \alpha_i(r_i^{s^*_{-i,1}}, r^{*,s_i}_{-i,1})[(y_i, c)] = 0$, which implies that $y_i = d$ since $\sum_{y_{-i}} r^{*,s_i}_{-i,1}[(c, y_{-i})] = 1$. In addition, for each j, $\mu^j(h^{*,1}_{-i}|h_i) = 0$ and, for each $h_{-i} \neq h^{*,1}_{-i}$, letting $\bar{s}_{-i} \neq s^*_{-i,1}$,

$$\mu^j(h_{-i}|h_i) = \frac{\alpha_i(r_i^{\bar{s}_{-i}}, 1_{(d,c)})[(d, y_{-i})] + \pi_i(h_i, h_{-i}, \hat{\sigma})j^{-(j-1)}}{\sum_{\hat{y}_{-i}} \alpha_i(r_i^{\bar{s}_{-i}}, 1_{(d,c)})[(d, \hat{y}_{-i})] + j^{-(j-1)} \sum_{\hat{h}_{-i}:(\hat{r}_{-i},\hat{s}_{-i}) \neq (1_{(c,d)}, \bar{s}_{-i})} \pi_i(h_i, \hat{h}_{-i}, \hat{\sigma})}$$

if $(r_{-i}, s_{-i}) = (1_{(c,d)}, \bar{s}_{-i})$ and

$$\mu^j(h_{-i}|h_i) = \frac{\pi_i(h_i, h_{-i}, \hat{\sigma})j^{-(j-1)}}{\sum_{\hat{y}_{-i}} \alpha_i(r_i^{\bar{s}_{-i}}, 1_{(d,c)})[(d, \hat{y}_{-i})] + j^{-(j-1)} \sum_{\hat{h}_{-i}:(\hat{r}_{-i},\hat{s}_{-i}) \neq (1_{(c,d)}, \bar{s}_{-i})} \pi_i(h_i, \hat{h}_{-i}, \hat{\sigma})}$$

otherwise. It then follows that $(1_{(c,d)}, \bar{s}_{-i}, c) \in \text{supp}(\mu^\nu(\cdot|h_i))$ by property 2 and that $\text{supp}(\mu^\nu(\cdot|h_i)) \subseteq \{(1_{(c,d)}, \bar{s}_{-i}, c), (1_{(c,d)}, \bar{s}_{-i}, d)\}$. For each $h_{-i} \in \{(1_{(c,d)}, \bar{s}_{-i}, c), (1_{(c,d)}, \bar{s}_{-i}, d)\}$, we have that $h_{-i} \in H^{*0}_{-i}$ since $\alpha_{-i}(1_{(c,d)}, 1_{(c,d)})[(y_{-i}, c)] = 0$ by property 1. Hence,

$\text{supp}(\mu^\nu(\cdot|h_i)) \subseteq H^{*0}_{-i}$. Furthermore, for each $h_{-i} \in H^1_{-i} \setminus \text{supp}(\mu^\nu(\cdot|h_i))$, $\lim_j (j^{j-2}\mu^j(h_{-i}|h_i)) = 0$; since $H^1_{-i} \setminus H^{*0}_{-i} \subseteq H^1_{-i} \setminus \text{supp}(\mu^\nu(\cdot|h_i))$, then

$$\lim_j (j^{j-2}\mu^j(h_{-i}|h_i)) = 0 \text{ for each } h_{-i} \in H_{-i} \setminus H^{*0}_{-i}.$$

Let $t > 1$ and assume that we have established that, for each $k = 1, \ldots, t-1$ and $h_i \in H^k_i \cap H^{*0}_i \cap H^{D0}_i$, $\text{supp}(\mu^\nu(\cdot|h_i)) \subseteq H^{*0}_{-i}$ and $\lim_j (j^{j-1-k}\mu^j(h_{-i}|h_i)) = 0$ for each $h_{-i} \in H^k_{-i} \setminus H^{*0}_{-i}$.

Let $h_i \in H^t_i \cap H^{*0}_i \cap H^{D0}_i$ and $h_{-i} \in H^t_{-i} \setminus H^{*0}_{-i}$. Note that $h^{t-1}_{-i} \in H^{t-1}_{-i} \setminus H^{*0}_{-i}$. We will show that $\lim_j (j^{j-1-t}\mu^j(h_{-i}|h_i)) = 0$ for each $h_{-i} \in H^t_{-i} \setminus H^{*0}_{-i}$.

Consider first the case where $h_{-i} = h^{*,t}_{-i}$. In this case, $j^{j-1-t}\mu^j(h_{-i}|h_i) = 0$ for each $j \in \mathbb{N}$ since $h_i \in H^{*0}_i$ and the result follows. Thus, we may assume that $h_{-i} \neq h^{*,t}_{-i}$. For convenience, let

$$B_j = \sum_{\hat{h}_{-i} \in H^{t-1}_{-i}} \sum_{(\hat{r}_{-i,t}, \hat{s}_{-i,t}, \hat{y}_{-i,t}) \in H^1_{-i}} \mu^j(\hat{h}_{-i}|h^{t-1}_i)\pi^t_i(h_i, \hat{h}_{-i} \cdot (\hat{r}_{-i,t}, \hat{s}_{-i,t}, \hat{y}_{-i,t}), \sigma^j).$$

We consider two cases.

Case (i): $h^{t-1}_i \in H^{*0}_i$.

