Moral Hazard and Repeated Games with Many Agents*

Takuo Sugaya — Stanford GSB
Alexander Wolitzky — MIT

May 21, 2024

Abstract

Motivated by the problem of sustaining cooperation in large groups with limited information, we analyze the relationship between group size, monitoring precision, and incentive instruments in multiagent moral hazard problems and repeated games with independent, player-level noise. We link the viability of cooperation to the *per-capita channel capacity* of the monitoring structure. In static moral hazard problems, cooperation is impossible if the per-capita channel capacity is much smaller than the squared inverse of the maximum feasible reward; conversely, cooperation is possible in some games where this relationship is reversed. In repeated games, cooperation is impossible if the per-capita channel capacity is much smaller than the discount rate; again, a converse holds for a class of games. If attention is restricted to *team equilibria*, where incentives are provided collectively, cooperation is possible only under much more severe parameter restrictions. Personalized incentives greatly outperform team incentives when information is scarce.

Keywords: moral hazard, repeated games, large teams, monitoring, mutual information, channel capacity, tail test

JEL codes: C72, C73, D86

---

*This paper supercedes our earlier paper, "Repeated Games with Many Players." For helpful comments, we thank Tommaso Denti, Drew Fudenberg, Stephen Morris, Andy Postlewaite, Yuliy Sannikov, Satoru Takahashi, and many seminar participants. Wolitzky acknowledges financial support from the NSF.

Two neighbours may agree to drain a meadow which they possess in common; because it is easy for them to know each other's mind; and each must perceive that the immediate consequence of his failing in his part is the abandoning of the whole project. But it is very difficult, and indeed impossible, that a thousand persons should agree in any such action; it being difficult for them to concert so complicated a design, and still more difficult for them to execute it; while each seeks pretext to free himself of the trouble and expense, and would lay the whole burden on others.

—David Hume, *A Treatise of Human Nature*

1 Introduction

Hume's intuition notwithstanding, large groups of individuals often have a remarkable capacity for cooperation. Large-group cooperation occurs within firms and other organizations governed by explicit contracts, as well as in long-run relationships where contractual enforcement is unavailable (Ostrom, 1990; Ellickson, 1991; Seabright, 2004). In both static settings with explicit contracts and dynamic settings with informal enforcement, large-group cooperation typically relies on accurate monitoring of individual agents' actions, together with sanctions that narrowly target deviators.[1] However, the principle that large-group cooperation requires precise monitoring and personalized sanctions is not clearly expressed in standard economic models of either static incentive problems (*moral hazard in teams*, following Holmström, 1982) or dynamic ones (*repeated games*, surveyed by Mailath and Samuelson, 2006). In particular, existing results do not quantify how the "amount" of information needed to support cooperation depends on the number of agents and the scale of the available incentive instruments. Standard models also do not investigate how the relationship between monitoring precision, group size, and incentive instruments varies depending on whether the group relies on *personalized incentives* (e.g., individual bonuses or fines) or *team incentives* (e.g., price wars; or, in Hume's example, "the abandoning of the whole project").

This paper extends standard models of moral hazard in teams and repeated games by letting the monitoring structure, the scale of the available incentive instruments (or in the

[1] Examples include the community resource management settings documented by Ostrom (1990); the local public goods setting studied by Miguel and Gugerty (2005); and the group lending settings studied by Karlan (2007) and Feigenberg, Field, and Pande (2013).

repeated game context, the discount factor), and the size of the group all vary simultaneously. These features can vary in a flexible manner: we assume only a uniform upper bound on the range of the stage game payoffs and a uniform lower bound on the amount of independent, player-level "monitoring noise." Our main results provide necessary and sufficient conditions for cooperation as a function of group size, maximum rewards/discounting, and a measure of monitoring precision. We also establish severe obstacles to cooperation under team incentives. Notably, personalized incentives greatly outperform team incentives even when information is scarce, so that precisely monitoring each individual is infeasible. In sum, we show that large-group cooperation requires precise monitoring and powerful incentive instruments, and that large-group cooperation cannot be based on team incentives for reasonable parameter values.

We now preview our model and results. We model independent, player-level noise by assuming that each player i's action a_i stochastically determines an *individual outcome* x_i, independently across players, and that the distribution of observed signals y depends on the action profile $a = (a_i)$ only through the outcome profile $x = (x_i)$.[2] As the following example illustrates, in the absence of noise there may be no tradeoff between group size, monitoring precision, and discounting or other incentive instruments.

Example 1. Suppose N players repeatedly play a prisoner's dilemma with a binary public signal $y \in \{0, 1\}$, where in each period $y = 0$ if *every* player cooperates, and $y = 1$ if *any* player defects. A player's payoff is the fraction of players who cooperate, less a constant (independent of N) if she cooperates herself. In this game, each player's action is pivotal in determining y when the others cooperate, so the range of values for the discount factor for which mutual cooperation is a sequential equilibrium outcome is independent of N. Thus, a single "bit" of information (the binary signal y) can form the basis for cooperation among an arbitrarily large group of players in a repeated game where the range of stage game payoffs, the cost of cooperation, and the discount factor are all fixed independent of N.

Now introduce noise. For example, let each player "tremble" in her choice of action with probability π, independently across players, with π fixed independent of N. Assume that the distribution of the public signal depends only on the players' realized actions, not their intended actions. Then a single bit of information can no longer motivate cooperation by

[2] This setup follows Fudenberg, Levine, and Pesendorfer (1998) and al-Najjar and Smorodinsky (2000, 2001).

a large group of players for a fixed discount factor. Moreover, a novel implication of the our results is that the required number of bits of information (e.g., the log of the number of possible signal realizations) is proportional to N.

Our key technical insight is that the analysis of games with independent noise is facilitated by basic tools from information theory. Most importantly, we find that a useful measure of monitoring precision is the *channel capacity*, C, of the conditional signal distribution $q(Y|X)$.[3] Channel capacity is a standard measure in information theory, which in our context is defined as the maximum *mutual information* $I(X;Y)$ between the profile of individual outcomes X and the signal Y: that is, the expected reduction in uncertainty about the outcome profile X that results from observing the signal Y.

Mutual information obeys three elementary inequalities that play important roles in our theory. First, from the perspective of single-agent moral hazard, the "influence" of a player i's action a_i on the distribution of the signal Y can be bounded in terms of $I^{\pi(\cdot|a)}(X_i;Y)$, the mutual information between player i's outcome X_i and the signal under the probability distribution that results when a is played. Specifically, letting $\mathrm{TV}(p(\cdot|a_i', a_{-i})\,||p(\cdot|a))$ denote the *total variation distance* between the signal distributions $p(\cdot|a_i', a_{-i})$ and $p(\cdot|a)$, and letting $\underline{\pi}$ denote the minimum "noise level" (e.g., the tremble probability in Example 1), it is straightforward to show that

$$(\mathrm{TV}(p(\cdot|a_i', a_{-i})\,||p(\cdot|a)))^2 \leq \kappa(\underline{\pi})\, I^{\pi(\cdot|a)}(X_i;Y), {}^4 \tag{1}$$

for some function κ. Inequality (1) is not especially useful for analyzing single-agent moral hazard problems, but it is useful for analyzing multiagent problems with independent noise. This is because of a second key inequality:

$$\sum_i I^{\pi(\cdot|a)}(X_i;Y) \leq I^{\pi(\cdot|a)}(X;Y) \leq C, \tag{2}$$

which holds because the individual outcomes X_i are assumed to be independent conditional on the action profile a. In particular, combining inequalities (1) and (2) bounds the *average*

[3] We use capital letters for random variables and lower-case letters for their realizations, so X and Y denote the (random) outcome profile and signal.

[4] Inequality (5) in the text also covers the χ^2-*divergence* of $p(\cdot|a_i', a_{-i})$ from $p(\cdot|a)$, in addition to the TV distance.

over players i of the influence of player i's action a_i on the distribution of the signal Y in terms of the *per-capita channel capacity* C/N. This simple observation is one of our key insights. Finally, a third inequality illustrates how our results improve on prior results in the literature: because the mutual information $I(X;Y)$ is bounded by the entropy of the signal Y, which in turn is bounded by the log of the number of possible signal realizations, $\log|\mathcal{Y}|$, we have

$$C \leq \log|\mathcal{Y}|. \tag{3}$$

Inequality (3) implies that bounds on the players' incentives in terms of channel capacity immediately imply corresponding bounds in terms of the number of possible signal realizations. This illustrates how our results based on channel capacity improve on prior results based on the number of possible signal realizations, especially the results of Fudenberg, Levine, and Pesendorfer (1998) and al-Najjar and Smorodinsky (2000, 2001), which we discuss below.

Turning to our results, we begin by considering static moral hazard in teams problems with many agents, independent noise, limited liability, and an exogenous upper bound on *rewards*, $\bar{w}$. Our first result (Theorem 1) shows that if $\bar{w}^2 C/N$—the product of the square of the maximum reward $\bar{w}$ and the per-capita channel capacity C/N—is small, then cooperation is impossible: all implementable outcomes are consistent with approximately myopic play. This shows that cooperation in a large team requires large rewards and/or precise monitoring. Moreover, our second result (Proposition 1) shows that Theorem 1 is tight, in that if $\bar{w}^2 C/N$ is large then cooperation is possible in some games.

We then restrict attention to *team contracts*, where incentives are provided collectively, in that the players' rewards are co-linear. This restriction makes a bang-bang reward structure optimal, so increasing C beyond $\log(2)$ (i.e., one bit) is no longer valuable. More importantly, we show that the per-capita expected cost of motivating cooperation with a team contract explodes with N, unless the maximum reward $\bar{w}$ is *exponentially* large relative to N (Theorem 2). The intuition is that the optimal team contract is a *tail test*, where the players are all paid $\bar{w}$ if the number of "good" outcomes x_i exceeds a threshold n^* (and are all paid 0 otherwise). It can then be shown that the ratio of the probability that one player's action is pivotal for the tail test and the probability that the test is passed converges to zero as $N \to \infty$, unless these probabilities are both very close to zero. But a tail test where the pivot probability is very close to zero provides only small incentives, unless $\bar{w}$ is very large. Thus, with limited liability and maximum rewards that are not exceedingly large, team incentives are much less cost-effective than personalized incentives, even if information is scarce. This result provides a novel rationalization for the use of personalized incentive schemes.

We then turn to repeated games. Here, we ask questions analogous to those in the static model, but with equilibrium continuation payoffs replacing exogenous, bounded rewards. Our first result here (Theorem 3) shows that cooperation is impossible if $(1-\delta)^{-1} C/N$ is small: that is, if the per-capita channel capacity C/N is much smaller than the discount rate $1-\delta$. This impossibility result for cooperation in a repeated game with a discount rate of $1-\delta$ is the same as that in a static game with a maximum reward of $\bar{w} = (1-\delta)^{-1/2}$, whereas naïve intuition would suggest that the maximum reward in a repeated game is $(1-\delta)^{-1}$. Theorem 3 builds on a general necessary condition for cooperation in repeated games that we establish in a companion paper (Sugaya and Wolitzky, 2023; henceforth SW). Compared to that result, the key difference is that here we assume independent noise, which allows a connection between the main information measure in SW (the χ^2-*divergence* of the signal distribution following a deviation from the equilibrium signal distribution) and channel capacity.

In the online appendix, we show that the relationship among N, δ, and C in Theorem 3 is tight up to a $\log(N)$ factor. This entails establishing a "folk theorem" in a setting where the discount factor, monitoring structure, and stage game (including the number of players N) all vary simultaneously.

Our final result (Theorem 4) considers the consequences of restricting attention to team incentives in repeated games. In a repeated game, a *team equilibrium* is one where all on-equilibrium-path continuation payoff vectors are co-linear. When the stage game is symmetric and continuation payoff vectors lie on the 45° line, team equilibria reduce to *strongly symmetric equilibria*, which are a standard model of collusion through the threat of price wars (Green and Porter, 1984; Abreu, Pearce, and Stacchetti, 1986; Athey, Bagwell, and Sanchirico, 2004). Here, we show that cooperation in a team equilibrium is impossible, unless $(1-\delta)^{-1}$ is exponentially large relative to N. This theorem is analogous to Theorem 2 in the static context, except that Theorem 2 concludes that supporting cooperation in a team equilibrium is extremely costly, whereas Theorem 4 concludes that it is impossible. This difference arises because the expected rewards required to support cooperation in Theorem 2 are too large to deliver as continuation payoffs in a repeated game.[5]

[5] It is well-known that strongly symmetric equilibria are typically less efficient than general perfect public

1.1 Related Literature

This paper contributes to the literatures on static moral hazard in teams problems and repeated games. The closest papers concern repeated games. Here, prior research has established folk theorems in the $\delta \to 1$ limit for fixed N, as well as "anti-folk" theorems in the $N \to \infty$ limit for fixed δ, but has not considered the case where N and δ vary together. The closest paper is our companion work, SW. That paper establishes general necessary and sufficient conditions for cooperation in repeated games as a function of discounting and monitoring precision. Relative to SW, the current paper introduces two features specific to large-population games: independent noise and the possibility that N varies together with discounting and monitoring. Independent noise is crucial for all of our results, while letting N vary with discounting and monitoring is the key novelty in our folk theorem (Theorem 5 in the online appendix).

