Bounding payoffs in repeated games with private monitoring: n-player games [☆]

Takuo Sugaya, Alexander Wolitzky [∗]

Stanford Graduate School of Business and MIT, United States

Abstract

We provide a simple upper bound on the Nash equilibrium payoff set at a fixed discount factor in repeated games with imperfect private monitoring. The bound admits a tractable recursive characterization and can thus be applied "off-the-shelf" to any repeated game. The bound is not tight in general, but it is tight if the stage game is "concave" or if a certain form of observable mixed actions is allowed. We illustrate our results with applications to the repeated prisoners' dilemma and to canonical public goods and oligopoly games.
© 2018 Elsevier Inc. All rights reserved.

1. Introduction

Repeated games with imperfect private monitoring have been a major topic of research for some time and have been used to model important economic settings such as collusion with secret price cuts (Stigler, 1964) and relational contacting with subjective performance evaluations (Levin, 2003; MacLeod, 2003; Fuchs, 2007). While extensive progress has been made in analyzing the equilibrium set in these games in the limit where players become very patient (see

Sugaya, 2016, and references therein), our understanding of equilibria at fixed discount factors—arguably the more relevant case for most economic applications—remains much more limited. In particular, no simple characterization of the sequential equilibrium payoff set at a fixed discount factor is available for repeated games with private monitoring, in contrast to the case of perfect public equilibria in games with public monitoring, where such a characterization is provided by Abreu et al. (1990; henceforth APS) and extended to mixed strategies by Fudenberg et al. (1994; henceforth FLM), who also prove the folk theorem with imperfect public monitoring. Kandori (2002) discusses the well-known difficulties involved in generalizing the results of APS and FLM to private monitoring.

In a recent paper (Sugaya and Wolitzky, 2017), we have shown that the equilibrium payoff set in a private monitoring repeated game is bounded by the equilibrium set in the same game played under perfect monitoring with the assistance of a mediator, so long as attention is restricted to two-player games where the discount factor exceeds a cutoff value δ^*. This bound is as tractable as APS's characterization for public monitoring games, and it is also tight (from the perspective of an observer who does not know the monitoring structure under which the game is played) insofar as mediated perfect monitoring can itself be viewed as a kind of private monitoring. However, the bound does not apply to games with more than two players or when the discount factor is too low.

In the current paper, we derive and analyze a more permissive bound on the equilibrium payoff set in private monitoring repeated games, which however applies to *any* game, regardless of the number of players or the discount factor. The bound may be interpreted as the equilibrium payoff set in an "information-free" version of the repeated game, where players cannot use information learned in the course of the game when deciding whether or not to deviate from equilibrium play. This bound again admits a simple, APS-style recursive characterization. Thus, while the bound we provide here is not always tight, it gives a simple, off-the-shelf method of bounding equilibrium payoffs in any repeated game with private monitoring.

We then ask when the bound is tight. We show that this is the case if the stage game is *concave*, in the sense that (i) payoffs are continuous and jointly concave in actions, and (ii) a player's maximum deviation gain is greater when she faces a "riskier" distribution of opposing actions. While this is a restrictive condition, we show that it is satisfied by Bertrand and Cournot competition with linear demand curves, as well as by additively separable public goods games.

We also show that the bound is tight if a certain form of observable mixed actions is allowed. Thus, in a precise sense, the slack in our bound is analogous to the slack introduced by assuming observable mixed actions when analyzing repeated games with perfect monitoring. In addition, to state this result we develop a model of observable mixed actions when actions can be correlated, which may be of some independent interest.[1]

Other than our earlier paper, we are not aware of any work that provides a finite-dimensional recursive bound on the entire sequential equilibrium payoff set in repeated games with private monitoring at a fixed discount factor.[2] That paper contains a review of the broader literature on recursive methods in private monitoring repeated games, which we do not repeat here. We may however note the conceptual connection between our approach in these papers and recent work on "informational robustness" in static incomplete information games, following Bergemann and Morris (2013): Bergemann and Morris characterize the set of payoffs that can arise in equilibrium

---

[1] This model was inspired by a comment from Jeff Ely, to whom we are grateful.

[2] Both in the current paper and in our previous work, the payoff set we characterize actually bounds the entire Nash equilibrium payoff set, and indeed the entire communication equilibrium payoff set (Forges, 1986; Myerson, 1986).

in a static incomplete information game for some information structure, while our results characterize (or bound, if the game is not concave and observable mixed actions are not allowed) the set of payoffs that can arise in equilibrium in a repeated complete information game for some monitoring structure. We also note that a couple recent papers (Gossner and Hörner, 2010; Awaya and Krishna, 2015, 2017; Pai et al., 2016) provide non-trivial upper bounds on equilibrium payoffs in repeated private monitoring games *as a function of the monitoring structure*; in contrast, our bound applies for any monitoring structure. Finally, Sekiguchi (2005) notes a recursive structure of correlated equilibria in repeated games that is somewhat reminiscent of our approach.

2. Repeated games with private monitoring

A stage game $G = (I, A, u) = \left(I, (A_i, u_i)_{i \in I}\right)$ is repeated in periods $t = 1, 2, \ldots$, where $I = \{1, \ldots, n\}$ is the set of players, A_i is the set of player i's actions, and $u_i : A \to \mathbb{R}$ is player i's payoff function. Assume that each A_i is a non-empty, compact metric space and that each u_i is continuous. A *mixed action* for player i (denoted $\alpha_i \in \Delta (A_i)$) is a probability distribution on the Borel subsets of A_i, and a *correlated action* (denoted $\alpha \in \Delta (A)$) is a probability distribution on the Borel subsets of A. Payoff functions are extended to mixed and correlated actions in the usual way. Let $\mathcal{F} = \mathrm{co} \left(u (A)\right)$ denote the convex hull of the feasible payoff set. Players maximize discounted expected payoffs with common discount factor $\delta \in (0, 1)$.

In each period t, the game proceeds as follows: Each player i takes an action $a_{i,t} \in A_i$. A signal $y_t = \left(y_{i,t}\right)_{i \in I} \in \prod_{i \in I} Y_i = Y$ is drawn from distribution $p (y_t | a_t)$, where Y_i is the set of player i's signals and (Y, p) is the monitoring structure. Player i observes $y_{i,t}$. For brevity, we usually index a monitoring structure (Y, p) by p alone, leaving Y defined implicitly as the range of p. We restrict attention to monitoring structures where Y is a non-empty, compact metric space and $p (Z | a)$ is a measurable function of a for all Borel sets $Z \subseteq Y$.

A period t history for player i is an element of $H_i^t = (A_i \times Y_i)^{t-1}$, with typical element $h_i^t = \left(a_{i,\tau}, y_{i,\tau}\right)_{\tau=1}^{t-1}$, where H_i^1 consists of the null history $\emptyset$. A (behavior) strategy for player i is a sequence $\sigma_i = \left(\sigma_i^t\right)_{t=1}^\infty$, where $\sigma_i^t : H_i^t \to \Delta (A_i)$ is a measurable function. As the composition of (Borel) measurable functions is measurable, the assumption that $p (Z | a)$ is measurable implies that, for any strategy profile $\sigma = (\sigma_i)_{i \in I}$, the induced distribution over period t histories and payoffs is well-defined.

Let $E (\delta, p)$ be the set of Nash equilibrium payoffs in the repeated game with discount factor δ and monitoring structure p. Let p_0 denote the perfect monitoring structure given by $Y = (A)_{i \in I}$ and $p (y | a) = 1_{\left\{y = (a)_{i \in I}\right\}}$, where $1_{\{\cdot\}}$ is the indicator function.

3. General bounds

3.1. The basic result

Let $\underline{u}_i = \min_{\alpha_{-i} \in \Delta (A_{-i})} \max_{a_i \in A_i} u_i (a_i, \alpha_{-i})$ be player i's correlated minimax payoff, and let $\left(a_i^{\min}, \alpha_{-i}^{\min}\right)$ be a solution for the problem of minimaxing player i. For $\alpha \in \Delta (A)$, let α_{-i} denote the marginal of α over A_{-i}, and let $d_i (\alpha) = \max_{a_i \in A_i} u_i (a_i, \alpha_{-i})$ be player i's maximum deviation payoff against α.

The following is the key definition:

Definition 1. Let $V^* (\delta)$ denote the set of payoff vectors $v \in \mathbb{R}^n$ with the following property: there exists $(\alpha_t)_{t=1}^\infty \in \Delta (A)^\infty$ such that $v = (1 - \delta) \sum_{t \geq 1} \delta^{t-1} u (\alpha_t)$ and, for all i and t, we have

$$(1 - \delta) \sum_{\tau \geq t} \delta^{\tau - t} u_i(\alpha_\tau) \geq (1 - \delta) d_i(\alpha_t) + \delta \underline{u}_i. \tag{1}$$

The interpretation of (1) is that, if in every period players can either follow the (correlated) action path $(\alpha_t)_{t=1}^\infty$ or deviate and then be minimaxed, (1) says that players prefer to follow $(\alpha_t)_{t=1}^\infty$. Note that this interpretation relies on "observable mixed actions," as in (1) $u_i(\alpha_t)$ is evaluated only at α_t and not separately for each $a_{i,t} \in \text{supp}\,\alpha_{i,t}$. We formalize the connection between the set $V^*(\delta)$ and the equilibrium payoff set with observable mixed actions in Section 5.

The set $V^*(\delta)$ admits a recursive characterization a la APS and FLM.

Proposition 1. *Given a set of payoff vectors $W \subseteq \mathbb{R}^n$, let $B^*(W)$ be the set of payoff vectors v such that there exist $\alpha \in \Delta(A)$ and $w \in W$ with $v = (1 - \delta) u(\alpha) + \delta w$ and*

$$(1 - \delta) u_i(\alpha) + \delta w_i \geq (1 - \delta) d_i(\alpha) + \delta \underline{u}_i \quad \text{for all } i. \tag{2}$$

Then $V^*(\delta)$ is the largest bounded fixed point of the operator B^*, and $V^*(\delta) = \lim_{m \to \infty} (B^*)^m(\mathcal{F})$. In addition, $V^*(\delta)$ is a compact and convex set.

Proof. Standard; see Theorems 1 and 2 of APS or Section 7.3 of Mailath and Samuelson (2006). Note in particular that B^* preserves convexity because $\Delta(A)$ is a convex set, u_i is linear in α, and d_i is convex in α (as the max operator is convex).[3] □

The set $V^*(\delta)$ offers an upper bound for the (convex hull of the) equilibrium payoff set, for any discount factor and monitoring structure. This simple result is the basis of all the subsequent analysis.

Theorem 1. *For every discount factor δ and monitoring structure p, $\text{co}(E(\delta, p)) \subseteq V^*(\delta)$.*

The intuition for Theorem 1 is that, if we fix a Nash equilibrium σ of the repeated game with some monitoring structure and let α_t be the induced distribution over period t actions, then (1) is the relaxed period t incentive constraint that would result if player i had to decide which action to deviate to in period t without observing her period t history.

Proof. Fix a Nash equilibrium σ of the repeated game with some monitoring structure and let α_t be the induced distribution over period t actions.[4] We show that $(\alpha_t)_{t=1}^\infty$ satisfies (1) for all i and t. As $V^*(\delta)$ is convex, this completes the proof.

Fix i and t. Since a player can guarantee her minimax payoff starting from any history, the fact that σ is a Nash equilibrium implies that, for almost every on-path history h_i^t,[5]

---

[3] The reader might also ask if the presence of continuum action sets introduces any complications. The answer is no: the difficulty in extending APS to continuous actions or mixed strategies arises when continuation payoffs must be defined following each of a continuum of signals, as the set of (infinite) vectors of continuation payoffs then fails to be sequentially compact, and hence APS's operator B may fail to preserve compactness. Here, a correlated action $\alpha \in \Delta(A)$ is enforced by only a single promised continuation payoff vector w, and it is straightforward to see that the operator B^* preserves compactness even with continuous actions.

[4] As will become clear, the proof remains valid if α_t is only a communication equilibrium distribution (Forges, 1986; Myerson, 1986) rather than a Nash equilibrium distribution.

