Performance Feedback in Long-Run Relationships: A Rate of Convergence Approach*

Takuo Sugaya Alexander Wolitzky

Stanford GSB MIT

August 2, 2023

Abstract

In repeated agency problems and games, witholding feedback about past performance from the agent/players relaxes incentive constraints and thereby expands the set of implementable outcomes. We investigate the value of withholding feedback by comparing equilibrium payoffs in repeated games with public signals and in games where signals are drawn from the same distribution but are observed only by a principal/mediator. Under an identifiability condition, we find that the value of withholding feedback is small, in that inefficiency is of the same $1-\delta$ power order in both cases. Thus, while private strategies or monitoring (e.g., "review strategies") can outperform public ones for a fixed discount factor, they cannot accelerate the rate of convergence to the efficient payoff frontier.

Keywords: repeated games, repeated agency, imperfect monitoring, feedback, review strategies, rate of convergence, martingales

JEL codes: C72, C73

*This paper supercedes our earlier paper, "Rate of Convergence in Repeated Games: A Universal Speed Limit." We thank Drew Fudenberg, Johannes Hörner, Stephen Morris, and Satoru Takahashi for helpful comments.

1 Introduction

The design of performance feedback rules is an important aspect of principal-agent relationships. While providing feedback can have many practical benefits, a well-known reason *not* to give feedback is that, by informing agents of their own and others' past performance, feedback lets agents game incentive schemes. In other words, withholding feedback pools agents' information sets, which relaxes incentive constraints and thereby expands the set of implementable outcomes. The goal of this paper is to systematically measure this benefit from withholding feedback in standard repeated agency problems and games.

To assess the value of withholding feedback, we consider a repeated game where, in each period, players take actions a, and a signal y is drawn from a distribution $p(y|a)$, which we assume has non-moving support. We compare the equilibrium payoff sets in a version of the game with full feedback (or *public monitoring*), where the signal y is publicly observed, and a version with no feedback, where the signal y is observed only by a principal or mediator, who recommends actions to the players. We call these two versions of the game the *public game* and the *blind game*. By the revelation principle (Forges, 1986), for any discount factor δ, the equilibrium payoff set is (weakly) larger in the blind game than in the public game. Our question is, how much larger?

For any fixed discount factor $\delta < 1$, this question is difficult to answer in any generality, because we lack an explicit characterization of the equilibrium payoff set in public games, and we know even less about the equilibrium payoff set in blind games.[1] We instead adopt a *rate of convergence* approach: under standard identification conditions that ensure that efficiency is attainable in the $\delta \to 1$ limit, how quickly does inefficiency vanish as $\delta \to 1$ in the most efficient equilibrium in the public game, as compared to that in the blind game?

Our main result is that inefficiency is of the same power order of $1 - \delta$ in both games. In this sense, the value of withholding feedback is small.

This result requires some unpacking. A key subtlety is that the order of inefficiency de-

[1] In a public game, the set of perfect public equilibrium payoffs admits a fixed-point characterization due to Spear and Srivastava (1987) (for agency problems) and Abreu, Pearce, and Stacchetti (1990) (for games). However, once we allow private strategies in games with public monitoring or consider blind games, the equilibrium payoff set at a fixed discount factor is intractable, as in repeated games with private monitoring (Kandori, 2002).

pends on the local geometry of the feasible payoff set. In a finite stage game, the boundary of the feasible payoff set is kinked, which implies inefficiency of power order $(1-\delta)^{1/2}$.[2] In a smooth stage game where the boundary of the feasible payoff set has positive quadratic curvature (as in, e.g., Green and Porter (1984), Sannikov (2007, 2008), or Sadzik and Stacchetti (2015)), inefficiency is of order $1-\delta$.[3] Nonetheless, we show that, regardless of the local geometry of the feasible payoff set, inefficiency is of the same power order in the public game and the blind game.

Another subtlety is that inefficiency in the public and blind games can differ by a factor of $(-\log(1-\delta))^{1/2}$ when the boundary of the feasible payoff set is kinked. In contrast, inefficiency differs only by a constant factor (i.e., the rate of convergence is identical) when the boundary is smooth. Thus, while the value of withholding feedback in always "small" (i.e., no improvement in the power rate of convergence), it is somewhat less small in the kinked case (where there can be a log-factor improvement) than in the smooth case (where there is at most a constant-factor improvement).

Whether the rate of convergence in $1-\delta$ is a good guide to the impact of feedback in practice must be assessed on a case-by-case basis, as the rate inevitably hides constant factors that depend on details of the stage game and the monitoring structure, and these constants might outweigh the rate when δ is not very close to 1. However, the fact that the rate does depend on the curvature of the feasible payoff set and can differ between the public and blind games (by a log factor) provides some reassurance that it is a reasonably discerning measure, so our finding that it is unaffected by feedback has some economic significance.

The key force behind our results is that pooling information across periods—which is facilitated by withholding feedback—improves monitoring precision, but also necessitates larger rewards and punishments, which reduces the scope for providing incentives by transferring surplus over time rather than destroying it. As we show, these two effects essentially cancel out. So, little is gained by withholding feedback.

To see the logic in more detail, consider first a finite stage game, where the boundary of the feasible payoff is kinked. With public monitoring, Hörner and Takahashi (2016, henceforth

[2] For public games, this was already observed by Hörner and Takahashi (2016).

[3] For a class of continuous-time principal-agent problems with public monitoring, this was already observed by Sannikov (2008).

HT) established that the rate of convergence toward a strictly individually rational payoff vector is $(1-\delta)^{1/2}$. This result builds on Fudenberg, Levine, and Maskin (1994, henceforth FLM), and similarly relies on *orthogonal enforcement*, where in every period continuation payoffs move along translated tangent hyperplanes. In contrast, in the blind game, one could hope to exceed this rate by (for example) employing a *review strategy* (Radner, 1985; Abreu, Milgrom, and Pearce, 1991; Matsushima, 2004), which aggregates signals over T periods—without providing feedback—before adjusting the players' continuation payoffs.[4]

It is therefore instructive to consider the possible advantage of review strategies over orthogonal enforcement. Heuristically, an efficient review strategy pools information for $T = O\left((1-\delta)^{-1}\right)$ periods—during which the players take constant actions—and then applies a penalty if the number of "good signals" over these periods falls short of a cutoff. Call the number of standard deviations by which the number of good signals falls short of its mean the *score*. Since the number of good signals, normalized by $1/\sqrt{T}$, is approximately normally distributed, for any cutoff score z the probability that a single signal is pivotal is $O\left(\phi(z)/\sqrt{T}\right) = O\left(\phi(z)(1-\delta)^{1/2}\right)$.[5] As stage game payoffs are $O(1-\delta)$, incentive compatibility requires that z is at most $O\left((-\log(1-\delta))^{1/2}\right)$. Thus, the cutoff score can increase only slowly as δ increases, or else the pivot probability decreases very quickly, which violates incentive compatibility. In particular, when $z = O\left((-\log(1-\delta))^{1/2}\right)$, the review strategy's "false positive rate" (and hence its minimum inefficiency) is $\Phi(-z) = O\left(((1-\delta)/(-\log(1-\delta)))^{1/2}\right)$.[6] Review strategies thus yield only a log-factor improvement over orthogonal enforcement when the boundary of the feasible payoff set is kinked. Moreover, no other strategies further improve the rate of convergence.

Next consider the case where the boundary of the feasible payoff set is smooth. Orthogonal enforcement is now more efficient than in the kinked case, as small payoff transfers along translated tangent hyperplanes are more efficient with a smooth boundary. In particular, we show that if the *order of curvature* of the boundary of the feasible payoff set at an exposed point is $\beta \in [1,2]$, then inefficiency under orthogonal enforcement is $O\left((1-\delta)^{\beta/2}\right)$.

[4] Indeed, HT observed that "It is certainly possible that regarding imperfect monitoring, allowing equilibria in private strategies could accelerate the rate of convergence beyond the results that we have derived... This is left for future research." The current paper resolves this question.

[5] Here and throughout the paper, ϕ and Φ denote the standard normal pdf and cdf, respectively.

[6] This follows from the standard normal Mills ratio approximation $\Phi(-z) \approx \phi(z)/z$ for $z \gg 0$.

For instance, in the positive quadratic curvature case (where $\beta = 2$), inefficiency under orthogonal enforcement is $O(1-\delta)$. However, since review strategies involve infrequent, large continuation payoff movements, their efficiency is the same whether the boundary is kinked or smooth: i.e., inefficiency under review strategies remains of power order $(1-\delta)^{1/2}$. Thus, for any $\beta > 1$ (i.e., whenever the boundary is smooth), orthogonal enforcement outperforms review strategies, and in fact attains the fastest possible rate of convergence to efficiency.

Methodologically, we develop a new technique for bounding equilibrium payoffs in repeated games with private monitoring. As in contract theory, the likelihood ratio difference $(p(y|a) - p(y|a'))/p(y|a)$ is a key quantity for providing incentives to take actions a rather than a'. We observe that the likelihood ratio difference is a martingale increment (i.e., the expected likelihood ratio difference under $p(\cdot|a)$ equals 0), so we can apply results from the large deviations theory for martingales to bound the probability that the cumulative likelihood ratio difference over T period grows faster than $O(T)$. This in turn can be used to bound the efficiency of any strategy, regardless of whether signals are public or private.

Relation to the literature. Our finding that the value of withholding feedback is small contrasts with two important strands of prior literature, which both find that this value is large. These strands share the feature that orthogonal enforcement is impossible. This feature reduces efficiency under public monitoring, and thereby generates a large value of withholding feedback.

First, Holmström and Milgrom (1987) study a dynamic principal-agent model where the agent exerts effort over T periods, but consumption occurs only at the end of the game. The value of withholding feedback is large: without feedback, first-best profits can be approximated as $T \to \infty$ using a review strategy that resembles the "penalty contract" of Mirrlees (1975); with feedback, optimal contracts are linear in the count of signal realizations, and profits are bounded away from the first best for all T. The key difference from our setup is that Holmström and Milgrom's model is not a repeated game (as consumption only occurs once), so there is no way to improve efficiency by transferring continuation payoffs over time. That is, orthogonal enforcement is impossible.[7]

[7] Relatedly, a recent paper by Frick, Iijima, and Ishii (2023) considers a one-shot principal-agent model and studies the rate at which profits converge to the first best as the number of signal observations increases.

Second, several papers study principal-agent problems or games that, while repeated, do not permit orthogonal enforcement. Abreu, Milgrom, and Pearce (1991) restrict attention to strongly symmetric equilibria, while Matsushima (2004) and Fuchs (2007) restrict attention to block belief-free equilibria. These classes of equilibria preclude orthogonal enforcement, and, consequently, these papers all find that efficiency is attainable as $\delta \to 1$ only when feedback is withheld.[8] Similarly, in Sannikov and Skrzypacz (2007), pairwise identifiability is violated, so deviations cannot be attributed, and hence orthogonal enforcement is impossible; while Rahman (2014) considers the same model with a mediator who can attribute deviations by randomizing the players' private action recommendations, which restores orthogonal enforcement. In Sannikov and Skrzypacz, the equilibrium set collapses to static Nash; in Rahman, the folk theorem holds.

In past work (Sugaya and Wolitzky, 2017, 2018), we showed that the value of withholding feedback (or "maintaining privacy") is large in some specific repeated and dynamic games when δ is small. For example, our 2018 paper examined how maintaining privacy can help sustain multi-market collusion. In contrast, the current paper shows that the value of privacy in repeated games is small when δ is close to 1.

We also relate to the broader literature on feedback in dynamic agency and games. We consider standard repeated games without payoff-relevant state variables, so feedback concerns only past performance, which is payoff-irrelevant in the continuation game. In contrast, most of the literature on feedback in dynamic agency involves dynamic (non-repeated) games with additional state variables, such as an agent's ability (Ederer, 2010; Smolin, 2021), other agents' progress in a tournament (Gershkov and Perry, 2009; Aoyagi, 2010; Ely et al., 2022), whether a project has been completed (Halac, Kartik, and Liu, 2017; Ely et al., 2023), or the evolution of an exogenous state variable (Ely and Szydlowski, 2020; Orlov, Skrzypacz, and Zryumov, 2020; Ball, 2023). An exception is Lizzeri, Meyer, and Persico (2002), who examine optimal two-period agency contracts with and without a "midterm review."

We also contribute to the literature on review strategies, introduced by Rubinstein (1979),

They find that this rate is much faster for review strategies than for linear contracts.

[8] More precisely, Matsushima considers two-player games where signals are conditionally independent, so each player does not learn about the status of her review. This form of lack of feedback is essential for supporting efficiency in a belief-free equilibrium. Sugaya (2022) shows how mixed strategies can be used to prevent learning with conditionally dependent signals.

Rubinstein and Yaari (1983), and Radner (1985), and developed by Abreu, Milgrom, and Pearce (1991) and Matsushima (2001, 2004). These papers all show that review strategies can support efficient outcomes in various settings when $\delta \to 1$ (or when there is no discounting at all). In contrast, we identify limitations of review strategies when $\delta < 1$, and show that review strategies cannot greatly outperform orthogonal enforcement when δ is close to 1.

Methodologically, the closest papers are HT, who show that inefficiency is $O\left((1-\delta)^{1/2}\right)$ in repeated finite games with public monitoring; and Sugaya and Wolitzky (2023, henceforth SW), who obtain bounds on the strength of players' equilibrium incentives in repeated finite games with arbitrary (e.g., private) monitoring. The arguments in SW are based on variance decomposition and can easily be adapted to show that inefficiency is at least $O(1-\delta)$ when $\beta = 2$ (i.e., the positive quadratic curvature case). For other values for β—in particular, for finite stage games, where $\beta = 1$—SW's bound does not tightly characterize the rate of convergence, and new techniques (i.e., martingale large deviations theory) are required.

Outline. The paper is organized as follows. Section 2 describes the model. Section 3 establishes upper bounds on equilibrium efficiency without feedback. Section 4 establishes that, for any $\beta \in [1,2]$, these bounds are attainable with feedback (excepting a log factor when $\beta = 1$). Combining these results implies that the gains from withholding feedback are small. Section 5 concludes and discusses some extensions.

2 Preliminaries

A *stage game* $G = (I, A, u)$ consists of a finite set of players $I = \{1, \ldots, N\}$, a product set of actions $A = \times_{i \in I} A_i$, and a payoff function $u_i : A \to \mathbb{R}$ for each $i \in I$. We assume that each A_i is a nonempty, compact metric space, and each u_i is continuous.[9] By the Debreu-Fan-Glicksberg theorem, this implies that the stage game admits a Nash equilibrium in mixed actions. We denote the sets of stage-game Nash and correlated equilibria by $\Sigma^{NE} \subseteq \times_{i \in I} \Delta(A_i)$ and $\Sigma^{CE} \subseteq \Delta(A)$, respectively. In addition, we denote the feasible payoff set by $F = \text{co}\left(\{u(a)\}_{a \in A}\right) \subseteq \mathbb{R}^N$, and denote the sets of stage-game Nash and

[9] As is standard, we linearly extend the payoff functions u_i to distributions $\alpha \in \Delta(A)$. Here and throughout, for any compact metric space X, $\Delta(X)$ denotes the set of Borel probability measures on X, endowed with the weak* topology.

correlated equilibrium payoffs by $V^{NE} = \left\{v : v = u(\alpha) \text{ for some } \alpha \in \Sigma^{NE}\right\}$ and $V^{CE} = \left\{v : v = u(\alpha) \text{ for some } \alpha \in \Sigma^{CE}\right\}$. We also let $d(\cdot,\cdot)$ and $\|\cdot\|$ denote the Euclidean metric and norm on $\mathbb{R}^N$, and let $\Lambda = \{\lambda \in \mathbb{R}^N : \|\lambda\| = 1\}$ denote the set of unit vectors (or *directions*) in $\mathbb{R}^N$. Finally, we denote the boundary of F by $\text{bnd}(F)$ and the set of exposed points of F by $\exp(F)$, and, for any $v \in \exp(F)$, denote $\Lambda_v = \{\lambda \in \Lambda : v = \operatorname{argmax}_{w \in F} \lambda \cdot w\}$.[10]

A *monitoring structure* (Y, p) consists of a finite set of possible signal realizations Y and a family of conditional probability distributions $p(y|a)$.[11] We assume that the probability of each signal realization is bounded away from zero: there exists $\omega > 0$ such that $p(y|a) > \omega$ for all $y \in Y, a \in A$. This non-moving support assumption is crucial: e.g., our analysis excludes perfect monitoring.