Let $\tilde{h}^{t-1}_{-i} \in \text{supp}(\mu^\nu(\cdot|h^{t-1}_i)) \subseteq H^{*0}_{-i}$; since $\sigma_{-i}(1_{(c,d)}, D|\tilde{h}^{t-1}_{-i}) = 1$, it follows that $\lim_j B_j > 0$ when $\alpha_i(r^D_{i,t}, 1_{(d,c)})[(y_{i,t}, \tilde{y}_{-i})] > 0$ for some $\tilde{y}_{-i} \in Y_{-i}$; in particular, $\lim_j B_j > 0$ when $y_{i,t} = d$ by property 2. In this case, $\lim_j (j^{j-1-t}\mu^j(h_{-i}|h_i)) = 0$ since $\lim_j (j^{j-1-(t-1)}\mu^j(h^{t-1}_{-i}|h^{t-1}_i)) = 0$ by Claim 3 and the inductive step.

If $y_{i,t} = c$ and $\alpha_i(r^D_{i,t}, 1_{(d,c)})[(c, \tilde{y}_{-i})] = 0$ for all $\tilde{y}_{-i} \in Y_{-i}$, then, since $\bar{\sigma}_{-i}(\tilde{h}^{t-1}_{-i}) = (1_{(d,c)}, D)$,

$$\begin{aligned}\lim_j (jB_j) = \lim_j & \sum_{\hat{h}_{-i} \in H^{t-1}_{-i} \cap H^{*0}_{-i}} \mu^j(\hat{h}_{-i}|h^{t-1}_i) \times \\ & \times \left(\frac{1}{j^{j-1}} \sum_{(\hat{r}_{-i,t}, \hat{s}_{-i,t}, \hat{y}_{-i,t}) : (\hat{r}_{-i,t}, \hat{s}_{-i,t}) \neq (1_{(d,c)}, D)} \pi^t_i(h_i, \hat{h}_{-i} \cdot (\hat{r}_{-i,t}, \hat{s}_{-i,t}, \hat{y}_{-i,t}), \hat{\sigma}) \right. \\ & \left. + \sum_{\hat{y}_{-i,t}} \alpha_i(r^D_{i,t}, 1_{(c,d)})[(c, \hat{y}_{-i,t})] \right)\end{aligned}$$

which is strictly positive since $\alpha_i(r^D_{i,t}, 1_{(c,d)})[(c,d)] > 0$ by property 2. Since, by Claim 3 and the inductive step,

$$\lim_j \left(j^{j-1-(t-1)} \mu^j(h^{t-1}_{-i}|h^{t-1}_i) \pi^t_i(h_i, h_{-i}, \sigma^j) \right) = 0,$$

it follows that $\lim_j (j^{j-1-t}\mu^j(h_{-i}|h_i)) = 0$ for each $h_{-i} \in H_{-i} \setminus H^{*0}_{-i}$.

Case (ii): $h^{t-1}_i \in H_i \setminus H^{*0}_i$.

In this case, we have $\alpha_i(r^{s^*_{-i,t}}_{i,t}, r^{*,s_{i,t}}_{-i,t})[(y_{i,t}, c)] = \alpha_i(r^{s^*_{-i,t}}_{i,t}, r^{*,s_{i,t}}_{-i,t})[(y_{i,t}, d)] = 0$ since, respectively, $h_i \in H^{*0}_i$ and $h_i \in H^{D0}_i$. Thus, $y_{i,t} = d$ and the argument in case (i) can be applied to conclude that $\lim_j B_j > 0$ provided that there is $\tilde{h}^{t-1}_{-i} \in \text{supp}(\mu^\nu(\cdot|h^{t-1}_i)) \cap H^{*0}_{-i}$. Claim 2 then implies that $\lim_j (j^{j-1-(t-1)}\mu^j(h^{t-1}_{-i}|h^{t-1}_i)) = 0$ and, hence, $\lim_j (j^{j-1-t}\mu^j(h_{-i}|h_i)) = 0$ if $h^{t-1}_{-i} \neq h^{*,t-1}_{-i}$.

If $h^{t-1}_{-i} = h^{*,t-1}_{-i}$, then, letting $\bar{s}_{-i} \neq s^*_{-i,t}$, it cannot be that $(r_{-i,t}, s_{-i,t}) = (1_{(c,d)}, \bar{s}_{-i}) = \bar{\sigma}_{-i}(h^{*,t-1}_{-i})$; indeed, $h_{-i} \notin H^{*0}_{-i}$ and, for each $y_{-i,t} \in Y$, $h^{*,t-1}_{-i} \cdot (r_{-i,t}, s_{-i,t}, y_{-i,t}) \in H^{*0}_{-i}$ as $r^{*,\bar{s}_{-i}}_{i,t} = 1_{(d,c)}$ and $\alpha_{-i}(1_{(c,d)}, 1_{(c,d)})[(y_{-i,t}, c)] = 0$ by property 1. Thus, $\pi^t_i(h_i, h_{-i}, \bar{\sigma}) = 0$. In addition, $\pi^t_i(h_i, h_{-i}, \sigma^\nu) = 0$ since $\sigma^\nu_{-i}(h^{t-1}_{-i}) = (r^*_{-i,t}, s^*_{-i,t})$ and $\alpha_i(r^{s^*_{-i,t}}_{i,t}, r^{*,s_{i,t}}_{-i,t})[(y_{i,t}, c)] = \alpha_i(r^{s^*_{-i,t}}_{i,t}, r^{*,s_{i,t}}_{-i,t})[(y_{i,t}, d)] = 0$. Thus, $\pi^t_i(h_i, h_{-i}, \sigma^j) < j^{-j}$. Thus, $\lim_j (j^{j-1-t}\mu^j(h_{-i}|h_i)) = 0$.