Other than those in SW, the most relevant prior results are those of Fudenberg, Levine, and Pesendorfer (1998), al-Najjar and Smorodinsky (2000, 2001), and Awaya and Krishna (2016, 2019). Following earlier work by Green (1980) and Sabourian (1990), these papers establish conditions under which play in a repeated game is approximately myopic as $N \to \infty$ for fixed δ.[6] These conditions can be adapted to the case where N, δ, and monitoring vary together, but the results so obtained are weaker than ours (and are not tight up to log terms). As we explain in Section 4.2, the key difference is that prior results rely on bounds on the strength of players' incentives with a worse order in the discount rate than that given in SW. In sum, prior work has established anti-folk theorems as $N \to \infty$ for fixed δ, while our paper tightly (up to log terms) characterizes the tradeoff among N, δ, and monitoring that is required for supporting cooperation.[7]

Since the monitoring structure varies with δ in our model, we also relate to repeated games with frequent actions, where the monitoring structure varies with δ in a particular,

equilibria in repeated games. Theorem 4 instead shows that the relationship between N and δ required for *any* non-trivial incentive provision differs dramatically between strongly symmetric (more generally, linear) equilibria and general ones.

[6] Pai, Roth, and Ullman (2014) establish another, similar anti-folk theorem. Awaya and Krishna instead establish conditions under which cheap talk is valuable. Green and Sabourian's papers impose a continuity condition on the mapping from action distributions to signal distributions. Continuity is implied by independent noise.

[7] Farther afield, there is also work suggesting that repeated game cooperation is harder to sustain in larger groups based on evolutionary models (Boyd and Richerson, 1988), simulations (Bowles and Gintis, 2011; Chapter 4), and experiments (Camera, Casari, and Bigoni, 2013).

parametric manner (e.g., Abreu, Milgrom, and Pearce, 1991; Fudenberg and Levine, 2007, 2009; Sannikov and Skrzypacz, 2007, 2010). The most relevant results here are Sannikov and Skrzypacz's (2007) theorem on the impossibility of collusion with frequent actions and Brownian noise, as well as a related result by Fudenberg and Levine (2007). These results relate to our impossibility theorem for team equilibrium (Theorem 4), as we explain in Section 4.3.

Most of the literature on static moral hazard in teams, following Alchian and Demsetz (1972) and Holmström (1982), does not focus on settings with a large number of agents. An exception is a set of papers that consider the implications of limited managerial attention for team size and organizational structure. For example, Calvo and Wellisz (1978) and Qian (1994) model the "span of control" of a manager in an organizational hierarchy as the number of immediate subordinates that she can monitor or control. Our results suggest an interpretation of the span of control in terms of information: by Theorem 1, the maximum number of subordinates that a manager can motivate with any contract with bounded rewards is proportional to the number of bits of information about the subordinates' performance that the manager can acquire and process. This connection could perhaps be pursued in future research.

Finally, we are not aware of prior papers that employ entropy methods in static moral hazard problems. In repeated games, entropy methods have been used to study issues including complexity and bounded recall (Neyman and Okada, 1999, 2000; Hellman and Peretz, 2020), communication (Gossner, Hernández, and Neyman, 2006), and reputation effects (Gossner, 2011; Ekmekci, Gossner, and Wilson, 2011; Faingold, 2020). However, other than sharing a reliance on entropy methods, these papers are not very related to ours.

2 Moral Hazard in Large Teams

We first consider static multiagent moral hazard problems with independent noise, limited liability, and bounded rewards.

The Game. There is a finite set of players $\mathcal{I} = \{1, \ldots, N\}$, a finite product set of actions $\mathcal{A} = \times_{i \in I} \mathcal{A}_i$, and a payoff function $u_i : \mathcal{A} \to \mathbb{R}$ for each $i \in I$. We assume that $|\mathcal{A}_i| \geq 2$ for all i, and denote the range of player i's payoff function by $\bar{u}_i = \max_{a,a'} u_i(a) - u_i(a')$. For any $\bar{u} > 0$, we say that payoffs are $\bar{u}$-*bounded* if $\bar{u}_i \leq \bar{u}$ for all i.

$$\begin{tikzcd} \mathcal{A} \arrow[r, "\pi(x|a)"] \arrow[rr, bend right, "p(y|a)"'] & \mathcal{X} \arrow[r, "q(y|x)"] & \mathcal{Y} \end{tikzcd}$$

Figure 1: The noise structure $(\pi, \mathcal{X})$ and the outcome monitoring structure $(q, \mathcal{Y})$ jointly determine the action monitoring structure $(p, \mathcal{Y})$.

Noise. There is a finite product set of *outcome profiles* $\mathcal{X} = \times_{i \in I} \mathcal{X}_i$, where $\mathcal{X}_i$ is the set of *individual outcomes* for player i. When player i takes action a_i, her individual outcome $x_i \in \mathcal{X}_i$ is drawn from a probability distribution $\pi_i(\cdot|a_i) \in \Delta(\mathcal{X}_i)$. When action profile $a \in$A is played, the outcome profile $x \in$X is drawn from the product distribution $\pi(\cdot|a) = \times_i \pi_i(\cdot|a_i)$. We call the pair $(\mathcal{X}, \pi)$ a *noise structure*. Let $\underline{\pi}_i = \min_{a_i, x_i} \pi_i(x_i|a_i)$ and assume that $\min_i \underline{\pi}_i > 0$. For any $\underline{\pi} > 0$, we say that noise is $\underline{\pi}$*-bounded* if $\underline{\pi}_i \geq \underline{\pi}$ for all i. Note that if noise is $\underline{\pi}$-bounded then $|\mathcal{X}_i| \leq 1/\underline{\pi}$ for all i. We assume that $|\mathcal{X}_i| \geq 2$ for at least one player i, which implies that noise can be $\underline{\pi}$-bounded only for $\underline{\pi} \leq 1/2$.

A simple example of a noise structure arises when there is independent noise in the execution of the players' actions, so that a_i is player i's intended action and x_i is her realized action. In this case, $\mathcal{X} = \mathcal{A}$ and $\pi_i(a_i'|a_i)$ is the probability that player i "trembles" to a_i' when she intends to take a_i.

Monitoring. An *outcome monitoring structure* $(\mathcal{Y}, q)$ consists of a finite set of possible signal realizations Y and a conditional probability distribution $q(\cdot|x) \in \Delta(\mathcal{Y})$ for each outcome profile x. The signal distribution thus depends only on the outcome profile and not directly on the action profile. In other words, if we view the action profile, the outcome profile, and the signal as random variables A, X, and Y, they form a Markov chain $A \to X \to Y$.

Given an outcome monitoring structure $(\mathcal{Y}, q)$, we denote the probability of signal profile y at action profile a by $p(y|a) = \sum_x \pi(x|a) q(y|x)$. We refer to the pair $(\mathcal{Y}, p)$ as the *action monitoring structure* induced by $(\mathcal{X}, \pi, \mathcal{Y}, q)$. Without loss, we assume that for every $y \in \mathcal{Y}$, there exists $x \in \mathcal{X}$ such that $q(y|x) > 0$. Since $\underline{\pi}_i > 0$ for each i, this implies that p has full support: $p(y|a) > 0$ for all a, y.

Figure 1 summarizes the relationship between the noise structure $(\mathcal{X}, \pi)$, the outcome monitoring structure $(\mathcal{Y}, q)$, and the action monitoring structure $(\mathcal{Y}, p)$.

Finally, for any action profile $a \in \mathcal{A}$, let $\varphi^a \in \Delta(\mathcal{X} \times \mathcal{Y})$ denote the joint distribution on $\mathcal{X} \times \mathcal{Y}$ when a is played, so that X has distribution $\pi(\cdot|a)$ and, conditional on each realization x, Y has distribution $q(\cdot|x)$.

Contracts. A *principal* specifies contracts for the players and privately recommends actions. A *contract for player* i is a function $w_i : \mathcal{A} \times \mathcal{Y} \to \mathbb{R}_+$ specifying a non-negative *reward* $w_i(a,y)$ for player i when action profile a is recommended and signal y realizes. A *contract* is a collection $w = (w_i)$. We assume that rewards are $\bar{w}$*-bounded*: there exists $\bar{w} > 0$ such that $w_i(a,y) \in [0, \bar{w}]$ for all i, a, y. Thus, the players are protected by limited liability and there is a maximum feasible reward. A contract is *public* if it depends only on y, so that $w(a,y) = w(\tilde{a}, y)$ for all $a, \tilde{a}, y$.

Equilibrium. A *manipulation* for a player i is a mapping $s_i : \mathcal{A}_i \to \Delta(\mathcal{A}_i)$. The interpretation is that when player i is recommended a_i, she instead plays $s_i(a_i)$. A distribution over action profiles $\alpha \in \Delta(\mathcal{A})$, together with a contract w, is a *correlated equilibrium* if, for any player i and manipulation s_i, we have

$$\sum_{a,y} \alpha(a)\left(u_i(a) + p(y|a)\, w_i(a,y)\right) \geq \sum_{a,y} \alpha(a)\left(u_i(s_i(a_i), a_{-i}) + p(y|s_i(a_i), a_{-i})\, w_i(a,y)\right).^{8}$$

A correlated equilibrium (α, w) is *public* if $\alpha \in \prod_i \Delta(\mathcal{A}_i)$ and w is a public contract. Finally, the *per-capita expected cost* incurred by the principal at a correlated equilibrium (α, w) equals

$$\frac{1}{N} \sum_{a,i} \alpha(a)\, p(y|a)\, w_i(a,y).$$

***ε*-Myopic Play.** Player i's *gain* from manipulation s_i at an action profile distribution $\alpha \in \Delta(\mathcal{A})$ is

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

Thus, the pair (α, w) is a correlated equilibrium if and only if

$$g_i(s_i, \alpha) \leq \sum_{a,y} \alpha(a)\left(p(y|a) - p(y|s_i(a_i), a_{-i})\right) w_i(a,y) \qquad \text{for all } i, s_i. \tag{4}$$

Player i's *maximum gain* at $\alpha \in \Delta(\mathcal{A})$ is $\bar{g}_i(\alpha) = \max_{s_i : \mathcal{A}_i \to \Delta(\mathcal{A}_i)} g_i(s_i, \alpha)$. Finally, for any

[8] Here and throughout, u_i and p linearly extend to mixed actions.

$\varepsilon > 0$, the set of *ε-myopic action distributions* is

$$\mathcal{A}(\varepsilon) = \left\{ \alpha \in \Delta(\mathcal{A}) : \frac{1}{N} \sum_i \bar{g}_i(\alpha) \leq \varepsilon \right\},$$

and the set of *ε-myopic payoff vectors* is

$$V(\varepsilon) = \left\{ v \in \mathbb{R}^N : v = u(\alpha) \text{ for some } \alpha \in \mathcal{A}(\varepsilon) \right\}.$$

Thus, an action distribution α is ε-myopic if the per-player average deviation gain at α is less than ε. If the game is symmetric and α is a symmetric distribution, this definition implies that all players have deviation gains smaller than ε. Otherwise, it allows a few players to have large gains. In Appendix A, we provide some results comparing $V(\varepsilon)$ with the (smaller) set of payoff vectors that are consistent with *all* players having small deviation gains.

Mutual Information and Channel Capacity. Given a distribution of outcomes $\xi \in \Delta(\mathcal{X})$, a standard measure of the informativeness of the signal Y about the outcome X is the *mutual information* between the random variables X and Y, defined as

$$I^{\xi}(X;Y) = \sum_{(x,y) \in \mathcal{X} \times \mathcal{Y}} \xi(x) q(y|x) \log \left(\frac{q(y|x)}{\sum_{x' \in \mathcal{X}} \xi(x') q(y|x')} \right).^{9}$$

Mutual information measures the expected reduction in uncertainty (entropy) about X that results from observing Y. The mutual information between X and Y is an endogenous object in our model, as it depends on the distribution ξ of X, which in turn is determined by the players' actions, a.

Next, denote the set of outcome distributions ξ that can arise for some action distribution α under noise structure $(\mathcal{X}, \pi)$ by

$$\vartheta = \left\{ \xi \in \Delta(\mathcal{X}) : \exists \alpha \in \Delta(\mathcal{A}) \text{ such that } \xi(x) = \sum_{a \in \mathcal{A}} \alpha(a) \pi(x|a) \text{ for all } x \in \mathcal{X} \right\}.$$

Finally, define the *channel capacity* of the tuple $(\mathcal{X}, \pi, \mathcal{Y}, q)$ as

$$C = \max_{\xi \in \vartheta} I^{\xi}(X;Y).$$

[9] In this paper, all logarithms are base e.

Channel capacity is an exogenous measure of the informativeness of Y about X, as it is a function of the noise structure $(\mathcal{X}, \pi)$ and the outcome monitoring structure $(\mathcal{Y}, q)$. Channel capacity plays a central role in information theory as the maximum rate at which information can be transmitted over a noisy channel (*Shannon's channel coding theorem*; Cover and Thomas, 2006, Theorem 7.7.1). Our analysis does not use this theorem; we only use channel capacity as an exogenous upper bound on mutual information. In turn, mutual information enters our analysis because it obeys inequalities (1)–(3) displayed in the introduction. We next explain these inequalities.

Mutual Information Inequalities. Among inequalities (1)–(3), the third one—$C \leq \log |\mathcal{Y}|$—is completely standard: see, e.g., Theorem 2.6.4 of Cover and Thomas (henceforth, CT). Similarly, if noise is $\underline{\pi}$-bounded then $C \leq -N \log \underline{\pi}$, because $|\mathcal{X}| = \prod_i |\mathcal{X}_i| \leq 1/\underline{\pi}^N$.