In a *repeated game with public monitoring* (Abreu, Pearce, and Stacchetti, 1990, henceforth APS; FLM), in each period $t \in \mathbb{N}$, each player i takes an action a_i, and then a signal y is drawn according to $p(y|(a_i))$ and is publicly observed. A history for player i at the beginning of period t takes the form $h_i^t = (a_{i,t'}, y_{t'})_{t'=1}^{t-1}$. A strategy σ_i for player i maps histories h_i^t to distributions over actions $a_{i,t}$. A strategy for player i is *public* if it depends on h_i^t only through its public component $y^t = (y_{t'})_{t'=1}^{t}$. Players choose strategies to maximize discounted expected payoffs, with common discount factor $\delta \in [0, 1)$. A *perfect public equilibrium* (*PPE*) is a profile of public strategies that, beginning at any period t and any public history y^t, forms a Nash equilibrium from that period on. We denote the repeated game with public monitoring with stage game G, monitoring structure (Y, p), and discount factor δ by $\Gamma^P(\delta)$, and we denote the corresponding set of PPE payoff vectors by $E^P(\delta) \subseteq \mathbb{R}^N$.

In a *blind repeated game* (Sugaya and Wolitzky, 2017, 2023), the players are assisted by a mediator. In each period $t \in \mathbb{N}$, (i) the mediator privately recommends an action $r_i \in A_i$ to each player i, (ii) each player i takes an action a_i, and (iii) a signal y is drawn according to $p(y|(a_i))$ and is observed only by the mediator. A history for the mediator at the beginning of period t takes the form $h_0^t = \left((r_{i,t'})_i, y_{t'}\right)_{t'=1}^{t-1}$, while a history for player i just before she takes an action in period t takes the form $h_i^t = \left((r_{i,t'}, a_{i,t'})_{t'=1}^{t-1}, r_{i,t}\right)$. A strategy σ_0 for the mediator maps histories h_0^t to distributions over recommendation profiles $(r_{i,t})_i$, while a

[10] Recall that Λ_v is non-empty iff $v \in \exp(F)$.

[11] Section 5.3 discusses the extension to the $|Y| = \infty$ case.

strategy σ_i for player i maps histories h_i^t to distributions over actions $a_{i,t}$. We denote the blind repeated game with stage game G, monitoring structure (Y, p), and discount factor δ by $\Gamma^B(\delta)$, and we denote the corresponding set of Nash equilibrium payoff vectors by $E^B(\delta) \subseteq \mathbb{R}^N$. Note that a player's payoff in the blind game is not measurable with respect to her own information. The blind game may thus withhold feedback from the players to an unrealistic extent—but this only strengthens our finding that withholding feedback has limited value.

By standard arguments (similar to Forges, 1986), any Nash equilibrium outcome $\mu \in \Delta((A \times Y)^\infty)$ (i.e., any equilibrium distribution over infinite paths of action profiles and signals) in $\Gamma^P(\delta)$ can also be implemented by a Nash equilibrium in $\Gamma^B(\delta)$ where the players follow the mediator's recommendations on path. In particular, $E^P(\delta) \subseteq E^B(\delta)$.

The goal of this paper is assessing the value of withholding feedback from the players. The maximum such value is described by the set of payoffs that are attainable with the smallest possible amount of feedback—i.e., are attainable in $\Gamma^B(\delta)$—but are not attainable with the largest possible amount of feedback—i.e., are not attainable in $\Gamma^P(\delta)$. Since our main result is that the value of withholding feedback is small when δ is close to 1, there is no harm in "over-estimating" the value by restricting attention to PPE in $\Gamma^P(\delta)$, while admitting all Nash equilibria in $\Gamma^B(\delta)$. That is, we estimate the value of withholding feedback by the size of the set $E^B(\delta) \setminus E^P(\delta)$.

Remark 1 *The model is easily adapted to allow a player with commitment power (such as the principal in a standard principal-agent model) or players with perfectly observed actions (such as a principal who offers contracts each period in a relational contracting model). A player with commitment power is treated like any other player, except that no incentive constraints are imposed on her strategy. For example, in a principal-agent model,* Σ^{NE} *is the set of* (principal, agent) *mixed action profiles where the agent does not have a profitable deviation. Moreover, it suffices to impose non-moving support only for the agent, so that* $\operatorname{supp} p(\cdot|a) = \operatorname{supp} p(\cdot|a')$ *for all* a, a' *that agree on the principal's action. We explain how to accommodate players with observable actions (without commitment power) in Section 5.1.*

3 Maximum Efficiency without Feedback

3.1 Main Result

Our first theorem gives an upper bound on the rate of convergence of $E^B(\delta)$ toward an exposed point $v \in \exp(F)$ that is not attainable as a static correlated equilibrium. The upper bound depends on the order of curvature of the boundary of F at v.

Definition 1 *Fix an exposed point $v \in \exp(F)$. For $\beta \geq 1$, the boundary of F has* max-curvature of order at least β at v *if, for all $\lambda \in \Lambda_v$, there exists $\eta > 0$ such that*

$$\lambda \cdot (v - w) \geq \eta d(v, w)^{\beta} \qquad \textit{for all } w \in \operatorname{bnd}(F).$$

The boundary of F has max-curvature of order β at v *if*

$$\beta = \inf\left\{\tilde{\beta} : \operatorname{bnd}(F) \;\textit{ has }\; \max\textit{-curvature of order at least } \tilde{\beta} \textit{ at } v\right\}.$$

This says that moving away from v in F entails an efficiency loss of order at least β, relative to Pareto weights λ. (Or, heuristically, $\operatorname{bnd}(F)$ is approximated by a power function of degree β at v.) To understand the definition, the key cases to consider are $\beta = 1$, $\beta = 2$, and the limit case $\beta = \infty$.

- The $\beta = 1$ case arises when the stage game G is finite, as in APS, FLM, or HT. Here, F is the convex hull of a finite collection of points, so the boundary of F is kinked at every extreme point. This implies a first-order loss from moving away from any extreme point.

- The $\beta = 2$ case arises when the boundary of F has positive quadratic curvature. This is the typical case in smooth games or agency models with continuous actions, such as Green and Porter (1984), Sannikov (2007, 2008), or Sadzik and Stacchetti (2015).

- The $\beta = \infty$ case arises when the boundary of F is linear at v. This is the case in repeated games with transferable utility, as in Athey and Bagwell (2001), Levin (2003), or Goldlücke and Kranz (2012).

- To appreciate the role of the max in the definition, suppose that $N = 2$, $(0,0) \in F$, and the local boundary of F at $(0,0)$ is given by $f(x) = -x$ if $x < 0$ and $f(x) = x^2$ if $x \geq 0$. Then the max-curvature of $\text{bnd}(F)$ at $(0,0)$ is 2.

Most of our insights can be obtained when $\beta \in \{1,2\}$. We cover the cases where $\beta \in (1,2)$ and $\beta > 2$ for completeness.

Theorem 1 *Fix an exposed point $v \in \exp(F) \backslash V^{CE}$ and a direction $\lambda \in \Lambda_v$. If* $\text{bnd}(F)$ *has* max-*curvature of order β at v, then there exists $c > 0$ such that*

$$\lambda \cdot (v - w) \geq c \times \begin{cases} \left(\frac{1-\delta}{\max\{-\log(1-\delta),1\}} \right)^{1/2} & \text{if } \beta = 1, \\ (1-\delta)^{\beta/2} & \text{if } \beta \in (1,2], \\ (1-\delta)^{\beta-1} & \text{if } \beta > 2, \end{cases} \qquad \text{for all } \delta \in [0,1) \text{ and } w \in E^B(\delta). \tag{1}$$

The key implications of Theorem 1 are as follows:

- For Pareto weights where welfare is maximized at a point where bnd(F) is kinked (i.e., $\beta = 1$), inefficiency is at least $O\left(\left((1-\delta) / (-\log(1-\delta))\right)^{1/2}\right)$.
- For Pareto weights where welfare is maximized at a point where bnd(F) has positive quadratic curvature (i.e., $\beta = 2$), inefficiency is at least $O(1-\delta)$.

We will see that both of these bounds—as well as the $(1-\delta)^{\beta/2}$ bound for $\beta \in (1,2)$—are tight. Moreover, with public monitoring, the bound in the kinked case remains tight up to log-factor slack, while the bound in the $\beta \in (1,2]$ case remains tight up to constant-factor slack. These results imply that the gains from withholding feedback are small.

In contrast, for Pareto weights where welfare is maximized at a point where bnd(F) is approximately linear, Theorem 1 allows inefficiency much smaller than $1-\delta$. This bound is tight in the $\beta \to \infty$ limit, as in some games with linear Pareto frontiers efficiency is exactly achieved at some $\delta < 1$ (e.g., Athey and Bagwell, 2001). We conjecture that the $(1-\delta)^{\beta-1}$ bound given by Theorem 1 is in fact tight for any $\beta > 2$—in that there exists some game and $v \in \text{ext}(F) \backslash V^{CE}$ with max-curvature of order β that can be approached at rate $(1-\delta)^{\beta-1}$—but we have not proved this.

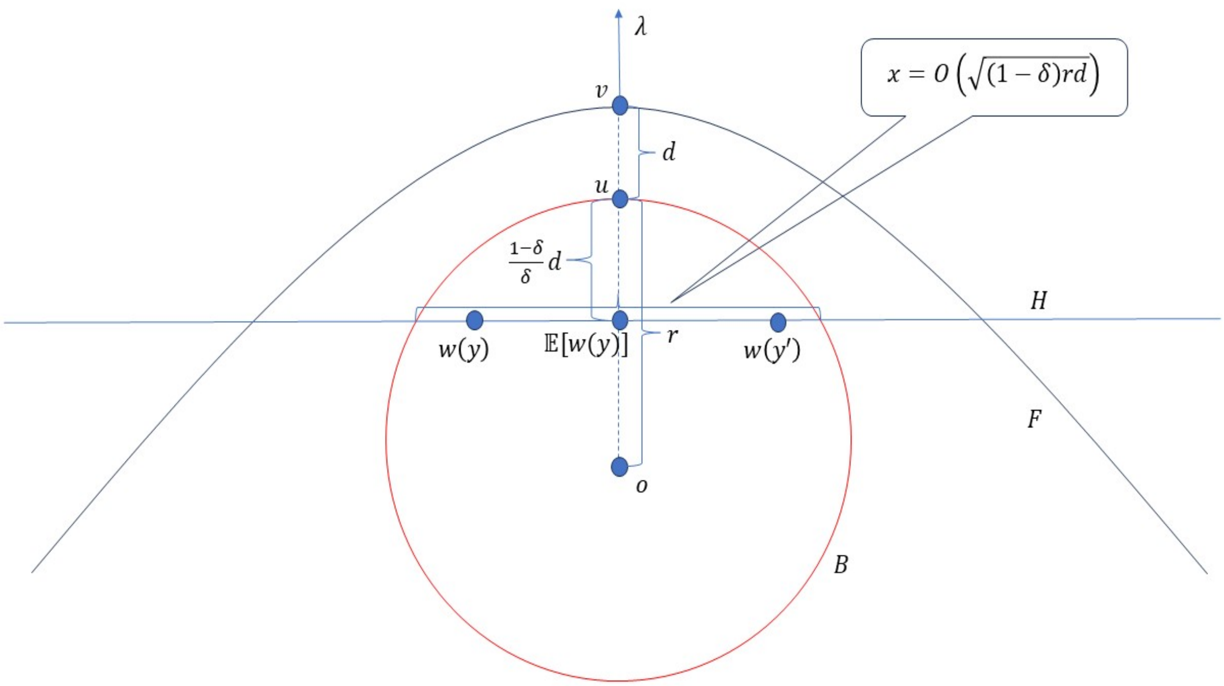

Figure 1: Self-Generating a Ball. To maximize efficiency, r and d must be chosen to minimize d subject to the constraints that $B \subseteq F$ and x is at least $O(1-\delta)$.

We can explain the intuition for Theorem 1 in two steps. First we explain why the conclusion of Theorem 1 holds under public monitoring: that is, why (1) holds for all $w \in E^P(\delta)$. Then we explain why the same conclusion holds in the blind game.

The logic for why (1) holds for all $w \in E^P(\delta)$ builds on APS, FLM, and HT. Following these authors, we ask how close a self-generating ball $B \subseteq F$ can be to an exposed point v. To answer this, fix $\lambda \in \Lambda_v$, let $d = d(B, v)$ be the desired distance, and (without loss) let $u = v - d\lambda$ be the closest point to v in B. (See Figure 1.) Consider decomposing u into instantaneous payoff v and continuation payoffs $(w(y))_y$ that lie on the translated tangent hyperplane H with normal vector λ passing through the point $\mathbb{E}[w(y)] = v - ((1-\delta)/\delta) d\lambda$. The diameter of $H \cap B$, which we denote by x, is then the largest available continuation payoff movement, which by incentive compatibility must be at least $O(1-\delta)$. At the same time, denoting the radius of the ball B by r, the Pythagorean theorem gives $(x/2)^2 + (r - ((1-\delta)/\delta) d)^2 = r^2$, and hence $x = O\left(\sqrt{(1-\delta) rd}\right)$. It follows that the product rd must be at least $O(1-\delta)$.

We are thus left with the following geometry question: for a point v where the (max-)curvature of $\text{bnd}(F)$ equals β, what is the smallest distance d such that the ball B with radius r satisfying $rd = O(1-\delta)$ and center $v - (r+d)\lambda$ lies below F (in the λ-direction)? We leave it to the reader to verify that, when $\beta \in [1,2]$, the answer is $O\left((1-\delta)^{\beta/2}\right)$. For example, when $\beta = 2$ the closest self-generating ball has radius $O(1)$ and distance $O(1-\delta)$ from v, while when $\beta = 1$ the closest self-generating ball has radius $O\left((1-\delta)^{1/2}\right)$ and distance $O\left((1-\delta)^{1/2}\right)$ from v.[12] Intuitively, whenever $\beta < 2$, the self-generation condition $B \subseteq F$ implies that the closest self-generating ball B must shrink as it approaches v (in particular, its radius r is $(1-\delta)^{1-\beta/2}$), and the distance d is then determined by the incentive compatibility condition $rd = O(1-\delta)$, which gives $d = O\left((1-\delta)^{\beta/2}\right)$.

In contrast, when $\beta > 2$, the self-generation condition $B \subseteq F$ is slack in a neighborhood of v: any ball B has finite quadratic curvature at v, while the quadratic curvature of F at v is infinite. Of course, B cannot be infinitely large, because the constraint $B \subseteq F$ binds somewhere away from v. However, since only continuation payoff movements of size $O(1-\delta)$ are required for incentives, it suffices to take B such that $rd = O(1-\delta)$ and every point in B at distance $O(1-\delta)$ from u is at distance at least d from $\text{bnd}(F)$. Another geometric argument (the details of which we omit) shows that the smallest such d is $O\left((1-\delta)^{\beta-1}\right)$.

The main insight of this paper is that nearly the same efficiency bounds apply without feedback. More precisely, in the smooth case ($\beta \in (1,2]$), minimum inefficiency is exactly the same for $E^B(\delta)$ and $E^P(\delta)$; while in the kinked case ($\beta = 1$), withholding feedback can reduce inefficiency by a log factor. These results hold even though the set $E^B(\delta)$ is not self-generating—so the geometric arguments just given do not apply—and few general bounds on equilibrium payoffs in repeated games with private monitoring or mediation are known.