[5] That is, for every history h_i^t contained in a set of histories $\bar{H}_i^t \subseteq H_i^t$ of measure 1 under the distribution over histories induced by σ.

$$(1 - \delta) \sum_{\tau \geq t} \delta^{\tau - t} \mathbb{E}^{\sigma} \left[u_i \left(a_{\tau} \right) | h_i^t \right] \geq (1 - \delta) \max_{a_i} \mathbb{E}^{\sigma} \left[u_i \left(a_i, a_{-i,t} \right) | h_i^t \right] + \delta \underline{u}_i,$$

where $\mathbb{E}^{\sigma} \left[\cdot | \cdot \right]$ denotes conditional expectation with respect to the sigma-algebra generated by player i histories h_i^t, under the distribution on complete histories $\prod_i h_i^t$ induced by σ. Taking the ex ante expectation of both sides of this inequality with respect to h_i^t, we have

$$(1 - \delta) \sum_{\tau \geq t} \delta^{\tau - t} \mathbb{E}^{\sigma} \left[\mathbb{E}^{\sigma} \left[u_i \left(a_{\tau} \right) | h_i^t \right] \right] \geq (1 - \delta) \mathbb{E}^{\sigma} \left[\max_{a_i} \mathbb{E}^{\sigma} \left[u_i \left(a_i, a_{-i,t} \right) | h_i^t \right] \right] + \delta \underline{u}_i.$$

By the law of iterated expectation, $\mathbb{E}^{\sigma} \left[\mathbb{E}^{\sigma} \left[u_i \left(a_t \right) | h_i^t \right] \right] = u_i \left(\alpha_t \right)$. Since the max operator is convex, Jensen's inequality implies

$$\mathbb{E}^{\sigma} \left[\max_{a_i} \mathbb{E}^{\sigma} \left[u_i \left(a_i, a_{-i,t} \right) | h_i^t \right] \right] \geq \max_{a_i} \mathbb{E}^{\sigma} \left[\mathbb{E}^{\sigma} \left[u_i \left(a_i, a_{-i,t} \right) | h_i^t \right] \right]$$
$$= \max_{a_i} \mathbb{E}^{\sigma} \left[u_i \left(a_i, a_{-i,t} \right) \right]$$
$$= d_i \left(\alpha_t \right),$$

where the second line again follows from the law of iterated expectation. Combining these observations yields (1). $\square$

A helpful way to understand the bound $V^* \left(\delta \right)$ is to compare it to the equilibrium payoff set under perfect monitoring, $E \left(\delta, p_0 \right)$. The set $E \left(\delta, p_0 \right)$ also admits a simple recursive characterization (due to APS and FLM), and Theorem 1 implies that $E \left(\delta, p_0 \right) \subseteq V^* \left(\delta \right)$. However, for a variety of reasons that we now discuss, it is not true that $E \left(\delta, p \right) \subseteq E \left(\delta, p_0 \right)$ for all monitoring structures p.[6] Thus, to bound the equilibrium payoff set for any monitoring structure, we are led to use the more permissive set $V^* \left(\delta \right)$.

Why does perfect monitoring fail to bound $E \left(\delta, p \right)$ for some p, and how does $V^* \left(\delta \right)$ avoid this shortcoming? First, observing noisy private signals gives players a way to correlate their actions, which is impossible under perfect monitoring. A more plausible bound for $E \left(\delta, p \right)$ is thus the equilibrium payoff set with perfect monitoring *and a mediator*, where the mediator observes all actions and issues private action recommendations to players each period.[7] Denoting the equilibrium payoff set under mediated perfect monitoring by $E_{\text{med}} \left(\delta \right)$, it is clear that $E_{\text{med}} \left(\delta \right) \subseteq V^* \left(\delta \right)$: a player can still guarantee her minimax payoff starting from any history in the mediated game, so the proof of Theorem 1 goes through. However, $E_{\text{med}} \left(\delta \right)$ is still not an upper bound on $E \left(\delta, p \right)$ for all p:

Claim 1 *(Sugaya and Wolitzky (2017), Proposition 1). For some game G, discount factor δ, and monitoring structure p, $E \left(\delta, p \right) \nsubseteq E_{\text{med}} \left(\delta \right)$.*

The intuition for this result is that, compared to perfect monitoring with mediation, imperfect private monitoring has the advantage of "pooling players' information sets." Under perfect monitoring of actions, a player can perfectly infer what recommendations were made to the other players on the equilibrium path, and can tailor potential deviations to these recommendations. However, under imperfect private monitoring, a player cannot infer her opponents' actions—if

---

[6] This is well-known. See, for example, Exercise 5.10 in Fudenberg and Tirole (1991).

[7] See Sugaya and Wolitzky (2017) for a formal exposition of repeated games with mediated perfect monitoring.

these actions are stochastic along the equilibrium path—and thus has access to coarser information when contemplating a deviation. This difference can make it easier to sustain a given stochastic equilibrium path under private monitoring.

In contrast, the construction of the set $V^*(\delta)$ is based only on players' ex ante incentive constraints: for any monitoring structure, in any equilibrium it must be unprofitable for a player to wait until some period t and then deviate to a fixed action a_i regardless of her signal observations. In Section 6, we explore how the payoff bound $V^*(\delta)$ can be tightened by imposing more refined incentive constraints.

3.2. Symmetric games and stationary action paths

An action path $(\alpha_t)_{t=1}^\infty$ is *stationary* if $\alpha_t = \alpha_{t'}$ for all t, t'. As we will see, in general it is not without loss to restrict attention to stationary action paths when computing $V^*(\delta)$. However, if the game is symmetric then the best symmetric payoff vector in $V^*(\delta)$ is always attained at a stationary (and symmetric) action path.[8] This result makes our approach particularly sharp for symmetric games.

Proposition 2. *If G is symmetric and v^* is the greatest symmetric vector in $V^*(\delta)$, then there exists a symmetric correlated action $\alpha \in \Delta(A)$ such that $u(\alpha) = v^*$ and, for all i,*

$$u_i(\alpha) \geq (1 - \delta) d_i(\alpha) + \delta \underline{u}_i. \tag{3}$$

Proof. Fix $(\alpha_t)_{t=1}^\infty$ such that (1) holds and $v^* = (1 - \delta) \sum_{t \geq 1} \delta^{t-1} u(\alpha_t)$. Let $w(\alpha_t) = \sum_{i \in I} u_i(\alpha_t)$, and let $w = \sup_t w(\alpha_t)$. Note that $(w/n, \ldots, w/n) \geq v^*$. It suffices to show that there exists $\bar{\alpha}$ such that $u_i(\bar{\alpha}) = w/n$ and $u_i(\bar{\alpha}) \geq (1 - \delta) d_i(\bar{\alpha}) + \delta \underline{u}_i$ for all i.

Note that, for all t,

$$w \geq \sum_i (1 - \delta) \sum_{\tau \geq t} \delta^{\tau - t} u_i(\alpha_\tau) \geq \sum_i \left[(1 - \delta) d_i(\alpha_t) + \delta \underline{u}_i \right]$$

(by (1)). As payoffs are continuous and bounded, there exists $\alpha \in \Delta(A)$ such that $\sum_i u(\alpha) = w$ and $w \geq \sum_i \left[(1 - \delta) d_i(\alpha) + \delta \underline{u}_i \right]$.

Let Π be the set of all permutations π on I. For each $\pi \in \Pi$, let α^π be the correlated action given by $\alpha^\pi(a_1, \ldots, a_n) = \alpha\left(a_{\pi(1)}, \ldots, a_{\pi(n)}\right)$. Finally, let $\bar{\alpha} = (1/n!) \sum_{\pi \in \Pi} \alpha^\pi$.

Then, for all i, $u_i(\bar{\alpha}) = w/n$ and (by convexity of the max operator) $d_i(\bar{\alpha}) \leq (1/n) \sum_i d_i(\alpha)$. Since $w \geq \sum_i \left[(1 - \delta) d_i(\alpha) + \delta \underline{u}_i \right]$, and since $\underline{u}_i$ is the same for all i, it follows that $u_i(\bar{\alpha}) \geq (1 - \delta) d_i(\bar{\alpha}) + \delta \underline{u}_i$ for all i. $\quad\square$

3.3. Examples

We illustrate Theorem 1 and Proposition 2 with examples. We first show that V^* is not a tight bound on the equilibrium payoff set in general, and then show that it is nonetheless useful for analyzing some canonical repeated games. Additional examples are given in Section 4.

---

[8] A game G is symmetric if $A_i = A_j$ for all $i, j \in I$ and $u_1(a_1, \ldots, a_n) = u_{\pi(1)}\left(a_{\pi(1)}, \ldots, a_{\pi(n)}\right)$ for every permutation π on I. A payoff vector is symmetric if its components are all the same. A distribution $\alpha \in \Delta(A)$ is symmetric if $\alpha(a_1, \ldots, a_n) = \alpha\left(a_{\pi(1)}, \ldots, a_{\pi(n)}\right)$ for every permutation π on I.

3.3.1. No tightness in general

There exist games for which $E(\delta, p)$ is a strict subset of V^* for every monitoring structure p. In other words, in general V^* is not a tight bound on the equilibrium payoff set, even from the perspective of an observer who does not know the monitoring structure.

The logic is similar to the reason why allowing observable mixed actions can expand the equilibrium set in repeated games with perfect monitoring. For example, consider the following game:

	L	R
T	1, 1	0, 2
B	1, −1	0, 0

If $\delta < 1/3$, then player 2 cannot play L in any equilibrium of the repeated game, since the deviation gain of $1 - \delta$ would exceed the greatest possible change in continuation payoff (from $\delta(2)$ to $\delta(0)$, noting that 0 is the minimax payoff). Hence, the best payoff in $E(\delta, p)$ for player 1 is 0, for any monitoring structure p.

On the other hand, the payoff vector $(1/2, 3/2)$ is included in $V^*(\delta)$ for all $\delta \geq 1/4$: this may be seen by taking $\alpha_t = \frac{1}{2}(T, L) + \frac{1}{2}(T, R)$ for all t. Intuitively, with observable mixtures, if $\frac{1}{2}(T, L) + \frac{1}{2}(T, R)$ is to be played each period and deviators are minimaxed, then player 2's deviation gain is $(1 - \delta)/2 \leq 3/8$, and his lost continuation payoff from a deviation is $\delta(3/2) \geq 3/8$.

3.3.2. Prisoners' Dilemma

Consider the canonical two-player repeated prisoners' dilemma:

	C	D
C	1, 1	−l, 1 + g
D	1 + g, −l	0, 0

where $g, l \geq 0$ and $g \leq l + 1$. If $\delta \geq g/(1 + g)$, repeated (C, C) is an equilibrium outcome under perfect monitoring, yielding the best symmetric payoff of $(1, 1)$. Sorin (1986) shows that, if $g = l$, then under perfect monitoring the equilibrium payoff set collapses to $\{(0, 0)\}$ as soon as δ falls below $g/(1 + g)$. Stahl (1991) extends this result to $g \leq l$, while assuming that public randomization is available and restricting attention to pure strategy equilibria. We use Proposition 2 to show that the same result holds for any monitoring structure (and without restricting to pure strategies).[9]

Claim 2. *If $g \leq l$ and $\delta < g/(1 + g)$, then $E(\delta, p) = \{(0, 0)\}$ for every monitoring structure p.*

Proof. Suppose $E(\delta, p) \neq \{(0, 0)\}$. Since the minimax payoff is 0, $E(\delta, p) \subseteq \mathbb{R}_+^2$. Since $E(\delta, p)$ is symmetric about the 45° line, by Theorem 1 there is a symmetric payoff vector $v \in V^*(\delta)$ with $v > 0$. Hence, by Proposition 2, there exists a symmetric correlated action α such that $u_i(\alpha) > 0$ and $u_i(\alpha) \geq (1 - \delta) d_i(\alpha)$. That is,

$$\alpha_{CC} + (1 + g)\alpha_{DC} - l\alpha_{CD} \geq (1 - \delta)(1 + g)(\alpha_{CC} + \alpha_{DC}). \tag{4}$$

---

[9] Note that allowing explicit public randomization would not affect the following results, as public randomization can be modeled as part of the monitoring structure.

By symmetry, $\alpha_{CD} = \alpha_{DC}$. If $g \leq l$, this implies

$$\alpha_{CC} + \alpha_{DC} \geq (1 - \delta)(1 + g)(\alpha_{CC} + \alpha_{DC}).$$

Finally, as $u_i(\alpha) > 0$ implies $\alpha_{CC} + \alpha_{DC} > 0$, it follows that $\delta \geq g/(1 + g)$. □

If instead $g \geq l$, Stahl shows that $E(\delta, p_0) = \{(0, 0)\}$ whenever $\delta < l/(1 + g)$. We can also extend this result to any monitoring structure, as well as characterize the best symmetric payoff vector in $V^*(\delta)$ more generally.[10]

Claim 3. *If $g \geq l$ and $\delta < l/(1 + g)$, then $E(\delta, p) = \{(0, 0)\}$ for every monitoring structure p. If $g \geq l$ and $\delta \in [l/(1 + g), g/(1 + g))$, then the best symmetric payoff vector in $V^*(\delta)$ is $\left(\frac{(1-\delta)(1+g)(g-l)}{2g-\delta(1+g)-l}, \frac{(1-\delta)(1+g)(g-l)}{2g-\delta(1+g)-l} \right)$.*

Proof. As in the $g \leq l$ case, if $E(\delta, p) \neq \{(0, 0)\}$ then there must exist $\alpha \in \Delta(A)$ with $u_i(\alpha) > 0$ and $\alpha_{CD} = \alpha_{DC}$ that satisfies (4). When $g \geq l$, $\alpha_{CD} = \alpha_{DC}$ and (4) imply

$$(1 + g - l)(\alpha_{CC} + \alpha_{DC}) \geq (1 - \delta)(1 + g)(\alpha_{CC} + \alpha_{DC}).$$

As $u_i(\alpha) > 0$ implies $\alpha_{CC} + \alpha_{DC} > 0$, it follows that $\delta \geq l/(1 + g)$.