Now consider inequality (1). As condition (4) indicates, the maximum gain $g_i(s_i, \alpha)$ that player i is willing to forgo in a correlated equilibrium α is related to the expectation over recommended action profiles a of the "distance" of the signal distribution following manipulation s_i, $p(\cdot|s_i(a_i), a_{-i})$, from the prescribed signal distribution, $p(\cdot|a)$. We give an upper bound (generalizing inequality (1)) in terms of the mutual information $I^{\pi(\cdot|a)}(X_i; Y)$ for two standard notions of this "distance": the *total variation distance*,

$$\mathrm{TV}\left(p\left(\cdot|a_i', a_{-i}\right)||p\left(\cdot|a\right)\right) := \frac{1}{2}\sum_y \left|p\left(y|a\right) - p\left(y|a_i', a_{-i}\right)\right|,$$

and the χ^2-*divergence,*

$$\chi^2\left(p\left(\cdot|a_i', a_{-i}\right)||p\left(\cdot|a\right)\right) := \sum_y \frac{\left(p\left(y|a\right) - p\left(y|a_i', a_{-i}\right)\right)^2}{p\left(y|a\right)}.$$

Lemma 1 *If noise is $\underline{\pi}$-bounded, then for any $a \in \mathcal{A}$, $i \in I$, and $a_i' \in \mathcal{A}_i$, we have*

$$\max\left\{\left(\mathrm{TV}\left(p\left(\cdot|a_i', a_{-i}\right)||p\left(\cdot|a\right)\right)\right)^2, \frac{\underline{\pi}}{2}\chi^2\left(p\left(\cdot|a_i', a_{-i}\right)||p\left(\cdot|a\right)\right)\right\} \leq \kappa\left(\underline{\pi}\right) I^{\pi(\cdot|a)}\left(X_i; Y\right), \quad (5)$$

where $\kappa(\underline{\pi}) = 2\left(1 - 2\underline{\pi}\right)^2/\underline{\pi}$.

An intuition for Lemma 1 is that, since player i's action affects the signal Y only through the outcome X_i, if a deviation from a_i to a_i' has a large effect on the distribution of Y, then Y must provide a large amount of information about X_i. Lemma 1 is related to standard *f-divergence inequalities* (Sason and Verdú, 2016), but it differs from standard results because of the $A \to X \to Y$ Markov chain structure of our model and the $\underline{\pi}$-bounded noise assumption.[10] The proof is deferred to the appendix.

We next restate inequality (2) and provide an intuition and proof.

Lemma 2 *For any action profile $a \in \mathcal{A}$, we have*

$$\sum_i I^{\pi(\cdot|a)}(X_i; Y) \leq I^{\pi(\cdot|a)}(X; Y) \leq C. \tag{2}$$

Inequality (2) is where we use the assumption that the X_i are independent conditional on a. An intuition for the inequality is that if $\sum_i I^{\pi(\cdot|a)}(X_i; Y) > I^{\pi(\cdot|a)}(X; Y)$ then there is some redundancy in the information that Y provides about the X_i, which is impossible when the X_i are conditionally independent. Note that inequality (2) can be strict: for example, if X_1 and X_2 are independent *Bernoulli* $(1/2)$ variables and Y is the parity of their sum, then $I(X_1; Y) = I(X_2; Y) = 0$ but $I((X_1, X_2); Y) > 0$.

Proof. We recall some basic concepts from information theory (see, e.g., Chapter 2 of CT). For any discrete random variable Z with distribution ζ, its *entropy* is $H(Z) = -\sum_z \zeta(z) \log \zeta(z)$. For any pair of discrete random variables (Z, Z') with joint distribution ζ, the mutual information $I(Z; Z')$ satisfies

$$I(Z; Z') = \sum_{z,z'} \zeta(z, z') \log\left(\frac{\zeta(z, z')}{\zeta(z)\zeta(z')}\right) = H(Z) - H(Z|Z'),$$

where the *conditional entropy* $H(Z|Z')$ is $H(Z|Z') = -\sum_{z,z'} \zeta(z, z') \log \zeta(z|z')$. We also recall the *independence bound on entropy* (CT, Theorem 2.6.6): if $Z = (Z_1, \ldots, Z_N)$ then $H(Z) \leq \sum_i H(Z_i)$, with equality if and only if the Z_i are independent.

We now prove inequality (2). Suppressing the superscript $\pi(\cdot|a)$, we have

$$\begin{aligned} \sum_i I(X_i; Y) &= \sum_i (H(X_i) - H(X_i|Y)) \\ &= \sum_i H(X_i) - \sum_i H(X_i|Y) \leq H(X) - H(X|Y) = I(X; Y), \end{aligned}$$

[10] In turn, Lemma 1 relies on *Pinsker's inequality* (CT, Lemma 11.1.1), which states that for any two probability distributions ζ and ζ' on a finite set $\mathcal{Z}$, we have TV $(\zeta||\zeta')^2 \leq (1/2) \sum_z \zeta(z) \log(\xi(z)/\xi'(z))$.

where the inequality follows because, by the independence bound on entropy and independence of the X_i, we have $H(X) = \sum_i H(X_i)$ and $H(X|Y) \leq \sum_i H(X_i|Y)$. Finally, $I(X;Y) \leq C$ by definition of channel capacity. ■

3 Cooperation and Non-Cooperation in Moral Hazard

We now present our results on the prospects for cooperation in static multiagent moral hazard problems. We first allow arbitrary contracts and then consider a restriction to team contracts, where the agents' rewards are co-linear.

3.1 Arbitrary Contracts

Our first result shows that if the per-capita channel capacity is much smaller than the squared inverse of the maximum reward, then all correlated equilibria are ε-myopic for small ε. Thus, cooperation in static moral hazard problems with many agents requires a large amount of information and/or large rewards.

Theorem 1 *Any correlated equilibrium in a multiagent moral hazard problem with N players, channel capacity C, $\underline{\pi}$-bounded noise, and $\bar{w}$-bounded rewards is ε-myopic, for*

$$\varepsilon = \sqrt{\kappa(\underline{\pi}) \frac{C}{N}} \bar{w}.$$

Theorem 1 restricts the set of correlated equilibria when noise is not too small, per-capita channel capacity is not too large, and maximum rewards is not too large. Conversely, it is vacuous in the limit where noise is small or the maximum reward is large.[11]

Theorem 1 is similar to earlier results by Fudenberg, Levine, and Pesendorfer (1998) and al-Najjar and Smorodinsky (2000, 2001). The key difference is measuring information by channel capacity rather than the number of possible signal realizations. The approach of these earlier papers would yield $\log|\mathcal{Y}|$ in place of $\sqrt{C}$ in Theorem 1, which gives a considerably weaker result as $\sqrt{C} \ll C \leq \log|\mathcal{Y}|$.[12]

[11] Since $C \leq -N \log \underline{\pi}$ when noise is $\underline{\pi}$-bounded, the per-capita channel capacity C/N cannot be very large unless noise is small.

[12] Another related result is Lemma 3 of Sugaya and Wolitzky (2021), which bounds the average influence of N independent binary random variables on an aggregate signal.

In addition to yielding a stronger result, information theory also allows a shorter proof.

Proof of Theorem 1. By Lemma 1, for any correlated equilibrium (α, w), any player i, and any manipulation s_i, we have

$$\begin{aligned} g_i(s_i, \alpha) &\leq \sum_{a,y} \alpha(a) \left(p(y|a) - p(y|s_i(a_i), a_{-i})\right) w_i(a, y) \\ &\leq \sum_a \alpha(a) \, \mathrm{TV}\left(p(\cdot|s_i(a_i), a_{-i}) \,||\, p(\cdot|a)\right) \bar{w} \\ &\leq \sum_a \alpha(a) \sqrt{\kappa(\underline{\pi}) \, I^{\pi(\cdot|a)}(X_i; Y)} \bar{w}. \end{aligned}$$

Theorem 1 follows, as we have

$$\begin{aligned} \frac{1}{N} \sum_i \bar{g}_i(\alpha) &\leq \frac{1}{N} \sum_i \sum_a \alpha(a) \sqrt{\kappa(\underline{\pi}) \, I^{\pi(\cdot|a)}(X_i; Y)} \bar{w} \\ &\leq \sqrt{\frac{1}{N} \sum_i \sum_a \alpha(a) \, \kappa(\underline{\pi}) \, I^{\pi(\cdot|a)}(X_i; Y)} \bar{w} \\ &= \sqrt{\kappa(\underline{\pi}) \frac{1}{N} \sum_a \alpha(a) \sum_i I^{\pi(\cdot|a)}(X_i; Y)} \bar{w} \\ &\leq \sqrt{\kappa(\underline{\pi}) \frac{C}{N}} \bar{w}, \end{aligned}$$

where the second inequality is by Jensen and the third is by Lemma 2. ■

We now show that, when per-capita channel capacity is small, Theorem 1 is tight up to a factor of 3.

Proposition 1 *Fix any $\underline{\pi} > 0$. There exists k such that, for any N and C satisfying $N/C > k$ and any $\bar{w} > 0$, there exists a multiagent moral hazard problem with N players, channel capacity C, $\underline{\pi}$-bounded noise, and $\bar{w}$-bounded rewards that has a correlated equilibrium that is not ε-myopic, for*

$$\varepsilon = \frac{1}{3} \sqrt{\kappa(\underline{\pi}) \frac{C}{N}} \bar{w}. \tag{6}$$

Note that if we consider a sequence where $N \to \infty$ while C/N, $\underline{\pi}$, and $\bar{w}$ are held fixed, Proposition 1 implies that not only does an equilibrium that is not ε-myopic exist (for ε given by (6)), but its per-capita expected cost remains bounded as $N \to \infty$, as this cost cannot exceed $\bar{w}$. This will contrast with the situation under a team contract, where we will see that the per-capita expected cost of motivating cooperation must explode as $N \to \infty$, unless the maximum reward $\bar{w}$ increases exponentially with N.

Proof. As in Example 1, consider an N-player prisoner's dilemma where the probability that a player's individual outcome is a *success* is $1-\underline{\pi}$ if she cooperates and $\underline{\pi}$ if she defects. Divide the N players into $C/\log 2$ teams, each of size $\lfloor (\log 2)\, N/C \rfloor$ or $\lceil (\log 2)\, N/C \rceil$. Suppose that, for each team, the signal Y reveals whether the number of successes in the team is at least $(1-\underline{\pi})(\log 2)\, N/C$ (the mean number of successes, modulo rounding): note that the channel capacity of this signal is at most $\log(|\mathcal{Y}|) = \log\left(2^{C/\log 2}\right) = C$. Now consider the contract that, for each team, pays everyone $\bar{w}$ if the number of successes is at least $(1-\underline{\pi})(\log 2)\, N/C$ and pays 0 otherwise. By the de Moivre–Laplace Theorem, the probability that a single success is pivotal is

$$\frac{1+o(1)}{\sqrt{2\pi\underline{\pi}(1-\underline{\pi})(\log 2)\, N/C}},$$

where $\pi = 3.14\ldots$ and $o(1)$ denotes a function that converges to 0 as $N/C \to \infty$. Thus, each player's "incentive strength" (the maximum static gain from defection that the player is willing to forgo), which equals the pivot probability multiplied by $(1-2\underline{\pi})\,\bar{w}$, is

$$\frac{1-2\underline{\pi}}{\sqrt{2\pi(1-\underline{\pi})(\log 2)}}\sqrt{\frac{C}{\underline{\pi}N}}\bar{w}(1+o(1)) = \frac{1}{\sqrt{4(\log 2)\pi(1-\underline{\pi})}}\sqrt{\kappa(\underline{\pi})\frac{C}{N}}\bar{w}(1+o(1)).$$

Since $\underline{\pi} > 0$ and $\sqrt{4(\log 2)\pi} \approx 2.95 < 3$, for sufficiently large N/C this exceeds $(1/3)\sqrt{\kappa(\underline{\pi})\, C/N}\bar{w}$.
■

3.2 Team Contracts

We now consider the implications of restricting attention to team contracts. A *team contract* is a public contract w where the players' rewards are co-linear: for each player $i \neq 1$, there exists a constant $b_i \in \mathbb{R}$ such that, for all signals y, y', we have $w_i(y') - w_i(y) = b_i(w_1(y') - w_1(y))$.[13] A public equilibrium (α, w) is a *team equilibrium* if w is a team contract. Team equilibria model collective incentive schemes, such as equity shares in a partnership, price wars in an oligopoly, or Hume's threat of "the abandoning of the whole project."

[13]We allow the possibility that some b_i are negative, which can be interpreted as dividing the agents are divided into two teams with opposing interests.

We show that, while team contracts can support cooperation under conditions similar to those in Theorem 1 and Proposition 1, the per-capita expected cost so incurred explodes with N, unless the maximum payment $\bar{w}$ is exponentially large relative to N.

Theorem 2 *Fix any $\underline{\pi} > 0$ and any sequence of multiagent moral hazard problems with N players, channel capacity C, $\underline{\pi}$-bounded noise, and $\bar{w}$-bounded rewards (where N, C, and $\bar{w}$ vary along the sequence, while $\underline{\pi}$ is fixed), with $N \to \infty$. The following hold:*

1. *If $N/\bar{w}^2 \to \infty$ then, for any $\varepsilon > 0$, eventually all team equilibria are ε-myopic.*[14]

2. *If there exists $\rho > 0$ such that* $\exp\left(N^{1-\rho}\right)/\bar{w} \to \infty$ *then, for any $\varepsilon > 0$ and $K > 0$, eventually all team equilibria are either ε-myopic or incur a per-capita expected cost greater than K.*

Unlike Theorem 1 and Proposition 1, Theorem 2 does not depend on channel capacity. Intuitively, the optimal team equilibrium takes a bang-bang form even when the realized outcome profile is perfectly observed, so a binary signal that indicates whether or not the maximum reward should be delivered is as effective as any more informative signal.