A simple intuition for these results relies on comparing orthogonal enforcement and review strategies, as described in the introduction. That argument explains why review strategies improve efficiency by a log factor in the kinked case and otherwise yield no improvement. Of course, the proof of Theorem 1 must account for arbitrary strategies. We outline the proof in the next subsection. The basic logic is that if a repeated game Nash equilibrium gives payoffs close to $v \in \exp(F)$, then the stage game payoff must be close to v almost all the time along the equilibrium path of play. Since signals have full support, this implies

[12] The latter fact was already observed by HT.

that payoffs must still be close to v almost all the time even after low-probability (but still on-path) signal realizations. This in turn implies that, on average, equilibrium continuation play does not vary much with the signal realizations. But then, if $v \notin V^{CE}$, we can conclude that δ must be so high that even small variations in continuation play can provide strong incentives.[13]

We mention a couple technical aspects of the statement of Theorem 1. First, generically, the condition $v \in \exp(F) \backslash V^{CE}$ is equivalent to $v \in \exp(F) \backslash V^{NE}$: since v is extremal, the distinction only matters in the non-generic case where v is attained at two different pure action profiles. Second, the condition $\lambda \in \Lambda_v$ (i.e., $v = \text{argmax}_{w \in F} \lambda \cdot w$) cannot be weakened to $v \in \text{argmax}_{w \in F} \lambda \cdot w$. To see this, consider the stage game

	L	R
C	$1,1$	$0,1$
$D1$	$2,0$	$-2,0$
$D2$	$-2,0$	$2,0$

The point $v = (1,1)$ is exposed and is not attainable as a static CE, $\text{bnd}(F)$ has curvature of order 1 (i.e., a kink) at v, and $v \in \text{argmax}_{w \in F} \lambda \cdot w$ for $\lambda = (0,1)$. But, the point $w = (0.5, 1)$ is attained by the static NE $\left(C, \frac{1}{2}L + \frac{1}{2}R\right)$ (so $w \in E^B(\delta)$ for all $\delta \in [0,1)$) and satisfies $\lambda \cdot w = \lambda \cdot v$, so the conclusion of Theorem 1 fails.

3.2 Proof Sketch

We sketch the proof of Theorem 1, deferring the details to the appendix. Fix some $v \in \exp(F) \backslash V^{CE}$ and $\lambda \in \Lambda_v$, and consider any $\delta \in [0,1)$ together with a Nash equilibrium σ in $\Gamma^B(\delta)$ with equilibrium payoff w. We wish to derive a lower bound for $\lambda \cdot (v - w)$.

We introduce some notation. Let $\mu \in \Delta((A \times Y)^\infty)$ be the repeated game outcome induced by σ. The outcome μ defines, in particular, the marginal distribution over period-t

[13] This logic is the same as that of Theorem 6.5 of FLM (who credit Madrigale, 1986), which says that an extremal non-static Nash payoff vector v cannot be exactly attained for any $\delta < 1$ under full-support monitoring. FLM state this result for PPE, but the same argument works for Nash. Theorem 1 is a quantitative version of this result.

action profiles, $\alpha_t \in \Delta(A)$, as well as the *occupation measure* $\alpha \in \Delta(A)$, defined as

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

Note that the occupation measure is a sufficient statistic for equilibrium payoffs, as, by linearity of u,

$$w = (1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}u(\alpha_t) = u\left((1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}\alpha_t\right) = u(\alpha).$$

Next, for each player i, let S_i denote the set of functions $s_i : A_i \to A_i$, which we call *manipulations*. For any $i \in I$, $\alpha \in \Delta(A)$, and $s_i \in S_i$, define the *deviation gain*

$$g_i(s_i,\alpha) = \sum_{a\in A}\alpha(a)\left(u_i(s_i(a_i),a_{-i}) - u_i(a)\right).$$

The interpretation is: if the recommended action profile a is drawn according to α and player i takes $s_i(a_i)$ when recommended a_i rather than obeying the recommendation, her expected payoff gain is $g_i(s_i,\alpha)$. Note that, since $v \in \exp(F)\backslash V^{CE}$, if w is close to v then there exist i and s_i such that $g_i(s_i,\alpha)$ is bounded away from 0. Fix such i and s_i.

Finally, for any complete history of play $h = (a_t,y_t)_{t=1}^{\infty}$, let

$$u_{i,t}(h) = u_i(a_t) \qquad \text{and} \qquad \ell_t(h) = \frac{p(y_t|a_t) - p(y_t|s_i(a_{i,t}),a_{-i,t})}{p(y_t|a_t)}.$$

That is, $u_{i,t}(h)$ is player i's realized period-t payoff at history h, and $\ell_t(s_i,h)$ is the realized *likelihood ratio difference* of the period-t signal y_t at the period-t action profile a_t, as compared to the action profile $(s_i(a_{i,t}),a_{-i,t})$ that results when player i manipulates according to s_i. Note that

$$\mathbb{E}^{\mu}\left[\ell_t(h)\,|\,(a_t,y_t)_{t'=1}^{t-1}\right] = 0,$$

so $\mathcal{L}_T = \sum_{t=1}^{T}\ell_t(h)$ is a martingale. Intuitively, the martingale $\mathcal{L}_T$ determines the informativeness of the first T signals for distinguishing equilibrium play by player i from actions taken according to the manipulation s_i.

A simple consequence of incentive compatibility (Lemma 4 in the appendix) is that, for each period t, we have

$$g_i(s_i, \alpha_t) \le \mathbb{E}^{\mu}\left[\ell_t(h) \sum_{t'=t+1}^{\infty} \delta^{t'-t} u_{i,t'}(h)\right].$$

This inequality holds because, if it were violated, player i could gain by obeying her recommendation in every period other than t, while manipulating according to s_i in period t. Given this inequality, since $\mathrm{bnd}(F)$ has max-curvature of order β at v, it follows that $\lambda \cdot (v - w)$ is no less than the value of the convex program

$$\begin{aligned}
\inf_{(u_{i,t}(h))_{t,h}} \quad \mathbb{E}^{\mu}\left[(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}\eta\left|u_{i,t}(h) - v_i\right|^{\beta}\right] & \qquad \text{subject to} \\
g_i(s_i, \alpha_t) \le \mathbb{E}^{\mu}\left[\ell_t(h)(1-\delta)\sum_{t'=t+1}^{\infty}\delta^{t'-t}\left(u_{i,t'}(h) - v_i\right)\right] & \qquad \text{for all } t, \qquad \text{and} \\
\left|u_{i,t}(h) - v_i\right| \le \bar{u} & \qquad \text{for all } t, h,
\end{aligned} \tag{2}$$

where $\bar{u}$ is the range of u_i. Intuitively, this program minimizes the β^{th} moment of the deviation of player i's realized repeated game payoff from v_i, subject to a relaxed version of incentive compatibility, and feasibility.

To prove the theorem, it remains to bound the value of the convex program (2) as a function of δ and β. This can be done by duality, using the fact that $g_i(s_i, \alpha)$ is bounded away from 0, and employing some martingale large deviations bounds for the moments of $\mathcal{L}_T$. Intuitively, these bounds reflect the fact that sequences of signals with large cumulative likelihood ratio differences—which are highly informative when they occur—also occur with low equilibrium probability, and hence do not provide a large amount of information on average. Incentive compatibility thus requires that player i's realized payoff varies substantially on-path, which in turn implies the desired lower bound for ex ante inefficiency.

3.3 Tightness of the Bound in the Kinked Case

We will see in the next section that inefficiency of order $(1-\delta)^{\beta/2}$ can be attained when $\beta \in [1,2]$ under public monitoring. Here we show that, when $\beta = 1$ (i.e., in the kinked case), the inefficiency bound of $((1-\delta)/-\log(1-\delta))^{1/2}$ cannot be improved in the blind game. Thus, withholding feedback can accelerate the rate of convergence by at most $(-\log(1-\delta))^{-1/2}$ when $\beta = 1$.

We consider a one-sided prisoners' dilemma, where the stage game is

	L	R
C	$2,2$	$0,0$
D	$3,0$	$1,1$

and the monitoring structure is given by $Y = \{0,1\}$ and

$$p(y=1|a) = \begin{cases} 1/2 & \text{if } a_1 = C, \\ 1/4 & \text{if } a_1 = D. \end{cases}$$

We investigate the possibility of attaining payoffs close to $(2,2)$. Note that signals do not depend on player 2's action, but this does not pose an obstacle to attaining payoffs close to $(2,2)$, because at action profile (C,L) player 2 is taking a static best response.

Proposition 1 *In the one-sided prisoner's dilemma, there exist $c > 0$ and $\bar{\delta} < 1$ such that, for all $\delta > \bar{\delta}$, there exists $v \in E^B(\delta)$ satisfying*

$$v_1 = v_2 > 2 - c\left(\frac{1-\delta}{-\log(1-\delta)}\right)^{1/2}.$$

Proof. We sketch the proof, providing the details in the appendix. Consider a review strategy where the game is divided into blocks of T consecutive periods. We take $T = \lfloor \rho/(1-\delta) \rfloor$, where $\rho > 0$ is a small number to be determined: note that $\rho \approx 1 - \delta^T$ when $\delta \approx 1$. In the first block, the players are prescribed (C,L) in every period. At the end of the first block—as well any subsequent block where (C,L) is prescribed—the mediator records the summary statistic

$$E = \mathbf{1}\left\{ \frac{1}{\sqrt{T}} \sum_{t=1}^{T} (2y_t - 1) \leq -z \right\},$$

where $z > 0$ is a large number to be determined. (Here periods are numbered from the start of the block.) If $E = 0$, the players "pass the review," and (C, L) is prescribed in the next block. If $E = 1$, then with some probability $q \in [0, 1]$ (which also remains to be determined), the players fail the review and (D, B) is prescribed forever. With the complementary probability $1 - q$, the players pass the review anyway, and (C, L) is prescribed in the next block.

We show that the parameters ρ, z, and q can be chosen so that this strategy profile is an equilibrium that yields payoff $v > 2 - c\sqrt{(1-\delta) / (-\log(1-\delta))}$ for each player.[14]

Let p be the probability that $E = 1$ when player 1 takes C throughout a block, let p_1 be the probability that $E = 1$ when player 1 takes D once and takes C $T - 1$ times, and let p_T be the probability that $E = 1$ when player 1 takes D throughout. Observe that v is given by

$$v = \left(1 - \delta^T\right) 2 + \delta^T (1 - pq) v \quad \Longleftrightarrow \quad v = 2 - \frac{\delta^T}{1 - \delta^T} pqv. \tag{PK}$$

At the same time, the incentive conditions that player 1 prefers to take C throughout a block where (C, L) is prescribed, rather than taking D in period 1 only, or always taking D, are

$$1 - \delta \leq \delta^T (p_1 - p) qv \qquad \text{and} \tag{IC$_1$}$$

$$1 - \delta^T \leq \delta^T (p_T - p) qv. \tag{IC$_T$}$$

Conditions (IC$_1$) and (IC$_T$) are obviously necessary for the review strategy to be an equilibrium; moreover, as shown by Matsushima (2004, p. 846), they are also sufficient.[15] It thus suffices to find ρ, z, and q that satisfy (PK), (IC$_1$), and (IC$_T$) with $v > 2 - c\sqrt{(1-\delta) / (-\log(1-\delta))}$.

Since the random variable $2y_t - 1$ has zero mean and unit variance when player 1 takes C, when $\delta \approx 1$ the central limit theorem implies that the test statistic $\left(1/\sqrt{T}\right) \sum_{t=1}^{T} (2y_t - 1)$

[14] We write v instead of v_i here, since the players' payoffs are the same.

[15] Matsushima considered repeated games with two players who receive conditionally independent signals. Conditional independence implies that a player does not learn about her opponents signals during a review block, just as players do not learn about the mediator's signals in Γ^B. The same argument thus applies here.

is approximately $N(0,1)$, so that $p \approx \Phi(-z)$ and $p_1 - p \approx \left(\frac{3}{4} - \frac{1}{2}\right) \frac{\phi(-z)}{\sqrt{T}} = \frac{\phi(-z)}{4\sqrt{T}}$. Therefore, the smallest value for q that satisfies (IC_1) is approximately $\frac{1-\delta}{\delta^T} \frac{4\sqrt{T}}{\phi(-z)v}$. For this value to be less than 1 when $v \approx 2$, we must have $\frac{1-\delta}{\delta^T} \frac{2\sqrt{T}}{\phi(-z)} \leq 1$. Since $(1-\delta)\sqrt{T}/\delta^T \approx \sqrt{1-\delta}$ and $\phi(-z) = \exp(-z^2/2)/\sqrt{2\pi}$, it follows that $z \leq c_0 \sqrt{-\log(1-\delta)}$ for some constant c_0. At the same time, by (PK) and (IC_1), we have $v \approx 2 - \frac{\delta^T}{1-\delta^T} \Phi(-z) qv$ and $1 - \delta \approx \delta^T \frac{\phi(-z)}{4\sqrt{T}} qv$, and hence

$$v \approx 2 - 4 \frac{(1-\delta)\sqrt{T}}{1-\delta^T} \frac{\Phi(-z)}{\phi(-z)} \approx 2 - \frac{4}{\sqrt{\rho}} \frac{\sqrt{1-\delta}}{z},$$

where the second approximation follows because, when ρ is small and z is large, $\frac{\sqrt{(1-\delta)T}}{1-\delta^T} \approx \frac{\sqrt{\rho}}{\rho} = \frac{1}{\sqrt{\rho}}$ and $\frac{\Phi(-z)}{\phi(-z)} \approx \frac{1}{z}$. Taking the largest possible value for z for which $q \leq 1$—i.e., $z = c_0 \sqrt{-\log(1-\delta)}$—now gives the desired bound for v. Finally, with this value for z we have $p_T \approx 1$ and $p \approx 0$, so when $\rho \approx 0$, $q \approx 1$, and $v \approx 2$, (IC_T) holds, as the LHS is close to 0 and the RHS is close to 2. The constructed strategies therefore form an equilibrium. ■

4 Attainable Efficiency with Feedback

We now ask whether the maximum efficiency levels identified in Theorem 1 can be attained under public monitoring. To this end, denote the set of feasible and strictly individually rational payoffs by $F^* = \left\{ v \in F : v_i > \underline{v}_i := \min_{\alpha_{-i} \in \times_{j \neq i} \Delta(A_j)} \max_{a_i \in A_i} u_i(a_i, \alpha_{-i}) \ \forall i \right\}$. For $v \in \mathrm{bnd}(F^*)$, define $\Lambda_v^* = \{\lambda \in \Lambda : v \in \mathrm{argmax}_{w \in F^*} \lambda \cdot w\}$.

Definition 2 *Fix a boundary point $v \in \mathrm{bnd}(F^*)$. For $\beta \geq 1$, the boundary of F^* has* min-curvature of order at most β *at v if, for all $\lambda \in \Lambda_v^*$, there exist $k > 0$ such that*

$$\lambda \cdot (v - w) < k d(v, w)^\beta \qquad \textit{for all } w \in \mathrm{bnd}(F^*).$$

The boundary of F has min-curvature of order β at v *if*

$$\beta = \sup \left\{ \tilde{\beta} : \mathrm{bnd}(F) \ \textit{has} \ \min\textit{-curvature of order at most } \tilde{\beta} \ \textit{at } v \right\}.$$

Definition 2 is a converse of Definition 1. It says that moving away from v along the boundary of F entails an efficiency loss of order at most β, relative to Pareto weights λ. For example, if $N = 2$, $(0, 0) \in F$, and the local boundary of F at $(0, 0)$ is given by $f(x) = -x$ if $x < 0$ and $f(x) = x^2$ if $x \geq 0$, then the min-curvature of $\text{bnd}(F)$ at $(0, 0)$ is 1. Note that, in general, the min-curvature of $\text{bnd}(F)$ at v is always at least 1 and at most the max-curvature.

The following assumption generalizes standard identification conditions for the public-monitoring folk theorem to the case where action sets can be infinite.