Hence, we are left with the case where $\text{supp}(\mu^\nu(\cdot|h^{t-1}_i)) = \{h^{*,t-1}_{-i}\}$. In this case, since $\bar{\sigma}_{-i}(h^{*,t-1}_{-i}) = (1_{(c,d)}, \bar{s}_{-i})$, where $\bar{s}_{-i} \neq s^*_{-i,t}$,

$$\begin{aligned}
\lim_j (jB_j) = \lim_j \mu^j(h^{*,t-1}_{-i}|h^{t-1}_i) \times \\
\times \Bigg(\frac{1}{j^{j-1}} \sum_{(\hat{r}_{-i,t}, \hat{s}_{-i,t}, \hat{y}_{-i,t}) : (\hat{r}_{-i,t}, \hat{s}_{-i,t}) \neq (1_{(c,d)}, \bar{s}_{-i})} \pi^t_i(h_i, h^{*,t-1}_{-i} \cdot (\hat{r}_{-i,t}, \hat{s}_{-i,t}, \hat{y}_{-i,t}), \hat{\sigma}) \\
+ \sum_{\hat{y}_{-i,t}} \alpha_i(r^{\bar{s}_{-i}}_{i,t}, 1_{(d,c)})[(d, \hat{y}_{-i,t})] \Bigg)
\end{aligned}$$

which is strictly positive since $\alpha_i(r^{\bar{s}_{-i}}_{i,t}, 1_{(d,c)})[(d,c)] > 0$ by property 2. Since, by Claim 2 when $h^{t-1}_{-i} \neq h^{*,t-1}_{-i}$ and because $\pi^t_i(h_i, h_{-i}, \sigma^j) < j^{-j}$ when $h^{t-1}_{-i} = h^{*,t-1}_{-i}$ as argued above,

$$\lim_j \left(j^{j-1-(t-1)} \mu^j(h^{t-1}_{-i}|h^{t-1}_i) \pi^t_i(h_i, h_{-i}, \sigma^j) \right) = 0,$$

it follows that $\lim_j (j^{j-1-t}\mu^j(h_{-i}|h_i)) = 0$ for each $h_{-i} \in H_{-i} \setminus H^{*0}_{-i}$. ■

It follows by Claims 3 and 4 that part (b) of Lemma 1 holds. Part (c) of Lemma 1 follows from Claim 2. Finally, part (d) of Lemma 1, i.e. the continuity of $\nu \mapsto \mu^\nu$, follows by part 1 of Claims 2 – 4 since the definition of H_i^{*0}, H_i^{D0} and $\bar{H}_i$ do not depend on ν.

6.1.4 The fixed point argument

Let $\nu \in \Sigma$ be given. For each $h \in H$ and $t \in \mathbb{N}$, let $\xi^t(h)$ be the probability measure on H^t induced by σ^ν and h. Specifically, set $\xi^1(h)[r, s, y] = \sigma^\nu(r, s|h)\gamma(y|r, s)$; assuming that $\xi^1(h), \ldots, \xi^{t-1}(h)$ have been defined, set, for each $\bar{h} \in H^t$,

$$\xi^t(h)[\bar{h}] = \xi^{t-1}(h)[\bar{h}^{t-1}]\sigma^\nu(\bar{r}_t, \bar{s}_t|h \cdot \bar{h}^{t-1})\gamma(\bar{y}_t|\bar{r}_t, \bar{s}_t).$$

Let $U_i^\nu(h)$ be player i's expected payoff following history $h \in H$:

$$U_i^\nu(h) = (1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}\sum_{\bar{h}\in H^{t-1}} u_i(\sigma_S^\nu(h \cdot \bar{h}))\xi^{t-1}(h)[\bar{h}],$$

where $\xi^0(h)[h^{*,0}] = 1$ and for each $h \in H$, $\sigma_S^\nu(h)$ denotes the marginal of $\sigma^\nu(h)$ on S. Let $U_i^\nu(h_i)$ be player i's expected payoff following history $h_i \in H_i$:

$$U_i^\nu(h_i) = \sum_{h_{-i}\in H_{-i}} \mu^\nu(h_{-i}|h_i)U_i^\nu(h_i, h_{-i}).$$

For each $i \in \{1, 2\}$, $h_i \in H_i$ and $(r_i, s_i) \in X^2 \times \{C, D\}$, let $U_i^{\nu,r_i,s_i}(h_i)$ be player i's expected payoff of an one-shot deviation from σ_i^ν to (r_i, s_i); formally, $U_i^{\nu,r_i,s_i}(h_i)$ is defined in the same way as $U_i^\nu(h_i)$ by changing, for each $h_{-i} \in H_{-i}$, $\xi^1(h)$ to $((r_i, s_i), \sigma_{-i}^\nu(h_{-i}))$, i.e.

$$\begin{aligned} U_i^{\nu,r_i,s_i}(h_i) = \sum_{h_{-i}\in H_{-i}} \mu^\nu(h_{-i}|h_i)\Big(&(1-\delta)u_i(s_i, \sigma_{-i,S_{-i}}^\nu(h_{-i})) \\ &+ \delta \sum_{y,\bar{r}_{-i},\bar{s}_{-i}} \sigma_{-i}^\nu(\bar{r}_{-i}, \bar{s}_{-i}|h_{-i})\gamma(y|r_i, s_i, \bar{r}_{-i}, \bar{s}_{-i})U_i^\nu(h \cdot ((r_i, s_i, y_i), (\bar{r}_{-i}, \bar{s}_{-i}, y_{-i})))\Big), \end{aligned}$$

where $\sigma_{-i,S_{-i}}^\nu(h_{-i})$ denotes the marginal of $\sigma_{-i}^\nu(h_{-i})$ on S_{-i}.