The point of Theorem 2 is that general ("personalized") incentives greatly outperform team incentives even if monitoring provides only a single bit of information, unless the maximum reward $\bar{w}$ is exponentially large relative to N. This point can be illustrated by comparing two simple incentive schemes in the context of the N-player prisoner's dilemma, where the probability that a player's individual outcome is a success is $1-\underline{\pi}$ if she cooperates and $\underline{\pi}$ if she defects.

First, consider *randomly monitoring one agent*: at the end of each period, a single random player is monitored, and is paid w if her outcome is a success and 0 if her outcome is a failure, while the other players are paid 0. Letting g denote the cost of cooperating, this scheme (which *is not* a team contract) supports cooperation if $(1-2\underline{\pi})\,w/N \geq g$, and each player's equilibrium expected reward is $(1-\underline{\pi})\,w/N$. Thus, whenever the maximum reward $\bar{w}$ exceeds $gN/(1-2\underline{\pi})$, randomly monitoring one agent can support cooperation at a constant per-capita expected cost of $g(1-\underline{\pi})/(1-2\underline{\pi})$, regardless of the number of agents N. Intuitively, the probability that a player's action is pivotal for her reward and

[14] Here and in all results stated in terms of sequences of games, "eventually" means "for games sufficiently far along the sequence."

the probability that she is rewarded are both of order $1/N$, so incentives can be provided at bounded cost whenever $\bar{w}$ is large relative to N.

Next, consider *collectively monitoring all agents:* there is a threshold number of successes n^* such that, at the end of each period, each player is paid w if the number of successes exceeds n^*, and everyone is paid 0 otherwise.[15] We argue that this scheme (which *is* a team contract) cannot support cooperation at a bounded per-capita expected cost as N increases, unless the maximum reward $\bar{w}$ is exponentially large relative to N.

Observe that, due to independent noise, the distribution of the number of successes n is approximately normal, with mean $(1-\underline{\pi})N$ and standard deviation $\sqrt{\underline{\pi}(1-\underline{\pi})N}$. Denote the z-score of the threshold number of successes n^* by $z^* = (n^* - (1-\underline{\pi})N)/\sqrt{\underline{\pi}(1-\underline{\pi})N}$, and let ϕ and Φ denote the standard normal pdf and cdf. Then each player is willing to cooperate under the collective monitoring scheme if and only if

$$\frac{(1-2\underline{\pi})\,\phi(z^*)\,w}{\sqrt{\underline{\pi}(1-\underline{\pi})N}} \geq g, \tag{7}$$

as the LHS of this inequality is the product of w and the probability that a player's action is pivotal for the event $\{n \geq n^*\}$, and the RHS is the cost of cooperating. At the same time, the contract's per-capita expected cost equals $(1-\Phi(z^*))\,w$. So, if the per-capita expected cost equals K, so that $w = K/(1-\Phi(z^*))$, we can rewrite (7) as

$$\frac{(1-2\underline{\pi})\,K}{\sqrt{\underline{\pi}(1-\underline{\pi})N}} \times \frac{\phi(z^*)}{1-\Phi(z^*)} \geq g. \tag{8}$$

Now, if (8) holds with K bounded as $N \to \infty$, the ratio $\phi(z^*)/(1-\Phi(z^*))$ must increase at least linearly with $\sqrt{N}$. However, this ratio is the standard normal *Mills ratio*, which is approximately equal to z^* when $z^* \gg 0$. Thus, z^* must increase at least linearly with $\sqrt{N}$. But, since $\phi(z^*)$ decreases exponentially with z^*, and thus must decrease exponentially with N, (7) then implies that w must increase exponentially with N. This argument establishes that a team contract can support cooperation at a bounded per-capita expected cost only if the maximum reward is exponentially large relative to N. Intuitively, the probability that a player's action is pivotal for the team reward is of order $\phi(z^*)/\sqrt{N}$ while the probability

[15] The analysis of such "tail tests" as optimal incentive contracts goes back to Mirrlees (1975). The proof of Theorem 2 implies that the size of the penalty in a Mirrleesian tail test must increase exponentially with the variance of the noise. We are not aware of a reference to this point in the literature.

that the team is rewarded is $1-\Phi(z^*)$, and the former is much smaller than the latter unless z^* is large, which is consistent with incentive compatibility only if $\bar{w}$ is exponentially large.

We also note a converse to Theorem 2: if the maximum reward is exponentially large in N, cooperation can be supported at a bounded per-capita expected cost by a team contract with threshold $n^* = N$. Under this contract, the agents are only paid in the unlikely event that *all* of them succeed.

4 Repeated Games with Many Players

We now turn to repeated games with independent noise. The setting is analogous to the static multiagent moral hazard problems considered above, but with equilibrium continuation payoffs replacing contracts.

4.1 Model

We describe our repeated game model and some associated concepts. A *repeated game with independent noise* $\Gamma = (I, \mathcal{A}, u, \mathcal{X}, \pi, \mathcal{Y}, q, \delta)$ is described by a stage game $(I, \mathcal{A}, u)$, a noise structure $(\mathcal{X}, \pi)$, an outcome monitoring structure $(\mathcal{Y}, q)$, and a discount factor $\delta \in [0, 1)$. In each period $t = 1, 2, \ldots$, (i) the players observe the outcome of a public randomizing device z_t drawn from the uniform distribution over $[0, 1]$, (ii) the players take actions a, (iii) the outcome x is drawn from distribution $\pi(\cdot|a)$, and (iv) the signal y is drawn from distribution $q(\cdot|x)$ and is publicly observed.[16] A *history* h_i^t for player i at the beginning of period t thus takes the form $h_i^t = \left((z_{t'}, a_{i,t'}, y_{t'})_{t'=1}^{t-1}, z_t\right)$. A *strategy* σ_i for player i maps histories h_i^t to distributions over actions $a_{i,t}$. A strategy σ_i is *public* if it depends on h_i^t only through the *public history* $h^t = \left((z_{t'}, y_{t'})_{t'=1}^{t-1}, z_t\right)$. A *Nash equilibrium* is a strategy profile where each player's strategy maximizes her discounted expected payoff. A *perfect public equilibrium* (PPE) is a profile of public strategies that, beginning at any period t and any public history h^t, forms a Nash equilibrium from that period on.

A *repeated game outcome* $\mu \in \Delta\left((\mathcal{A} \times \mathcal{X} \times \mathcal{Y})^\infty\right)$ (not to be confused with a single profile of individual outcomes x) is a distribution over infinite paths of actions, individual outcomes, and signals. Each strategy profile σ induces a unique outcome μ. In turn, each outcome

[16] It is natural to require that players' realized payoffs depend only on their own actions and the signal. However, this assumption is not necessary for our analysis.

μ induces a marginal distribution over period t action profiles $\alpha_t^\mu \in \Delta(\mathcal{A})$, as well as an *occupation measure* over action profiles, defined as

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

Intuitively, the occupation measure captures how the game is played "on average." Note that the players' payoffs are determined by the occupation measure, as

$$(1-\delta)\sum_t \delta^{t-1}\sum_a \alpha_t^\mu(a)\,u(a) = \sum_a (1-\delta)\sum_t \delta^{t-1}\alpha_t^\mu(a)\,u(a) = \sum_a \alpha^\mu(a)\,u(a) = u(\alpha^\mu).$$

4.2 Non-Cooperation in Repeated Games

The following theorem is the analogue of Theorem 1 for repeated games.

Theorem 3 *Any Nash equilibrium occupation measure in a repeated game with N players, channel capacity C, $\underline{\pi}$-bounded noise, and $\bar{u}$-bounded payoffs is ε-myopic (and hence any Nash equilibrium payoff vector is ε-myopic), for*

$$\varepsilon = \sqrt{\frac{2\kappa(\underline{\pi})}{\underline{\pi}}\frac{\delta}{1-\delta}\frac{C}{N}}\bar{u}.$$

In particular, for any fixed noise level $\underline{\pi}$, if the per-capita channel capacity C/N is much smaller than the discount rate $1-\delta$, then equilibrium play (i.e., the equilibrium occupation measure) is ε-myopic for small ε. Thus, cooperation in a repeated game with many players requires a large amount of information and/or low discounting.

Theorem 3 is analogous to Theorem 1 with a maximum reward of $\bar{w} = (1-\delta)^{-1/2}\,\bar{u}$. This result may be counterintuitive, as continuation payoffs in a repeated game are weighted by $(1-\delta)^{-1}$, not $(1-\delta)^{-1/2}$. However, with imperfect monitoring it is impossible to make each period fully responsible for determining continuation play for the rest of the game, so an average "incentive strength" of $(1-\delta)^{-1}$ cannot be attained. This last point is formalized by Theorem 1 of SW, which uses a recursive variance decomposition argument to establish an incentive strength bound of order $(1-\delta)^{-1/2}$.[17] In turn, Theorem 3 follows easily from

[17] An intuition for Theorem 1 of SW (which holds even with private monitoring) can be seen by considering a strategy where an agent's performance is reviewed every T periods, with continuation play determined by the outcome of the review. For the agent to put weight independent of δ on the outcome of the review,

Theorem 1 of SW and Lemmas 1 and 2 of the current paper.[18]

Proof of Theorem 3. By Theorem 1 of SW, for any Nash equilibrium outcome μ, any player i, and any manipulation s_i, we have

$$g_i\left(s_i, \alpha^{\mu}\right) \leq \sqrt{\frac{\delta}{1-\delta} \sum_a \alpha^{\mu}(a) \chi^2\left(p\left(\cdot | s_i\left(a_i\right), a_{-i}\right) || p\left(\cdot | a\right)\right) \bar{u}}.$$

Hence, by Lemma 1,

$$g_i\left(s_i, \alpha^{\mu}\right) \leq \sqrt{\frac{2\kappa\left(\underline{\pi}\right)}{\underline{\pi}} \frac{\delta}{1-\delta} \sum_a \alpha^{\mu}(a) I^{\pi(\cdot|a)}\left(X_i; Y\right) \bar{u}}.$$

Theorem 3 follows, as we have

$$\begin{aligned}
\frac{1}{N} \sum_i \bar{g}_i(\alpha) &\leq \frac{1}{N} \sum_i \sqrt{\frac{2\kappa\left(\underline{\pi}\right)}{\underline{\pi}} \frac{\delta}{1-\delta} \sum_a \alpha^{\mu}(a) I^{\pi(\cdot|a)}\left(X_i; Y\right) \bar{u}} \\
&\leq \sqrt{\frac{1}{N} \sum_i \frac{2\kappa\left(\underline{\pi}\right)}{\underline{\pi}} \frac{\delta}{1-\delta} \sum_a \alpha^{\mu}(a) I^{\pi(\cdot|a)}\left(X_i; Y\right) \bar{u}} \\
&= \sqrt{\frac{2\kappa\left(\underline{\pi}\right)}{\underline{\pi}} \frac{\delta}{1-\delta} \frac{1}{N} \sum_a \alpha^{\mu}(a) \sum_i I^{\pi(\cdot|a)}\left(X_i; Y\right) \bar{u}} \\
&\leq \sqrt{\frac{2\kappa\left(\underline{\pi}\right)}{\underline{\pi}} \frac{\delta}{1-\delta} \frac{C}{N} \bar{u}},
\end{aligned}$$

where the second inequality is by Jensen and the third is by Lemma 2. ■

Theorem 3 also holds for repeated games with private monitoring. Indeed, the same result holds for the *blind repeated game*, where in each period the signal y is observed only by a mediator (rather than being directly observed by the players themselves), who then privately recommends actions to the players. This shows that Theorem 3 depends only on the precision of the signal y (measured by channel capacity), not how information about the

---

a review must occur every $O\left((1-\delta)^{-1}\right)$ periods, so the standard deviation of the count of each signal realization over the course of the review is $O\left((1-\delta)^{-1/2}\right)$, and hence the probability that a single signal is pivotal for the review is $O\left((1-\delta)^{1/2}\right)$. Since the gain from deviating in a single period is $O(1-\delta)$, the agent's "incentive strength" is $O\left((1-\delta)^{1/2} / (1-\delta)\right) = O\left((1-\delta)^{-1/2}\right)$.

[18] A subtle difference between the proofs of Theorems 1 and 3 is that the former uses TV distance, while the latter uses χ^2-divergence. TV distance gives a sharper bound in static games, while χ^2-divergence is required for SW's recursive approach, which gives a sharper bound in repeated games.

signal is distributed across the players.[19]

In large groups, the necessary condition for cooperation implied by Theorem 3—that $(1-\delta)\, N/C$ is not too large—is easier to satisfy in some classes of repeated games than in others. For example, if the space of possible signal realizations $\mathcal{Y}$ is fixed independently of N, then, since $C \leq \log |\mathcal{Y}|$, the necessary condition implies that $(1-\delta)^{-1}$ must be at least proportional to N, which is a restrictive condition. This negative conclusion applies for traditional applications of repeated games with public monitoring where the signal space is fixed independent of N, such as when the public signal is the market price facing Cournot competitors, the level of exploitation of a common-pool resource, the output of team production, or some other aggregate statistic.