Assumption 1 There exists $\bar{x} > 0$ such that the following conditions hold:

i. For each i, there exists a minmax profile against i, $\alpha^i \in \times_{j \neq i} \Delta(A_j) \times A_i$, and $x_j : Y \to [-\bar{x}, \bar{x}]$ for each $j \neq i$, such that

$$a_j \in \underset{a_j'}{\text{argmax}}\, u_j\left(a_j', \alpha_{-j}\right) + \mathbb{E}\left[x_j(y)\,|a_j', \alpha_{-j}\right] \text{ for all } j \neq i \text{ and } a_j \in \text{supp}(\alpha_j). \tag{3}$$

ii. For each $a \in A$, $\nu \in \{-1, +1\}$, and (i, j) with $i \neq j$, there exists $x_i : Y \to [-\bar{x}, \bar{x}]$ such that

$$a_i \in \underset{a_i'}{\text{argmax}}\, u_i\left(a_i', \alpha_{-i}\right) + \mathbb{E}\left[x_i(y)\,|a_i', \alpha_{-i}\right], \tag{4}$$

$$a_j \in \underset{a_j'}{\text{argmax}}\, \mathbb{E}\left[\nu x_i(y)\,|a_j', a_{-j}\right]. \tag{5}$$

Intuitively, when payoff transfers of magnitude at most $\bar{x}$ are available, Assumption 1(i) says that players $-i$ can be incentivized to minmax player i, and Assumption 1(ii) says that player i can be incentivized to take a_i via transfers from player j without affecting player j's incentive to take a_j. These conditions are similar to assumptions (A1)–(A3) of Kandori and Matsushima (1998). The difference is that, since we allow $|A| = \infty$, Assumption 1 is stated directly in terms of the existence of transfers x that satisfy (3)–(5), while Kandori and Matsushima assume that $|A| < \infty$ and hence can state their conditions in terms of the convex hull of the set of vectors of signal probabilities generated by different actions, from which the existence of transfers x satisfying (3)–(5) follows from the separating hyperplane theorem.

We mention a class of infinite games where Assumption 1(ii) holds. (If Assumption 1(ii) holds, then even if Assumption 1(i) fails a Nash-threat folk theorem still holds, i.e., Theorem 2 holds with F^* replaced by the set of feasible payoffs that Pareto dominate a convex combination of static Nash payoffs.) Say that the game is *linear-concave* if (i) for each i, A_i is a compact interval $[\underline{a}_i, \bar{a}_i] \subseteq \mathbb{R}$, and $u_i(a_i, a_{-i})$ is differentiable and concave in a_i for every a_{-i} and satisfies the Inada conditions $\lim_{a_i \to \underline{a}_i} \frac{d}{da_i} u_i(a_i, a_{-i}) = \infty$ and $\lim_{a_i \to \bar{a}_i} \frac{d}{da_i} u_i(a_i, a_{-i}) = -\infty$; and (ii) the public signal is a D-dimensional real variable, $Y = \times_{d=1}^{D} Y^d \subseteq \mathbb{R}^D$, and $\mu^d(a) = \mathbb{E}\left[y^d | a\right]$ is a linear function of a for each dimension d. In a linear-concave game, let $M^i(a) = \left(\frac{d}{da_i} \mu^d(\hat{a}) \Big|_{\hat{a}=a} \right)_d$ be a D-dimensional vector representing the sensitivity of the mean public signal to player i's action. Say that *pairwise identifiability* holds if for any a and $i \neq j$, $M^i(a) \neq 0$ and the spans of $M^i(a)$ and $M^j(a)$ intersect only at the origin.[16]

Proposition 2 *In any linear-concave game satisfying pairwise identifiability, Assumption 1(ii) holds.*

Under Assumption 1, we examine the rate of convergence of $E^P(\delta)$ toward a strictly individually rational payoff vector $v \in \text{bnd}(F^*)$. For finite games (where $\beta = 1$), HT show that the rate of convergence of $E^P(\delta)$ equals $(1-\delta)^{1/2}$. Thus, withholding feedback can accelerate the rate of convergence by at most $(-\log(1-\delta))^{-1/2}$ in the kinked case. We now show that, for $\beta \in [1, 2]$, the rate of convergence of $E^P(\delta)$ toward a strictly individual payoff vector $v \in \text{bnd}(F^*)$ equals $(1-\delta)^{\beta/2}$. (We discuss the $\beta > 2$ case below.) Thus, withholding feedback cannot accelerate the rate of convergence for $\beta \in (1, 2]$.

Our result requires the usual assumption that $\dim F^* = N$ and further excludes payoff vectors where some player obtains her maximum feasible payoff.

Theorem 2 *Assume that* $\dim F^* = N$*, and fix any* $v \in \text{bnd}(F^*)$ *such that* $v_i < \max_a u_i(a)$ *for all* i*. If* $\text{bnd}(F^*)$ *has* min-*curvature of order* $\beta \geq 1$ *at* v *and Assumption 1 holds, then there exist* $c > 0$ *and* $\bar{\delta} < 1$ *such that* $d\left(v, E^P(\delta)\right) \leq c(1-\delta)^{\min\{\beta,2\}/2}$ *for all* $\delta > \bar{\delta}$*.*

Theorem 2 builds on FLM, HT, and SW. The idea is that, under Assumption 1, continuation payoff movements of magnitude $O(1-\delta)$ suffice for orthogonal enforcement, so the

[16] This condition is the same as Assumption 1 of Sannikov (2007).

logic given following Theorem 1 implies that minimum inefficiency is $O\left((1-\delta)^{\beta/2}\right)$ when $\beta \in [1,2]$, and is at most $O(1-\delta)$ when $\beta > 2$.

In light of Theorem 1, when $\beta > 2$ one might hope to find conditions under which $d\left(v, E^P(\delta)\right) = O\left((1-\delta)^{\beta-1}\right)$, or at least $d\left(v, E^P(\delta)\right) = O\left((1-\delta)^{\rho}\right)$ for some $\rho > 1$. While this may be possible, we do not pursue such a result here. The difficulty is that, as explained following Theorem 1, the ball B that is closest to an exposed point v and satisfies the self-generation condition $B \subseteq F$ in a neighborhood of v has radius $O(1-\delta)^{1-\beta/2}$, and thus expands as $\delta \to 1$ when $\beta > 2$. As $\delta \to 1$, this ball eventually violates self-generation at some point far from v. Thus, conditions under which $d\left(v, E^P(\delta)\right)$ can be less than $O(1-\delta)$ must involve the global geometry of the feasible payoff set. Investigating such conditions is left for future work.

5 Discussion

5.1 Players with Observable Actions

Our non-moving support assumption excludes perfect monitoring, but our results easily extend to the case where some players' actions are perfectly observed. Let $I^* \subseteq I$ be the set of players with observable actions, and assume that deviations by players $i \in I \backslash I^*$ do not affect the support of p, so that $\operatorname{supp} p(\cdot|a) = \operatorname{supp} p(\cdot|a_i', a_{-i})$ for all $a \in A$, $i \in I \backslash I^*$, and $a_i' \in A_i$. Theorem 1 then continues to apply for any $v \in \exp(F)$ that cannot be attained by an action profile distribution α such that $g_i(s_i, \alpha) = 0$ for each player $i \in I \backslash I^*$ and each manipulation s_i. Moreover, Theorem 2 does not rely on non-moving support and thus holds verbatim when some players' actions are observed.[17]

[17] In fact, it suffices to impose Assumption 1(i) for $j \in I \backslash I^*$ and Assumption 1(ii) for $(i,j) \in (I \backslash I^*)^2$. This is because (3) holds automatically for $j \in I^*$; and Assumption 1(ii) is used to deter deviations from a pure strategy profile without destroying surplus, and observable deviations can always be so deterred when δ is sufficiently high.

5.2 Intrinsic Discounting and Frequent Actions

The current paper focuses on the rate at which inefficiency vanishes as $\delta \to 1$, for a fixed monitoring structure. In SW, we showed that if one considers the double limit where $\delta \to 1$ at the same time as monitoring precision degrades, whether a folk or anti-folk theorem holds depends on a ratio of discounting and monitoring precision, which can be viewed as a measure of discounting relative to an intrinsic timescale. This double limit arises, for example, in the "frequent action limit" considered by Abreu, Milgrom and Pearce (1991), Fudenberg and Levine (2007), and Sannikov and Skrzypacz (2010), where signals are parameterized by an underlying continuous-time process, actions and signal observations occur simultaneously every Δ units of time, and the analysis concerns the $\Delta \to 0$ limit.

The results in the current paper can be extended to the low-discounting/low-monitoring double limit. We sketch one such extension. Define the *maximum detectability* of any manipulation by any player as

$$\chi^2 = \max_{a \in A, i \in I, s_i : A \to A} \sum_y p(y|a) \left(\frac{p(y|a) - p(y|s_i(a_i), a_{-i})}{p(y|a)} \right)^2 .$$

The notation is explained by noting that this is the χ^2-divergence of the signal distribution under the manipulation from that under equilibrium play. Fix a finite stage game (so we are in the kinked case where $\beta = 1$), and fix an exposed point $v \in \exp(F) \backslash V^{CE}$ and a direction $\lambda \in \Lambda_v$. It can be shown that there exists $c > 0$ such that

$$\lambda \cdot (v - w) \geq c \left(\frac{\chi^2}{1-\delta} \max \left\{ \log \frac{\chi^2}{1-\delta}, 1 \right\} \right)^{1/2}$$

for all $\delta \in [0, 1)$ and all monitoring structures with maximum detectability χ^2.[18] Thus, when $\beta = 1$, Theorem 1 extends to the low-discounting/low-monitoring double limit by simply replacing $1 - \delta$ with the ratio $(1 - \delta) / \chi^2$. This result can be proved by replacing the martingale large deviations bound used in the proof of Theorem 1 (Azuma's inequality) with a bound that also tracks the variance of the martingale increments (e.g., the inequality

[18] This result also requires that our full support assumption holds uniformly over monitoring structures: i.e., for some $\omega > 0$, attention is restricted to monitoring structures satisfying $p(y|a) > \omega$ for all y, a.

of Freedman, 1975). Similar extensions may also be possible when $\beta > 1$.

5.3 Summary and Directions for Future Research

This paper has taken a rate-of-convergence approach to studying the value of withholding feedback in standard repeated agency problems and games with patient players. The main result is that this value is "small": in finite-action settings where the feasible payoff set is kinked, withholding feedback accelerates convergence to efficiency by at most a log factor, while in smooth settings, withholding feedback improves efficiency by at most a constant factor. The key economic force underlying this result is that, while pooling information across many periods leads to more precise monitoring, it also entails larger rewards and punishments, which reduces the scope for providing incentives by transferring continuation surplus rather than destroying it.

A basic lesson is that the value of withholding feedback is very different in a one-off production process that unfolds gradually over time (as in Holmström and Milgrom, 1987) as compared to a genuinely repeated interaction. Since continuation payoff transfers are impossible in one-shot interactions, the monitoring benefit of withholding feedback dominates, so withholding feedback can be very valuable. But in repeated interactions, this benefit is offset by the cost of using larger rewards and punishments, which limit continuation payoff transfers.

We mention some possible extensions of our results. First, as discussed in Section 4, further analyzing rates of convergence toward exposed points with curvature of order $\beta > 2$ is a challenging open question, which involves non-local geometric properties of the feasible payoff set. Second, it would be interesting to allow an infinite set of signal realizations with unbounded likelihood ratios: for example, perhaps a player's action is observed with normal noise. Introducing infinite signals could increase the rate of convergence, because as $\delta \to 1$ it becomes possible to base incentives on rarer but more informative signal realizations (e.g., "tail tests"), so there is a sense in which increasing δ now endogenously increases monitoring precision. We conjecture that whether such an acceleration occurs depends on the tail behavior of the signal distribution. Third, the rate of convergence when discounting and monitoring vary simultaneously could be studied in detail. For example, it remains to analyze the rate of convergence in the frequent-action limit in the case where different actions of player 1 generate signals of player 2's action of very unequal precision.

We also believe that the rate of convergence to efficiency can be a useful lens for analyzing other questions about long-run economic relationships, besides the impact of feedback. This may be particularly true in settings with private monitoring, where analyzing equilibrium payoffs for a fixed discount factor is typically intractable.

Appendix: Omitted Proofs

A Proof of Theorem 1

We first bound the deviation gain at any $\alpha \in \Delta(A)$ that attains payoffs close to v.

Lemma 1 *There exist $\varepsilon > 0$ and $\gamma > 0$ such that, for all $\alpha \in \Delta(A)$ satisfying $\lambda \cdot (v - u(\alpha)) < \varepsilon$, there exist $i \in I$ and $s_i \in S_i$ such that $g_i(s_i, \alpha) > \gamma$.*

Proof. Since $v \in \exp(F) \setminus V^{CE}$, for all $\alpha \in \Delta(A)$ such that $v = u(\alpha)$, there exist i and s_i such that $g_i(s_i, \alpha) > 0$. Let

$$\gamma = \frac{1}{2} \inf_{\alpha \in \Delta(A): v = u(\alpha)} \sup_{i, s_i} g_i(s_i, \alpha).$$

Note that $\gamma > 0$. To see this, note that $g_i(\mathrm{Id}, \alpha) = 0$ for all i, α, so $\gamma \geq 0$, and suppose toward a contradiction that there exists a sequence α^n such that $v = u(\alpha^n)$ for all n and $\sup_{i, s_i} g_i(s_i, \alpha^n) \to 0$. Since $\Delta(A)$ is weak*-compact by Alaoglu's theorem, taking a subsequence if necessary, $\alpha^n \to \alpha \in \Delta(A)$. Moreover, since each u_i is continuous, $u(\alpha) = v$; and since each A_i is compact, by the maximum theorem, $\sup_{s_i} g_i(s_i, \alpha) = \lim_n \sup_{s_i} g_i(s_i, \alpha^n) = 0$ for all i, contradicting $v \notin V^{CE}$.

Now suppose that for all $\varepsilon > 0$ there exists $\alpha^\varepsilon \in \Delta(A)$ satisfying $\lambda \cdot (v - u(\alpha^\varepsilon)) < \varepsilon$ and $g_i(s_i, \alpha^\varepsilon) < \gamma$ for all i, s_i. Taking a subsequence if necessary, $\alpha^\varepsilon \to \alpha \in \Delta(A)$. Moreover, we have $u(\alpha) = \lim_\varepsilon u(\alpha^\varepsilon) = v$ (since $u(\alpha^\varepsilon) \in F$ and $v \in \exp(F)$), and $\sup_{s_i} g_i(s_i, \alpha) = \lim_\varepsilon \sup_{s_i} g_i(s_i, \alpha^\varepsilon) \leq \gamma$ for all i (by the maximum theorem), so $\sup_{i, s_i} g_i(s_i, \alpha) \leq \gamma$, contradicting the definition of γ. ■

Fix such ε and γ. We now fix a sufficiently small constant $c > 0$. First, for $\psi \geq 1$, let

$$k_1(\psi) = \left(8(\psi - 1)\max\left\{2^{\psi-3}, 1\right\}\right)^{\psi}, \tag{6}$$

and let $k_2(\psi) \geq 1$ satisfy

$$\sum_{t=1}^{\infty} \delta^t t^{\psi} \leq \frac{k_2(\psi)}{(1-\delta)^{\psi+1}} \qquad \text{for all } \delta. \tag{7}$$

(The existence of such $k_2(\psi)$ follows from the standard fact that $\sum_{t=1}^{\infty} \delta^t t^{\psi} = \Gamma(\psi+1)(1-\delta)^{-(\psi+1)} + O\left((1-\delta)^{-\psi}\right)$: see, e.g., Wood, 1992, eqn. (6.4).) Next, define

$$\begin{aligned}
\zeta(\delta) &= \begin{cases} \sqrt{\frac{1-\delta}{\max\{-\log(1-\delta),1\}}} & \text{if } \beta = 1, \\ (1-\delta)^{\beta/2} & \text{if } \beta \in (1,2], \\ (1-\delta)^{\beta-1} & \text{if } \beta \in (2,\infty), \end{cases} \\
T(\delta) &= \left\lceil \frac{\log 2}{-\log \delta} \right\rceil, \qquad \text{and} \\
f(c,\delta) &= \begin{cases} \frac{4\sqrt{\pi}\bar{u}}{\eta\omega^2} \frac{\sqrt{c}\zeta(\delta)\delta}{1-\delta} \exp\left(-\frac{\eta^2\omega^2}{2c\zeta(\delta)^2 T(\delta)}\right) & \text{if } \beta = 1, \\ \left(\frac{\sqrt{c}}{\eta\omega^{\beta}} \frac{\beta^{\beta-1}-1}{\beta^{\beta}}\right)^{\frac{1}{\beta-1}} k_1\left(\frac{\beta}{\beta-1}\right) k_2\left(\frac{\beta}{2(\beta-1)}\right) & \text{if } \beta > 1. \end{cases}
\end{aligned}$$

Note that $\zeta(\delta) \in [0,1]$, $\zeta(\delta)$ is decreasing, $\lim_{\delta \to 1} \zeta(\delta) = 0$, $T(\delta)$ is increasing and $f(c,\delta)$ is increasing in c for all δ. We fix $c = \min\{\varepsilon/2, c_0\}$, where c_0 satisfies the following lemma.