Define $\Phi : \Sigma \rightrightarrows \Sigma$ by setting, for each $\nu \in \Sigma$,

$$\Phi_{i,h_i}(\nu) = \{\lambda_i \in \Delta(X^2 \times \{C, D\}) : \ (r_i, s_i) \text{ solves } \max_{(r_i', s_i')\in X^2\times\{C,D\}} U_i^{\nu,r_i,s_i}(h_i) \text{ for each } (r_i, s_i) \in \operatorname{supp}(\lambda_i)\}$$

for each $(i, h_i) \in \Theta$ and

$$\Phi(\nu) = \prod_{(i,h_i)\in\Theta} \Phi_{i,h_i}(\nu).$$

Let $\mathbb{R}^{X^2\times\{C,D\}\times\Theta}$ be endowed with the product topology. The following claim establishes some properties of this topological space and of its subset Σ.

Claim 5 *The set Σ is a nonempty, convex and compact subset of $\mathbb{R}^{X^2\times\{C,D\}\times\Theta}$, which itself is a locally convex Hausdorff topological vector space.*

Proof. Note that $X^2 \times \{C, D\} \times \Theta$ is countable and, thus, $\mathbb{R}^{X^2\times\{C,D\}\times\Theta}$ is first countable by Kelley (1955, Theorem 6, p. 92). This implies that addition and scalar multiplication in $\mathbb{R}^{X^2\times\{C,D\}\times\Theta}$ are continuous since a sequence in a product space converges to a point c if and only if its projection in each coordinate space converges to the projection of c by Kelley (1955, Theorem 4, p. 91) and both addition and scalar multiplication are continuous in each coordinate space.[17,18] We have that $\mathbb{R}^{X^2\times\{C,D\}\times\Theta}$ is Hausdorff by Kelley (1955, Theorem 5, p. 92) since each coordinate space is Hausdorff. It is also locally convex since, writing B_r for the open ball of radius $r > 0$ around zero in $\mathbb{R}$, the collection of sets $U_{F,\{r_k\}_{k\in F}} = \{c : c_k \in B_{r_k} \text{ for each } k \in F\}$ where F is a finite subset of $X^2 \times \{C, D\} \times \Theta$ and $r_k > 0$ for each $k \in F$ is a local base whose members are convex; this follows by Kelley (1955, p.90) and the fact that $\{B_r : r > 0\}$ is a local base of $\mathbb{R}$. Finally, Σ is nonempty and convex because the product of nonempty (resp. convex) sets is nonempty (resp. convex) and it is compact by Tychonoff Theorem e.g. Kelley (1955, Theorem 13, p. 143). ■

It is clear that Φ is convex-valued and, since $\nu \mapsto \mu^\nu$ and $(\nu, r_i, s_i) \mapsto U_i^{\nu,r_i,s_i}(h_i)$ are continuous, it follows that Φ is closed. It then follows by the Fan-Glicksberg Fixed Point Theorem that there is $\nu^* \in \Sigma$ such that $\nu^* \in \Phi(\nu^*)$.

[17] Addition and scalar multiplication are defined coordinate-wise as usual i.e. the kth coordinate of $a+b$ and λa are, respectively, $a_k + b_k$ and λa_k for each $a, b \in \mathbb{R}^{X^2\times\{C,D\}\times\Theta}$, $k \in X^2 \times \{C, D\} \times \Theta$ and $\lambda \in \mathbb{R}$.

[18] That $\mathbb{R}^{X^2\times\{C,D\}\times\Theta}$ is first countable allows us to use sequences to show that addition and scalar multiplication are continuous but it is not needed since the same argument applies to nets.

6.1.5 Sequential rationality

Let $\sigma = \sigma^{\nu^*}$, $\mu = \mu^{\nu^*}$, $U_i(h) = U_i^{\nu^*}(h)$, $U_i(h_i) = U_i^{\nu^*}(h_i)$ and $U_i^{r_i,s_i}(h_i) = U_i^{\nu^*,r_i,s_i}(h_i)$ for each $i \in \{1,2\}$, $h \in H$, $h_i \in H_i$ and $(r_i, s_i) \in X^2 \times \{C, D\}$. We show that

$$U_i(h_i) \geq U_i^{r_i,s_i}(h_i) \text{ for each } i \in \{1,2\}, h_i \in H_i \text{ and } (r_i, s_i) \in X^2 \times \{C, D\} \quad (4)$$

from which the sequential rationality of (σ, μ) follows.

Let $i \in \{1,2\}$, $h_i \in H_i$ and $(r_i, s_i) \in X^2 \times \{C, D\}$. If $h_i \in H_i \backslash (H_i^{*0} \cup \{h_i^{*,t} : t \in \mathbb{N}_0\}$, then (4) holds since $\nu^* \in \Phi(\nu^*)$. If $h_i \in H_i^{*0}$, then $\mathrm{supp}(\mu(\cdot|h_i)) \subseteq H_{-i}^{*0}$ and, hence, $U_i(h_i) = 0 \geq U_i^{r_i,s_i}(h_i)$ since player i plays D in every history following h_i and so does player $-i$ in every history following $h_{-i} \in H_{-i}^{*0}$. Thus, it remains to show that σ is sequentially rational following $h_i^{*,t}$ for each $t \geq 0$, i.e. to consider $h_i = h_i^{*,t}$.