However, in other settings C scales linearly with N, so $(1-\delta)\, N/C$ is small whenever players are patient—regardless of group size. In repeated games with random matching (Kandori, 1992; Ellison, 1994; Deb, Sugaya, and Wolitzky, 2020), the players match in pairs each period, and each player observes her partner's action. In these games, $C = N \log |\mathcal{A}_i|$, so per-capita channel capacity is independent of N. Intuitively, in a random matching game each player gets a distinct signal of the overall action profile, so the total amount of information available to society is proportional to N. Similarly, channel capacity scales linearly with N in public monitoring games where the public signal is a vector that includes a distinct signal of each player's action, as in the ratings systems used by online platforms like eBay and AirBnB. In general, C/N can be expected to be roughly independent of the population size in settings where players are monitored "separately," rather than being monitored jointly through an aggregate statistic.

Remark 1 *In applications like Cournot competition, resource exploitation, or team production, the signal space may be modeled as a continuum, in which case the cardinality bound* $C \leq \log |\mathcal{Y}|$ *is vacuous. However, our results extend to the case where* $\mathcal{Y}$ *is a compact metric space and there exists another compact metric space* $\mathcal{Z}$ *and a function* $f^N : \mathcal{X}^N \to \mathcal{Z}$ *(which can vary with* N*) such that the signal distribution admits a conditional density of the form* $q_{\mathcal{Y}|\mathcal{Z}}(y|z)$*, where* $\mathcal{Y}$*,* $\mathcal{Z}$*, and* $q_{\mathcal{Y}|\mathcal{Z}}$ *are fixed independent of* N*. (For example, in Cournot competition* z *is industry output and* y *is the market price, which depends on* z *and a noise*

[19] For more on blind games, see SW.

term with variance fixed independent of N.) In this case,

$$C = \max_{\xi \in \vartheta} \int_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} \xi (x) \, q_{\mathcal{Y}|\mathcal{Z}} \left(y | f^N (x) \right) \log \left(\frac{q_{\mathcal{Y}|\mathcal{Z}} \left(y | f^N (x) \right)}{\sum_{x' \in \mathcal{X}} \xi (x') \, q_{\mathcal{Y}|\mathcal{Z}} \left(y | f^N (x') \right)} \right),$$

which is bounded by

$$\bar{C} = \max_{q_{\mathcal{Z}} \in \Delta(\mathcal{Z})} \int_{y \in \mathcal{Y}} \int_{z \in \mathcal{Z}} q_{\mathcal{Z}} (z) \, q_{\mathcal{Y}|\mathcal{Z}} (y|z) \log \left(\frac{q_{\mathcal{Y}|\mathcal{Z}} (y|z)}{\int_{z' \in \mathcal{Z}} q_{\mathcal{Z}} (z') \, q_{\mathcal{Y}|\mathcal{Z}} (y|z')} \right).$$

Since $\bar{C}$ is independent of N, it follows that C is bounded independent of N.

Theorem 3 can be compared to prior results by Fudenberg, Levine, and Pesendorfer (1998), al-Najjar and Smorodinsky (2000, 2001), and Pai, Roth, and Ullman (2014), which establish anti-folk theorems as $N \to \infty$ for fixed δ. When N and δ vary together, the arguments in these papers can be used to show that cooperation is impossible if $(1 - \delta)^2 N \to \infty$. The same result can also be obtained by directly applying Theorem 1 with $\bar{w} = (1 - \delta)^{-1} \bar{u}$, as continuation payoffs in a repeated game are weighted by $(1 - \delta)^{-1}$. In contrast, Theorem 3 gives the qualitatively stronger result that cooperation is impossible if $(1 - \delta) N \to \infty$. The improvement comes from applying Theorem 1 of SW, which bounds incentive strength by a multiple of $(1 - \delta)^{-1/2}$, rather than $(1 - \delta)^{-1}$.

Indeed, the bound in Theorem 3 is essentially the best possible. To show this, in the online appendix we establish a folk theorem in a setting where the discount factor, monitoring structure, and stage game (including the number of players N) all vary simultaneously. The theorem has the following corollary.

Corollary 1 *For any $\varepsilon > 0$, there exists a sequence of repeated games with public monitoring with N players, channel capacity C, $\underline{\pi}$-bounded noise, and $\bar{u}$-bounded payoffs (where N, δ, and C vary along the sequence, while $\underline{\pi}$ and $\bar{u}$ are fixed), satisfying $N \to \infty$ and $(1 - \delta) N \log (N) / C \to 0$, and a corresponding sequence of Nash equilibria that, eventually, are not ε-myopic.*

As Theorem 3 shows that, for any $\varepsilon > 0$, all equilibria are eventually ε-myopic along any sequence of repeated games where $(1 - \delta) N/C \to \infty$, Corollary 1 implies that Theorem 3 is tight up to a factor of $\log N$.[20]

4.3 Team Equilibria

Finally, we consider the consequences of restricting to collective incentives in repeated games. A *team equilibrium* is a PPE where the players' continuation payoffs at all public histories are co-linear: for each player $i \neq 1$, there exists a constant $b_i \in \mathbb{R}$ such that, for all public histories h, h', we have $\omega_i(h') - \omega_i(h) = b_i(\omega_1(h') - \omega_1(h))$, where $\omega_i(h)$ denotes player i's equilibrium continuation payoff at history h. This notion generalizes that of *strongly symmetric equilibrium* (SSE) in symmetric games, where $b_i = 1$ for all i.

Our result for team equilibria is as follows.

Theorem 4 *Fix any $\underline{\pi} > 0$ and any sequence of repeated games with public monitoring with N players, channel capacity C, $\underline{\pi}$-bounded noise, and $\bar{u}$-bounded rewards (where N, δ, C, and $\bar{u}$ vary along the sequence, while $\underline{\pi}$ is fixed), with $N \to \infty$. If there exists $\rho > 0$ such that $(1-\delta)\exp(N^{1-\rho})/\bar{u} \to \infty$ then, for any $\varepsilon > 0$, eventually all team equilibrium payoff vectors are ε-myopic.*

Theorem 4 is the repeated game analogue of Theorem 2. The main difference between these results is that unbounded per-capita expected rewards cannot be delivered as continuation payoffs in a repeated game. Hence, whereas Theorem 2 allows the possibility that cooperation can be motivated at an unbounded per-capita cost in static moral hazard problems where $\exp(N^{1-\rho})/\bar{w} \to \infty$, Theorem 4 shows that cooperation cannot be motivated at all in repeated games where $(1-\delta)\exp(N^{1-\rho})/\bar{u} \to \infty$. To see the logic, note that player i's continuation payoff at history h in a repeated game is given by

$$\omega_i(h) = (1-\delta)\, u_i(a(h)) + \delta \mathbb{E}[\omega_i(h, y)],$$

where $a(h)$ is the equilibrium action profile at period h (assumed to be pure to ease notation) and $\omega_i(h, y)$ is player i's next-period continuation payoff following signal realization y. The

[20] While the relationship among N, δ, and C in Theorem 3 is tight, the extra $\underline{\pi}^{-1}$ term in Theorem 3 (as compared to Theorem 1) is unnecessary: it can be shown that Theorem 3 remains valid when $\kappa(\underline{\pi})/\underline{\pi}$ is replaced by $2\kappa(\underline{\pi})$. The proof of this stronger version of Theorem 3 is more intricate, and we omit it.

analogue of player i's static reward upon signal realization y is then

$$W_i(y) := \frac{\delta}{1-\delta}\left(\omega_i(h,y) - \omega_i(h)\right),$$

as with this definition player i chooses her current-period action a_i to maximize

$$u_i(a_i, a_{-i}(h)) + \mathbb{E}[W_i(y)] = u_i(a_i, a_{-i}(h)) + \frac{\delta}{1-\delta}\mathbb{E}[\omega_i(h,y)] - \frac{\delta}{1-\delta}\omega_i(h),$$

where the last term is a constant independent of a_i. But note that, since $\omega_i(h)$ is a feasible stage game payoff, we have

$$\begin{aligned}\mathbb{E}[W_i(y)] &= \frac{\delta}{1-\delta}\left(\mathbb{E}[\omega_i(h,y)] - \omega_i(h)\right) \\ &= \omega_i(h) - u_i(\alpha(h)) \in [-\bar{u}, \bar{u}].\end{aligned}$$

Thus, expected rewards in a repeated game are bounded by the range of the stage game payoffs, independent of δ. This observation explains the differing conclusions of Theorems 2 and 4.

Theorem 4 is related to Proposition 1 of Sannikov and Skrzypacz (2007), which is an anti-folk theorem for SSE in a two-player repeated game where actions are observed with additive, normally distributed noise, with variance proportional to $(1-\delta)^{-1}$.[21] As a tail test is optimal in their setting, the proof of Theorem 4 implies that non-vanishing incentives can be provided only if $(1-\delta)^{-1}$ increases exponentially with the variance of the noise. Since in their model $(1-\delta)^{-1}$ increases with variance only linearly, they likewise obtain an anti-folk theorem. Similarly, Proposition 2 of Fudenberg and Levine (2007) is an anti-folk theorem in a game with one patient player and a myopic opponent, where the patient player's action is observed with additive, normal noise, with variance proportional to $(1-\delta)^{-\rho}$ for some $\rho > 0$; and their Proposition 3 is a folk theorem when the variance is constant in δ. Theorem 4 suggest that their anti-folk theorem extends whenever variance asymptotically dominates $(-\log(1-\delta))^{1/(1-\rho)}$ for some $\rho > 0$, while their folk theorem extends whenever variance is asymptotically dominated by $(-\log(1-\delta))^{1/(1+\rho)}$ for some $\rho > 0$.

[21] Their interpretation is that the players change their actions every Δ units of time, where $\delta = e^{-r\Delta}$ for fixed $r > 0$ and variance is inversely proportional to Δ, for example as a consequence of observing the increments of a Brownian process.

5 Conclusion

This paper has developed a theory of large-group cooperation in moral hazard problems and repeated games. Our key assumption is that monitoring is imperfect and respects a degree of independence across players. Our main results establish necessary and sufficient conditions for cooperation in terms of the ratio of the squared inverse of the maximum feasible reward (in static problems) or the discount rate (in repeated games) and the *per-capita channel capacity* of the monitoring structure. We also show that cooperation in a *team equilibrium*, where the players' rewards are co-linear, is possible only under much more stringent conditions. This result demonstrates a sense in which large-group cooperation must rely on personalized sanctions. Notably, this result holds even when information is scarce, so that precisely monitoring all players is infeasible.

Our results raise several questions for future theoretical and applied research. On the theory side, this paper has focused on insufficient monitoring precision as an obstacle to large-group cooperation. In reality, insufficient precision coexists with other obstacles to cooperation, such as decentralized monitoring (as in community enforcement models) and the possibility that a small fraction of players may be irrational or fail to understand the equilibrium being played (as in Sugaya and Wolitzky 2020, 2021). Combining these features may help develop a richer and more realistic perspective on the determinants of large-group cooperation. We also believe it could be interesting to explore the implications of independent noise and limited monitoring precision for organizational design, for example the design of large hierarchies. Finally, another open question is whether some version of our results survives under an appropriate relaxation of independent noise.

As for applied work, more systematic empirical or experimental evidence on the determinants of large-group cooperation under imperfect monitoring would be valuable.[22] For example, a novel prediction of our paper is that personalized incentive contracts are much more cost-effective than team contracts in large groups, even when the total amount of available information about the agents' performance is small. It would be interesting to test this prediction.

[22] Camera and Casari (2009) and Duffy and Ochs (2009), among others, run experiments on repeated games with random matching and private monitoring, i.e., community enforcement. Community enforcement raises additional issues beyond the ones we focus on, which arise even under public monitoring. Camera, Casari, and Bigoni (2013) include a treatment with public monitoring, and find that larger groups cooperate less.

Appendix

A How Large is $V(\varepsilon)$?

Theorems 1 and 3 give conditions under which all equilibrium payoffs lie in the set

$$V(\varepsilon) = \left\{ v \in \mathbb{R}^N : v = u(\alpha) \text{ for some } \alpha \text{ such that } \frac{1}{N} \sum_i \bar{g}_i(\alpha) \leq \varepsilon \right\}.$$

Payoffs in $V(\varepsilon)$ are attained by action distributions where the per-player average deviation gain is less than ε; however, a few players can have large deviation gains. A more standard notion of "ε-myopic play" requires that *all* players' deviations gains are less than ε. The corresponding payoff vectors are the *static ε-correlated equilibrium payoffs*, given by

$$CE(\varepsilon) = \left\{ v \in \mathbb{R}^N : v = u(\alpha) \text{ for some } \alpha \text{ such that } \bar{g}_i(\alpha) \leq \varepsilon \text{ for all } i \right\}.$$

Here we compare the sets $V(\varepsilon)$ and $CE(\varepsilon)$. We first give an example where $V(\varepsilon)$ and $CE(\varepsilon)$ are very different (and $V(\varepsilon)$ cannot be replaced by $CE(\varepsilon)$ in Theorems 1 and 3). We then give a condition under which maximum per-capita utilitarian welfare $\sum_i v_i/N$ in $V(\varepsilon)$ is little greater than that in $CE(c\sqrt{\varepsilon})$, for a constant c. Intuitively, $V(\varepsilon)$ and $CE(\varepsilon)$ can be very different if incentive constraints bind for only a few players and these players' actions have large effects on others' payoffs; while maximum utilitarian welfare in $V(\varepsilon)$ and $CE(c\sqrt{\varepsilon})$ is similar if each player's action has only a small effect on every opponent's payoff.