Moreover, if $\delta \geq l/(1 + g)$ then the best symmetric payoff in $V^*(\delta)$ is given by choosing α_{CC} and $\alpha_{CD} = \alpha_{DC}$ so as to maximize the left-hand side of (4), subject to the constraints $\alpha_{CC} \geq 0$, $\alpha_{DC} \geq 0$, $\alpha_{CC} + 2\alpha_{DC} \leq 1$, and (4). It is straightforward to check that the solution to this linear program is given by

$$\left(\alpha_{CC} = \frac{\delta(1+g) - l}{2g - \delta(1+g) - l}, \alpha_{DC} = \frac{g - \delta(1+g)}{2g - \delta(1+g) - l} \right)$$

if $\delta < g/(1 + g)$, and by $\alpha_{CC} = 1$ if $\delta \geq g/(1 + g)$. The resulting payoffs can then be read off of (4). □

Claims 2 and 3 completely characterize the best symmetric payoff vector in $V^*(\delta)$ for the prisoners' dilemma. For asymmetric payoffs, the APS-style algorithm given by Proposition 1 can be simplified substantially in the $g \leq l$ case. Stahl shows that the exact folk theorem holds with pure strategies and public randomization when $\delta \geq l/(1 + l)$; we therefore assume $\delta \in [g/(1 + g), l/(1 + l))$. In this case, we can show that the extreme points of the set $B^*(W)$ are always generated by distributions α satisfying the following properties (in what follows, without loss we focus on extreme points v with $v_1 \geq v_2$): (i) Either $\alpha_{CC} = 1$ or player 2's incentive constraint is binding. (ii) $\alpha_{CD} = 0$. (iii) Either $\alpha_{CC} = 0$ or $\alpha_{DD} = 0$. Using these facts, it is simple to calculate $V^*(\delta)$ for any vector of parameters (g, l, δ) with $g \leq l$. For example, with $g = 0.8$, $l = 1$, and $\delta = 0.475$, we find

$$V^*(\delta) = \text{co}\{(0, 0), (0, 1.295), (0.45, 1.22), (1, 1), (1.22, 0.45), (1.295, 0)\}.$$
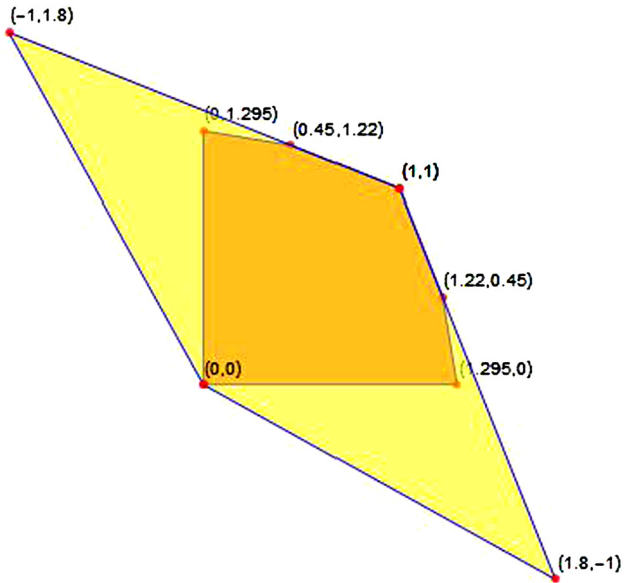
See Fig. 1.

Fig. 1. $\mathcal{F}$ and $V^*(\delta)$ for the prisoners' dilemma with $g = 0.8$, $l = 1$, and $\delta = 0.475$.

In contrast, when $g \geq l$, an analogous simplification of the APS-style algorithm does not seem to be available. The difficulty is that increasing the probability of defection now has the benefit of reducing players' deviation gains, so extreme points of $B^*(W)$ can be generated by a richer set of action distributions.

As an additional remark, we note that if $g \geq l$ and $\delta \geq g/(1+g)$, then the point in $V^*(\delta)$ that maximizes player 1's payoff cannot be attained by a stationary action path. To see this, first observe that any action path that maximizes player 1's payoff subject to (1) involves (D, C) with probability 1 in period 1: this follows because having player 1 play D increases her payoff while reducing player 2's deviation gain. But it is clear that such an action path cannot be stationary, as the infinite repetition of (D, C) leaves player 2 with less than his minimax payoff.[11]

3.3.3. A public good game

Consider the n-player symmetric game with $A_i = \mathbb{R}_+$ and

$$u_i(a) = f\left(\sum_{j=1}^{n} a_j\right) - a_i, \qquad (5)$$

where f is an increasing and concave function with $f(0) = 0$ and $f'(x) < 1$ for all x.[12] Note that $a_i = 0$ is dominant in the stage game.

We wish to find an upper bound on symmetric equilibrium payoffs in this game that remains valid regardless of the monitoring structure under which the game is played. We proceed by

---

[11] This is similar to the non-stationary of optimal asymmetric equilibria in repeated prisoners' dilemma-type games uncovered by Abreu (1986).

[12] The set of actions in this game is not compact and payoffs are unbounded. However, it will become clear that imposing a sufficiently high upper bound on actions would have no effect on the analysis.

characterizing v^* (the best symmetric payoff vector in $V^*(\delta)$), which gives one such bound. We then compare this bound to an alternative, more obvious bound, and show that our bound is tighter whenever the discount factor is above a cutoff.

By Proposition 2, we know that v^* is attained by the infinite repetition of a symmetric action $\alpha \in \Delta(A)$. We first show that this α puts weight only on vectors a with $a_i = 0$ for all but a single player i: that is, it is never optimal to have two players work at the same time.

To see this, let $\tilde{a}_{-i} = \sum_{j \neq i} a_j$, and note that player i's deviation gain from α is given by

$$\mathbb{E}[a_i] + \mathbb{E}\left[f(\tilde{a}_{-i})\right] - \mathbb{E}\left[f(a_i + \tilde{a}_{-i})\right].$$

A necessary condition for α to be optimal is thus that it maximizes $\mathbb{E}\left[f(a_i + \tilde{a}_{-i}) - f(\tilde{a}_{-i})\right]$ for given $\mathbb{E}[a_i]$ (for each i). As f is concave, $f(a_i + \tilde{a}_{-i}) - f(\tilde{a}_{-i})$ is submodular in a_i and $\tilde{a}_{-i}$, so α must have the property that $a_i > 0$ if and only if $\tilde{a}_{-i} = 0$.[13]

It remains only to find the optimal effort level a^* such that randomly having one player work at level a^* each period is optimal. (Of course, a player does not know if she will be the one to work when we calculate (1).) Assuming a^* is below the first-best level (defined by $nf'(na) = 1$), this is given by

$$f(a^*) - \frac{1}{n}a^* = (1 - \delta)\frac{n-1}{n}f(a^*),$$

or

$$a^* = f(a^*) + \delta(n-1)f(a^*).$$

The best symmetric payoff vector in $V^*(\delta)$ thus gives each player payoff $f(a^*) - a^*/n$, with this value of a^*. This payoff is therefore an upper bound on the best symmetric payoff attainable in a repeated game Nash equilibrium for any monitoring structure.

While this payoff bound is permissive, it is not obvious (at least to us) how to give a tighter bound in this game. One obvious upper bound on each player's action is given by the value $\bar{a}$ that solves

$$(1 - \delta)[\bar{a} - f(\bar{a})] = \delta f((n-1)\bar{a}),$$

or

$$\bar{a} = f(\bar{a}) + \frac{\delta}{1 - \delta}f((n-1)\bar{a}).$$

This bound follows because a player's minimax payoff is 0, the instantaneous gain from deviating from $a_i = \bar{a}$ to $a_i = 0$ is at least $(1 - \delta)[\bar{a} - f(\bar{a})]$, and—if $\bar{a}$ is the least upper bound on a player's equilibrium action—a player's continuation payoff is at most $\delta f((n-1)\bar{a})$. Comparing a^* to $\bar{a}$, we see that $a^* < \bar{a}$ if and only if δ is above some cutoff $\bar{\delta}$ (noting that $(n-1)f(a) > f((n-1)a)$ for all $a > 0$, as f is concave and $f(0) = 0$). Thus, for sufficiently patient players the payoff bound coming from Proposition 2 is tighter than the obvious bound.

4. Concave games

We now introduce a class of games for which $V^*(\delta)$ is a tight bound on the equilibrium payoff set under perfect monitoring, at least for the Pareto frontier (in contrast to the first example

---

[13] This may be shown formally using a standard rearrangement inequality, such as Lorentz (1953).

above). A corollary is that perfect monitoring is the optimal monitoring structure in this class of games. Knowing that $V^*(\delta)$ is tight bound is useful from several perspectives. From the perspective of robust prediction, $V^*(\delta)$ characterizes the set of payoffs that can arise in equilibrium from the viewpoint of an observer who knows the physical environment but not the information structure. From the perspective of information design, perfect monitoring is the optimal information structure under which to play the game.

For $\alpha \in \Delta(A)$, define player i's maximum deviation gain at α by

$$\bar{d}_i(\alpha) := d_i(\alpha) - u_i(\alpha).$$

Definition 2. A game G is *concave* if the following conditions hold for all $i \in I$:

- A_i is a compact and convex subset of $\mathbb{R}^m$ for some $m \in \mathbb{N}$.
- $u_i(a)$ is continuous and jointly concave: for $a, a' \in A$ and $\lambda \in (0, 1)$, $u_i(\lambda a + (1 - \lambda) a') \geq \lambda u_i(a) + (1 - \lambda) u_i(a')$.
- $\bar{d}_i(\alpha)$ increases with risk: for $\alpha \in \Delta(A)$, $\bar{d}_i(\mathbb{E}^\alpha[a]) \leq \bar{d}_i(\alpha)$, where $\mathbb{E}^\alpha[a]$ denotes the Dirac distribution on the pure action profile with $a_i = \sum_{a'_i} \alpha_i(a'_i) a'_i$ for all i.[14]

Our main result for concave games is as follows:

Theorem 2. *Suppose G is concave, and suppose that, with perfect monitoring, for each player i there exists a subgame perfect equilibrium that gives player i payoff $\underline{u}_i$. Then, for every non-negative Pareto weight $\lambda \in \Lambda_+ := \{\lambda \in \mathbb{R}^n_+ : \|\lambda\| = 1\}$, we have*

$$\max_{v \in E(\delta, p_0)} \lambda \cdot v = \max_{v \in V^*(\delta)} \lambda \cdot v.$$

There are three parts to the reasoning underlying Theorem 2. First, deterministic action paths are optimal in concave games, as replacing the ex ante distribution over period t actions with its expectation increases payoffs (as u_i is concave) and reduces the temptation to deviate (as $\bar{d}_i$ increases with risk). Second, for given punishment payoffs, perfect monitoring is the optimal monitoring structure for supporting deterministic action paths, as revealing actions gives no new information on the equilibrium path. Third, the availability of minimax punishments under perfect monitoring ensures that punishments can be as harsh as possible.

Proof. As $E(\delta, p_0) \subseteq V^*(\delta)$, it suffices to show that $\max_{v \in E(\delta, p_0)} \lambda \cdot v \geq \max_{v \in V^*(\delta)} \lambda \cdot v$ for all $\lambda \in \Lambda_+$. To see this, fix $v \in V^*(\delta)$, and fix $(\alpha_t)_{t=1}^\infty$ such that $v = (1 - \delta) \sum_{t \geq 1} \delta^{t-1} u(\alpha_t)$ and (1) holds for all i and t. Consider the following strategy profile under perfect monitoring:

- On path, player i plays $\mathbb{E}^{\alpha_t}[a_{i,t}] \in A_i$ in period t.
- If player i unilaterally deviates (on or off path), switch to a subgame perfect equilibrium that gives her payoff $\underline{u}_i$. Ignore simultaneous deviations.

---

[14] This terminology is by analogy with expected utility theory: an expected utility function $U : \Delta(\mathbb{R}) \to \mathbb{R}$ represents risk-loving preferences if and only if $U(\mathbb{E}^\alpha[a]) \leq U(\alpha)$ for every distribution $\alpha \in \Delta(\mathbb{R})$. Note also that $\mathbb{E}^\alpha[a] \in A$ because A is convex.

Since u_i is concave for all i, we have $(1 - \delta) \sum_{t \geq 1} \delta^{t-1} u \left(\mathbb{E}^{\alpha_t} \left[a_{1,t} \right], \ldots, \mathbb{E}^{\alpha_t} \left[a_{n,t} \right] \right) \geq v$, and hence $\lambda \cdot (1 - \delta) \sum_{t \geq 1} \delta^{t-1} u \left(\mathbb{E}^{\alpha_t} \left[a_{1,t} \right], \ldots, \mathbb{E}^{\alpha_t} \left[a_{n,t} \right] \right) \geq \lambda \cdot v$ for all $\lambda \in \Lambda_+$. It thus suffices to show that this strategy profile is a subgame perfect equilibrium.

Subtracting $(1 - \delta) u_i (\alpha_t)$ from both sides of (1), we have

$$(1 - \delta) \sum_{\tau \geq t+1} \delta^{\tau - t} u_i (\alpha_\tau) \geq (1 - \delta) \left[d_i (\alpha_t) - u_i (\alpha_t) \right] + \delta \underline{u}_i.$$

Since $u_i (\alpha_\tau) \leq u_i (\mathbb{E}^{\alpha_\tau} [a_\tau])$ (by concavity of u_i) and $d_i (\alpha_t) - u_i (\alpha_t) \geq d_i (\mathbb{E}^{\alpha_t} [a_t]) - u_i (\mathbb{E}^{\alpha_t} [a_t])$ (as $\bar{d}_i$ increases with risk), we have

$$(1 - \delta) \sum_{\tau \geq t+1} \delta^{\tau - t} u_i \left(\mathbb{E}^{\alpha_\tau} [a_\tau] \right) \geq (1 - \delta) \left[d_i \left(\mathbb{E}^{\alpha_t} [a_t] \right) - u_i \left(\mathbb{E}^{\alpha_t} [a_t] \right) \right] + \delta \underline{u}_i.$$

This is precisely the condition for the constructed strategy profile to be a subgame perfect equilibrium under perfect monitoring. $\square$

We also introduce the notion of a *symmetric-concave* game. (Recall that symmetric games and distributions are defined in footnote 8.)