We have that $\mu(h_{-i}^{*,t}|h_i^{*,t}) = 1$ and that $U_i(h^{*,t}) = (1-\delta)\sum_{k=t+1}^{\infty} \delta^{k-1} u_i(s_k^*)$ since, for each $k \in \mathbb{N}$, $\xi^k(h^{*,t})$ assigns probability one to $h^{*,k}$.

For each $\tau \in \mathbb{N}_0$, let $V_{i,\tau}^* = \sup_{h_i \in H_i^\tau} U_i(h_i, h_{-i}^{*,\tau})$.

Claim 6 *For each $\tau \in \mathbb{N}_0$, $V_{i,\tau}^* = (1-\delta)\sum_{k=\tau+1}^{\infty} \delta^{k-1} u_i(s_k^*)$.*

Proof. Note that

$$(1-\delta)\sum_{k=\tau+1}^{\infty} \delta^{k-1} u_i(s_k^*) \leq V_{i,\tau}^* \leq 1 + g \quad (5)$$

since $U_i(h_i^{*,\tau}, h_{-i}^{*,\tau}) = (1-\delta)\sum_{k=\tau+1}^{\infty} \delta^{k-1} u_i(s_k^*)$ as shown above. We now show that

$$V_{i,\tau}^* \leq \max\{(1-\delta)u_i(s_{\tau+1}^*) + \delta V_{i,\tau+1}^*, (1-\delta)u_i(\bar{s}_i, s_{-i,\tau+1}^*) + \delta(1-\underline{\alpha}_i)V_{i,\tau+1}^*\},$$

where $\bar{s}_i \neq s_{i,\tau+1}^*$ and, recall, $\underline{\alpha}_i = \min_{\rho \in \Delta Y, y \in Y} \alpha_i(\rho, 1_y)[y]$. Indeed, for each $h_i \in H_i^\tau$ and some $(\hat{r}_i, \hat{s}_i) \in X^2 \times \{C, D\}$, letting

$$y(\hat{s}_i) = \begin{cases} c & \text{if } \hat{s}_i = s_{i,\tau+1}^*, \\ d & \text{otherwise} \end{cases}$$

and $\alpha_{i,Y_{-i}}(\rho,\rho')[y_{-i}] = \sum_{y_i} \alpha_i(\rho,\rho')[y_i, y_{-i}]$ for each $(\rho,\rho') \in (\Delta Y)^2$ and $y_{-i} \in \{c,d\}$, we have that

$$\begin{aligned} U_i(h_i, h^{*,\tau}_{-i}) &= (1-\delta)u_i(\hat{s}_i, s^*_{-i,\tau+1}) \\ &\quad + \delta \sum_{y_i} \Big(\alpha_i(\hat{r}_i^{s^*_{-i,\tau+1}}, 1_{(c,y(\hat{s}_i))})[(y_i,c)] U_i(h_i \cdot (\hat{r}_i, \hat{s}_i, y_i), h^{*,\tau+1}_{-i}) \\ &\quad + \alpha_i(\hat{r}_i^{s^*_{-i,\tau+1}}, 1_{(c,y(\hat{s}_i))})[(y_i,d)] U_i(h_i \cdot (\hat{r}_i, \hat{s}_i, y_i), h^{*,\tau}_{-i} \cdot (r^*_{-i,\tau+1}, s^*_{-i,\tau+1}, d)) \Big) \\ &\quad \leq (1-\delta)u_i(\hat{s}_i, s^*_{-i,\tau+1}) + \delta(1 - \min_\rho \alpha_{i,Y_{-i}}(\rho, 1_{(c,y(\hat{s}_i))})[d]) V^*_{i,\tau+1} \\ &\quad \leq \max\{(1-\delta)u_i(s^*_{\tau+1}) + \delta V^*_{i,\tau+1}, (1-\delta)u_i(\bar{s}_i, s^*_{-i,\tau+1}) + \delta(1-\underline{\alpha}_i)V^*_{i,\tau+1}\}, \end{aligned} \tag{6}$$

which holds because $U_i(h_i \cdot (\hat{r}_i, \hat{s}_i, y_i), h^{*,\tau}_{-i} \cdot (r^*_{-i,\tau+1}, s^*_{-i,\tau+1}, d)) \leq 0$ as $h^{*,\tau}_{-i} \cdot (r^*_{-i,\tau+1}, s^*_{-i,\tau+1}, d) \in H^{*0}_{-i}$, $\min_\rho \alpha_{i,Y_{-i}}(\rho, 1_{(c,y(\hat{s}_i))})[d] = 0$ when $\hat{s}_i = s^*_{i,\tau+1}$ with $\rho = 1_{(c,c)}$ and, when $\hat{s}_i = \bar{s}_i$, for some ρ',

$$\begin{aligned} &\min_\rho \alpha_{i,Y_{-i}}(\rho, 1_{(c,d)})[d] = \alpha_i(\rho', 1_{(c,d)})[(c,d)] + \alpha_i(\rho', 1_{(c,d)})[(d,d)] \\ &\geq \alpha_i(\rho', 1_{(c,d)})[(c,d)] \geq \min_{\rho,y} \alpha_i(\rho, 1_y)[y] = \underline{\alpha}_i. \end{aligned}$$