Lemma 2 *There exists $c_0 > 0$ such that, for all $c \leq c_0$, we have*

$$f(c,\delta) + \sqrt{c} \leq \frac{\gamma}{2} \qquad \textit{for all } \delta \in [0,1). \tag{8}$$

Proof. If $\beta > 1$ then $f(c,\delta)$ is independent of δ and satisfies $\lim_{c \to 0} f(c,\delta) = 0$, so this is immediate.

If $\beta = 1$, let $c_1 = \min\left\{\frac{\eta^2\omega^2}{4\log 3}, \frac{\gamma^2}{5}\right\}$. Since $\sqrt{c_1} < \gamma/2$ and $f(c,\delta)$ is increasing in c and satisfies $\lim_{\delta \to 0} f(c_1,\delta) = 0$, there exists $\delta_1 > 0$ such that (8) holds for all $c < c_1$ and all $\delta < \delta_1$. Moreover, there exists $\delta_2 > 0$ such that (8) holds for all $c < c_1$ and all $\delta > \delta_2$. To see this, for sufficiently high δ, we have

$$T(\delta) \leq \frac{\log 2}{-\log \delta} + 1 \leq \frac{\log 3}{-\log \delta} \leq \frac{\log 3}{1-\delta},$$

and hence

$$f(c,\delta) \leq \frac{4\sqrt{\pi}\bar{u}}{\eta\omega^2} \frac{\sqrt{c_1}\zeta(\delta)\delta}{1-\delta} \exp\left(\frac{\eta^2\omega^2 \log(1-\delta)}{2c_1 \log 3}\right) \leq \frac{4\sqrt{\pi}\bar{u}}{\eta\omega^2} \sqrt{c_1}\zeta(\delta)\delta.$$

Hence, since $\lim_{\delta\to 1} \zeta(\delta) = 0$, $\lim_{\delta\to 1} f(c,\delta) + \sqrt{c} = \sqrt{c} \leq \sqrt{c_1} < \gamma/2$, so δ_2 is well-defined. Finally, for all $c > 0$ and all $\delta \in [\delta_1, \delta_2]$, since $\zeta(\delta)$ is decreasing and $T(\delta)$ is increasing, we have

$$f(c,\delta) \leq \frac{4\sqrt{\pi}\bar{u}}{\eta\omega^2} \frac{\sqrt{c}\zeta(\delta_1)\delta_2}{1-\delta_2} \exp\left(-\frac{\eta^2\omega^2}{2c\zeta(\delta_1)^2 T(\delta_2)}\right).$$

Since the RHS converges to zero as $c \to 0$, there exists $c_2 > 0$ such that (8) holds for all $c < c_2$ and all $\delta \in [\delta_1, \delta_2]$. Taking $c_0 = \min\{c_1, c_2\}$ completes the proof. ■

Now fix $\delta \in [0,1)$ and fix a Nash equilibrium σ in $\Gamma^B(\delta)$ with equilibrium payoff w. Denote the induced repeated game outcome by $\mu \in \Delta((A \times Y)^\infty)$, the distribution over period-t action profiles by $\alpha_t \in \Delta(A)$, and the occupation measure over the first T periods by

$$\alpha^T = \frac{1-\delta}{1-\delta^T} \sum_{t=1}^{T} (1-\delta)\delta^{t-1}\alpha_t.$$

For any i and s_i, we also let

$$g_{i,t}(s_i) = g_i(s_i, \alpha_t) \qquad \text{and} \qquad g_i^T(s_i) = \frac{1-\delta}{1-\delta^T} \sum_{t=1}^{T} \delta^{t-1} g_{i,t} = g_i\left(s_i, \alpha^T\right),$$

and, for any complete history $h = (a_t, y_t)_{t=1}^\infty$, we let

$$u_{i,t}(h) = u_i(a_t) \qquad \text{and} \qquad \ell_{i,t}(s_i, h) = \frac{p(y_t|a_t) - p(y_t|s_i(a_{i,t}), a_{-i,t})}{p(y_t|a_t)}.$$

Finally, as c and δ are now fixed, we reduce notation by letting $\zeta = \zeta(\delta)$, $T = T(\delta)$, and $f = f(c,\delta)$. To complete the proof, we show that $\lambda \cdot (v - w) \geq c\zeta$.

We first consider the case where $\sup_{i,s_i} g_i^T(s_i) \leq \gamma$.

Lemma 3 *If* $\sup_{i,s_i} g_i^T(s_i) \leq \gamma$, *then* $\lambda \cdot (v - w) \geq c\zeta$.

Proof. Since $\delta^T \leq 1/2$ by construction, we have

$$\begin{aligned}
\lambda \cdot (v - w) &= (1-\delta)\sum_{t=1}^{\infty} \delta^{t-1} \lambda \cdot (v - u(\alpha_t)) \\
&\geq (1-\delta)\sum_{t=1}^{T} \delta^{t-1} \lambda \cdot (v - u(\alpha_t)) \\
&= \left(1-\delta^T\right) \lambda \cdot \left(v - u\left(\alpha^T\right)\right) \geq \frac{1}{2} \lambda \cdot \left(v - u\left(\alpha^T\right)\right).
\end{aligned}$$

By construction of (ε, γ), if $\lambda \cdot \left(v - u\left(\alpha^T\right)\right) < \varepsilon$ then $\sup_{i,s_i} g_i^T(s_i) > \gamma$. Hence, $\sup_{i,s_i} g_i^T(s_i) \leq \gamma$ implies $\lambda \cdot \left(v - u\left(\alpha^T\right)\right) \geq \varepsilon \geq 2c\zeta$, and therefore $\lambda \cdot (v - w) \geq c\zeta$. ■

The rest of the proof considers the case where $\sup_{i,s_i} g_i^T(s_i) > \gamma$. We fix i and s_i such that $g_i^T(s_i) \geq \gamma$, and reduce notation by letting $g_t = g_{i,t}(s_i)$, $g^T = g_i^T(s_i)$, and $\ell_t(h) = \ell_{i,t}(s_i, h)$. We first use player i's period-t incentive constraint to relate g_t, $u_{i,t}$, and ℓ_t.

Lemma 4 *For every* $t \in \mathbb{N}$, *we have* $g_t \leq \mathbb{E}^{\mu}\left[\ell_t(h) \sum_{t'=t+1}^{\infty} \delta^{t'-t} u_{i,t'}(h)\right]$.

Proof. For any sequence of action profiles $(a_t)_{t=1}^{\infty}$ and any period t, let $w_t = \sum_{t'=t}^{\infty} \delta^{t'-t} u_i(a_{t'})$. Since μ is an equilibrium outcome, for every $t \in \mathbb{N}$ we have

$$g_t \leq \int_{h^t, a_t, y_t} \left(p(y_t|a_t) - p(y_t|s_i(a_{i,t}), a_{-i,t})\right) \delta \mathbb{E}\left[w_{t+1}|h^t, a_t, y_t\right] d\mu\left(h^t, a_t\right).$$

This holds because, if she follows her recommendation in every period $t' \neq t$ while manipulating according to s_i in period t, player i obtains an expected continuation payoff of $\int_{h^t, a_t, y_t} p(y_t|s_i(a_{i,t}), a_{-i,t}) \mathbb{E}[w_{t+1}|h^t, a_t, y_t] d\mu(h^t, a_t)$ in period $t+1$, and this deviation must be unprofitable. The lemma follows as

$$\begin{aligned}
& \int_{h^t, a_t, y_t} \left(p(y_t|a_t) - p(y_t|s_i(a_{i,t}), a_{-i,t})\right) \delta \mathbb{E}\left[w_{t+1}|h^t, a_t, y_t\right] d\mu\left(h^t, a_t\right) \\
= & \int_{h^t, a_t, y_t} p(y_t|a_t) \ell_t(h) \delta \mathbb{E}\left[w_{t+1}|h^t, a_t, y_t\right] d\mu\left(h^t, a_t\right) \\
= & \int_h \ell_t(h) \sum_{t'=t+1}^{\infty} \delta^{t'-t} u_{i,t'}(h) d\mu(h),
\end{aligned}$$

where the last line follows by iterated expectation. ■

Since $\operatorname{bnd}(F)$ has max-curvature of order β at v, Lemma 4 implies that $\lambda \cdot (v - w)$ is no less than the value of program (2). By weak duality, letting $x_t(h) = u_{i,t}(h) - v_i$, the value of this program is no less than

$$\sup_{(\xi_t)_t \geq 0} \inf_{(x_t(h))_{t,h} \in [-\bar{u}, \bar{u}]} (1-\delta)\, \mathbb{E}^{\mu} \begin{bmatrix} \sum_{t=1}^{\infty} \delta^{t-1} \eta \left| x_t(h) \right|^{\beta} \\ - \sum_{t=1}^{\infty} \xi_t \ell_t(h) \sum_{t'=t+1}^{\infty} \delta^{t'-1} x_{t'}(h) \end{bmatrix} + (1-\delta) \sum_{t=1}^{\infty} \delta^{t-1} \xi_t g_t.$$

To bound the value from below, let

$$\xi_t = \begin{cases} \sqrt{c}\zeta & \text{if } t \leq T, \\ 0 & \text{if } t > T. \end{cases}$$

We then have

$$\sum_{t=1}^{\infty} \xi_t \ell_t(h) \sum_{t'=t+1}^{\infty} \delta^{t'-1} x_{t'}(h) = \sqrt{c}\zeta \sum_{t=1}^{T} \ell_t(h) \sum_{t'=t+1}^{\infty} \delta^{t'-1} x_{t'}(h) = \sqrt{c}\zeta \sum_{t=2}^{\infty} \delta^{t-1} \left(\sum_{t'=1}^{\min\{t-1,T\}} \ell_{t'}(h) \right) x_t(h),$$

as well as

$$(1-\delta) \sum_{t=1}^{\infty} \delta^{t-1} \xi_t g_t = \sqrt{c}\zeta (1-\delta) \sum_{t=1}^{T} \delta^{t-1} g_t = \sqrt{c}\zeta \left(1 - \delta^T\right) g^T \geq \frac{\sqrt{c}\zeta\gamma}{2}.$$

In total, we have

$$\begin{aligned} & \lambda \cdot (v - w) \\ \geq \; & \frac{\sqrt{c}\zeta\gamma}{2} + \inf_{(x_t(h))_{t \geq 2, h} \in [-\bar{u}, \bar{u}]} (1-\delta) \sum_{t=2}^{\infty} \delta^{t-1} \mathbb{E}^{\mu} \left[\eta \left| x_t(h) \right|^{\beta} - \sqrt{c}\zeta \left(\sum_{t'=1}^{\min\{t-1,T\}} \ell_{t'}(h) \right) x_t(h) \right]. \end{aligned} \tag{9}$$

It remains to bound the last term in (9).

Lemma 5 *For any $\mu \in \Delta\left((A \times Y)^{\infty}\right)$, we have*

$$\inf_{(x_t(h))_{t \geq 2, h} \in [-\bar{u}, \bar{u}]} (1-\delta) \sum_{t=2}^{\infty} \delta^{t-1} \mathbb{E}^{\mu} \left[\eta \left| x_t(h) \right|^{\beta} - \sqrt{c}\zeta \left(\sum_{t'=1}^{\min\{t-1,T\}} \ell_{t'}(h) \right) x_t(h) \right] \geq -\sqrt{c}\zeta f. \tag{10}$$

This will complete the proof, as (9), together with $\frac{\gamma}{2} - f \geq \sqrt{c}$ by (8), implies that

$$\lambda \cdot (v - w) \geq \sqrt{c}\zeta \left(\frac{\gamma}{2} - f\right) \geq \sqrt{c}\zeta\sqrt{c} = c\zeta.$$

We now prove Lemma 5. We consider separately the cases where $\beta = 1$ and $\beta > 1$.

Case 1: $\beta = 1$**.** When $\beta = 1$, the LHS of (10) is linear in $x_t(h) \in [-\bar{u}, \bar{u}]$, so its value is no less than

$$(1-\delta)\sum_{t=2}^{\infty}\delta^{t-1}\int_{h:\sqrt{c}\zeta\left|\Sigma_{t'=1}^{\min\{t-1,T\}}\ell_{t'}(h)\right|\geq\eta}\left(\eta - \sqrt{c}\zeta\left|\sum_{t'=1}^{\min\{t-1,T\}}\ell_{t'}(h)\right|\right)\bar{u}d\mu(h).$$

We prove that this value is no less than $-\sqrt{c}\zeta f$.

Note that

$$\begin{aligned}
\mathbb{E}\left[\ell_t | (a_{t'}, y_{t'})_{t'=1}^{t-1}, a_t\right] &= \sum_{y_t} p(y_t|a_t)\frac{p(y_t|a_t) - p(y_t|s_i(a_{i,t}), a_{-i,t})}{p(y_t|a_t)} = 0 \qquad \text{and} \\
|\ell_t(h)| &\leq \frac{1}{\omega} \qquad \text{for all } t, h.
\end{aligned}$$

Applying the Azuma-Hoeffding inequality to the martingale increments $(\ell_t)_t$, for any $x \geq 0$, we have

$$\Pr\left(\sqrt{c}\zeta\left|\sum_{t'=1}^{\min\{t-1,T\}}\ell_{t'}(h)\right| \geq x\right) \leq 2\left(\exp\left(-\frac{x^2\omega^2}{2c\zeta^2\min\{t-1,T\}}\right)\right). \tag{11}$$

We thus have

$$\begin{aligned}
&\int_{h:\sqrt{c}\zeta\left|\sum_{t'=1}^{\min\{t-1,T\}}\ell_{t'}(h)\right|\geq\eta}\left(\eta - \sqrt{c}\zeta\left|\sum_{t'=1}^{\min\{t-1,T\}}\ell_{t'}(h)\right|\right)d\mu(h) \\
= \;& \Pr\left(\sqrt{c}\zeta\left|\sum_{t'=1}^{\min\{t-1,T\}}\ell_{t'}(h)\right| \geq \eta\right)\eta - \mathbb{E}\left[\mathbf{1}\left\{\sqrt{c}\zeta\left|\sum_{t'=1}^{\min\{t-1,T\}}\ell_{t'}(h)\right| \geq \eta\right\}\left|\sqrt{c}\zeta\sum_{t'=1}^{\min\{t-1,T\}}\ell_{t'}(h)\right|\right] \\
= \;& -\int_{x\geq\eta}\Pr\left(\sqrt{c}\zeta\left|\sum_{t'=1}^{\min\{t-1,T\}}\ell_{t'}(h)\right| \geq x\right)dx \geq -2\int_{x\geq\eta}\exp\left(-\frac{x^2\omega^2}{2c\zeta^2\min\{t-1,T\}}\right)dx,
\end{aligned}$$

where the second equality is by integration by parts. Now note that

$$\begin{aligned}\int_{x\geq\eta}\exp\left(-\frac{x^2\omega^2}{2c\zeta^2\min\{t-1,T\}}\right)dx &= \frac{\sqrt{2c}\zeta\sqrt{\min\{t-1,T\}}}{\omega}\int_{y\geq\frac{\eta\omega}{\sqrt{2c}\zeta\sqrt{\min\{t-1,T\}}}}\exp\left(-y^2\right)dy \\ &\leq \frac{2\sqrt{\pi}c\zeta^2}{\eta\omega^2}\min\{t-1,T\}\exp\left(-\frac{\eta^2\omega^2}{2c\zeta^2\min\{t-1,T\}}\right) \\ &\leq \frac{2\sqrt{\pi}c\zeta^2}{\eta\omega^2}(t-1)\exp\left(-\frac{\eta^2\omega^2}{2c\zeta^2 T}\right),\end{aligned}$$

where the first inequality uses the Mills ratio inequality $\phi(-x)/\Phi(-x)\geq x$ for $x\geq 0$. Hence, we have

$$\begin{aligned}&(1-\delta)\sum_{t=2}^{\infty}\delta^{t-1}\int_{h:\sqrt{c}\zeta\left|\Sigma_{t'=1}^{\min\{t-1,T\}}\ell_{t'}(h)\right|\geq\eta}\left(\eta-\sqrt{c}\zeta\left|\sum_{t'=1}^{t-1}\ell_{t'}(h)\right|\right)\bar{u}d\mu(h) \\ \geq\ & -\frac{4\sqrt{\pi}\bar{u}}{\eta\omega^2}c\zeta^2(1-\delta)\sum_{t=2}^{\infty}\delta^{t-1}(t-1)\exp\left(-\frac{\eta^2\omega^2}{2c\zeta^2T}\right) \\ =\ & -\frac{4\sqrt{\pi}\bar{u}}{\eta\omega^2}\frac{c\zeta^2\delta}{1-\delta}\exp\left(-\frac{\eta^2\omega^2}{2c\zeta^2T}\right)=-\sqrt{c}\zeta f(c,\delta),\end{aligned}$$

where the first equality uses $(1-\delta)\sum_{t=2}^{\infty}\delta^{t-1}(t-1)=\delta(1-\delta)^{-1}$.