Definition 3. A game G is *symmetric-concave* if it is symmetric and the following conditions hold for all $i \in I$:

- A_i is a compact and convex subset of $\mathbb{R}^m$ for some $m \in \mathbb{N}$.
- $u_i (a)$ is continuous.
- If $\alpha \in \Delta (A)$ is symmetric, then $u_i (\mathbb{E}^\alpha [a]) \geq u_i (\alpha)$.
- If $\alpha \in \Delta (A)$ is symmetric, then $\bar{d}_i (\mathbb{E}^\alpha [a]) \leq \bar{d}_i (\alpha)$.

The property that a game is symmetric-concave is slightly weaker than the condition that it is both symmetric and concave, in that the required concavity/convexity conditions on u_i and $\bar{d}_i$ are imposed only for symmetric distributions α. This distinction can be important: as we will see, both Bertrand and Cournot competition with linear demand curves are symmetric-concave but not concave. Note that, for symmetric G and α, the last two bullet points in the definition are equivalent to $\sum_i u_i (\mathbb{E}^\alpha [a]) \geq \sum_i u_i (\alpha)$ and $\sum_i \bar{d}_i (\mathbb{E}^\alpha [a]) \leq \sum_i \bar{d}_i (\alpha)$, respectively.

If G is symmetric-concave, optimal symmetric equilibria are both stationary and deterministic (i.e., in pure strategies). As we will see, this makes computing optimal symmetric equilibria easy.

Proposition 3. *If G is symmetric-concave and v^* is the greatest symmetric vector in $V^* (\delta)$, then there exists a symmetric pure action profile $a \in A$ such that $u (a) = v^*$ and, for all i,*

$$u_i (a) \geq (1 - \delta) d_i (a) + \delta \underline{u}_i. \tag{6}$$

If it is also the case that, with perfect monitoring, for each player i there exists a subgame perfect equilibrium that gives player i payoff $\underline{u}_i$, then the infinite repetition of such an action profile a is a subgame perfect equilibrium outcome with perfect monitoring.

Proof. By Proposition 2, there exists a symmetric correlated action $\alpha \in \Delta (A)$ such that $u (\alpha) = v^*$ and, for all i,

$$\delta u_i (\alpha) \geq (1 - \delta) \left[d_i (\alpha) - u_i (\alpha) \right] + \delta \underline{u}_i$$

(where we have subtracted $(1 - \delta) u_i (\alpha)$ from both sides of (3)). Letting $a = \mathbb{E}^\alpha [a']$, the conditions that $u_i (a) \geq u_i (\alpha)$ and $\bar{d}_i (a) \leq \bar{d}_i (\alpha)$ yield (6).

Finally, if deviators can be minimaxed under perfect monitoring, the stationary grim trigger strategy profile that implements the repetition of a on-path and minimaxes deviators is a subgame perfect equilibrium if and only if (6) holds.   □

How restrictive is the condition in Theorem 2 and Proposition 3 that there is an equilibrium that minimaxes a deviator? This condition is of course satisfied if the stage game admits a mutual minimax Nash equilibrium, and it is satisfied more generally if the discount factor is high enough to admit a minimaxing optimal penal code, as in Abreu (1986, 1988). We will see examples of both cases below. Another case in which deviators can be minimaxed is when the players have access to a mediator and $\delta > \delta^*$. See Lemma 1 of Sugaya and Wolitzky (2017) or the proof of Theorem 3 below for details.

In addition, if the worst subgame perfect equilibrium payoff under perfect monitoring is some $\hat{u}_i > \underline{u}_i$, the same arguments as above show that the conclusions of Theorem 2 and Proposition 3 continue to apply within the class of equilibria where players' continuation payoffs never fall below $\hat{u}_i$. Given the finding below that linear Bertrand and Cournot oligopoly games are symmetric-concave, this observation resolves an issue in the literature on collusion following Abreu (1986) (e.g., Lambson, 1987; Wernerfelt, 1989; Chang, 1991; Ross, 1992; Häckner, 1996; Compte et al., 2002), which, while focusing on computing optimal punishment paths under various assumptions, typically assumes without formal justification that the firms' goal is sustaining the best possible *stationary* path of play. At least for the linear demand curves typically considered in the literature, Proposition 3 provides a justification for this approach.

Finally, an antecedent for the conditions of concavity and symmetric-concavity are the curvature assumptions in Kloosterman (2015). Kloosterman proves that, under his conditions, giving players more information about future states shrinks the pure-strategy subgame perfect equilibrium payoff set in Markov games with perfect information about the current state and perfect monitoring of actions. Our conditions are not nested with Kloosterman's, and the two papers consider different settings (repeated games with imperfect monitoring here; Markov games with perfect information and monitoring in Kloosterman) and prove different results (optimality of perfect monitoring here; optimality of less information about future states in Kloosterman).

4.1. Examples

The following examples show that Theorem 2 and Proposition 3 can be useful for characterizing optimal equilibria in some canonical repeated games.

4.1.1. An additively separable public good game
Consider the symmetric game with $A_i = \mathbb{R}_+$ and payoffs

$$u_i (a) = \sum_{j=1}^{n} f (a_j) - a_i,$$

where f is an increasing and concave function with $f (0) = 0$ and $f' (x) < 1$ for all x. The game is thus an additively separable version of the public good game in Section 3.

This game is concave, because $u_i (a)$ is concave and

$$\bar{d}_i (\alpha) = u_i (a_i = 0, \alpha_{-i}) - u_i (\alpha) = \mathbb{E}^\alpha [a_i - f (a_i)] \geq \mathbb{E}^\alpha [a_i] - f (\mathbb{E}^\alpha [a_i]).$$

In addition, $a = (0)_{i \in I}$ is a mutual minimax Nash equilibrium in the stage game, so $\underline{u}_i = 0$ is a subgame perfect equilibrium payoff with perfect monitoring. Hence, Proposition 3 ensures that perfect monitoring is the optimal monitoring structure for maximizing symmetric payoffs, and that the optimal symmetric equilibrium is stationary and in pure strategies. Letting $\bar{a} \in \mathbb{R}_+$ satisfy the incentive constraint

$$nf(\bar{a}) - \bar{a} = (1 - \delta)(n - 1) f(\bar{a})$$

and letting $a^{FB} = (f')^{-1}(1/n)$ denote the first-best level of a, it follows that the optimal symmetric equilibrium is given by repeating the action profile $\left(\min\{\bar{a}, a^{FB}\}\right)_{i \in I}$.

(In contrast, note that the original public good game with payoffs $u_i(a) = f\left(\sum_{j=1}^{n} a_j\right) - a_i$ is not concave, as in that game

$$\bar{d}_i(\alpha) = u_i(a_i = 0, \alpha_{-i}) - u_i(\alpha) = \mathbb{E}^\alpha[a_i] + \mathbb{E}^\alpha\left[f(\bar{a}_{-i})\right] - \mathbb{E}^\alpha\left[f(a_i + \bar{a}_{-i})\right],$$

which does not increase with risk.[15] Indeed, $V^*(\delta)$ is not a tight bound in that game.)

4.1.2. Linear Cournot oligopoly

Consider n-player linear Cournot competition with zero production costs, where firm i sets quantity $q_i \in \mathbb{R}_+$, and the market price at quantity vector q equals

$$\max\left\{b - \sum_i q_i, 0\right\}$$

for a constant $b > 0$. Payoffs are thus given by

$$u_i(q) = \max\left\{\left(b - \sum_j q_j\right) q_i, 0\right\}.$$

We claim that, in computing $V^*(\delta)$, it is without loss to restrict attention to distributions with support consisting of quantity vectors that yield positive prices. To see this, first take an arbitrary distribution over quantity vectors α, and let $\bar{\alpha}$ be the modified distribution that results when, whenever a vector q with $\sum_i q_i \geq b$ is drawn from α, this vector is replaced with the vector $\bar{q} := (b, \ldots, b)$. Clearly, $u_i(\alpha) = u_i(\bar{\alpha})$ for all i, and $d_i(\alpha) \geq d_i(\bar{\alpha})$ for all i because other firms' quantities have increased. So it is without loss to restrict to distributions over vectors where $q_i = b$ for all i whenever $\sum_i q_i \geq b$. Finally, for any such distribution α, let $\beta = \Pr^\alpha\left(\sum_i q_i < b\right)$ and define a distribution $\hat{\alpha}$ by $\hat{\alpha}(q) = \alpha(q)/\beta$ if $\sum_i q_i < b$ and $\hat{\alpha}(q) = 0$ otherwise. Then $u_i(\alpha) = \beta u_i(\hat{\alpha})$ and $d_i(\alpha) = \beta d_i(\hat{\alpha})$. As $\underline{u}_i = 0$, (1) implies that $u_i(\hat{\alpha}) \in V^*(\delta)$ whenever $u_i(\alpha) \in V^*(\delta)$. Hence, it is without loss to restrict attention to distributions with $\beta = 1$ when computing $V^*(\delta)$.

We now show that, given the restriction to quantity vectors yielding positive prices, the game is symmetric-concave. To see this, note that replacing a symmetric distribution over q with its expectation increases u_i if and only if it increases $\sum_i u_i(q) = \left(b - \sum_i q_i\right) \sum_i q_i$. As $\sum_i u_i(q)$ depends on q only through the sum $\sum_i q_i$ and is concave in this sum, the required condition

---

[15] For example, if $f(x) = \log(1 + x)$ and we compare $\frac{1}{2}\left\{(a_i, \bar{a}_{-i}) = \left(\frac{1}{2}, 0\right)\right\} + \frac{1}{2}\left\{(a_i, \bar{a}_{-i}) = \left(\frac{1}{2}, 1\right)\right\}$ with $\left\{(a_i, \bar{a}_{-i}) = \left(\frac{1}{2}, \frac{1}{2}\right)\right\}$, the latter gives a higher value for $\bar{d}_i(\alpha)$.

on u_i holds. Similarly, letting $\tilde{q}$ denote a symmetric distribution over q, replacing $\tilde{q}$ with its expectation decreases $\bar{d}_i$ if and only if it decreases $\sum_i \bar{d}_i$. Note that

$$\bar{d}_i(\tilde{q}) = \frac{1}{4}\left(b - \mathbb{E}^{\tilde{q}_{-i}}\left[\sum_{j\neq i} q_j\right]\right)^2 - \mathbb{E}^{\tilde{q}}[u_i(q)],$$

so

$$\sum_i \bar{d}_i(\tilde{q}) = \frac{1}{4}\sum_i\left(b - \mathbb{E}^{\tilde{q}_{-i}}\left[\sum_{j\neq i} q_j\right]\right)^2 - \mathbb{E}^{\tilde{q}}\left[\sum_i u_i(q)\right].$$

Replacing $\tilde{q}$ with its expectation does not affect $\left(b - \mathbb{E}^{\tilde{q}_{-i}}\left[\sum_{j\neq i} q_j\right]\right)^2$ for each i and increases $\mathbb{E}^{\tilde{q}}\left[\sum_i u_i(q)\right]$, so the required condition on $\bar{d}_i$ is also satisfied.

In addition, in the current formulation without production costs, the game admits a mutual minimax Nash equilibrium, so there is always a subgame perfect equilibrium that minimaxes a deviator.[16]

Given these observations, Proposition 3 implies that perfect monitoring is the optimal monitoring structure for maximizing total industry profits in linear Cournot oligopoly, and that the optimal equilibrium is stationary and in pure strategies. Letting q^* satisfy the incentive constraint

$$(b - nq^*)q^* = (1 - \delta)\frac{1}{4}(b - (n-1)q^*)^2$$

and letting $q^m = b/2$ denote monopoly output, it follows that this equilibrium is given by the infinite repetition of production at the symmetric level $q_i = \max\{q^*, q^m/n\}$.

4.1.3. Linear Bertrand oligopoly

Consider n-player linear, differentiated-product Bertrand competition with zero production costs, where firm i sets price $p_i \in \mathbb{R}_+$, and firm i's demand at price vector p equals

$$\max\left\{b + c\sum_{j\neq i} p_j - p_i, 0\right\}$$

for constants $b, c > 0$. Payoffs are thus given by

$$u_i(p) = \max\left\{b + c\sum_{j\neq i} p_j - p_i, 0\right\} p_i.$$

Unlike in the Cournot case, it is no longer completely without loss to restrict attention to price vectors yielding positive quantities: for example, for some very high price $\bar{p} \gg b$, firms 1 and 2 might benefit from alternating between price vectors $(p_1 = \bar{p}, p_2 = (b + c\bar{p})/2)$ and $(p_1 = (b + c\bar{p})/2, p_2 = \bar{p})$. Such equilibria clearly rely on an unrealistic aspect of the linear demand system and would be ruled out if we instead specified demand as

---

[16] In the more realistic version with production costs, the game does not admit a mutual minimax static Nash equilibrium, but it nonetheless admits a subgame perfect equilibrium yielding payoff $\underline{u}$ if δ is sufficiently high. The argument is as in the Bertrand case (or see Abreu, 1986). The game with production costs also remains concave, as long as we restrict attention to equilibria with positive prices.

$$\max \left\{ b + c \sum_{j \neq i} \min \{ p_j, b \} - p_i, 0 \right\}.$$

With this modified demand system, in computing $V^*(\delta)$ it is clearly without loss to restrict attention to distributions where $p_i \in [0, b]$ for all i, in which case the game reduces to the original one with a restriction to price vectors yielding positive quantities. We therefore proceed under this restriction.