Note that $(1-\delta)u_i(s^*_{\tau+1}) + \delta V^*_{i,\tau+1} \geq (1-\delta)u_i(\bar{s}_i, s^*_{-i,\tau+1}) + \delta(1-\underline{\alpha}_i)V^*_{i,\tau+1}$ if and only if $\delta \underline{\alpha}_i V^*_{i,\tau+1} \geq (1-\delta)\big(u_i(\bar{s}_i, s^*_{-i,\tau+1}) - u_i(s^*_{\tau+1})\big)$. The latter condition holds since

$$\delta \underline{\alpha}_i V^*_{i,\tau+1} \geq \delta \underline{\alpha}_i (1-\delta) \sum_{k=\tau+1}^{\infty} \delta^{k-1} u_i(s^*_k) \geq \delta \underline{\alpha}_i (v_i - \varepsilon) \geq (1-\delta)\big(u_i(\bar{s}_i, s^*_{-i,\tau+1}) - u_i(s^*_{\tau+1})\big)$$

since

$$\delta \geq \delta^* \geq \frac{\max\{g,l\}}{\max\{g,l\} + (v_i - \varepsilon)\underline{\alpha}_i} \geq \frac{u_i(\bar{s}_i, s^*_{-i,\tau+1}) - u_i(s^*_{\tau+1})}{u_i(\bar{s}_i, s^*_{-i,\tau+1}) - u_i(s^*_{\tau+1}) + (v_i - \varepsilon)\underline{\alpha}_i}.$$

Thus,

$$(1-\delta)u_i(s^*_{\tau+1}) + \delta V^*_{i,\tau+1} \geq (1-\delta)u_i(\bar{s}_i, s^*_{-i,\tau+1}) + \delta(1-\underline{\alpha}_i)V^*_{i,\tau+1}$$

and, therefore,

$$V^*_{i,\tau} \leq (1-\delta)u_i(s^*_{\tau+1}) + \delta V^*_{i,\tau+1}. \tag{7}$$

Since (7) holds for each $\tau \in \mathbb{N}$, it follows that $V^*_{i,\tau} \le (1-\delta)\sum_{k=\tau+1}^{T} \delta^{k-1} u_i(s^*_k) + \delta^T V^*_{i,\tau+T} \le (1-\delta)\sum_{k=\tau+1}^{T} \delta^{k-1} u_i(s^*_k) + \delta^T(1+g)$ for each $T \in \mathbb{N}$ using (5) and, hence,

$$V^*_{i,\tau} \le (1-\delta) \sum_{k=\tau+1}^{\infty} \delta^{k-1} u_i(s^*_k).$$

This, together with (5), then implies that $V^*_{i,\tau} = (1-\delta)\sum_{k=\tau+1}^{\infty} \delta^{k-1} u_i(s^*_k)$. ■

Let $\bar{s}_i \neq s^*_{i,t+1}$, $y(s_i) = c$ if $s_i = s^*_{i,t+1}$ and $y(s_i) = d$ if $s_i \neq s^*_{i,t+1}$. Thus, as in (6),

$$\begin{aligned}
U_i^{r_i,s_i}(h_i^{*,t}) &= (1-\delta)u_i(s_i, s^*_{-i,t+1}) \\
&\quad + \delta \sum_{y_i} \Big(\alpha_i(r_i^{s^*_{-i,t+1}}, 1_{(c,y(s_i))})[(y_i,c)] U_i(h_i^{*,t}\cdot(r_i,s_i,y_i), h_{-i}^{*,t+1}) \\
&\quad + \alpha_i(r_i^{s^*_{-i,t+1}}, 1_{(c,y(s_i))})[(y_i,d)] U_i(h_i^{*,t}\cdot(r_i,s_i,y_i), h_{-i}^{*,t}\cdot(r^*_{-i,t+1}, s^*_{-i,t+1}, d))\Big) \\
&\le \max\{(1-\delta)u_i(s^*_{t+1}) + \delta V^*_{i,t+1}, (1-\delta)u_i(\bar{s}_i, s^*_{-i,t+1}) + \delta(1-\underline{\alpha}_i)V^*_{i,t+1}\} \\
&= (1-\delta)u_i(s^*_{t+1}) + \delta V^*_{i,t+1} \\
&= (1-\delta)\sum_{k=t+1}^{\infty} \delta^{k-1} u_i(s^*_k) = U_i(h_i^{*,t}).
\end{aligned}$$

6.2 Proof of Remark 1

Let i be the player who controls the monitoring structure, $j \neq i$ and (σ, μ) be a sequential equilibrium. Since player i controls the monitoring structure, $\gamma(y|r,s) = r_i^{s_j}(y)$ for each $(r,s,y) \in R \times S \times Y$.

We first show that $\sigma_{i,S_i}(h_i) = 1_D$ for each $h_i \in H_i$ such that $\prod_{k=1}^{t} \sigma(r_k, s_k|h^{k-1}) r_{i,k}^{s_{j,k}}(y_k) > 0$ for some $h_j \in H_j^t$ where $t = \ell(h_i)$ and $\sigma_{i,S_i}(h_i)$ denotes the marginal of $\sigma_i(h_i)$ on S_i; in words, player i plays D with probability 1 at every on-path history h_i given σ.

Let $h_i \in H_i$ be such that $\prod_{k=1}^{t} \sigma(r_k, s_k|h^{k-1}) r_{i,k}^{s_{j,k}}(y_k) > 0$ for some $h_j \in H_j^t$ where $t = \ell(h_i)$ and suppose that $\sigma_{i,S_i}(D|h_i) < 1$.