For an example where $V(\varepsilon)$ and $CE(\varepsilon)$ differ, consider a "product choice" game where player 1 is a seller who chooses high or low quality (H or L), and the other $N-1$ players are buyers who choose whether to buy or not (B or D). If the seller takes $a_1 \in \{H, L\}$ and a buyer i takes $a_i \in \{B, D\}$, this buyer's payoff is given by

$$\mathbf{1}\{a_i = B\}(-1 + 2 \times \mathbf{1}\{a_1 = H\}),$$

while the seller's payoff is given by

$$\frac{2k}{N} - \mathbf{1}\{a_1 = H\},$$

where $k \in \{0, 1, \ldots, N\}$ is the number of buyers who take B. Suppose also that the players tremble with independent, uniform noise $\underline{\pi} \in (0, 1/3)$. Note that in this game the payoff range is bounded by 3 and noise is bounded by $\underline{\pi}$.

In this game, for any $\varepsilon > 0$, when N is sufficiently large, we have $(H, B, \ldots, B) \in \mathcal{A}(\varepsilon)$, and hence $(1, 1, \ldots, 1) \in V(\varepsilon)$. This follows because the per-player average deviation gain at action profile $(H, B, \ldots, B)$ equals $1/N$: the seller has a deviation gain of 1, while each buyer has a deviation gain of 0. Thus, Theorems 1 and 3 do not preclude $(1, 1, \ldots, 1)$ (or any convex combination of $(1, 1, \ldots, 1)$ and $(0, 0, \ldots, 0)$) as an equilibrium payoff vector, even when $(1 - \delta) N/C$ is very large. This is reassuring, because the monitoring structure given by perfect monitoring of the seller's realized action (i.e., $\mathcal{Y} = \{H, L\}$, $q(y|x) = \mathbf{1}\{y = x_1\}$) has channel capacity $\log 2$ and supports the payoff vector $(1, \ldots, 1)$ for all $\bar{w} \geq 1/(1 - 2\underline{\pi})$ and all $N \geq 2$ in static moral hazard; and supports the payoff vector $((1 - 3\underline{\pi})/(1 - 2\underline{\pi}), \ldots (1 - 3\underline{\pi})/(1 - 2\underline{\pi}))$ for all $\delta \geq 1/(2 - 3\underline{\pi})$ and all $N \geq 2$ in repeated games.[23] In contrast, the greatest symmetric payoff vector in $CE(\varepsilon)$ is $(\varepsilon, \varepsilon, \ldots, \varepsilon)$, because the seller's deviation gain equals the probability that she takes H.

Intuitively, even though the efficient action profile $(H, B, \ldots, B)$ is not a static ε-correlated equilibrium, it can be supported with "not very informative" monitoring. The reason is that only the seller is tempted to deviate at the efficient action profile, so monitoring one player suffices to support this action profile regardless of the number of buyers.

Next, for any $d \in (0, \bar{u})$, say that *per-capita externalities are bounded by* d if $\left|u_i\left(a_j', a_{-j}\right) - u_i(a)\right| \leq d/N$ for all $i \neq j, a_j', a$. For example, in a repeated random matching game, d can be taken as the maximum impact of a player's action on her partner's payoff, which is independent of N. In contrast, in the product choice game, per-capita externalities cannot be bounded uniformly in N, because the seller exerts an externality of 2 on each buyer who purchases.

In games with bounded per-capita externalities, any level of per-capita utilitarian welfare that is atttainable in $V(\varepsilon)$ can also be approximated in $CE\left(\sqrt{8d\varepsilon}\right)$.

Proposition 2 *Assume that per-capita externalities are bounded by* d. *Then, for any* $\varepsilon \in$

[23]This is a standard calculation, which results from considering "tolerant trigger strategies" that prescribe Nash reversion with probability ϕ when $y = L$. The smallest value of ϕ that induces the seller to take H is given by $\phi = (1 - \delta)/(\delta - 3\delta\pi)$, and substituting this into the value recursion $v = (1 - \delta)(1) + \delta(1 - \pi\phi)v$ yields $v = (1 - 3\underline{\pi})/(1 - 2\underline{\pi})$.

$(0, 2d)$ *and any* $v \in V(\varepsilon)$, *there exists* $v' \in CE\left(\sqrt{8d\varepsilon}\right)$ *such that*

$$\left|\frac{1}{N}\sum_i (v_i - v_i')\right| \leq \sqrt{\frac{2\varepsilon}{d}}\bar{u}.$$

Proof. We establish the stronger conclusion that, for any $v \in V(\varepsilon)$ and any $c \geq \sqrt{8d/\varepsilon}$, there exists $v' \in CE(c\varepsilon)$ such that $\left|\sum_i (v_i - v_i')/N\right| \leq 4\bar{u}/c$. (The stated conclusion follows by taking $c = \sqrt{8d/\varepsilon}$.) Fix $\varepsilon \in (0, d)$ and $\alpha \in \mathcal{A}(\varepsilon)$. Let $J = \{i : \bar{g}_i(\alpha) > c\varepsilon/2\}$, and note that $|J| \leq 2N/c$. Let $\tilde{\alpha} \in \Delta(\mathcal{A})$ be an action distribution that has the same marginal on $\mathcal{A}_{I\backslash J}$ as α and that satisfies $\bar{g}_i(\tilde{\alpha}) \leq c\varepsilon$ for all $i \in J$: for example, take a Nash equilibrium in the game among the players in J where the action distribution among the players in $I\backslash J$ is held fixed. Since $\left|u_i\left(a_j', a_{-j}\right) - u_i(a)\right| \leq d/N$ for all $i \neq j, a_j', a$, and the actions of at most $2N/c$ players differ between $\tilde{\alpha}$ and α, we have $\bar{g}_i(\tilde{\alpha}) \leq \bar{g}_i(\alpha) + 4d/c$ for each $i \in I\backslash J$. Since $\bar{g}_i(\alpha) \leq c\varepsilon/2$ (as $i \in I\backslash J$) and $4d/c \leq c\varepsilon/2$ (as $c \geq \sqrt{8d/\varepsilon}$), we have $\bar{g}_i(\tilde{\alpha}) \leq c\varepsilon$. Since we also assumed that $\bar{g}_i(\tilde{\alpha}) \leq c\varepsilon$ for all $i \in J$, we have $\bar{g}_i(\tilde{\alpha}) \leq c\varepsilon$ for all $i \in I$, and hence $u(\tilde{\alpha}) \in CE(c\varepsilon)$. Finally, since the actions of at most $2N/c$ players differ between $\tilde{\alpha}$ and α, we have $|u_i(\tilde{\alpha}) - u_i(\alpha)| \leq 2d/c \leq 2\bar{u}/c$ for all $i \in I\backslash J$, and by definition of $\bar{u}$ we have $|u_i(\tilde{\alpha}) - u_i(\alpha)| \leq \bar{u}$ for all $i \in J$. Since $c > 2$ (as $\varepsilon < 2d$) and $|J| \leq 2N/c$, we have $\left|\sum_{i \in I}(u_i(\tilde{\alpha}) - u_i(\alpha))\right| \leq (N - 2N/c)\, 2\bar{u}/c + (2N/c)\,\bar{u} \leq 4N\bar{u}/c$. ■

B Proof of Lemma 1

We first show that

$$\mathrm{TV}\left(p\left(\cdot|a_i', a_{-i}\right)||p\left(\cdot|a\right)\right)^2 \leq \kappa\left(\underline{\pi}\right) I^{\pi(\cdot|a)}\left(X_i; Y\right). \tag{9}$$

Let $\mathcal{Y}^+ = \{y : p(y|a) \geq p(y|a_i', a_{-i})\}$, so that $\text{TV}(p(\cdot|a_i', a_{-i})||p(\cdot|a)) = p(\mathcal{Y}^+|a) - p(\mathcal{Y}^+|a_i', a_{-i})$, and let $\mathcal{X}_i^+ = \{x_i : \pi_i(x_i|a_i) \geq \pi_i(x_i|a_i')\}$. For any a, i, and a_i', we have

$$
\begin{aligned}
& \text{TV}\left(p\left(\cdot|a_i', a_{-i}\right)||p\left(\cdot|a\right)\right) \\
= \ & \sum_{x_i} \left(\pi_i\left(x_i|a_i\right) - \pi_i\left(x_i|a_i'\right)\right) \varphi^a\left(\mathcal{Y}^+|x_i\right) \\
= \ & \sum_{x_i} \left(\pi_i\left(x_i|a_i\right) - \pi_i\left(x_i|a_i'\right)\right) \left(\varphi^a\left(\mathcal{Y}^+|x_i\right) - \varphi^a\left(\mathcal{Y}^+\right)\right) \\
\leq \ & \left(\pi_i\left(\mathcal{X}_i^+|a_i\right) - \pi_i\left(\mathcal{X}_i^+|a_i'\right)\right) \left(\max_{x_i}\left(\varphi^a\left(\mathcal{Y}^+|x_i\right) - \varphi^a\left(\mathcal{Y}^+\right)\right) - \min_{x_i}\left(\varphi^a\left(\mathcal{Y}^+|x_i\right) - \varphi^a\left(\mathcal{Y}^+\right)\right)\right) \\
\leq \ & \left(1 - 2\underline{\pi}\right) \left(\max_{x_i}\left(\varphi^a\left(\mathcal{Y}^+|x_i\right) - \varphi^a\left(\mathcal{Y}^+\right)\right) - \min_{x_i}\left(\varphi^a\left(\mathcal{Y}^+|x_i\right) - \varphi^a\left(\mathcal{Y}^+\right)\right)\right), \qquad (10)
\end{aligned}
$$

where first equality holds because $\varphi^a\left(\mathcal{Y}^+|x_i\right) = \varphi^{a_i', a_{-i}}\left(\mathcal{Y}^+|x_i\right)$, and the second inequality holds because $\pi_i\left(\mathcal{X}_i^+|a_i\right) \leq 1 - \underline{\pi}$ and $\pi_i\left(\mathcal{X}_i^+|a_i'\right) \geq \underline{\pi}$. Next, for any x_i, we have

$$
\begin{aligned}
\varphi^a\left(\mathcal{Y}^+|x_i\right) - \varphi^a\left(\mathcal{Y}^+\right) & \leq \sqrt{\frac{1}{2}\sum_y \varphi^a\left(y|x_i\right)\log\left(\frac{\varphi^a\left(y|x_i\right)}{\varphi^a\left(y\right)}\right)} \\
& = \sqrt{\frac{1}{2\pi_i\left(x_i|a_i\right)}\sum_y \varphi^a\left(x_i, y\right)\log\left(\frac{\varphi^a\left(x_i, y\right)}{\varphi^a\left(x_i\right)\varphi^a\left(y\right)}\right)} \\
& \leq \sqrt{\frac{1}{2\underline{\pi}}\sum_{x_i', y} \varphi^a\left(x_i', y\right)\log\left(\frac{\varphi^a\left(x_i', y\right)}{\varphi^a\left(x_i'\right)\varphi^a\left(y\right)}\right)} = \sqrt{\frac{I^{\pi(\cdot|a)}\left(X_i; Y\right)}{2\underline{\pi}}},
\end{aligned}
$$

where the first inequality is by Pinsker; the second inequality holds because, for each x_i', $\pi_i\left(x_i'|a_i\right) \geq \underline{\pi}$ and

$$
\sum_y \varphi^a\left(x_i', y\right)\log\left(\frac{\varphi^a\left(x_i', y\right)}{\varphi^a\left(x_i'\right)\varphi^a\left(y\right)}\right) = \frac{1}{\varphi^a\left(x_i'\right)}\sum_y \varphi^a\left(y|x_i'\right)\log\left(\frac{\varphi^a\left(y|x_i'\right)}{\varphi^a\left(y\right)}\right) \geq 0,
$$

by the information inequality (CT, Theorem 2.6.3); and the last equality is by the definition of mutual information. Hence, $\max_{x_i}\left(\varphi^a\left(\mathcal{Y}^+|x_i\right) - \varphi^a\left(\mathcal{Y}^+\right)\right) \leq \sqrt{I^{\pi(\cdot|a)}\left(X_i; Y\right)/2\underline{\pi}}$, and, similarly, $\min_{x_i}\left(\varphi^a\left(\mathcal{Y}^+|x_i\right) - \varphi^a\left(\mathcal{Y}^+\right)\right) \geq -\sqrt{I^{\pi(\cdot|a)}\left(X_i; Y\right)/2\underline{\pi}}$. Substituting these inequalities in (10) and squaring both sides establishes inequality (9).