In particular, given the restriction to price vectors yielding positive quantities, the game is symmetric-concave if $c \leq 1/(n-1)$.[17] The argument is similar to the Cournot case. Note that

$$\sum_i u_i(p) = b \sum_i p_i + c \sum_i \sum_{j \neq i} p_i p_j - \sum_i p_i^2.$$

The Hessian matrix of $\sum_i u_i(p)$ thus consists of -2's on the diagonal and $2c$'s off the diagonal, so this matrix is negative semi-definite if and only if

$$-\left(\sum_i x_i^2 - c \sum_i \sum_{j \neq i} x_i x_j \right) = -(1 + c) \sum_i x_i^2 + c \left(\sum_i x_i \right)^2 \leq 0 \text{ for all } x \in \mathbb{R}^n.$$

This in turn holds if and only if $c \leq 1/(n-1)$. The required condition on u_i thus holds if and only if $c \leq 1/(n-1)$. In addition, letting $\tilde{p}$ denote a distribution over p, we have

$$\sum_i \bar{d}_i(\tilde{p}) = \frac{1}{4} \sum_i \left(b + c \mathbb{E}^{\tilde{p}_{-i}} \left[\sum_{j \neq i} p_j \right] \right)^2 - \mathbb{E}^{\tilde{p}} \left[\sum_i u_i(p) \right].$$

Replacing $\tilde{p}$ with its expectation does not affect $\left(b + c \mathbb{E}^{\tilde{p}_{-i}} \left[\sum_{j \neq i} p_j \right] \right)^2$ for each i and increases $\mathbb{E}^{\tilde{p}} \left[\sum_i u_i(p) \right]$ if and only if $c \leq 1/(n-1)$, so the required condition on $\bar{d}_i$ is satisfied if $c \leq 1/(n-1)$.

Unlike in the Cournot case, this game does not admit a mutual minimax Nash equilibrium. However, it follows from standard arguments (similar to Abreu, 1986) that it admits a subgame perfect equilibrium yielding payoff $\underline{u}_i$ whenever δ is sufficiently high. For example, this can be attained by a "stick-and-carrot" equilibrium path consisting of one period of zero prices followed by an infinite stream of constant, positive prices. Furthermore, we verify in Appendix B that (under an additional restriction on c) such a minimaxing stick-and-carrot equilibrium exists at a discount factor low enough such that first-best industry profits are unattainable, so that the conclusion of Proposition 3 is not trivial.

Therefore, for such a discount factor, Proposition 3 implies that perfect monitoring is the optimal monitoring structure for maximizing total industry profits in linear Bertrand oligopoly, and that the optimal equilibrium is stationary and in pure strategies. Letting p^* satisfy the incentive constraint

$$(b + c(n-1)p^* - p^*)p^* = (1 - \delta)\frac{1}{4}(b + c(n-1)p^*)^2$$

---

[17] This is a natural upper bound for c: as shown by Amir et al. (2017), our demand function arises from a representative consumer's utility maximization problem if and only if $c \in (-1, 1/(n-1))$. We thank an anonymous referee for the reference.

and letting $p^m = b/[2(1 - c(n - 1))]$ denote monopoly price, it follows that this equilibrium is given by the infinite repetition of symmetric pricing at $p_i = \min\{p^*, p^m\}$.

5. Observable mixed actions and delegation games

This section shows that the bound $V^*(\delta)$ is always tight under (mediated) perfect monitoring of *mixed* actions. While the assumption that mixed actions are observable is typically unrealistic, it is noteworthy that allowing observable mixtures is enough to encompass $E(\delta, p)$ for any monitoring structure p. This result also has the benefit of formalizing the interpretation of $V^*(\delta)$ sketched in Section 3.

To make these statements precise, we need to extend the notion of observable mixed actions to a setting where actions can be correlated across players. A natural way of doing this is through a game of repeated action delegation to a mediator, where in every period the mediator suggests a correlated action α, and players have the choice of *approving* the mediator's recommendation or *disapproving* it and playing any action in A_i. Note that, if actions are independent, so that $\alpha = \prod_i \alpha_i$, then this game corresponds to the usual notion of observable mixed actions, where in every period player i either plays her on-path mixed action α_i or deviates to some action in A_i.

Formally, assume in this section that each A_i is a finite set. Let $C = (C_i)_{i \in I} := (A_i \cup \{approve\})_{i \in I}$ with typical element $c = (c_i)_{i \in I} \in (A_i \cup \{approve\})_{i \in I}$. Given the stage game G, the corresponding *repeated action delegation game* proceeds as follows in each period t:

1. The mediator publicly recommends a correlated action $\alpha_t \in \Delta(A)$.
2. After observing α_t, each player i simultaneously chooses an alternative from the set $C_i = A_i \cup \{approve\}$, where the alternative *approve* is interpreted as approving the mediator's recommendation for that player, and the alternative $a_i \in A_i$ is interpreted as disapproving the mediator's recommendation and instead playing action a_i.
3. If the set of players $J \subseteq I$ approves and the set $-J = I \setminus J$ does not approve and chooses actions a_{-J}, then the realized action profile is drawn from the distribution (α_J, a_{-J}), where α_J denotes the marginal of α on A_J. The realized action profile is perfectly observed, but only the mediator and player i observe player i's choice $c_{i,t}$. Formally, the set of period t histories for the mediator is $H_m^t := (\Delta(A) \times C \times A)^{t-1}$, with typical element $h_m^t = (\alpha_\tau, c_\tau, a_\tau)_{\tau=1}^{t-1}$; while the set of period t histories for player i is $H_i^t := (\Delta(A) \times C_i \times A)^{t-1} \times \Delta(A)$, with typical element $h_i^t := \left((\alpha_\tau, c_{i,\tau}, a_\tau)_{\tau=1}^{t-1}, \alpha_t\right)$.

Let $E_{del}(\delta)$ be the set of sequential equilibrium payoffs in the repeated action delegation game.

(Two details regarding this solution concept: (i) In specifying an equilibrium, we define beliefs and impose sequential rationality only at histories consistent with the mediator's strategy. The interpretation is that the mediator is not a player in the game but rather a "machine" that cannot tremble. Without this assumption, we would have to consider trembles over the mediator's infinite action set. (ii) We endow the space of beliefs and strategies with the product topology. This assumption is for concreteness only; using a stronger topology would not affect our results.)

We require the assumption that the discount factor is above the cutoff

$$\delta^* := \min_{v \in \mathcal{F}: v \geq \underline{u}} \max_{i \in I, \alpha \in \Delta(A)} \frac{\bar{d}_i(\alpha)}{\bar{d}_i(\alpha) + v_i - \underline{u}_i},$$

with convention $0/0 = 1$. For example, in the prisoners' dilemma $\delta^* = \max\{g/(1+g),\ l/(1+l)\}$, and in linear Cournot $\delta^* = n/(1+n)$. The role of this assumption is to allow the minimaxing of deviators in the action delegation game. The intuition is that, if $\delta > \delta^*$, then players are willing to play their minimaxing actions in a single period if they believe they will be rewarded in the future. The mediator can then introduce noise into her recommendations so that, when a deviation occurs, the other players do not become aware of this fact and thus continue to be willing to play their minimaxing actions.[18]

Theorem 3. *If $\delta > \delta^*$, then* $\overline{E_{del}(\delta)} = V^*(\delta)$.

Proof. See Appendix A. □

The conclusion of Theorem 3 can fail if $\delta \leq \delta^*$.

Proposition 4. *For some game G and discount factor $\delta \leq \delta^*$, $V^*(\delta) \not\subseteq \overline{E_{del}(\delta)}$.*

Proof. See Appendix A. □

Finally, we show that $V^*(\delta)$ is also tight under mediated perfect monitoring of mixed actions if $\delta \leq \delta^*$ but a slightly more powerful form of mediation is allowed. The key insight is that, if the mediator recommends a *history-contingent* mixed action, then deviators can be minimaxed in the repeated delegation game even if $\delta \leq \delta^*$.

In particular, the *repeated mapping delegation game* differs from the repeated action delegation game, in that now the mediator recommends a mapping from histories to correlated actions. Formally, recursively define sets M_t for $t = 1, 2, \ldots$ by letting $M_1 = \Delta(A)$ and, for $t > 1$, letting M_t be the set of mappings $\mu_t : H_m^t \to \Delta(A)$, where $H_m^t := \left(\prod_{\tau=1}^{t-1} M_\tau\right) \times (C \times A)^{t-1}$ with typical element $h_m^t = (\mu_\tau, c_\tau, a_\tau)_{\tau=1}^{t-1}$ is the mediator's history (and $C_i = A_i \cup \{approve\}$ as above). The game proceeds as follows in each period t:

1. The mediator publicly recommends a mapping from histories to correlated actions $\mu_t : H_m^t \to \Delta(A)$.
2. After observing μ_t, each player i simultaneously chooses an alternative from the set $C_i = A_i \cup \{approve\}$.
   (Note that the mediator does *not* announce the history h_m^t. As we will see, players typically face uncertainty regarding h_m^t. Thus, players must approve or disapprove the mapping μ_t without knowing the resulting correlated action $\mu\left(h_m^t\right)$.)
3. If the set of players $J \subseteq I$ approves and the set $-J = I \setminus J$ does not approve and chooses actions a_{-J}, then given the mediator's history h_m^t the realized action profile is drawn from the distribution $\left(\mu_J\left(h_m^t\right), a_{-J}\right)$, where $\mu_J\left(h_m^t\right)$ denotes the marginal of $\mu\left(h_m^t\right)$ on A_J. Again, the realized action profile is perfectly observed, but only the mediator and player i observe player i's choice $c_{i,t}$. Thus, the set of histories for player i is $H_i^t := \left(\prod_{\tau=1}^{t-1} M_\tau\right) \times (C_i \times A)^{t-1} \times M_t$, with typical element $h_i^t := \left(\left(\mu_\tau, c_{i,\tau}, a_\tau\right)_{\tau=1}^{t-1}, \mu_t\right)$.

---

[18] Similarly, Sugaya and Wolitzky (2017) show that deviators can be minimaxed under mediated perfect monitoring whenever $\delta > \delta^*$. The observation that miscoordination off the equilibrium path can lead to harsher punishments under imperfect monitoring was previously exploited by Kandori (1991), Sekiguchi (2002), and Mailath et al. (2002).

Let $E_{map}(\delta)$ be the set of sequential equilibrium payoffs in the repeated mapping delegation game.

The point of introducing mapping delegation is that we can show that $\overline{E_{map}(\delta)} = V^*(\delta)$ under a much weaker condition on the discount factor than $\delta > \delta^*$ (although $\delta > \delta^*$ is already a fairly weak condition). In particular, we require only the following assumption:

Assumption 1. There exists $\hat{\alpha} \in \Delta(A)$ such that

$$u_i(\hat{\alpha}) > (1 - \delta) d_i(\hat{\alpha}) + \delta \underline{u}_i \text{ for all } i. \tag{7}$$

This assumption holds whenever G admits a static correlated equilibrium $\hat{\alpha}$ with $u_i(\hat{\alpha}) > \underline{u}_i$ for all i. More generally, it is equivalent to the condition that

$$\delta > \hat{\delta} := \inf_{\alpha \in \Delta(A)} \max_{i \in I} \frac{d_i(\alpha) - u_i(\alpha)}{d_i(\alpha) - \underline{u}_i}$$

(with convention $0/0 = 1$). Note that $\hat{\delta} \leq \delta^*$, so this is a weaker condition than the one needed for Theorem 3; for example, in the prisoners' dilemma $\hat{\delta} = \min\{g, l\}/(1 + g)$, and in linear Cournot $\hat{\delta} = 0$ (as there exists a static Nash equilibrium with payoffs above minimax). Moreover, it can be shown that $\delta \geq \hat{\delta}$ is a necessary condition for the existence of a repeated game equilibrium where all players' continuation payoffs are bounded away from their minimax payoffs in every period. We include a proof of this fact in Appendix A.

Theorem 4. *If Assumption 1 holds, then* $\overline{E_{map}(\delta)} = V^*(\delta)$.

Proof. The simple proof that $\overline{E_{map}(\delta)} \subseteq V^*(\delta)$ is the same as for action delegation; see Appendix A.

The proof of the reverse inclusion follows from two lemmas:

Lemma 1. *Under Assumption 1, there exists $\bar{\varepsilon} > 0$ such that, for all $\varepsilon, \eta \leq \bar{\varepsilon}$, the following hold:*

First, it is a sequential equilibrium strategy profile in the mapping delegation game for the players to always approve the mediator's recommendation and for the mediator to recommend the following mapping in every period t:

- *If all players have always approved prior to time t, then $\mu(h_m^t) = \hat{\alpha}^\varepsilon := (1 - \varepsilon)\hat{\alpha} + (\varepsilon/|A|) \sum_{a \in A} a$.*
- *If the set of players who have ever disapproved prior to time t is the singleton $\{i\}$ for some $i \in I$, then $\mu(h_m^t) = (a_i^\eta, \alpha_{-i}^\eta)$, where $\alpha_{-i}^\eta = (1 - \eta)\alpha_{-i}^{\min} + \eta \sum_{a_{-i} \in A_{-i}} (a_{-i}/|A_{-i}|)$ is a full-support perturbation of the minimax strategy and a_i^η is any best response to α_{-i}^η.*
- *If the set of players who have ever disapproved prior to time t is $J \subseteq I$ with $|J| \geq 2$, then $\mu(h_m^t) = \left(\left(a_j^\eta\right)_{j \in J}, \alpha_{-J}^{\min}\right)$, where $\alpha_{-J}^{\min}$ is the correlated action that assigns probability $1/|A_{-J}|$ to each action profile $a_{-J} \in A_{-J}$.*

Second, under this strategy profile, on-path incentives to approve hold with $\bar{\varepsilon}$ slack: for all $\varepsilon, \eta \leq \bar{\varepsilon}$ and all i, we have

$$u_i(\hat{\alpha}^\varepsilon) > (1 - \delta) d_i(\hat{\alpha}^\varepsilon) + \delta \underline{u}_i^\eta + \bar{\varepsilon}, \tag{8}$$

where $\underline{u}_i^\eta = u_i(a_i^\eta, \alpha_{-i}^\eta)$.