We have that

$$\begin{aligned}
U_i(h_i) &= (1-\delta)u_i\big(\sigma_{i,S_i}(h_i), \sum_{h_j} \mu(h_j|h_i)\sigma_{j,S_j}(h_j)\big) \\
&\quad + \delta \sum_{h_j} \mu(h_j|h_i) \sum_{r,s,y} \xi^1(h)[r,s,y] U_i(h\cdot(r,s,y))
\end{aligned}$$

where the definitions of $U_i(h_i)$, $U_i(h)$ and $\{\xi^t(h)\}_{t=1}^{\infty}$ are as in Section 6.1.4 and $\sigma_{j,S_j}(h_j)$ denotes the marginal of $\sigma_j(h_j)$ on S_j. For each $(r_i, y_i) \in R_i \times Y_i$, let

$$\theta(r_i, y_i) = \frac{\sum_{h_j \in H_j^t} \sum_{(\hat{r}_j, \hat{s}_j, \hat{y}_j) \in H_j^1} \prod_{k=1}^t \sigma_j(r_{j,k}, s_{j,k}|h_j^{k-1}) r_{i,k}^{s_{j,k}}(y_k) \sigma_j(\hat{r}_j, \hat{s}_j|h_j) r_i^{\hat{s}_j}(y_i, \hat{y}_j)}{\sum_{h_j \in H_j^t} \prod_{k=1}^t \sigma_j(r_{j,k}, s_{j,k}|h_j^{k-1}) r_{i,k}^{s_{j,k}}(y_k)}$$

be the probability of player i receiving signal y_i in period $t+1$ given that he has observed h_i and chosen r_i in period $t+1$. Then

$$\sum_{h_j} \mu(h_j|h_i) \sum_{r,s,y} \xi^1(h)[r,s,y] U_i(h \cdot (r,s,y)) =$$
$$\sum_{r_i,s_i,y_i} \sum_{h_j} \sum_{r_j,s_j,y_j} \theta(r_i, y_i) \sigma_i(r_i, s_i|h_i) \mu(h_j \cdot (r_j, s_j, y_j)|h_i \cdot (r_i, s_i, y_i)) U_i(h \cdot (r,s,y)).$$

Claim 7 *For each $y_i \in Y_i$ and $r_i \in R_i$ such that $\theta(r_i, y_i) > 0$, $U_i(h_i \cdot (r_i, C, y_i)) = U_i(h_i \cdot (r_i, D, y_i))$.*

Proof. Let $y_i \in Y_i$ and $r_i \in R_i$ be such that $\theta(r_i, y_i) > 0$ and suppose that $U_i(h_i \cdot (r_i, C, y_i)) > U_i(h_i \cdot (r_i, D, y_i))$ (the case where $U_i(h_i \cdot (r_i, C, y_i)) < U_i(h_i \cdot (r_i, D, y_i))$ is analogous). Let $\tilde{h}_i = h_i \cdot (r_i, C, y_i)$ and $\bar{h}_i = h_i \cdot (r_i, D, y_i)$. Consider a deviation at $\bar{h}_i$ to $\sigma_i|\tilde{h}_i$ and let $\{\bar{\xi}^k\}_{k=1}^{\infty}$ and $\bar{U}_i(\bar{h}_i) = \sum_{h_j \in H_j^{t+1}} \mu(h_j|\bar{h}_i) \bar{U}_i(\bar{h}_i, h_j)$ be the corresponding sequence of probability measures and payoff. We have that $\bar{\xi}^k(\bar{h}_i, h_j)[\hat{h}] = \xi^k(\tilde{h}_i, h_j)[\hat{h}]$ for each $k \in \mathbb{N}$, $h_j \in H_j^{t+1}$ and $\hat{h} \in H^k$. Thus, $\bar{U}_i(\bar{h}_i, h_j) = U_i(\tilde{h}_i, h_j)$ for each $h_j \in H_j^{t+1}$. Moreover,

$$\mu(h_j|\tilde{h}_i) = \mu(h_j|\bar{h}_i) = \frac{\prod_{k=1}^{t+1} \sigma_j(r_{j,k}, s_{j,k}|h_j^{k-1}) r_{i,k}^{s_{j,k}}(y_{i,k}, y_{-i,k})}{\sum_{\hat{h}_j \in H_j^{t+1}} \prod_{k=1}^{t+1} \sigma_j(\hat{r}_{j,k}, \hat{s}_{j,k}|\hat{h}_j^{k-1}) r_{i,k}^{\hat{s}_{j,k}}(y_{i,k}, \hat{y}_{-i,k})}$$

for each $h_j \in H_j^{t+1}$. Hence,

$$\bar{U}_i(\bar{h}_i) = \sum_{h_j \in H_j^{t+1}} \mu(h_j|\bar{h}_i) \bar{U}_i(\bar{h}_i, h_j) = \sum_{h_j \in H_j^{t+1}} \mu(h_j|\tilde{h}_i) U_i(\tilde{h}_i, h_j) = U_i(\tilde{h}_i) > U_i(\bar{h}_i).$$

But this is a contradiction since (σ, μ) is a sequential equilibrium. ■

For each $y_i \in Y_i$ and $r_i \in R_i$ such that $\theta(r_i, y_i) > 0$, let $U_i(r_i, y_i)$ denote the common value of $U_i(h_i \cdot (r_i, C, y_i))$ and $U_i(h_i \cdot (r_i, D, y_i))$ given by Claim 7. It then follows that

$$U_i(h_i) = (1-\delta) u_i(\sigma_{i,S_i}(h_i), \bar{\sigma}_{j,S_j}) + \delta \sum_{r_i, y_i} \big(\sigma_i(r_i, C|h_i) + \sigma_i(r_i, D|h_i)\big) \theta(r_i, y_i) U_i(r_i, y_i),$$

where $\bar{\sigma}_{j,S_j} = \sum_{h_j} \mu(h_j|h_i)\sigma_{j,S_j}(h_j)$.