We next show that

$$\chi^2\left(p\left(\cdot|a_i',a_{-i}\right)||p\left(\cdot|a\right)\right) \le \frac{2\kappa\left(\underline{\pi}\right)}{\underline{\pi}} I^{\pi(\cdot|a)}\left(X_i;Y\right).$$

For any $x_i \in \mathcal{X}_i$ and $y \in \mathcal{Y}$, we have $\varphi^a\left(x_i,y\right) = \pi_i\left(x_i|a_i\right)\varphi^a\left(y|x_i\right) = p\left(y|a\right)\varphi^a\left(x_i|y\right)$. Hence, since $\pi_i\left(x_i|a_i\right) \ge \underline{\pi}$, we have

$$\left(\varphi^a\left(y|x_i\right) - p\left(y|a\right)\right)^2 = \left(\frac{p\left(y|a\right)}{\pi_i\left(x_i|a_i\right)}\left(\varphi^a\left(x_i|y\right) - \pi_i\left(x_i|a_i\right)\right)\right)^2 \le \left(\frac{p\left(y|a\right)}{\underline{\pi}}\left(\varphi^a\left(x_i|y\right) - \pi_i\left(x_i|a_i\right)\right)\right)^2. \tag{11}$$

Now, for any a, i, and a_i', we have

$$\begin{aligned}
\chi^2\left(p\left(\cdot|a_i',a_{-i}\right)||p\left(\cdot|a\right)\right) &= \sum_y \frac{\left(\sum_{x_i}\left(\pi_i\left(x_i|a_i\right) - \pi\left(x_i|a_i'\right)\right)\varphi^a\left(y|x_i\right)\right)^2}{p\left(y|a\right)} \\
&= \sum_y \frac{\left(\sum_{x_i}\left(\pi_i\left(x_i|a_i\right) - \pi_i\left(x_i|a_i'\right)\right)\left(\varphi^a\left(y|x_i\right) - p\left(y|a\right)\right)\right)^2}{p\left(y|a\right)} \\
&\le \sum_{x_i}\left(\pi_i\left(x_i|a_i\right) - \pi_i\left(x_i|a_i'\right)\right)^2 \sum_y \frac{\sum_{x_i}\left(\varphi^a\left(y|x_i\right) - p\left(y|a\right)\right)^2}{p\left(y|a\right)} \\
&\le \frac{2\left(1-2\underline{\pi}\right)^2}{\underline{\pi}^2}\sum_y p\left(y|a\right)\sum_{x_i}\left(\varphi^a\left(x_i|y\right) - \pi_i\left(x_i|a_i\right)\right)^2, \\
&\le \frac{4\left(1-2\underline{\pi}\right)^2}{\underline{\pi}^2}\sum_y p\left(y|a\right)\sum_{x_i}\varphi^a\left(x_i|y\right)\log\left(\frac{\varphi^a\left(x_i|y\right)}{\pi_i\left(x_i|a_i\right)}\right) \\
&= \frac{2\kappa\left(\underline{\pi}\right)}{\underline{\pi}} I^{\pi(\cdot|a)}\left(X_i;Y\right),
\end{aligned}$$

where the first inequality follows by Cauchy-Schwarz; the second follows by (11) and

$$\sum_{x_i}\left(\pi_i\left(x_i|a_i\right) - \pi_i\left(x_i|a_i'\right)\right)^2 \le \begin{array}{l}\left(\pi_i\left(\mathcal{X}^+|a_i\right) - \pi_i\left(\mathcal{X}^+|a_i'\right)\right)^2 \\ +\left(\pi_i\left(\mathcal{X}\backslash\mathcal{X}^+|a_i\right) - \pi_i\left(\mathcal{X}\backslash\mathcal{X}^+|a_i'\right)\right)^2\end{array} \le 2\left(1-2\underline{\pi}\right)^2;$$

and the third follows by Pinsker.

C Proof of Theorem 2

Fix a team equilibrium (α, w) with coefficients $(b_1, b_2, \ldots, b_N)$, where (without loss) $|b_i| \leq 1$ for all i. Let $w(y) = w_1(y)$. For any player i with $b_i \geq 0$ and any manipulation s_i, we have

$$\begin{aligned} g_i(s_i, \alpha) &\leq \sum_r \alpha(r) \sum_y \left(p(y|r) - p(y|s_i(r_i), r_{-i})\right) b_i w(y) \\ &\leq \sum_r \alpha(r) \sum_y \max_{a_i} \left(p(y|r) - p(y|a_i, r_{-i})\right) w(y). \end{aligned}$$

Since a symmetric inequality holds for players with $b_i < 0$, and $w(y) \in [0, \bar{w}]$ for all y, we see that $\sum_i \bar{g}_i(\alpha)/N$ is bounded by the solution to the following program, which is parameterized by N, $\bar{w}$, and $\underline{\pi}$:

$$\max_{(\mathcal{X},\pi),(\mathcal{Y},q),r,a,w} \frac{2}{N} \sum_i \sum_y \left(p(y|r) - p(y|a_i, r_{-i})\right) w(y) \quad \text{s.t.}$$

$$p(y|\tilde{a}) = \sum_x \left(\prod_i \pi_i(x_i|\tilde{a}_i)\right) q(y|x) \quad \text{for all } y, \tilde{a},$$

$$\pi_i(x_i|\tilde{a}_i) \geq \underline{\pi} \quad \text{for all } i, x_i, \tilde{a}_i,$$

$$w(y) \in [0, \bar{w}] \quad \text{for all } y.$$

To prove the first statement in the theorem, we show that, for any $\underline{\pi} > 0$, the value of this program converges to 0 along any sequence $(N, \bar{w})$ where $\min\{N, N/\bar{w}^2\} \to \infty$. To prove the second statement, we show that, for any $\underline{\pi} > 0$ and $\rho > 0$, the value of this program with the additional constraint $\sum_y p(y|r) w(y) \leq K$ converges to 0 along any sequence $(N, \bar{w})$ where $\min\{N, \exp(N^{1-\rho})/\bar{w}\} \to \infty$ (and, hence, any team equilibrium that incurs a per-capita expected cost of at most K must be ε-myopic for small ε).

The following lemma establishes that the solution is a tail test (with or without the additional constraint $\sum_y p(y|r) w(y) \leq K$).

Lemma 3 *The above program is solved by a tail test, where* $\mathcal{Y} = \mathcal{X} = \mathcal{A} = \{0,1\}^N$; $q(y|x) = \mathbf{1}\{y = x\}$ *for all* y, x; $r_i \neq a_i$ *for all* i; $\pi_i(r_i|r_i) = 1 - \underline{\pi}$ *and* $\pi_i(a_i|r_i) = \underline{\pi}$ *for all* $i, r_i \neq a_i$; *and, letting* $n = |\{i : y_i = r_i\}|$,

$$w(y) = \mathbf{1}\{n \geq (1 - \underline{\pi}) N\} \bar{w}.$$

Moreover, the program with the additional constraint $\sum_y p(y|r)\, w(y) \leq K$ *is also solved by a tail test, where now*

$$w(y) = \begin{cases} \bar{w} & \textit{if } n > n^*, \\ \beta \bar{w} & \textit{if } n = n^*, \\ 0 & \textit{if } n < n^*, \end{cases}$$

for some $n^* \in \{0, 1, \ldots, N\}$ *(not necessarily equal to* $(1-\underline{\pi})\, N$*) and* $\beta \in [0,1]$.

We prove Lemma 3 in the next subsection. Lemma 3 implies that the value of the program without the constraint $\sum_y p(y|r)\, w(y) \leq K$ is

$$2(1-2\underline{\pi}) \binom{N-1}{(1-\underline{\pi})\, N - 1} (1-\underline{\pi})^{(1-\underline{\pi})N-1}\, \underline{\pi}^{\underline{\pi}N}\, \bar{w}.$$

This equals $2(1-2\underline{\pi})\, \bar{w}$ times the maximum of the *Binomial* $(N-1, 1-\underline{\pi})$ probability mass function. The first statement in the theorem follows as the latter quantity is proportional to $N^{-1/2}$, by the De Moivre-Laplace theorem.

For the second statement, by Lemma 3, the program becomes

$$\begin{aligned} \max_{n^* \in \{0,1,\ldots,N\}, \beta \in [0,1]} \quad & 2\bar{w}(1-2\underline{\pi})\left(\beta \Pr(n_{-i} = n^* - 1) + (1-\beta)\Pr(n_{-i} = n^*)\right) \\ \text{s.t.} \quad & \beta \Pr(n = n^*) + \Pr(n \geq n^* + 1) \leq \frac{K}{\bar{w}}, \end{aligned}$$

where $n_{-i} = |\{j \neq i : y_j = r_j\}|$ and the probabilities are binomial with parameter $1-\underline{\pi}$. We show that, for any $\rho > 0$, there exist $c_0, c_1 > 0$ such that, for each N, the value of the program is at most

$$\max\left\{2\bar{w}\exp\left(-c_0 N^{1-\rho}\right), c_1 N^{-\rho/2} K\right\}.$$

This completes the proof, as if K is fixed and $\min\{N, \exp(N^{1-\rho})/\bar{w}\} \to \infty$ then both terms go to 0.

We bound the program separately for n^* such that $|n^* - (1-\underline{\pi})\, N| > N^{1-\rho/2}$ and n^* such that $|n^* - (1-\underline{\pi})\, N| \leq N^{1-\rho/2}$. In the first case, by Hoeffding's inequality, there exists $c_0 > 0$ such that

$$\min\{\Pr(n_{-i} \geq n^* - 1), \Pr(n_{-i} \leq n^*)\} \leq \exp\left(-c_0 N^{1-\rho}\right).$$

Since the value of the program is at most $2\bar{w}\min\{\Pr(n_{-i}\geq n^*-1),\Pr(n_{-i}\leq n^*)\}$, this gives the desired bound when $|n^*-(1-\underline{\pi})N|>N^{1-\rho/2}$.

For the second case, the value of the program is at most

$$\begin{aligned}
& 2\bar{w}\frac{\beta\Pr(n_{-i}=n^*-1)+(1-\beta)\Pr(n_{-i}=n^*)}{\beta\Pr(n=n^*)+\Pr(n\geq n^*+1)}\frac{K}{\bar{w}} \\
\leq\ & 2\frac{(\beta\Pr(n_{-i}=n^*-1)+(1-\beta)\Pr(n_{-i}=n^*))}{\beta\Pr(n\geq n^*)+(1-\beta)\Pr(n\geq n^*+1)}K \\
\leq\ & 2\left(\frac{\Pr(n_{-i}=n^*-1)}{\Pr(n\geq n^*)}+\frac{\Pr(n_{-i}=n^*)}{\Pr(n\geq n^*+1)}\right)K.
\end{aligned}$$

By McKay (1989, Theorem 2), for any $m\geq(1-\underline{\pi})N$, we have

$$\Pr(n\geq m)\geq\sqrt{N\underline{\pi}(1-\underline{\pi})}\Pr(n_{-i}=m-1)\frac{1-\Phi\left((m-\underline{\pi}N)/\sqrt{N\underline{\pi}(1-\underline{\pi})}\right)}{\phi\left((m-\underline{\pi}N)/\sqrt{N\underline{\pi}(1-\underline{\pi})}\right)}.$$

If $(1-\underline{\pi})N\leq n^*\leq(1-\underline{\pi})N+N^{1-\rho/2}$, applying this inequality for $m\in\{n^*,n^*+1\}$, together with the inverse Mills ratio inequality $(1-\Phi(x))/\phi(x)\geq 1/(1+x)$, we have

$$\begin{aligned}
\frac{\Pr(n_{-i}=n^*-1|r_{-i})}{\Pr(n\geq n^*|r)}+\frac{\Pr(n_{-i}=n^*|r_{-i})}{\Pr(n\geq n^*+1|r)} &\leq 2\frac{1}{\sqrt{N\underline{\pi}(1-\underline{\pi})}}\left(\frac{n^*+1-\underline{\pi}N}{\sqrt{N\underline{\pi}(1-\underline{\pi})}}+1\right) \\
&\leq 2\left(\frac{N^{-\rho/2}}{\sqrt{\underline{\pi}(1-\underline{\pi})}}+\frac{1}{N\underline{\pi}(1-\underline{\pi})}+\frac{1}{\sqrt{N\underline{\pi}(1-\underline{\pi})}}\right).
\end{aligned}$$

Thus, there exists $c_1>0$ such that the value of the program is at most $c_1N^{-\rho/2}K$. Symmetrically, the same bound applies when $(1-\underline{\pi})N-N^{1-\rho/2}\leq n^*\leq(1-\underline{\pi})N$.

C.1 Proof of Lemma 3

First, consider the sub-program where $(\mathcal{X},\pi,\mathcal{Y},q)$ is fixed, so the objective is maximized over (r,a,w). By Blackwell's theorem, the value of the sub-program with signal distribution p is greater than that with signal distribution $\hat{p}$, if $\hat{p}$ is a *garbling* of p. (That is, viewing p and $\hat{p}$ as $|\mathcal{Y}|\times|\mathcal{A}|$ matrices, there is a $|\mathcal{Y}|\times|\mathcal{Y}|$ Markov matrix M such that $\hat{p}=Mp$.) For any noise structure $(\mathcal{X},\pi)$, the action monitoring structure $(\mathcal{Y},p)$ induced by any outcome monitoring structure $(\mathcal{Y},q)$ is clearly a garbling of that induced by the outcome monitoring structure where $\mathcal{Y}=\mathcal{X}$ and $q(y|x)=\mathbf{1}\{y=x\}$ for all y,x, so that $p(y|a)=\pi(y|a)$ for all

y, a. It is thus without loss to focus on this $(\mathcal{Y}, q)$.