Proof. Since the on-path correlated action $\hat{\alpha}^\varepsilon$ has full support, a player who has always previously approved believes that (i) if she approves then $\hat{\alpha}^\varepsilon$ will be played forever, and (ii) if she disapproves and plays a_i then $(a_i, \hat{\alpha}^\varepsilon_{-i})$ will be played in the current period and $(a_i^\eta, \alpha_{-i}^\eta)$ will be played in every future period. By (7), there exists $\bar{\varepsilon}_1 > 0$ such that

$$\bar{\varepsilon}_1 < \frac{1}{2}\left[u_i\left(\hat{\alpha}\right) - (1 - \delta) d_i\left(\hat{\alpha}\right) - \delta \underline{u}_i\right].$$

As $u_i\left(\hat{\alpha}^\varepsilon\right)$ and $d_i\left(\hat{\alpha}^\varepsilon\right)$ are continuous in ε and $\underline{u}_i^\eta$ is continuous in η, there exists $\bar{\varepsilon}_2 > 0$ such that, for all $\varepsilon, \eta \le \bar{\varepsilon}_2$,

$$\left| u_i\left(\hat{\alpha}\right) - (1 - \delta) d_i\left(\hat{\alpha}\right) - \delta \underline{u}_i - \left[u_i\left(\hat{\alpha}^\varepsilon\right) - (1 - \delta) d_i\left(\hat{\alpha}^\varepsilon\right) - \delta \underline{u}_i^\eta\right]\right|$$

$$< \frac{1}{2}\left[u_i\left(\hat{\alpha}\right) - (1 - \delta) d_i\left(\hat{\alpha}\right) - \delta \underline{u}_i\right].$$

Taking $\bar{\varepsilon} := \min \{\bar{\varepsilon}_1, \bar{\varepsilon}_2\}$, (8) holds.

Since the off-path distribution of opposing action α_{-i}^η also has full support, if a player has ever disapproved, she believes that her opponents will play α_{-i}^η in every subsequent period regardless of her own behavior. Since a_i^η is a best response to α_{-i}^η, approval is optimal.

Thus, by the one-shot deviation principle, it is optimal to always approve the mediator's recommendation. □

Lemma 2. *If Assumption 1 holds and $(\alpha_t)_{t=1}^\infty$ satisfies (1) for all i and t, then*

$$(1 - \delta) \sum_{t \ge 1} \delta^{t-1} u\left(\alpha_t\right) \in \overline{E_{map}\left(\delta\right)}.$$

Proof. Let $v = (1 - \delta) \sum_{t \ge 1} \delta^{t-1} u\left(\alpha_t\right)$. It suffices to show that, for any $\varepsilon' > 0$, there exists an incentive-compatible strategy for the mediator μ such that

$$\max_i \left| v_i - (1 - \delta) \sum_{t \ge 1} \delta^{t-1} \mathbb{E}^\mu \left[u_i\left(a_t\right)\right]\right| < \varepsilon'.$$

Fix $\bar{\varepsilon} > 0$ such that the conclusion of Lemma 1 applies, and fix $\varepsilon \le \bar{\varepsilon}$ such that

$$\max_i \left| v_i - (1 - \delta) \sum_{t \ge 1} \delta^{t-1}\left[(1 - \varepsilon) u\left(\alpha_t\right) + \varepsilon \hat{\alpha}^{\bar{\varepsilon}}\right]\right| < \varepsilon'. \tag{9}$$

For each $\eta \in (0, \bar{\varepsilon}]$, consider the following strategy for the mediator: For every t, the mediator recommends the mapping $\mu_t : H_m^t \to \Delta\left(A\right)$ defined as follows:

- If all players have always approved prior to time t, then $\mu_t\left(h_m^t\right) = (1 - \varepsilon) \alpha_t + \varepsilon \hat{\alpha}^{\bar{\varepsilon}}$.
- If the set of players who have ever disapproved prior to time t is the singleton $\{i\}$ for some $i \in I$, then $\mu_t\left(h_m^t\right) = \left(a_i^\eta, \alpha_{-i}^\eta\right)$.
- If the set of players who have ever disapproved prior to time t is $J \subseteq I$ with $|J| \ge 2$, then

$$\mu_t\left(h_m^t\right) = \left(\left(a_j^\eta\right)_{j \in J}, \alpha_{-J}^{\text{mix}}\right).$$

Given (9), it remains to show that there exists $\eta > 0$ such that these recommendations are incentive-compatible.

Since the on-path correlated action has full support, a player who has always approved believes that all players have always approved. Similarly, since the off-path distribution of opposing actions following a single deviation has full support, a player who has ever disapproved believes that she is the only player who has ever disapproved.

If a player has ever disapproved, it is optimal for her to approve in the future by the same argument as in Lemma 1. We are thus left to consider a player's incentives when she has always approved.

In this case, for every i and t,

$$(1 - \delta) \sum_{\tau \geq t} \delta^{\tau - t} \mathbb{E}^\mu \left[u_i (a_\tau) | h_i^t \right]$$

$$= (1 - \varepsilon) (1 - \delta) \sum_{\tau \geq t} \delta^{\tau - t} u_i (\alpha_t) + \varepsilon (1 - \delta) \sum_{\tau \geq t} \delta^{\tau - t} u_i \left(\hat{\alpha}^{\bar{\varepsilon}} \right)$$

$$\geq (1 - \varepsilon) \left[(1 - \delta) d_i (\alpha_t) + \delta \underline{u}_i \right] + \varepsilon \left[(1 - \delta) d_i \left(\hat{\alpha}^{\bar{\varepsilon}} \right) + \delta \underline{u}_i^\eta + \bar{\varepsilon} \right]$$

$$\geq (1 - \delta) d_i \left((1 - \varepsilon) \alpha_t + \varepsilon \hat{\alpha}^{\bar{\varepsilon}} \right) + \delta \left[(1 - \varepsilon) \underline{u}_i + \varepsilon \underline{u}_i^\eta \right] + \varepsilon \bar{\varepsilon},$$

where the first inequality follows by (1) and Lemma 1 (given the assumption that $\eta \leq \bar{\varepsilon}$), and the second inequality follows because the function $d_i (\cdot)$ is convex. Hence, taking η small enough so that $\underline{u}_i > \underline{u}_i^\eta - \varepsilon \bar{\varepsilon}$, we have

$$(1 - \delta) \sum_{\tau \geq t} \delta^{\tau - t} \mathbb{E}^\mu \left[u_i (a_\tau) | h_i^t \right]$$

$$\geq (1 - \delta) d_i \left((1 - \varepsilon) \alpha_t + \varepsilon \hat{\alpha}^{\bar{\varepsilon}} \right) + \delta \left[\underline{u}_i^\eta - \varepsilon \bar{\varepsilon} + \varepsilon \left(\underline{u}_i^\eta - \underline{u}_i \right) \right] + \varepsilon \bar{\varepsilon}$$

$$\geq (1 - \delta) d_i \left((1 - \varepsilon) \alpha_t + \varepsilon \hat{\alpha}^{\bar{\varepsilon}} \right) + \delta \underline{u}_i^\eta.$$

This shows that approving the mediator's recommendations is optimal for a player who has always approved in the past, which completes the proof. $\square$

Remark. In the formulation of the mapping delegation game, it is crucial that the mediator recommends a mapping *before* players choose whether to approve or deviate. If instead players decided whether to delegate their choice of action to the mediator before the mediator's choice, then any feasible and individually rational payoff vector could be sustained in equilibrium for *any* discount factor, as the mediator could immediately minimax any player who failed to delegate.[19] One would thus typically have $V^* (\delta) \subsetneq \overline{E_{map} (\delta)}$, in contrast to Theorem 4.

6. Tightening the bound

So far, we have focused on the information-free payoff bound that results from considering only players' ex ante incentive constraints. We favor this bound because it has a natural interpretation, is simple to compute and useful in applications, and coincides with the equilibrium payoff set in the game with observable mixed strategies. But it is also possible to give tighter

---

[19] This logic is similar to, for example, Myerson's (1991; Section 6.1) discussion of games with contracts.

bounds. Here, we pursue a particularly natural approach: refining the incentive constraints by letting players access some information about the history of play. The obvious trade-off is that, as we give players more information, the resulting payoff bound becomes tighter but also more complicated and harder to compute.[20]

The simplest way of tightening the bound is to modify Definition 1 by letting players condition deviations on their recommended actions. For $\alpha \in \Delta(A)$, let

$$d_i'(\alpha) = \max_{\phi_i : A_i \to A_i} \sum_{a \in A} \alpha(a) u_i(\phi_i(a_i), a_{-i})$$

be player i's maximum deviation payoff against α when she can condition her deviation on her recommended action. Clearly, $d_i'(\alpha) \geq d_i(\alpha)$. Note that the difference between d_i' and d_i is analogous to the difference between correlated equilibrium and coarse correlated equilibria (Moulin and Vial, 1979; Awaya and Krishna, 2017).

Definition 4. Let $V^{**}(\delta)$ denote the set of payoff vectors $v \in \mathbb{R}^n$ with the following property: there exists $(\alpha_t)_{t=1}^{\infty} \in \Delta(A)^{\infty}$ such that $v = (1-\delta)\sum_{t \geq 1}\delta^{t-1}u(\alpha_t)$ and, for all i and t, we have

$$(1-\delta)\sum_{\tau \geq t}\delta^{\tau-t}u_i(\alpha_\tau) \geq (1-\delta)d_i'(\alpha_t) + \delta\underline{u}_i. \tag{10}$$

Note that $V^{**}(\delta) \subseteq V^*(\delta)$. Analogues of Proposition 1 and Theorem 1 for $V^{**}(\delta)$ are immediate. In particular, (i) for $W \subseteq \mathbb{R}^n$, letting $B^{**}(W)$ be the set of payoff vectors v such that there exist $\alpha \in \Delta(A)$ and $w \in W$ with $v = (1-\delta)u(\alpha) + \delta w$ and

$$(1-\delta)u_i(\alpha) + \delta w_i \geq (1-\delta)d_i'(\alpha) + \delta\underline{u}_i \text{ for all } i,$$

we have $V^{**}(\delta) = \lim_{m \to \infty}(B^{**})^m(\mathcal{F})$, and (ii) for every discount factor δ and monitoring structure p, $\text{co}(E(\delta, p)) \subseteq V^{**}(\delta)$. Thus, $V^{**}(\delta)$ gives a tighter bound than $V^*(\delta)$, and it is only slightly more complicated.[21]

The bound can be tightened further by letting on-path continuation payoffs vary with equilibrium actions. However, this approach threatens the bound's recursive structure, because the set of available continuation payoffs may depend on players' past actions through their past incentive constraints. One way of letting continuation payoffs vary with equilibrium actions while retaining a recursive structure is to consider the bound that results when players "forget" their past actions. This yields the following:

Proposition 5. *Given a set of payoff vectors $W \subseteq \mathbb{R}^n$, let $B^0(W)$ be the set of payoff vectors v such that there exist $\alpha \in \Delta(A)$ and $\gamma : A \to \mathcal{F}$ with $v = (1-\delta)u(\alpha) + \delta\gamma(\alpha)$, $\gamma(\alpha) \in W$, and*

$$(1-\delta)u_i(\alpha|a_i) + \delta\gamma_i(\alpha|a_i) \geq (1-\delta)d_i(\alpha|a_i) + \delta\underline{u}_i \text{ for all } i, a_i \in \text{supp}(\alpha_i),$$

where $\alpha|a_i$ stands for the conditional probability $(\alpha|a_i)(a_i, a_{-i}) = \alpha(a_i, a_{-i})/\sum_{a_{-i}'}\alpha(a_i, a_{-i}')$, and $\gamma(\alpha) = \sum_a \alpha(a)\gamma(a)$. Let $V^0(\delta) = \lim_{m \to \infty}(B^0)^m(\mathcal{F})$. Then $V^0(\delta) \subseteq V^{**}(\delta)$, and, for every discount factor δ and monitoring structure p, $\text{co}(E(\delta, p)) \subseteq V^0(\delta)$.

---

[20] Most of the results in this section were suggested by an anonymous referee, whom we thank.

[21] However, the two bounds coincide for all examples considered in the current paper.

The proof is again analogous to that of Proposition 1 and Theorem 1. Note that, in the construction of the set $B^0(W)$, player i relies on the information that her continuation payoff following equilibrium action a_i equals $\gamma_i(\alpha|a_i)$ when deciding whether to deviate from a_i, but $\gamma_i(\alpha|a_i)$ is not itself required to be an equilibrium payoff.