Case 2: $\beta>1$. When $\beta>1$, the LHS of (10) is convex in $x_t(h)$. Relaxing the constraint $x_t(h)\in[-\bar{u},\bar{u}]$ and minimizing over $x_t(h)\in\mathbb{R}$ gives

$$x_t(h)=\left(\frac{\sqrt{c}\zeta}{\eta\beta}\right)^{\frac{1}{\beta-1}}\operatorname{sign}\left(\sum_{t'=1}^{\min\{t-1,T\}}\ell_{t'}(h)\right)\left|\sum_{t'=1}^{\min\{t-1,T\}}\ell_{t'}(h)\right|^{\frac{1}{\beta-1}}\qquad\text{for all } t\geq 2.$$

Hence, the LHS of (10) is no less than

$$-\sqrt{c}\zeta\left(\frac{\sqrt{c}\zeta}{\eta}\frac{\beta^{\beta-1}-1}{\beta^{\beta}}\right)^{\frac{1}{\beta-1}}(1-\delta)\sum_{t=2}^{\infty}\delta^{t-1}\mathbb{E}\left[\left|\sum_{s=1}^{t-1}\ell_s(h)\right|^{\frac{\beta}{\beta-1}}\right].$$

In turn, this expression is no less than $-\sqrt{c}\zeta f$ iff

$$(1-\delta)\sum_{t=2}^{\infty}\delta^{t-1}\mathbb{E}\left[\left|\sum_{s=1}^{t-1}\ell_s(h)\right|^{\frac{\beta}{\beta-1}}\right]\leq\frac{k_1\left(\frac{\beta}{\beta-1}\right)k_2\left(\frac{\beta}{2(\beta-1)}\right)}{\omega^{\frac{\beta}{\beta-1}}\zeta^{\frac{1}{\beta-1}}}.\tag{12}$$

It thus suffices to establish (12). We consider separately the cases where $\beta \in (1,2]$ and $\beta > 2$.

Case 2(a): $\beta \in (1,2]$**.** When $\beta \in (1,2]$, recalling the definition of $k_1(\psi)$, (6), by Dharmadhikari, Fabian, and Jogdeo (1968, eqn. (1.1)),

$$\mathbb{E}^{\mu}\left[\left|\sum_{s=1}^{t-1}\ell_s(h)\right|^{\frac{\beta}{\beta-1}}\right] \leq \frac{k_1\left(\frac{\beta}{\beta-1}\right)(t-1)^{\frac{\beta}{2(\beta-1)}}}{\omega^{\frac{\beta}{\beta-1}}}.$$

Hence, by (7),

$$\begin{aligned}(1-\delta)\sum_{t=2}^{\infty}\delta^{t-1}\mathbb{E}^{\mu}\left[\left|\sum_{s=1}^{t-1}\ell_s(h)\right|^{\frac{\beta}{\beta-1}}\right] &\leq \frac{k_1\left(\frac{\beta}{\beta-1}\right)}{\omega^{\frac{\beta}{\beta-1}}}(1-\delta)\sum_{t=2}^{\infty}\delta^{t-1}(t-1)^{\frac{\beta}{2(\beta-1)}} \\ &\leq \frac{k_1\left(\frac{\beta}{\beta-1}\right)k_2\left(\frac{\beta}{2(\beta-1)}\right)}{\omega^{\frac{\beta}{\beta-1}}(1-\delta)^{\frac{\beta}{2(\beta-1)}}} = \frac{k_1\left(\frac{\beta}{\beta-1}\right)k_2\left(\frac{\beta}{2(\beta-1)}\right)}{\omega^{\frac{\beta}{\beta-1}}\zeta^{\frac{1}{\beta-1}}}.\end{aligned}$$

Case 2(b): $\beta > 2$**.** When $\beta > 2$, by Pinelis (2015, eqn. (1.11)),

$$\mathbb{E}^{\mu}\left[\left|\sum_{s=1}^{t-1}\ell_s(h)\right|^{\frac{\beta}{\beta-1}}\right] \leq \frac{2(t-1)}{\omega^{\frac{\beta}{\beta-1}}}.$$

In addition, $k_1(\beta/(\beta-1)) \geq 2$ and $k_2(\beta/(\beta-1)) \geq 1$. Hence,

$$\begin{aligned}(1-\delta)\sum_{t=2}^{\infty}\delta^{t-1}\mathbb{E}^{\mu}\left[\left|\sum_{s=1}^{t-1}\ell_s(h)\right|^{\frac{\beta}{\beta-1}}\right] &\leq \frac{2}{\omega^{\frac{\beta}{\beta-1}}}(1-\delta)\sum_{t=2}^{\infty}\delta^{t-1}(t-1) \\ &= \frac{2}{\omega^{\frac{\beta}{\beta-1}}}\frac{\delta}{1-\delta} \leq \frac{k_1\left(\frac{\beta}{\beta-1}\right)k_2\left(\frac{\beta}{2(\beta-1)}\right)}{\omega^{\frac{\beta}{\beta-1}}\zeta^{\frac{1}{\beta-1}}}.\end{aligned}$$

B Proof of Proposition 1

Consider the strategies described in the text, with $z = \sqrt{-\log(1-\delta)}$ (i.e., $c_0 = 1$). Define

$$v = 2 - \frac{1-\delta}{1-\delta^T}\frac{p}{p_1 - p} \qquad \text{and} \qquad q = \left(\frac{2\delta^T}{1-\delta}(p_1 - p) - \frac{\delta^T}{1-\delta^T}p\right)^{-1}. \tag{13}$$

With v and q so defined, (PK) and (IC_1) hold with equality. We show that

$$\begin{aligned}
\lim_{\delta\to 1} \frac{\frac{1-\delta}{1-\delta^T}\frac{p}{p_1-p}}{\sqrt{\frac{1-\delta}{-\log(1-\delta)}}} &< \frac{5\sqrt{\rho}e^\rho}{e^\rho-1} \text{ for all } \rho>0, && (14)\\
\lim_{\rho\to 0}\lim_{\delta\to 1} \frac{\delta^T}{1-\delta}\left(p_1-p\right) &> 1, \qquad \text{and} && (15)\\
\lim_{\rho\to 0}\lim_{\delta\to 1} \frac{\delta^T}{1-\delta^T} q\left(p_T-p\right) &> 1. && (16)
\end{aligned}$$

Given these results, the proof is completed by first taking $\rho>0$ and $\bar{\delta}_1>0$ such that the inequalities in (15) and (16) hold for ρ and all $\delta>\bar{\delta}_1$, then taking $\bar{\delta}_2>0$ such that the inequality in (14) holds for ρ for all $\delta>\bar{\delta}_2$, and finally taking $c=5\sqrt{\rho}e^\rho/\left(e^\rho-1\right)$ and $\bar{\delta}=\max\left\{\bar{\delta}_1,\bar{\delta}_2\right\}$.

We now establish (14)–(16). Let $k\in\mathbb{N}$ be the unique integer satisfying $k\in[\frac{\sqrt{T}}{2}\left(\sqrt{T}-z\right)-1, \frac{\sqrt{T}}{2}\left(\sqrt{T}-z\right))$. Note that

$$\begin{aligned}
p &= \Pr\left(\sum_{t=2}^{T} y_t<k\right)+\frac{1}{2}\Pr\left(\sum_{t=2}^{T} y_t=k\right)<\Pr\left(\sum_{t=2}^{T} y_t\le k\right), \qquad \text{and}\\
p_1-p &= \frac{1}{4}\Pr\left(\sum_{t=2}^{T} y_t=k\right)=\frac{(T-1)!}{k!\,(T-1-k)!}\left(\frac{1}{2}\right)^{T+1}\ge\frac{(T)!}{k!\,(T-k)!}\left(\frac{1}{2}\right)^{T+2}, \quad (17)
\end{aligned}$$

where the last inequality holds because $k\le T/2$.

We first establish (14). Recall that the y_t are independent Bernoulli random variables. As shown by Zhu, Li, and Hayashi (2022, Theorem 2.1),

$$\frac{\Pr\left(\sum_{t=2}^{T} y_t\le k\right)}{\Pr\left(\sum_{t=2}^{T} y_t=k\right)}\le k+1-\frac{T}{2}+\sqrt{\left(k-1-\frac{T}{2}\right)^2+2k}.$$

Since

$$\frac{p}{p_1-p}<4\frac{\Pr\left(\sum_{t=2}^{T} y_t\le k\right)}{\Pr\left(\sum_{t=2}^{T} y_t=k\right)} \qquad \text{and} \qquad \lim_{\delta\to 1}\frac{\frac{1-\delta}{1-\delta^T}\left(k+1-\frac{T}{2}+\sqrt{\left(k-1-\frac{T}{2}\right)^2+2k}\right)}{\sqrt{\frac{1-\delta}{-\log(1-\delta)}}}=\frac{\sqrt{\rho}e^\rho}{e^\rho-1},$$

where the second line follows by l'Hopital's rule, we have

$$\lim_{\delta\to 1}\frac{\frac{1-\delta}{1-\delta^T}\frac{p}{p_1-p}}{\sqrt{\frac{1-\delta}{-\log(1-\delta)}}}\le \lim_{\delta\to 1}\frac{\frac{1-\delta}{1-\delta^T}4\frac{\Pr\left(\sum_{t=2}^T y_t\le k\right)}{\Pr\left(\sum_{t=2}^T y_t= k\right)}}{\sqrt{\frac{1-\delta}{-\log(1-\delta)}}}=\frac{4\sqrt{\rho}e^\rho}{e^\rho-1}<\frac{5\sqrt{\rho}e^\rho}{e^\rho-1},$$

which establishes (14).

We next establish (15). Applying Stirling's formula to (17), we have

$$p_1-p\ge\frac{\sqrt{2\pi(T-1)}}{4e^2\sqrt{k(T-1-k)}}\left(\frac{T-1}{2k}\right)^k\left(\frac{T-1}{2(T-1-k)}\right)^{T-1-k}. \tag{18}$$

Therefore,

$$\lim_{\rho\to 0}\lim_{\delta\to 1}\frac{\delta^T}{1-\delta}(p_1-p)\ge \lim_{\rho\to 0}\lim_{\delta\to 1}\frac{\delta^T}{1-\delta}\frac{\sqrt{2\pi(T-1)}}{4e^2\sqrt{k(T-1-k)}}\left(\frac{T-1}{2k}\right)^k\left(\frac{T-1}{2(T-1-k)}\right)^{T-1-k}=\infty,$$

which establishes (15).

Finally, we establish (16). We will show that $\lim_{\delta\to 1}p=0$ and $\lim_{\delta\to 1}p_T=1$. Hence, for sufficiently large δ, $p_T-p\ge 1/2$. This implies (16), as we have

$$\begin{aligned}
& \lim_{\rho\to 0}\lim_{\delta\to 1}\frac{\delta^T}{1-\delta^T}q(p_T-p)\\
= {} & \lim_{\rho\to 0}\lim_{\delta\to 1}\frac{\delta^T}{1-\delta^T}\frac{1}{\delta^T}\frac{p_T-p}{2\frac{p_1-p}{1-\delta}-\frac{p}{1-\delta^T}} && \text{by (13)}\\
\ge {} & \lim_{\rho\to 0}\lim_{\delta\to 1}\frac{1}{1-\delta^T}\frac{\frac{1}{2}}{2\frac{p_1-p}{1-\delta}}\\
\ge {} & \lim_{\rho\to 0}\lim_{\delta\to 1}\frac{1-\delta}{1-\delta^T}\left(\frac{\sqrt{2\pi(T-1)}}{e^2\sqrt{k(T-1-k)}}\left(\frac{\frac{1}{2}(T-1)}{k}\right)^k\left(\frac{\frac{1}{2}(T-1)}{T-1-k}\right)^{T-1-k}\right)^{-1} && \text{by (18)}\\
= {} & \infty.
\end{aligned}$$

It remains to show that $\lim_{\delta\to 1}p=0$ and $\lim_{\delta\to 1}p_T=1$. Recall that $2y_t-1$ has zero mean and unit variance when player 1 takes C. Thus, by the Berry–Esseen theorem, there exists an absolute constant C_0 such that

$$
\begin{aligned}
p &= \Pr^{\text{player 1 takes } C}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(2y_t-1) \leq -\sqrt{-\log(1-\delta)}\right) \\
&\leq \Phi\left(-\sqrt{-\log(1-\delta)}\right) + C_0\left(\frac{\mathbb{E}^{\text{player 1 takes } C}\left[|2y_t-1|^3\right]}{\sqrt{T}}\right) \xrightarrow{\delta\to 1} 0.
\end{aligned}
$$

On the other hand, $(4y_t-1)/\sqrt{3}$ has zero mean and unit variance when player 1 takes D. Thus, again by Berry–Esseen,

$$
\begin{aligned}
p_T &= \Pr^{\text{player 1 takes } D}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(2y_t-1) \leq -\sqrt{-\log(1-\delta)}\right) \\
&= \Pr^{\text{player 1 takes } D}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\frac{4y_t-1}{\sqrt{3}} \leq \frac{\sqrt{T}-2\sqrt{-\log(1-\delta)}}{\sqrt{3}}\right) \\
&\geq \Phi\left(\frac{\sqrt{T}-2\sqrt{-\log(1-\delta)}}{\sqrt{3}}\right) + C_0\left(\frac{\mathbb{E}^{\text{player 1 takes } D}\left[\left|(4y_t-1)/\sqrt{3}\right|^3\right]}{\sqrt{T}}\right) \xrightarrow{\delta\to 1} 1,
\end{aligned}
$$

completing the proof.

C Proof of Theorem 2

Throughout the proof, we let $\beta^* = \min\{\beta, 2\}$.

The following definition and lemma are due to APS.

Definition 3 *A bounded set $W \subseteq \mathbb{R}^N$ is* self-generating *if for all $\hat{v} \in W$, there exist $\alpha \in \times_i \Delta(A_i)$ and $w : Y \to \mathbb{R}^N$ satisfying*

1. Promise keeping (PK): $\hat{v} = (1-\delta)u(\alpha) + \delta\sum_y p(y|\alpha)w(y)$.
2. Incentive compatibility (IC): $\text{supp}(\alpha_i) \subseteq \text{argmax}_{a_i}(1-\delta)u_i(a_i,\alpha_{-i}) + \delta\sum_y p(y|a_i,\alpha_{-i})w_i(y)$ *for all i.*
3. Self-generation (SG): $w(y) \in W$ *for all y.*

When (PK), (IC), and (SG) hold, we say that the pair (α, w) decomposes $\hat{v}$ on W.

Lemma 6 *Any bounded, self-generating set W is contained in $E^P(\delta)$.*

It thus suffices to find a bounded, self-generating set W such that $d(v, W) = O\left((1-\delta)^{\beta^*/2}\right)$. We first state a sufficient condition for a ball B to be self-generating.[19] We then prove that there exists B with $d(v, B) = O\left((1-\delta)^{\beta^*/2}\right)$ that satisfies this condition.