Consider an one-shot deviation at h_i to $\bar{\sigma}_i \in \Delta(X^2 \times \{C, D\})$, where $\bar{\sigma}_i(r_i, D) = \sigma_i(r_i, D|h_i) + \sigma_i(r_i, C|h_i)$ for each $r_i \in R_i$, and corresponding payoff $\bar{U}_i(h_i)$. Then

$$\begin{aligned}
\bar{U}_i(h_i) &= (1-\delta)u_i(D, \bar{\sigma}_{j,S_j}) + \delta \sum_{r_i, y_i} \bar{\sigma}_i(r_i, D)\theta(r_i, y_i)U_i(r_i, y_i) \\
&= (1-\delta)u_i(D, \bar{\sigma}_{j,S_j}) + \delta \sum_{r_i, y_i} \big(\sigma_i(r_i, C|h_i) + \sigma_i(r_i, D|h_i)\big)\theta(r_i, y_i)U_i(r_i, y_i) \\
&> (1-\delta)u_i(\sigma_{i,S_i}(h), \bar{\sigma}_{j,S_j}) + \delta \sum_{r_i, y_i} \big(\sigma_i(r_i, C|h_i) + \sigma_i(r_i, D|h_i)\big)\theta(r_i, y_i)U_i(r_i, y_i) \\
&= U_i(h_i).
\end{aligned}$$

But this is a contradiction since (σ, μ) is a sequential equilibrium. This contradiction shows that $\sigma_{i,S_i}(h_i) = 1_D$.

It then follows that $\sigma_{j,S_j}(h_j) = 1_D$ for each $h_j \in H_j$ such that, for some $h_i \in H_i^t$, $\prod_{k=1}^t \sigma(r_k, s_k|h^{k-1})r_{i,k}^{s_{j,k}}(y_k) > 0$. Indeed, if $\sigma_{j,S_j}(D|h_j) < 1$ for some $h_j \in H_j$, then player j can profitably deviate at h_j and each history after h_j, i.e. to $\bar{\sigma}_j$ such that $\bar{\sigma}_j(r_i, D|h_j \cdot h'_j) = \sigma_j(r_i, C|h_j \cdot h'_j) + \sigma_j(r_i, D|h_j \cdot h'_j)$ for each $h'_j \in H_j$.

In conclusion, $\sigma_S(h) = 1_{(D,D)}$ for each $h \in H$ such that $\prod_{k=1}^t \sigma(r_k, s_k|h^{k-1})r_{i,k}^{s_{j,k}}(y_k) > 0$. Hence, for each $t \in \mathbb{N}$, the marginal $\xi_{S^t}^t(h^{*,0})$ of $\xi^t(h^{*,0})$ on S^t assigns probability one to $\overbrace{(D, D), \ldots, (D, D)}^{t \text{ periods}}$.

References

Bhaskar, V., and I. Obara (2002): "Belief-Based Equilibria in the Repeated Prisoners' Dilemma with Private Monitoring," *Journal of Economic Theory*, 102, 40–69.

Fudenberg, D., D. Levine, and E. Maskin (1994): "The Folk Theorem with Imperfect Public Information," *Econometrica*, 62, 997–1039.

Fudenberg, D., and E. Maskin (1986): "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," *Econometrica*, 54, 533–554.

——— (1991): "On the Dispensability of Public Randomization in Discounted Repeated Games," *Journal of Economic Theory*, 53, 428–438.

Hörner, J., and W. Olszewski (2006): "The Folk Theorem for Games with Private Almost-Perfect Monitoring," *Econometrica*, 74, 1499–1544.

Kelley, J. (1955): *General Topology*. Springer, New York.

Mailath, G., and L. Samuelson (2006): *Repeated Games and Reputations: Long-Run Relationships*. Oxford University Press, Oxford.

Marshall, R. C., and L. M. Marx (2012): *The Economics of Collusion: Cartels and Bidding Rings*. The MIT Press.

Matsushima, H. (2004): "Repeated Games with Private Monitoring: Two Players," *Econometrica*, 72, 823–852.

Miyagawa, E., Y. Miyahara, and T. Sekiguchi (2008): "The Folk Theorem for Repeated Games with Observation Costs," *Journal of Economic Theory*, 139, 192–221.

Myerson, R., and P. Reny (2020): "Perfect Conditional ε-Equilibria of Multi-Stage Games with Infinite Sets of Signals and Actions," *Econometrica*, 88, 495–531.

Piccione, M. (2002): "The Repeated Prisoner's Dilemma with Imperfect Private Monitoring," *Journal of Economic Theory*, 102, 70–83.

Sekiguchi, T. (1997): "Efficiency in Repeated Prisoner's Dilemma with Private Monitoring," *Journal of Economic Theory*, 76, 345–361.

Sugaya, T. (2022): "Folk Theorem in Repeated Games with Private Monitoring," *Review of Economic Studies*, 89, 2201–2256.

Yamamoto, Y. (2012): "Characterizing Belief-Free Review-Strategy Equilibrium Payoffs under Conditional Independence," *Journal of Economic Theory*, 147, 1998–2027.