Finally, it is possible to attain yet a tighter bound by letting players remember their last T actions, for some $T > 0$. This approach generates a bound $V^T(\delta)$ for every T, where $V^T(\delta) \supseteq V^{T+1}(\delta)$ and $V^0(\delta)$ is the set constructed above. The construction of the bounds $V^T(\delta)$ for $T > 0$ is however qualitatively more complicated than the construction of $V^0(\delta)$, because the set of available continuation payoffs W will depend on the last T correlated actions. Thus, instead of being able to compute the payoff bound by iterating an operator B (or B^*, or B^0) from subsets of $\mathbb{R}^n$ to subsets of $\mathbb{R}^n$, one has to iterate an operator from subsets of T-period distributions of actions to subsets of $(\mathbb{R}^n)^{|A|^T}$ (corresponding to sets of continuation payoffs for each realization of the previous T action profiles). Presumably, this quickly becomes intractable as T increases. A formal construction of the resulting bound is available from the authors on request.

As a last remark, we note that this approach of bounding the true equilibrium payoff set in terms of equilibrium payoffs in auxiliary games where players have longer and longer memories is reminiscent of Amarante's (2003) approach of considering equilibrium payoffs in T-period finitely repeated games with longer and longer horizons and arbitrary terminal payoffs. Our proposed approach relaxes the self-generation constraint $\gamma(a) \in W$ by requiring that it is satisfied only in expectation given players' previous T actions, while Amarante allows for arbitrary continuation payoffs after T periods of play.

Appendix A. Omitted proofs for Section 5

A.1. Proof of Theorem 3

Proof. For any Nash equilibrium σ of the repeated action delegation game and any on-path history h_i^t, player i can guarantee herself a continuation payoff of at least $(1-\delta)\max_{a_i} \mathbb{E}^\sigma\left[u_i(a_i, a_{-i})|h_i^t\right] + \delta\underline{u}_i$. The proof that $E_{del}(\delta) \subseteq V^*(\delta)$ is therefore the same as the proof of Theorem 1. Since $V^*(\delta)$ is closed by Proposition 1, taking closures gives $\overline{E_{del}(\delta)} \subseteq V^*(\delta)$.

For the reverse inclusion, fix $(\alpha_t)_{t=1}^\infty$ satisfying (1), and let $v_t = (1-\delta)\sum_{\tau \geq t} \delta^{\tau-t} u(\alpha_\tau)$. We show that $v_1 \in \overline{E_{del}(\delta)}$.

Note that

$$\delta\left(v_{i,t+1} - \underline{u}_i\right) \geq (1-\delta)\bar{d}_i(\alpha_t) \text{ for each } i \text{ and } t. \tag{11}$$

On the other hand, as $\delta > \delta^*$, there exists a distribution $\bar{\alpha} \in \Delta(A)$ with full support ($\bar{\alpha}(a) > 0$ for each $a \in A$) and a constant $\Delta > 0$ such that, for all $\alpha \in \Delta(A)$,

$$\delta\left(u_i(\bar{\alpha}) - \underline{u}_i\right) > (1-\delta)\bar{d}_i(\alpha) + \Delta \text{ for each } i. \tag{12}$$

(That is, if on-path continuation payoffs are given by $u(\bar{\alpha})$ and a deviator is minimaxed forever, then no player wants to deviate from any distribution α.) Now, fix a constant $\varepsilon \in (0, 1)$, let $\alpha_t^\varepsilon = (1-\varepsilon)\alpha_t + \varepsilon\bar{\alpha}$, and let $v_t^\varepsilon = (1-\delta)\sum_{\tau \geq t} \delta^{\tau-t} u(\alpha_\tau^\varepsilon)$. Since $\bar{d}_i$ is convex and u_i is linear, combining (11) and (12) yields

$$\delta\left(v_{i,t+1}^\varepsilon - \underline{u}_i\right) > (1-\delta)\bar{d}_i(\alpha_t^\varepsilon) + \varepsilon\Delta \text{ for each } i \text{ and } t. \tag{13}$$

In addition, for any $\eta \in (0, 1)$, let $\alpha^{\eta}_{-i} = (1 - \eta) \alpha^{\min}_{-i} + \eta \sum_{a_{-i} \in A_{-i}} \frac{a_{-i}}{|A_{-i}|}$ denote a full support perturbation of the minimaxing distribution $\alpha^{\min}_{-i}$; let a^{η}_i be any best response to α^{η}_{-i}; and let $\alpha^{i,\eta} = (a^{\eta}_i, \alpha^{\eta}_{-i})$.

Consider the following strategy for the mediator: The mediator's *period t state* ω_t is an element of the set $\{normal\} \cup \{reward\} \cup I$. The mediator's recommendations in each state are determined as follows:

- If $\omega_t = normal$, then with probability $1 - \eta$ the mediator recommends α^{ε}_t, and with probability η/n the mediator recommends $\alpha^{i,\eta}$, for each $i \in I$.
- If $\omega_t = reward$, then with probability $1 - \eta$ the mediator recommends $\bar{\alpha}$, and with probability η/n the mediator recommends $\alpha^{i,\eta}$, for each $i \in I$.
- If $\omega_t = i$, then the mediator recommends $\alpha^{i,\eta}$ with probability 1.

The initial state is $\omega_1 = normal$. The state transition rule is as follows:

- If $\omega_t = normal$, the mediator recommends α^{ε}_t, and all players approve, then the state transitions to $\omega_{t+1} = normal$ with probability $1 - \eta$ and transitions to $\omega_{t+1} = reward$ with probability η. If a unique player i disapproves, then the state transitions to $\omega_{t+1} = i$ with probability $1 - \eta$ and transitions to $\omega_{t+1} = reward$ with probability η. If more than one player disapproves, then the state transitions to $\omega_{t+1} = reward$ with probability 1.
- If $\omega_t = normal$, the mediator recommends $\alpha^{i,\eta}$ for some i, and either all players approve, only player i disapproves, or two or more players disapprove, then the state transitions to $\omega_{t+1} = reward$. If a unique player $j \neq i$ disapproves, then the state transitions to $\omega_{t+1} = j$ with probability $1 - \eta$ and transitions to $\omega_{t+1} = reward$ with probability η.
- If $\omega_t = reward$ and either all players approve, $\alpha^{i,\eta}$ is recommended and only player i disapproves, or two or more players disapprove, then the state transitions to $\omega_{t+1} = reward$.
- If $\omega_t = reward$ and a unique player i disapproves the mediator's recommendation (whether it is $\bar{\alpha}$ or $\alpha^{j,\eta}$ for some $j \neq i$), then the state transitions to $\omega_{t+1} = i$ with probability $1 - \eta$ and transitions to $\omega_{t+1} = reward$ with probability η.
- If $\omega_t = i$, then the state transitions to $\omega_{t+1} = i$ with probability $1 - \eta$ and transitions to $\omega_{t+1} = reward$ with probability η (regardless of the players' actions).

Each player's strategy is to approve the mediator's recommendation at every history. We claim that there exists $\bar{\eta} > 0$ such that, for all $\eta < \bar{\eta}$, the resulting strategy profile (together with any consistent belief system) is a sequential equilibrium.

To see this, it is helpful to call player i's strategy *essentially obedient* if player i approves the mediator's recommendation at every history, except possibly at histories where $\alpha^{i,\eta}$ is recommended (that is, at histories where a unilateral deviation by player i is ignored). Note that any history h^t_i consistent with the mediator's strategy is also consistent with player i's opponents' playing essentially obedient strategy, even if player i has deviated in the past: that is, for any history h^t_i, there exists a realization of $(\alpha_\tau)^{t-1}_{\tau=1}$ that results in history h^t_i with positive probability when, for all $\tau \leq t - 1$, $j \neq i$, and $\alpha_\tau \neq \alpha^{j,\eta}$, $c_{j,\tau} = approve$. To see this, consider the following cases[22]:

---

[22] The following discussion implicitly assumes that $\alpha^{\varepsilon}_t \neq \bar{\alpha}$. If instead $\alpha^{\varepsilon}_t = \bar{\alpha}$, a player has less information when contemplating a deviation, so it remains optimal for her to approve the mediator's recommendations.

1. Player i observes α_t^ε. This recommendation only occurs in state *normal*, and the state can be *normal* only if no players have ever disapproved. Hence, this recommendation is consistent *only* with all players following their (fully obedient) equilibrium strategies.
2. Player i has never disapproved the mediator's recommendation and observes $\alpha_t \neq \alpha_t^\varepsilon$ for the first time. This history is consistent with all players following their equilibrium strategies and the current state being *reward* (or possibly *normal*, but we are claiming only that there is at least one history consistent with player i's opponents' playing essentially obedient strategies).
3. Player i observes $\alpha_t \neq \alpha_t^\varepsilon$ and has previously either disapproved the mediator's recommendation or observed $\alpha_\tau \neq \alpha_\tau^\varepsilon$ (for $\tau < t$). At such a history, $\omega_t \in \{reward\} \cup I$. Note that $\omega_t = reward$ is reached with positive probability after any ω_{t-1}, regardless of the players' actions. Moreover, note that any recommendation that occurs with positive probability in state i also occurs with positive probability in state *reward*, and that every action profile occurs with positive probability after each recommendation, except when the recommendation is $\alpha_\tau = \alpha_\tau^{j,\eta}$ for some j and player j does not play a_j^η. Hence, this history is consistent with the current state being *reward* and player i's opponents' playing essentially obedient strategies.

Hence, at every history h_i^t, player i assigns probability 1 to the event that for all $\tau \leq t - 1$, $j \neq i$, and $\alpha_\tau \neq \alpha^{j,\eta}$, $c_{j,\tau} = approve$.

Consequently, if player i has never disapproved the recommendation, then player i assigns probability 1 to the event that the state is either *normal* or *reward*. Furthermore, note that the state is revealed to be *normal* if the mediator recommends α_t^ε; the state is revealed to be *reward* if the mediator recommends $\bar{\alpha}$; and, if the mediator recommends $\alpha^{j,\eta}$ for some j, then the transition rule is independent of whether the current state is *normal* or *reward*. Hence, the relative probability of *normal* and *reward* is irrelevant for player i's incentives. Similarly, if player i has ever disapproved the recommendation, then player i assigns probability 1 to the event that the state is either i or *reward*. In this case, any recommendation other than $\alpha^{i,\eta}$ reveals the state to be *reward*, and player i clearly has no incentive to disapprove the recommendation $\alpha^{i,\eta}$.

It thus remains only to show that there exists $\bar{\eta} > 0$ such that, for all $\eta < \bar{\eta}$, (i) approving α_t^ε is optimal in state $\omega_t = normal$, (ii) approving $\bar{\alpha}$ is optimal in state $\omega_t = reward$, and (iii) approving $\alpha^{j,\eta}$ is optimal in state $\omega_t = normal$ or $\omega_t = reward$ (for any j).

For (i), note that approving α_t^ε is optimal in state $\omega_t = normal$ if and only if

$$\delta \left(v_{i,t+1}^{\varepsilon,\eta} - \underline{u}_{i,t+1}^{\eta} \right) \geq (1 - \delta) \, \bar{d}_i \left(\alpha_t^\varepsilon \right), \tag{14}$$

where $v_{i,t+1}^{\varepsilon,\eta}$ is player i's continuation payoff when the state is $\omega_{t+1} = normal$ with probability $1 - \eta$ and $\omega_{t+1} = reward$ with probability η, and $\underline{u}_{i,t+1}^{\eta}$ is the continuation payoff when the state is $\omega_{t+1} = i$ with probability $1 - \eta$ and $\omega_{t+1} = reward$ with probability η. For all $\varepsilon > 0$, $\sup_t \left| v_{i,t+1}^{\varepsilon,\eta} - v_{i,t+1}^\varepsilon \right| \to 0$ and $\sup_t \left| \underline{u}_{i,t+1}^{\eta} - \underline{u}_i \right| \to 0$ as $\eta \to 0$. Hence, by (13), for all $\varepsilon > 0$ there exists $\bar{\eta} > 0$ such that (14) holds for all i, t, and $\eta < \bar{\eta}$.

For (ii) and (iii), the desired conclusion follows similarly from (12) and the uniform continuity of continuation payoffs in η.

We have shown that, for all $\varepsilon > 0$, there exists $\bar{\eta} > 0$ such that, for all $\eta < \bar{\eta}$, $\bar{v}_{i,1}^{\varepsilon,\eta} \in E_{del}(\delta)$. As $\left| \bar{v}_{i,1}^{\varepsilon,\eta} - v_{i,1}^\varepsilon \right| \to 0$ as $\eta \to 0$ and $\left| v_1^\varepsilon - v_1 \right| \to 0$ as $\varepsilon \to 0$, it follows that $v_1 \in \overline{E_{del}(\delta)}$. $\quad\square$

A.2. Proof of Proposition 4

Proof. Consider the game

$$
\begin{array}{ccc}
 & L & R \\
T & 1,1 & 0,2 \\
B & -10,-1 & -10,0
\end{array}
$$

Let $\delta = 2/3$. (Note that $\delta^* = 11/12$ in this game, so $\delta < \delta^*$.)

Note that $(1,1) \in V^*(\delta)$. This follows by taking $\alpha_t = (T,L)$ for all t and checking (1) for player 2, which holds as $1 \geq (1-\delta)(2) + \delta(0)$.