Let $\bar{u} = \max_{u,u' \in F} \|u - u'\| > 0$.[20]

Definition 4 *The* maximum score in direction $\lambda \in \Lambda$ with reward bound $\bar{x} > 0$ *is*

$$k(\lambda, \bar{x}) := \sup_{\alpha \in \times_i \Delta(A_i), x: Y \to \mathbb{R}^N} \lambda \cdot \left(u(\alpha) + \sum_y p(y|\alpha) x(y) \right)$$

subject to

1. Incentive compatibility (IC): $\operatorname{supp}(\alpha_i) \subseteq \operatorname{argmax}_{a_i} u_i(a_i, \alpha_{-i}) + \sum_y p(y|a_i, \alpha_{-i}) x_i(y)$ *for all i.*

2. Half-space decomposability with reward bound $\bar{x}$ (HS$\bar{x}$): $\lambda \cdot x(y) \leq 0$ *and* $\|x(y)\| \leq \bar{x}$ *for all y.*

Lemma 7 *For any $\bar{x} > \bar{u}$ and $\varepsilon > 0$, if a ball B of radius r satisfies*

$$k(\lambda, \bar{x}) \geq \max_{v' \in B} \lambda \cdot v' + \varepsilon \quad \text{for all } \lambda \in \Lambda, \quad \text{and} \tag{19}$$

$$\bar{x}^2 \leq \frac{\delta}{1-\delta} \frac{\varepsilon r}{16}, \tag{20}$$

then B is self-generating.

We finally show that, for high enough $\bar{x}$ and δ, there exists a ball B satisfying (19) and (20) as well as $d(v, B) = O\left((1-\delta)^{\beta^*/2}\right)$.

[19] This condition, given in Definition 4 and Lemma 7 below, is similar to Definition 2 and Lemma 6 of SW, but is simpler because the monitoring structure varies together with δ in SW while it is is fixed in the current paper, so less control over the relationship between δ and the reward bound $\bar{x}$ is required.

[20] This is a slight abuse of notation, as in the proof of Theorem 1 we took $\bar{u}$ to be the range of u_i for a particular player i.

Lemma 8 *There exist $\bar{x} > \bar{u}$, $c > 0$ and $\bar{\delta} < 1$ such that, for all $\delta > \bar{\delta}$, there exist $\varepsilon > 0$ and a ball B of radius r satisfying*

$$\begin{aligned} k(\lambda, \bar{x}) &\geq \max_{v' \in B} \lambda' \cdot v' + \varepsilon \quad \textit{for all } \lambda' \in \Lambda, \\ \bar{x}^2 &\leq \frac{\delta}{1-\delta} \frac{\varepsilon r}{16}, \quad \textit{and} \\ d(v, B) &\leq c(1-\delta)^{\beta^*/2}. \end{aligned}$$

Taking $\bar{x}$, c, and $\bar{\delta}$ as in Lemma 8 completes the proof of Theorem 2.

C.1 Proof of Lemma 7

The proof is similar to (but simpler than) the proof of Lemma 6 of SW. To show that B is self-generating, it suffices to show that the extreme points of any ball $B' \subseteq B$ of radius $r/2$ are decomposable on B'.

Lemma 9 (SW, Lemma 10) *Suppose that for any ball $B' \subseteq B$ with radius $r/2$ and any direction $\lambda \in \Lambda$, the point $\hat{v} = \operatorname{argmax}_{v' \in B'} \lambda \cdot v'$ is decomposable on B'. Then B is self-generating.*

We thus fix a ball $B' \subseteq B$ of radius $r/2$ and a direction $\lambda \in \Lambda$, and let $\hat{v} = \operatorname{argmax}_{v' \in B'} \lambda \cdot v'$. We construct (α, w) that decompose $\hat{v}$ on B'.

Since $k(\lambda, \bar{x}) \geq \max_{v' \in B} \lambda \cdot v' + \varepsilon$ by hypothesis, there exist α and $x : Y \to \mathbb{R}^N$ satisfying (IC), (HS$\bar{x}$), and

$$\lambda \cdot \left(u(\alpha) + \sum_y p(y|\alpha) x(y) \right) \geq \max_{v' \in B} \lambda \cdot v' + \varepsilon/2 \geq \max_{v' \in B'} \lambda \cdot v' + \varepsilon/2. \tag{21}$$

To construct w, for each y, let

$$w(y) = \hat{v} + \frac{1-\delta}{\delta} \left(\hat{v} - u(\alpha) + x(y) - \sum_{y'} p(y'|\alpha) x(y') \right).$$

We show that (α, w) decomposes $\hat{v}$ on B' by verifying (PK), (IC), and (SG).

(PK): This holds by construction: we have $\sum_y p(y|\alpha) w(y) = (1/\delta)(\hat{v} - (1-\delta) u(\alpha))$, and hence $(1-\delta) u(\alpha) + \delta \sum_y p(y|\alpha) w(y) = \hat{v}$.

(IC): Setting aside the constant terms in $w(y)$, we see that an action a_i maximizes $(1-\delta) u_i(a_i, \alpha_{-i}) + \delta \sum_y p(y|a_i, \alpha_{-i}) w_i(y)$ iff it maximizes $u_i(a_i, \alpha_{-i}) + \sum_y p(y|a_i, \alpha_{-i}) x_i(y)$, which follows from (IC).

(SG): We start with a simple geometric observation.

Lemma 10 (SW, Lemma 11) *For each $w \in \mathbb{R}^N$, we have $w \in B'$ if $\lambda \cdot (\hat{v} - w) \geq 0$ and*

$$d(\hat{v}, w) \leq \sqrt{(r/2)\, \lambda \cdot (\hat{v} - w)}. \tag{22}$$

We thus show that, for each y, $w(y)$ satisfies $\lambda \cdot (\hat{v} - w(y)) \geq 0$ and (22). Note that

$$\hat{v} - w(y) = \frac{1-\delta}{\delta} \left(u(\alpha) + \sum_{y'} p(y'|\alpha) x(y') - \hat{v} - x(y) \right).$$

By (HS$\bar{x}$) and (21), we have $\lambda \cdot (\hat{v} - w(y)) \geq (\delta/(1-\delta))\, \varepsilon/2$, and therefore

$$\sqrt{(r/2)\, \lambda \cdot (\hat{v} - w(y))} \geq \frac{1-\delta}{\delta} \sqrt{\frac{\delta}{1-\delta} \frac{\varepsilon r}{4}}. \tag{23}$$

Similarly, we have

$$d(\hat{v}, w) \leq \frac{1-\delta}{\delta} \left(d(\hat{v}, u(\alpha)) + d\left(\sum_{y'} p(y'|\alpha) x(y'), x(y) \right) \right) \leq \frac{1-\delta}{\delta} (\bar{u} + \bar{x}) \leq \frac{1-\delta}{\delta} 2\bar{x}. \tag{24}$$

Comparing (23) and (24), we see that $w(y)$ satisfies (22) whenever $2\bar{x} \leq \sqrt{(\delta/(1-\delta))\, \varepsilon r/4}$, which holds by (20).

C.2 Proof of Lemma 8

Fix any $\bar{\eta} > 0$ and $\bar{F} \subseteq F^*$ such that $\{w \in F^* : d(v, w) \leq \bar{\eta}\} \subseteq \bar{F}$, $\dim \bar{F} = N$, and $w_i < \max_a u_i(a)$ for all i and $w \in \bar{F}$.

Lemma 11 *There exists $\bar{x} > \bar{u}$ such that $k(\lambda, \bar{x}) \geq \max_{v' \in \bar{F}} \lambda \cdot v'$.*

Proof. This is the same as Lemma 5 of HT. The only difference is that we allow $|A| = \infty$ and impose Assumption 1, while HT instead assume $|A| < \infty$ together with individual and pairwise full rank, which imply Assumption 1. ■

Now fix any $k > 16\bar{x}^2$. By Lemma 11, it suffices to find $c > 0$ and $\bar{\delta} < 1$ such that, for all $\delta > \bar{\delta}$, there exist $\varepsilon > 0$ and a ball B with radius $r > 0$ such that

$$\begin{aligned} \max_{v' \in \bar{F}} \lambda' \cdot v' &\geq \max_{v' \in B} \lambda' \cdot v' + \varepsilon \quad \text{ for all } \lambda' \in \Lambda, & (25) \\ r\varepsilon &\geq k(1-\delta), \quad \text{ and } & (26) \\ d(v, B) &\leq c(1-\delta)^{\beta^*/2}. & (27) \end{aligned}$$

If $\beta^* = 1$ then, as in Lemma 3 of HT, it suffices to take any $o \in \operatorname{int}(\bar{F})$ and any $\ell > 0$ sufficiently large compared to k, let $r = (1-\delta)^{1/2}$, and let B be the ball of radius r with center $(1 - \ell r) v + \ell r o$.

For the rest of the proof, we assume that $\beta^* > 1$. We first derive a geometric condition for $w \in F^*$, similar to Lemma 10.

Lemma 12 *There exist $\lambda \in \Lambda_v^*$, $\rho > 0$, and $\kappa > 0$ such that, if $d(v, w) < \rho$ and $\kappa d(v, w)^{\beta^*} < \lambda \cdot (v - w)$, then $w \in \operatorname{int}(F^*)$.*

Proof. Since F^* is full-dimensional and has min-curvature of order at most β at v, there exist $\bar{\varepsilon} > 0$ and $k < \infty$ such that, for all $w \in \operatorname{bnd}(F^*)$ satisfying $d(v, w) < \bar{\varepsilon}$, we have $\lambda \cdot (v - w) < kd(v, w)^{\beta} \leq kd(v, w)^{\beta^*}$ for all $\lambda \in \Lambda_v^*$. Let $B_{\varepsilon'}(v) = \{w \in \mathbb{R}^N : d(v, w) = \varepsilon'\}$. Since F^* is full-dimensional, there exists $\lambda \in \Lambda_v^*$, $\varepsilon' > 0$, and $t > 0$ such that $C := B_{\varepsilon'}(v) - t\lambda \subseteq F^*$. Fix such λ, ε', and t, and let $\varepsilon = \min\{\bar{\varepsilon}, \varepsilon', t\}$.

Now fix any $\kappa > k$, $\rho < \min\{\varepsilon, (t/\kappa)^{1/\beta^*}\}$, and $d < \rho$, and let $G = \{w \in B_d(v) : \lambda \cdot (v - w) \geq \kappa d^{\beta^*}\}$. We wish to show that $G \subseteq F^*$ (and in particular $G \subseteq \operatorname{int}(F^*)$, since $G \cap \operatorname{bnd}(F^*) = \emptyset$).

To see this, let $W = B_d(v) \cap \operatorname{bnd}(F^*)$, $H = \{w : \lambda \cdot (v - w) = kd^{\beta^*}\}$, $H' = \{w : \lambda \cdot (v - w) = t\}$, and $D = C \cap H'$. Since $\kappa > k$ and $d < \rho < \min\{\varepsilon, (t/\kappa)^{1/\beta^*}\}$, G lies in between H and H'. In addition, the projection of G onto H is a subset of the projection of W onto H, and the projection of G onto H' is a subset of D. Hence, we have $G \subseteq \operatorname{co}(W \cup D)$. Finally, since $W \subseteq F^*$ and $D \subseteq C \subseteq F^*$, and F^* is convex, we have $\operatorname{co}(W \cup D) \subseteq F^*$, so $G \subseteq F^*$. ■

Lemma 13 *There exist $\bar{c} > 0$, $\eta > 0$, and $\bar{\delta} < 1$ such that, for all $\delta > \bar{\delta}$, there exists a ball $B \subseteq \bar{F}$ of radius $r = \eta (1-\delta)^{1-\beta^*/2}$ satisfying $d(v, B) = \bar{c}(1-\delta)^{\beta^*/2}$.*

Proof. Fix $\lambda \in \Lambda_v^*$, $\rho > 0$, and $\kappa > 0$ as in Lemma 12. Given $\bar{c}$ and η to be determined, let B be the ball with radius $r = \eta(1-\delta)^{1-\beta^*/2}$ and center $o = v - (r+d)\lambda$, where $d = \bar{c}(1-\delta)^{\beta^*/2}$, and take any $\hat{w} \in \partial B$. Let $x = \lambda \cdot (\hat{w} - o)$, so that $x\lambda$ is the projection of $\hat{w} - o$ on λ. Then,

$$\begin{aligned} \|v - \hat{w}\|^2 &= \|v - o - x\lambda\|^2 + \|\hat{w} - o - x\lambda\|^2 = (r + d - x)^2 + r^2 - x^2, \qquad \text{and} \\ \lambda \cdot (v - \hat{w}) &= r + d - x. \end{aligned}$$

Recall that, by construction, $\{w \in F^* : d(v, w) \leq \bar{\eta}\} \subseteq \bar{F}$. Since $d(v, w) \leq d(v, o) + d(o, w) \leq 2r + d$ for all $w \in B$, it suffices to show that $2r + d \leq \bar{\eta}$ and $B \subseteq F^*$. By Lemma 12, if $d(v, w) < \rho$ and $\kappa d(v, w)^{\beta^*} \leq \lambda \cdot (v - w)$ then $w \in F^*$. Since $x \in [-r, r]$, it suffices to find $\bar{c}$, η, and $\bar{\delta}$ such that, for all $\delta > \bar{\delta}$, we have

$$\begin{aligned} 2r + d &\leq \bar{\eta}, && (28) \\ ((r + d) - x)^2 + r^2 - x^2 &\leq \rho^2 \qquad \text{for all } x \in [-r, r], \qquad \text{and} && (29) \\ \max_{x \in [-r, r]} f(x, \delta, \beta^*) &\leq 0, && (30) \end{aligned}$$

where

$$f(x, \delta, \beta^*) := \kappa \left((r + d - x)^2 + r^2 - x^2\right)^{\beta^*/2} - (r + d - x).$$

We consider separately the cases where $\beta^* = 2$ and $\beta^* \in (1, 2)$. First consider $\beta^* = 2$. Let $\hat{\eta} > 0$ be such that (29) holds whenever $\max\{r, d\} \leq \hat{\eta}$, and let any $\bar{c} = 1$ and $\eta = \min\{\hat{\eta}, \bar{\eta}/4, \kappa/4\}$, so that $r = \eta$ and $d = 1 - \delta$. For sufficiently large δ, we have $2r + d \leq \bar{\eta}$ and $d \leq \hat{\eta}$, and hence (28) and (29) hold. In addition, since $f(x, \delta, 2)$ is linear in x when $\beta^* = 2$, (30) holds whenever $f(r, \delta, 2) \leq 0$ and $f(-r, \delta, 2) \leq 0$. In turn, these inequalities hold for sufficiently large δ, since $f(r, \delta, 2) = d(\kappa d - 1)$ and $\lim_{\delta \to 1} \kappa d - 1 < 0$, and $f(-r, \delta, \beta^*) = (2r + d)(\kappa(2r + d) - 1)$ and $\lim_{\delta \to 1} \kappa(2r + d) - 1 = 2\kappa\eta - 1 < 0$.