However, $(1,1) \notin \overline{E_{del}(\delta)}$. To see this, note that player 1's minimax payoff is 0 and her maximum payoff is 1, so she will never approve a distribution α where the weight α_B on B satisfies

$$
(1-\delta)[\alpha_B(-10) + (1-\alpha_B)(1)] + \delta(1) < 0 \iff \alpha_B > \frac{3}{11}.
$$

Similarly, player 1 will never disapprove the mediator's recommendation and play B. Now, suppose toward a contradiction that there is an equilibrium in the action delegation game where player 1's payoff exceeds $1 - \varepsilon$. In such an equilibrium, player 1's (ex ante expected) continuation payoff starting in period 2, v_1^2, must satisfy

$$
(1-\delta)(1) + \delta\left(v_1^2\right) \geq 1 - \varepsilon \iff v_1^2 \geq 1 - \frac{\varepsilon}{\delta}.
$$

Hence, by feasibility, player 2's equilibrium continuation payoff starting in period 2 is at most $1 + \varepsilon/\delta$. Therefore, noting that player 2's instantaneous deviation gain from playing R rather than L is always 1, a necessary condition for player 2 to prefer his equilibrium strategy to always playing R is

$$
\delta\left(1 + \frac{\varepsilon}{\delta}\right) \geq (1-\delta)(1) + \delta\left(\frac{3}{11}(0) + \frac{8}{11}(2)\right).
$$

This condition fails for all sufficiently small $\varepsilon > 0$. Hence, $(1,1) \notin \overline{E_{del}(\delta)}$. □

A.3. Claim supporting Assumption 1

Claim 4. *For every discount factor δ and monitoring structure p, if there exist $(\alpha_t)_{t=1}^{\infty} \in \Delta(A)^{\infty}$ and $\eta > 0$ such that, for all i and t, we have*

$$
(1-\delta)\sum_{\tau \geq t} \delta^{\tau-t} u_i(\alpha_\tau) \geq \max\left\{(1-\delta) d_i(\alpha_t) + \delta\underline{u}_i, \underline{u}_i + \eta\right\},
$$

then $\delta \geq \hat{\delta}$.

Proof. Fix such a sequence $(\alpha_t)_{t=1}^{\infty} \in \Delta(A)^{\infty}$ and $\eta > 0$, and fix $\delta' > \delta$. We show that there exists $\hat{\alpha} \in \Delta(A)$ such that

$$
u(\hat{\alpha}) > (1-\delta') d(\hat{\alpha}) + \delta'\underline{u}.
$$

As this is established for any $\delta' > \delta$, it follows that $\delta \geq \hat{\delta}$.

To prove this, for every $\beta \in (\delta, 1)$, define an action distribution $\alpha(\beta)$ by

$$\alpha(\beta) = (1 - \beta) \sum_{t=1}^{\infty} \beta^{t-1} \alpha_t.$$

Letting $\alpha^t = (1 - \delta) \sum_{\tau \geq t} \delta^{\tau - t} \alpha_\tau$, one can express $\alpha(\beta)$ as a weighted average of the α^t's according to

$$\alpha(\beta) = \frac{1 - \beta}{1 - \delta} \left[\alpha^1 + (\beta - \delta) \sum_{t \geq 1} \beta^{t-1} \alpha^{t+1} \right].$$

Note that, as u_i is bounded for all i, there exists $\bar{\beta} < 1$ such that, if $\beta > \bar{\beta}$, then $\left| u_i(\alpha(\beta)) - u_i\left((1 - \beta) \sum_{t \geq 1} \beta^{t-1} \alpha^t\right) \right| < \left(\delta' - \delta\right) \eta$ for all i. Moreover, note that $u(\alpha^t) \geq \underline{u} + \eta$ for all t by hypothesis, and hence $u(\alpha(\beta)) \geq \underline{u} + \eta$.

As (1) holds for every t, for every β we have

$$(1 - \beta) \sum_{t \geq 1} \beta^{t-1} u\left(\alpha^t\right) \geq (1 - \beta) \sum_{t \geq 1} \beta^{t-1} \left[(1 - \delta) d(\alpha_t) + \delta \underline{u}\right].$$

As u is linear and d is convex, this implies

$$u\left((1 - \beta) \sum_{t \geq 1} \beta^{t-1} \alpha^t\right) \geq (1 - \delta) d(\alpha(\beta)) + \delta \underline{u}.$$

Hence, if $\beta > \bar{\beta}$, we have

$$u(\alpha(\beta)) \geq (1 - \delta) d(\alpha(\beta)) + \delta \underline{u} - \left(\delta' - \delta\right) \eta,$$

or equivalently

$$u(\alpha(\beta)) \geq d(\alpha(\beta)) - \frac{\delta}{1 - \delta} \left(u(\alpha(\beta)) - \underline{u}\right) - \frac{\delta' - \delta}{1 - \delta} \eta. \tag{15}$$

Next, note that

$$\frac{\delta'}{1 - \delta'} \left(u(\alpha(\beta)) - \underline{u}\right) = \frac{\delta}{1 - \delta} \left(u(\alpha(\beta)) - \underline{u}\right) + \frac{\delta' - \delta}{(1 - \delta)(1 - \delta')} \left(u(\alpha(\beta)) - \underline{u}\right)$$

$$> \frac{\delta}{1 - \delta} \left(u(\alpha(\beta)) - \underline{u}\right) + \frac{\delta' - \delta}{1 - \delta} \eta.$$

Hence, (15) implies

$$u(\alpha(\beta)) > d(\alpha(\beta)) - \frac{\delta'}{1 - \delta'} \left(u(\alpha(\beta)) - \underline{u}\right),$$

or equivalently

$$u(\alpha(\beta)) > \left(1 - \delta'\right) d(\alpha(\beta)) + \delta' \underline{u}.$$

Taking $\hat{\alpha} = \alpha(\beta)$ completes the proof. $\quad \square$

Appendix B. Stick-and-Carrot equilibria in the Bertrand example

Let δ_1 be lowest discount factor such that one can minimax a deviator with a stick-and-carrot equilibrium, where after a deviation all firms price at zero for one period and then price at some level $\bar{p}$ forever. Let δ_2 be lowest discount factor such that price p^m is sustainable when deviators can be minimaxed. If $\delta_1 < \delta_2$, then Proposition 3 applies (and yields a non-trivial conclusion) whenever $\delta \in (\delta_1, \delta_2)$.

Proposition 6. *If $c\,(n-1) > 2/3$ then $\delta_1 < \delta_2$.*

Proof. Note that the minimax payoff is $b^2/4$. In a stick-and-carrot equilibrium, the incentive compatibility constraint in the "carrot" state (pricing at $\bar{p}$) is

$$(b + c\,(n-1)\,\bar{p} - \bar{p})\,\bar{p} \geq (1 - \delta) \left(\frac{b + c\,(n-1)\,\bar{p}}{2} \right)^2 + \delta \frac{b^2}{4}. \tag{16}$$

The condition that the stick-and-carrot equilibrium attains the minimax payoff in the "stick" state is

$$\delta\,(b + c\,(n-1)\,\bar{p} - \bar{p})\,\bar{p} = \frac{b^2}{4}. \tag{17}$$

Note that this condition also implies incentive compatibility in the stick state, which is simply the condition that the left-hand side of (17) is no less than the right-hand side. Let δ_1 be the minimum discount factor for which there exists a price $\bar{p}$ that satisfies (16) and (17).

On the other hand, the condition for a constant price of p^m to be sustainable when deviators can be minimaxed is

$$\left(b + c\,(n-1)\,p^m - p^m \right) p^m \geq (1 - \delta) \left(\frac{b + c\,(n-1)\,p^m}{2} \right)^2 + \delta \frac{b^2}{4}. \tag{18}$$

Let δ_2 be the minimum discount factor that satisfies (18).

We wish to show that $\delta_1 < \delta_2$. First note that, if (17) holds with strict inequality at $\delta = \delta_2$ and $\bar{p} = p^m$, then $\delta_1 < \delta_2$. This follows because reducing $\bar{p}$ relaxes (16), so if (17) holds with strict inequality at $\bar{p} = p^m$ then there exists $\varepsilon > 0$ and $\delta < \delta_2$ such that both (16) and (17) are satisfied at discount factor δ when $\bar{p} = p^m - \varepsilon$.

Thus, it suffices to show that, when $\bar{p} = p^m$, the minimum discount factor at which (17) holds is less than the minimum discount factor at which (18) holds. The minimum discount factor at which (17) holds is given by

$$\delta_1 = \frac{b^2}{4\,(b + c\,(n-1)\,p^m - p^m)\,p^m}$$
$$= 1 - c\,(n-1)$$

(recalling that $p^m = b/\left[2\,(1 - c\,(n-1))\right]$). The minimum discount factor at which (18) holds (δ_2) is given by

$$\delta_2 = \frac{(b + c\,(n-1)\,p^m)^2 - 4\,(b + c\,(n-1)\,p^m - p^m)\,p^m}{(b + c\,(n-1)\,p^m)^2 - b^2}$$
$$= \frac{c\,(n-1)}{4 - 3c\,(n-1)}.$$

Hence, a sufficient condition for $\delta_1 < \delta_2$ is

$$(1 - c(n-1))(4 - 3c(n-1)) < c(n-1),$$

or $c(n-1) > 2/3$.   □

References

Abreu, D., 1986. Extremal equilibria of oligopolistic supergames. J. Econ. Theory 39, 191–225.

Abreu, D., 1988. On the theory of infinitely repeated games with discounting. Econometrica 56, 383–396.

Abreu, D., Pearce, D., Stacchetti, E., 1990. Toward a theory of discounted repeated games with imperfect monitoring. Econometrica 58, 1041–1063.

Amarante, M., 2003. Recursive structure and equilibria in games with private monitoring. Econ. Theory 22, 353–374.

Amir, R., Erickson, P., Jin, J., 2017. On the microeconomic foundations of linear demand for differentiated products. J. Econ. Theory 169, 641–665.

Awaya, Y., Krishna, V., 2015. On communication and collusion. Am. Econ. Rev. 106, 285–315.

Awaya, Y., Krishna, V., 2017. Communication and Cooperation in Repeated Games. Working paper.

Bergemann, D., Morris, S., 2013. Robust predictions in games with incomplete information. Econometrica 81, 1251–1308.

Chang, Myong-Hun, 1991. The effects of product differentiation on collusive pricing. Int. J. Ind. Organ. 9 (3), 453–469.

Compte, O., Jenny, F., Rey, P., 2002. Capacity constraints, mergers and collusion. Eur. Econ. Rev. 46, 1–29.

Forges, F., 1986. An approach to communication equilibria. Econometrica 54, 1375–1385.

Fuchs, W., 2007. Contracting with repeated moral hazard and private evaluations. Am. Econ. Rev. 97, 1432–1448.

Fudenberg, Drew, Tirole, Jean, 1991. Game Theory. The MIT Press.

Fudenberg, D., Levine, D., Maskin, E., 1994. The folk theorem with imperfect public information. Econometrica 62, 997–1039.

Gossner, O., Hörner, Johannes, 2010. When is the lowest equilibrium payoff in a repeated game equal to the minmax payoff? J. Econ. Theory 145, 63–84.

Häckner, J., 1996. Optimal symmetric punishments in a bertrand differentiated products duopoly. Int. J. Ind. Organ. 14, 611–630.

Kandori, M., 1991. Cooperation in Finitely Repeated Games with Imperfect Private Information. Working paper.

Kandori, M., 2002. Introduction to repeated games with private monitoring. J. Econ. Theory 102, 1–15.

Kloosterman, A., 2015. Public information in Markov games. J. Econ. Theory 157, 28–48.

Lambson, V.E., 1987. Optimal penal codes in price-setting supergames with capacity constraints. Rev. Econ. Stud. 54, 385–397.

Levin, J., 2003. Relational incentive contracts. Am. Econ. Rev. 93, 835–857.

Lorentz, G.G., 1953. An inequality for rearrangements. Am. Math. Mon. 60, 176–179.

MacLeod, B.W., 2003. Optimal contracting with subjective evaluation. Am. Econ. Rev. 93, 216–240.

Mailath, G.J., Matthews, S.A., Sekiguchi, T., 2002. Private strategies in finitely repeated games with imperfect public monitoring. Contrib. Theor. Econ. 2.

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

Moulin, H., Vial, J.-P., 1979. Strategically zero-sum games: the class of games whose mixed strategy equilibria cannot be improved upon. Int. J. Game Theory 7, 201–221.

Myerson, R.B., 1986. Multistage games with communication. Econometrica 54, 323–358.

Myerson, R.B., 1991. Game Theory: Analysis of Conflict. Harvard University Press, Cambridge.

Pai, M.M., Roth, A., Ullman, J., 2016. An anti-folk theorem for large repeated games with imperfect monitoring. ACM Trans. Econ. Comput. 5.

Ross, T., 1992. Cartel stability and product differentiation. Int. J. Ind. Organ. 10, 1–13.

Sekiguchi, T., 2002. Existence of nontrivial equilibria in repeated games with imperfect private monitoring. Games Econ. Behav. 40, 299–321.

Sekiguchi, T., 2005. Uniqueness of equilibrium payoffs in finitely repeated games with imperfect monitoring. Jpn. Econ. Rev. 56, 317–331.

Sorin, S., 1986. On repeated games with complete information. Math. Oper. Res. 11, 147–160.

Stahl, D.O., 1991. The graph of Prisoners' dilemma supergame payoffs as a function of the discount factor. Games Econ. Behav. 3, 368–384.

Stigler, G.J., 1964. A theory of oligopoly. J. Polit. Econ. 72, 44–61.

Sugaya, T., 2016. The Folk Theorem in Repeated Games with Private Monitoring. Working paper.

Sugaya, T., Wolitzky, A., 2017. Bounding equilibrium payoffs in repeated games with private monitoring. Theor. Econ. 12, 691–729.

Wernerfelt, B., 1989. Tacit collusion in differentiated Cournot games. Econ. Lett. 29, 303–306.