Next, consider $\beta^* \in (1,2)$. Let $\bar{c} = 4\kappa^{2/(2-\beta^*)}\beta^{*\beta^*/(2-\beta^*)}$ and $\eta = 1$, so that $r = (1-\delta)^{1-\beta^*/2}$ and $d = \bar{c}(1-\delta)^{\beta^*/2}$. Since $\max\{r,d\} \to 0$ as $\delta \to 1$, (28) and (29) hold for sufficiently large δ. In addition, $f(x,\delta,\beta^*)$ is concave in x and is maximized over $x \in [-r,r]$ at

$$x^* = \frac{2r^2 + 2dr + d^2 - (\kappa(r+d)\beta^*)^{\frac{2}{2-\beta^*}}}{2(r+d)}.$$

It thus suffices to show that $f(x^*,\delta,\beta^*) \le 0$ for sufficiently large δ. By algebra,

$$f(x^*,\delta,\beta^*) = -\frac{2r+d}{r+d}\frac{d}{2} + \left(\beta^{*\frac{\beta^*}{2-\beta^*}} - \frac{1}{2}\beta^{*\frac{2}{2-\beta^*}}\right)\kappa^{\frac{2}{2-\beta^*}}(r+d)^{\frac{\beta^*}{2-\beta^*}}.$$

Finally, since $r = (1-\delta)^{1-\beta^*/2} \ge \bar{c}(1-\delta)^{\beta^*/2} = d$ for sufficiently large δ, we have

$$\begin{aligned} f(x^*,\delta,\beta^*) &\le -\frac{d}{2} + 2\kappa^{\frac{2}{2-\beta^*}}\beta^{*\frac{\beta^*}{2-\beta^*}} r^{\frac{\beta^*}{2-\beta^*}} \\ &= -\frac{\bar{c}(1-\delta)^{\frac{\beta^*}{2}}}{2} + 2\kappa^{\frac{2}{2-\beta^*}}\beta^{*\frac{\beta^*}{2-\beta^*}}(1-\delta)^{\left(1-\frac{\beta^*}{2}\right)\frac{\beta^*}{2-\beta^*}} \\ &= (1-\delta)^{\frac{\beta^*}{2}}\left(-\frac{\bar{c}}{2} + 2\kappa^{\frac{2}{2-\beta^*}}\beta^{*\frac{\beta^*}{2-\beta^*}}\right) = 0. \end{aligned}$$

■

We now complete the proof of Lemma 8. Take $\bar{c}$, η, $\bar{\delta}$, B, and r as in Lemma 13. Let B' be the radial contraction of B by a factor of $1 - 2k(1-\delta)^{\beta^*/2}/(\eta r)$, and define $\varepsilon = 2k(1-\delta)^{\beta^*/2}/\eta$ and $c = \bar{c} + 2k/\eta$. Since $d(v,B) = \bar{c}(1-\delta)^{\beta^*/2}$, we have $d(v,B') = (\bar{c} + 2k/\eta)(1-\delta)^{\beta^*/2} = c(1-\delta)^{\beta^*/2}$, so (27) holds. Moreover, denoting the radius of B' by r', we have

$$r'\varepsilon = \left(1 - \frac{2k(1-\delta)^{\beta^*/2}}{\eta^2(1-\delta)^{1-\beta^*/2}}\right)\eta(1-\delta)^{1-\beta^*/2} \times \frac{2k(1-\delta)^{\beta^*/2}}{\eta} = \left(1 - \frac{2k(1-\delta)^{\beta^*-1}}{\eta^2}\right)2k(1-\delta).$$

For sufficiently large δ, this is greater than $k(1-\delta)$, so (26) holds. Finally, since $B \subseteq \bar{F}$, for all $\lambda' \in \Lambda$ we have $\max_{v' \in \bar{F}} \lambda' \cdot v' \ge \max_{v' \in B} \lambda' \cdot v' = \max_{v' \in B'} \lambda' \cdot v' + 2k(1-\delta)^{\beta^*/2}/\eta = \max_{v' \in B'} \lambda' \cdot v' + \varepsilon$, so (25) holds.

D Proof of Proposition 2

To define $\bar{x}$, we first observe that for each pair of players $i \neq j$ and each action profile a, we can take $\left(x_i^j(d;a)\right)_d$ such that (i) $\sum_d x_i^j(d;a)\, y^d$ has mean 0 and bounded Euclidean norm; (ii) rewards $\sum_d x_i^j(d;a)\, y^d$ induce player i to take a_i when her opponents take a_{-i}; and (iii) $\mathbb{E}\left[\sum_d x_i^j(d;a)\, y^d | a\right]$ is independent of player j's action.

Lemma 14 *There exists $\hat{x}$ such that, for each pair of players $i \neq j$ and action profile $a \in A$, there exist $\left(x_i^j(d;a)\right)_d$ such that $\mathbb{E}\left[\sum_d x_i^j(d;a)\, y^d | a\right] = 0$, $\frac{d}{da_i}\mathbb{E}\left[\sum_d x_i^j(d;a)\, y^d | a\right] = 1$, $\frac{d}{da_j}\mathbb{E}\left[\sum_d x_i^j(d;a)\, y^d | a\right] = 0$, and $\left|\sum_d x_i^j(d;a)\, y^d\right| \leq \hat{x}$ for all y.*

Proof. For each a and (i,j), let $f^{ij}(a)$ be the value of the program

$$\inf_{\beta \in \mathbb{R}^D} |\beta| \qquad \text{subject to}$$

$$\begin{aligned}
\sum_d \beta_d \frac{d}{da_i}\mu(a_i, a_{-i}) &= 1, \text{ or equivalently } \beta M_i(a) = 1, \\
\sum_d \beta_d \frac{d}{da_j}\mu(a_i, a_{-i}) &= 0, \text{ or equivalently } \beta M_j(a) = 0.
\end{aligned}$$

(Here β is a row vector while $M_i(a)$ and $M_j(a)$ are column vectors.)

Since $A \ni a$ is compact and N is finite, it suffices to prove that, for each (i,j), (i) $f^{ij}(a) < \infty$ for all a, and (ii) $f^{ij}(a)$ is upper-semicontinuous.

We first prove (i). As in Lemma 1 of Sannikov (2007), Assumption 1 implies that the columns of $[M^i(a); M^j(a)]$ are linearly independent, so there exists $L(a)$ such that $[M^i(a); M^j(a); L(a)]$ is a D-dimensional invertible matrix. For

$$Q(a) = [M_i(a); 0; 0]\left[M^i(a); M^j(a); L(a)\right]^{-1},$$

we have $Q(a) M^i(a) = M^i(a)$ and $Q(a) M^j(a) = 0$. Moreover, since $M^i(a)$ is non-degenerate, there exists $\bar{\beta}$ such that $\bar{\beta} M^i(a) = 1$. Since $\beta = \bar{\beta} Q(a)$ satisfies the constraints, we have $f^{ij}(a) < \infty$.

We next prove (ii). Fix any a and η_0. There exists β such that $|\beta| \leq f^{ij}(a) + \frac{\eta_0}{2}$. Take $L(a)$ as in the proof of (i). Taking η_1 sufficiently small, we can guarantee that

$[M^i(a'); M^j(a'); L(a)]$ is a D-dimensional invertible matrix for each a' with $|a - a'| \leq \eta_1$. Define a D-dimensional vector $\Delta_{a'}$ by

$$\Delta_{a'} = [\beta(M_i(a') - M_i(a)), \beta(M_j(a') - M_j(a)), 0] \left[M^i(a'); M^j(a'); L(a)\right]^{-1}.$$

By definition,

$$\begin{aligned}(\beta + \Delta_{a'}) M_i(a') &= \beta M_i(a') - \beta(M_i(a') - M_i(a)) = \beta M_i(a) = 1, \\ (\beta + \Delta_{a'}) M_j(a') &= \beta M_j(a') - \beta(M_j(a') - M_j(a)) = \beta M_j(a) = 0.\end{aligned}$$

Thus, $\beta - \Delta_{a'}$ satisfies the constraint for a', and hence $f^{ij}(a') \leq |\beta| + |\Delta_{a'}|$. Since $\limsup_{\eta_1 \to 0} \sup_{a':|a-a'| \leq \eta_1} |\Delta_{a'}| = 0$, for sufficiently small $\eta_1 > 0$, we have $f^{ij}(a') \leq |\beta| + |\Delta_{a'}| \leq |\beta| + \frac{1}{2}\eta_0 \leq f^{ij}(a) + \eta_0$ for all a' with $|a - a'| \leq \eta_1$, establishing upper-semicontinuity. ■

Given Lemma 14, Assumption 1(ii) holds with $\bar{x} = \bar{u}\hat{x}$. To see why, for any i and a, let $\partial u_i = \frac{\partial}{\partial a_i'} u_i(a_i', a_{-i})\Big|_{a_i' = a_i}$ and $x_i(y) = -\partial u_i^\lambda \sum_d x_i^j(d; a^\lambda) y^d$. Then,

$$\begin{aligned}\frac{\partial}{\partial a_i'}(u_i(a_i', a_{-i}) + \mathbb{E}[x_i(y) | a_i', a_{-i}])\Big|_{a_i' = a_i} &= 0 \quad \text{and} \\ \frac{\partial}{\partial a_j'}\mathbb{E}\left[x_j(y) | a_j', a_{-i}\right]\Big|_{a_j' = a_j} &= 0 \quad \text{for all } j \neq i.\end{aligned}$$

Since u_i is concave in a_i and satisfies the Inada conditions, $\mathbb{E}[x_i(y) | a_i', a_{-i}]$ is linear in a_i', and $\mathbb{E}\left[x_j(y) | a_j', a_{-j}\right]$ is linear in a_j', we have (4) and (5).

References

[1] Abreu, Dilip, David Pearce, and Ennio Stacchetti. "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring." *Econometrica* 58.5 (1990): 1041-1063.

[2] Abreu, Dilip, Paul Milgrom, and David Pearce. "Information and Timing in Repeated Partnerships." *Econometrica* 59.6 (1991): 1713-1733.

[3] Aoyagi, Masaki. "Information Feedback in a Dynamic Tournament." *Games and Economic Behavior* 70.2 (2010): 242-260.

[4] Athey, Susan, and Kyle Bagwell. "Optimal Collusion with Private Information." *RAND Journal of Economics* 32.3 (2001): 428-465.

[5] Ball, Ian. "Dynamic Information Provision: Rewarding the Past and Guiding the Future." *Econometrica*, Forthcoming (2023).

[6] Dharmadhikari, S. W., V. Fabian, and K. Jogdeo. "Bounds on the Moments of Martingales." *Annals of Mathematical Statistics* 39.5 (1968): 1719-1723.

[7] Ederer, Florian. "Feedback and Motivation in Dynamic Tournaments." *Journal of Economics & Management Strategy* 19.3 (2010): 733-769.

[8] Ely, Jeffrey, George Georgiadis, and Luis Rayo, "Feedback Design in Dynamic Moral Hazard," *Working Paper* (2023).

[9] Ely, Jeffrey, George Georgiadis, Sina Khorasani, and Luis Rayo, "Optimal Feedback in Contests." *Review of Economic Studies*, Forthcoming (2023).

[10] Ely, Jeffrey C., and Martin Szydlowski. "Moving the Goalposts." *Journal of Political Economy* 128.2 (2020): 468-506.

[11] Forges, Francoise. "An Approach to Communication Equilibria." *Econometrica* 54.6 (1986): 1375-1385.

[12] Freedman, David A. "On Tail Probabilities for Martingales." *Annals of Probability* 3.1 (1975): 100-118.

[13] Frick, Mira, Ryota Iijma, and Yuhta Ishii. "Monitoring with Rich Data." *Working Paper* (2023).

[14] Fuchs, William. "Contracting with Repeated Moral Hazard and Private Evaluations." *American Economic Review* 97.4 (2007): 1432-1448.

[15] Fudenberg, Drew, and David K. Levine. "Continuous Time Limits of Repeated Games with Imperfect Public Monitoring." *Review of Economic Dynamics* 10.2 (2007): 173-192.

[16] Fudenberg, Drew, David Levine, and Eric Maskin. "The Folk Theorem with Imperfect Public Information." *Econometrica* 62.5 (1994): 997-1039.

[17] Gershkov, Alex, and Motty Perry. "Tournaments with Midterm Reviews." *Games and Economic Behavior* 66.1 (2009): 162-190.

[18] Goldlücke, Susanne, and Sebastian Kranz. "Infinitely Repeated Games with Public Monitoring and Monetary Transfers." *Journal of Economic Theory* 147.3 (2012): 1191-1221.

[19] Green, Edward J., and Robert H. Porter. "Noncooperative Collusion under Imperfect Price Information." *Econometrica* 52.1 (1984): 87-100.

[20] Halac, Marina, Navin Kartik, and Qingmin Liu. "Contests for Experimentation." *Journal of Political Economy* 125.5 (2017): 1523-1569.

[21] Holmström, Bengt, and Paul Milgrom. "Aggregation and Linearity in the Provision of Intertemporal Incentives," *Econometrica* 55.2 (1987): 303-328.

[22] Hörner, Johannes, and Satoru Takahashi. "How Fast do Equilibrium Payoff Sets Converge in Repeated Games?" *Journal of Economic Theory* 165 (2016): 332-359.

[23] Kandori, Michihiro. "Introduction to Repeated Games with Private Monitoring." *Journal of Economic Theory* 102.1 (2002): 1-15.

[24] Kandori, Michihiro, and Hitoshi Matsushima. "Private Observation, Communication and Collusion." *Econometrica* 66.3 (1998): 627-652.

[25] Levin, Jonathan. "Relational Incentive Contracts." *American Economic Review* 93.3 (2003): 835-857.

[26] Lizzeri, Alessandro, Margaret Meyer, and Nicola Persico. "The Incentive Effects of Interim Performance Evaluations." *Working Paper* (2002).

[27] Madrigal, Vicente. "On the Non-existence of Efficient Equilibria of Repeated Principal Agent Games with Discounting." *Working Paper* (1986).

[28] Matsushima, Hitoshi. "Multimarket Contact, Imperfect Monitoring, and Implicit Collusion." *Journal of Economic Theory* 98.1 (2001): 158-178.

[29] Matsushima, Hitoshi. "Repeated Games with Private Monitoring: Two Players." *Econometrica* 72.3 (2004): 823-852.

[30] Mirrlees, James A. "The Theory of Moral Hazard and Unobservable Behaviour: Part I." *Working Paper* (1975) (published in *Review of Economic Studies* 66.1 (1999): 3-21).

[31] Orlov, Dmitry, Andrzej Skrzypacz, and Pavel Zryumov. "Persuading the Principal to Wait." *Journal of Political Economy* 128.7 (2020): 2542-2578.

[32] Pinelis, Iosif, "Best Possible Bounds of the Von Bahr-Esseen Type." *Annals of Functional Analysis* 6.4 (2015): 1-29.

[33] Radner, Roy. "Repeated Principal-Agent Games with Discounting," *Econometrica* 53.5 (1985): 1173-1198.

[34] Rahman, David. "The Power of Communication." *American Economic Review* 104.11 (2014): 3737-3751.

[35] Rubinstein, Ariel, and Menahem E. Yaari. "Repeated Insurance Contracts and Moral Jazard." *Journal of Economic Theory* 30.1 (1983): 74-97.

[36] Rubinstein, Ariel, "An Optimal Conviction Policy for offenses that May have been Committed by Accident," in *Applied Game Theory*, ed. by Brams, Schotter, and Schwodiauer, (1979) 406–413, Physical-Verlag: Heidleberg.

[37] Sadzik, Tomasz, and Ennio Stacchetti. "Agency Models with Frequent Actions." *Econometrica* 83.1 (2015): 193-237.

[38] Sannikov, Yuliy. "Games with Imperfectly Observable Actions in Continuous Time." *Econometrica* 75.5 (2007): 1285-1329.

[39] Sannikov, Yuliy. "A Continuous-Time Version of the Principal-Agent Problem." *Review of Economic Studies* 75.3 (2008): 957-984.

[40] Sannikov, Yuliy, and Andrzej Skrzypacz. "Impossibility of Collusion under Imperfect Monitoring with Flexible Production." *American Economic Review* 97.5 (2007): 1794-1823.

[41] Sannikov, Yuliy, and Andrzej Skrzypacz. "The Role of Information in Repeated Games with Frequent Actions." *Econometrica* 78.3 (2010): 847-882.

[42] Smolin, Alex. "Dynamic Evaluation Design." *American Economic Journal: Microeconomics* 13.4 (2021): 300-331.

[43] Spear, Stephen E., and Sanjay Srivastava. "On Repeated Moral Hazard with Discounting." *Review of Economic Studies* 54.4 (1987): 599-617.

[44] Sugaya, Takuo. "Folk Theorem in Repeated Games with Private Monitoring." *Review of Economic Studies* 89.4 (2022): 2201-2256.

[45] Sugaya, Takuo, and Alexander Wolitzky. "Bounding Equilibrium Payoffs in Repeated Games with Private Monitoring." *Theoretical Economics* 12 (2017): 691-729.

[46] Sugaya, Takuo, and Alexander Wolitzky. "Maintaining Privacy in Cartels." *Journal of Political Economy* 126.6 (2018): 2569-2607.

[47] Sugaya, Takuo, and Alexander Wolitzky. "Monitoring vs. Discounting in Repeated Games." *Econometrica*, Forthcoming (2023).

[48] Wood, David C. "The Computation of Polylogarithms." *Working Paper* (1992).

[49] Zhu, Huangjun, Zihao Li, and Masahito Hayashi. "Nearly Tight Universal Bounds for the Binomial Tail Probabilities." *Working Paper* (2022).