In addition, if we let $\mathcal{X} = \mathcal{A}$ and, for each $r, a \in \mathcal{A}$ and i, let

$$\bar{\pi}_i(y_i|r_i) = \begin{cases} 1-\underline{\pi} & \text{if } y_i = r_i, \\ \underline{\pi} & \text{if } y_i = a_i, \\ 0 & \text{otherwise,} \end{cases} \qquad \bar{\pi}_i(y_i|a_i) = \begin{cases} 1-\underline{\pi} & \text{if } y_i = a_i, \\ \underline{\pi} & \text{if } y_i = r_i, \\ 0 & \text{otherwise,} \end{cases}$$
$$\text{and} \qquad \bar{\pi}_i(y_i|\tilde{a}_i) = \mathbf{1}\{y_i = \tilde{a}_i\} \text{ for } \tilde{a}_i \notin \{a_i, r_i\},$$

then π_i is a garbling of $\bar{\pi}_i$ for each i, and hence π is a garbling of $\bar{\pi}$. To see this, since $\underline{\pi} < 1/2$, the matrix $\bar{\pi}_i$ is invertible, and the matrix inverse $\bar{\pi}_i^{-1}$ is given by

$$\bar{\pi}_i^{-1}(y_i|r_i) = \begin{cases} \frac{1-\underline{\pi}}{1-2\underline{\pi}} & \text{if } y_i = r_i, \\ -\frac{\underline{\pi}}{1-2\underline{\pi}} & \text{if } y_i = a_i, \\ 0 & \text{otherwise,} \end{cases} \qquad \bar{\pi}_i^{-1}(\hat{a}_i|a_i) = \begin{cases} \frac{1-\underline{\pi}}{1-2\underline{\pi}} & \text{if } y_i = a_i, \\ -\frac{\underline{\pi}}{1-2\underline{\pi}} & \text{if } y_i = r_i, \\ 0 & \text{otherwise,} \end{cases}$$
$$\text{and} \qquad \bar{\pi}_i^{-1}(y_i|\tilde{a}_i) = \mathbf{1}\{y_i = \tilde{a}_i\} \text{ for } \tilde{a}_i \notin \{a_i, r_i\}.$$

We can then calculate the matrix $M_i := \pi_i \bar{\pi}_i^{-1}$ as

$$M_i(y_i|\tilde{a}_i) = \begin{cases} \frac{(1-\underline{\pi})\pi_i(y_i|\tilde{a}_i) - \underline{\pi}(1-\pi_i(y_i|\tilde{a}_i))}{1-2\underline{\pi}} & \text{if } \tilde{a}_i \in \{a_i, r_i\}, \\ \pi_i(y_i|\tilde{a}_i) & \text{otherwise,} \end{cases}$$

and note that, for $\tilde{a}_i \in \{a_i, r_i\}$,

$$\sum_{y_i} M_i(y_i|\tilde{a}_i) = \frac{|\mathcal{A}_i| - 1 - \underline{\pi}}{|\mathcal{A}_i| - 1 - |\mathcal{A}_i|\,\underline{\pi}} - (|\mathcal{A}_i| - 1)\frac{\underline{\pi}}{|A_i| - 1 - \mathcal{A}\underline{\pi}} = 1,$$

and $\sum_{y_i} M_i(y_i|\tilde{a}_i) = 1$ for $\tilde{a}_i \notin \{a_i, r_i\}$; and that, since $\pi_i(y_i|\tilde{a}_i) \geq \underline{\pi}$ for all $y_i, \tilde{a}_i$,

$$\frac{(1-\underline{\pi})\,\pi_i(y_i|\tilde{a}_i) - \underline{\pi}\,(1 - \pi_i(y_i|\tilde{a}_i))}{1 - 2\underline{\pi}} \geq \frac{(1-\underline{\pi})\,\underline{\pi} - \underline{\pi}\,(1-\underline{\pi})}{|\mathcal{X}_i| - 1 - |\mathcal{X}_i|\,\underline{\pi}} = 0,$$

and $M_i(y_i|\tilde{a}_i) \leq 1$ for all $y_i, \tilde{a}_i$. Thus, M_i is a Markov matrix satisfying $\pi_i = M_i \bar{\pi}_i$.

It is thus without loss to take $(\mathcal{X}, \pi) = (\mathcal{A}, \bar{\pi})$. The program then simplifies to

$$\max_{r,a,w} \frac{2}{N} \sum_i \sum_y \bar{\pi}(y|r) \left(1 - \frac{\bar{\pi}_i(y_i|a_i)}{\bar{\pi}_i(y_i|r_i)}\right) w(y) \quad \text{s.t.}$$
$$w(y) \in [0, \bar{w}] \quad \text{for all } y.$$

Here it is without loss to take $a_i \neq r_i$ for all i, as if $a_i = r_i$ then the same value can be attained by taking $w(y)$ independent of y_i, at which point a_i can then be taken different from r_i without affecting the value. Next, note that, for each y, the objective is increasing in $w(y)$ (and hence is maximized at $w(y) = \bar{w}$) if

$$\frac{1}{N} \sum_i \frac{\bar{\pi}_i(y_i|a_i)}{\bar{\pi}_i(y_i|r_i)} \leq 1,$$

and is decreasing in $w(y)$ (maximized at $w(y) = 0$) otherwise. Finally, since $\bar{\pi}_i(r_i|r_i) = 1 - \underline{\pi} \geq \underline{\pi} = \bar{\pi}_i(r_i|a_i)$ for all i, this inequality holds if and only if $|\{i : y_i = r_i\}| \geq (1 - \underline{\pi}) N$. This proves the first part of the lemma.

Now add the constraint $\sum_y p(y|r) w(y) \leq K$. Letting $\lambda \geq 0$ denote the corresponding multiplier, for each y, the Lagrangian is increasing in $w(y)$ if and only if

$$\frac{1}{N} \sum_i \frac{\bar{\pi}_i(y_i|a_i)}{\bar{\pi}_i(y_i|r_i)} \leq 1 - \lambda.$$

It follows that $w(y)$ takes the prescribed form.

D Proof of Theorem 4

Fix a team equilibrium with coefficients $b = (1, b_2, \ldots, b_N)$, where $|b_i| \leq 1$ for all i. Let $I^+ = \{i : b_i \geq 0\}$ and $I^- = \{i : b_i < 0\}$. Define

$$\underline{v}_i = \begin{cases} \inf_h \omega_i(h) & \text{if } i \in I^+, \\ \sup_h \omega_i(h) & \text{if } i \in I^-, \end{cases} \quad \text{and} \quad \bar{v}_i = \begin{cases} \sup_h \omega_i(h) & \text{if } i \in I^+, \\ \inf_h \omega_i(h) & \text{if } i \in I^-. \end{cases}$$

Since $V(\varepsilon)$ is convex, it suffices to show that $\underline{v}, \bar{v} \in V(\varepsilon)$, where $\underline{v} = (\underline{v}_i)_{i \in I}$ and $\bar{v} = (\bar{v}_i)_{i \in I}$.

In the following lemma, for any $\alpha \in \Delta(\mathcal{A})$ and $f : \mathcal{A} \times \mathcal{Y} \to \mathbb{R}$, $\mathbb{E}^\alpha[f(r, y)]$ denotes expectation where $r \sim \alpha$ and then $y \sim p(\cdot|r)$, and $\mathbb{E}^{\alpha,a_i'}[f(r,y)]$ denotes expectation where $r \sim \alpha$ and then $y \sim p(\cdot|a_i', r_{-i})$.

Lemma 4 *There exist $\alpha \in \Delta(\mathcal{A})$ and $\omega : \mathcal{A} \times \mathcal{Y} \to \mathbb{R}$ such that*

$$\begin{aligned}
\bar{v} &= \mathbb{E}^{\alpha}[u(r) - b\omega(r,y)], \\
\mathbb{E}^{\alpha}[u_i(r) - b_i\omega(r,y)|r_i = a_i] &\geq \mathbb{E}^{\alpha,a_i'}[u_i(a_i', r_{-i}) - b_i\omega(r,y)|r_i = a_i] \quad \textit{for all } i, a_i \in \operatorname{supp}\alpha_i, a_i', \\
\omega(r,y) &\in \left[0, \frac{\delta}{1-\delta}\bar{u}\right] \quad \textit{for all } r, y, \\
\mathbb{E}^{\alpha}[\omega(r,y)] &\leq \bar{u} \quad \textit{for all } r, y.
\end{aligned}$$

Moreover, if the constraints $\omega(r,y) \in \left[0, \frac{\delta}{1-\delta}\bar{u}\right]$ and $\mathbb{E}^{\alpha}[\omega(r,y)] \leq \bar{u}$ are replaced by $\omega(r,y) \in \left[-\frac{\delta}{1-\delta}\bar{u}, 0\right]$ and $\mathbb{E}^{\alpha}[\omega(r,y)] \geq -\bar{u}$, then the same statement holds with $\underline{v}$ in place of $\bar{v}$.

Proof. Let $E = \{(1-\beta)\underline{v} + \beta\bar{v} : \beta \in [0,1]\}$. By standard arguments, E is self-generating: for any $v \in E$, there exist $\alpha \in \Delta(\mathcal{A})$ and $\omega : \mathcal{A} \times \mathcal{Y} \to E$ such that

$$\begin{aligned}
v &= \mathbb{E}^{\alpha}[(1-\delta)u(r) + \delta\omega(r,y)] \quad \text{and} \\
\mathbb{E}^{\alpha}[(1-\delta)u_i(r) + \delta\omega_i(r,y)|r_i = a_i] &\geq \mathbb{E}^{\alpha,a_i'}[(1-\delta)u_i(a_i', r_{-i}) + \delta\omega_i(r,y)|r_i = a_i],
\end{aligned}$$

for all $i, a_i \in \operatorname{supp}\alpha_i, a_i' \in \mathcal{A}_i$. Since $v \in E$ and $\omega(r,y) \in E$ for all r, y, we have $v_i - \omega_i(r,y) = b_i(v_1 - \omega_1(r,y))$ for all i, r, y. Since $\bar{v}_1 \geq v_1$ for all $v \in E$, if $v = \bar{v}$ then $\omega_1(r,y) \leq v_1$ for all r, y. Hence, taking $v = \bar{v} = (1-\delta)u(\alpha) + \delta\mathbb{E}^{\alpha}[\omega(r,y)]$ and defining $\omega(r,y) = \frac{\delta}{1-\delta}(\bar{v}_1 - \omega_1(r,y))$, we have $\omega(r,y) \in \left[0, \frac{\delta}{1-\delta}\bar{u}\right]$ and

$$\mathbb{E}^{\alpha}[|\omega(r,y)|] = \frac{\delta}{1-\delta}\mathbb{E}^{\alpha}[|(1-\delta)u_1(\alpha) + \delta\mathbb{E}^{\alpha}[\omega_1(r,y)] - \omega_1(r,y)|] = \delta|u_1(\alpha) - \omega_1(r,y)| \leq \bar{u}.$$

Moreover, we have

$$\begin{aligned}
\mathbb{E}^{\alpha}[u(r) - b\omega(r,y)] &= \mathbb{E}^{\alpha}\left[u(r) - b\frac{\delta}{1-\delta}(\bar{v}_1 - \omega_1(r,y))\right] \\
&= u(\alpha) - \mathbb{E}^{\alpha}\left[\frac{\delta}{1-\delta}(\bar{v} - \omega(r,y))\right] \\
&= u(\alpha) - \mathbb{E}^{\alpha}\left[\frac{\delta}{1-\delta}((1-\delta)u(\alpha) + \delta\mathbb{E}^{\alpha}[\omega(r,y)] - \omega(r,y))\right] \\
&= (1-\delta)u(\alpha) + \delta\mathbb{E}^{\alpha}[\omega(r,y)] = v,
\end{aligned}$$

and, for all $i, a_i \in \operatorname{supp} \alpha_i, a_i' \in \mathcal{A}_i$,

$$\begin{array}{ccc}
\mathbb{E}^{\alpha}\left[(1-\delta) u_i(r) + \delta \omega_i(r,y) | r_i = a_i\right] & \geq & \mathbb{E}^{\alpha, a_i'}\left[(1-\delta) u_i(a_i', r_{-i}) + \delta \omega_i(r,y) | r_i = a_i\right] \\
& \Longleftrightarrow & \\
\mathbb{E}^{\alpha}\left[u_i(r) + \frac{\delta}{1-\delta}\left(\omega_i(r,y) - \bar{v}_i\right) | r_i = a_i\right] & \geq & \mathbb{E}^{\alpha, a_i'}\left[u_i(a_i', r_{-i}) + \frac{\delta}{1-\delta}\left(\omega_i(r,y) - \bar{v}_i\right) | r_i = a_i\right] \\
& \Longleftrightarrow & \\
\mathbb{E}^{\alpha}\left[u_i(r) - b_i \omega(r,y) | r_i = a_i\right] & \geq & \mathbb{E}^{\alpha, a_i'}\left[u_i(a_i', r_{-i}) - b_i \omega(r,y) | r_i = a_i\right].
\end{array}$$

Similarly, if $v = \underline{v}$ then $\omega_1(r,y) \geq v_1$ for all r, y, and the symmetric conclusion holds. ■

Taking α and ω as in Lemma 4, we see that $\sum_i \bar{g}(\alpha)/N$ is bounded by the solution to the program

$$\begin{aligned}
\max_{(\mathcal{Y},p),r,a,\omega} \frac{1}{N} \sum_i \left(\mathbb{E}^r[\omega(y)] - \mathbb{E}^{(a_i,r_{-i})}[\omega(y)]\right) \quad &\text{s.t.} \\
\omega(y) \in \left[0, \frac{\delta}{1-\delta}\bar{u}\right] \quad &\text{for all } y, \\
\mathbb{E}^r[\omega(y)] \leq \bar{u}.&
\end{aligned}$$

This is identical to the program in Theorem 2, with $\bar{w} = (\delta/(1-\delta))\bar{u}$ and $K = \bar{u}$. The result therefore follows from Theorem 2.

References

[1] Abreu, Dilip, David Pearce, and Ennio Stacchetti. "Optimal Cartel Equilibria with Imperfect Monitoring." *Journal of Economic Theory* 39.1 (1986): 251-269.

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

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

[4] Alchian, Armen A., and Harold Demsetz. "Production, Information Costs, and Economic Organization." *American Economic Review* 62.5 (1972): 777-795.

[5] Al-Najjar, Nabil I., and Rann Smorodinsky. "Pivotal Players and the Characterization of Influence." *Journal of Economic Theory* 92.2 (2000): 318-342.

[6] Al-Najjar, Nabil I., and Rann Smorodinsky. "Large Nonanonymous Repeated Games." *Games and Economic Behavior* 37.1 (2001): 26-39.

[7] Athey, Susan, Kyle Bagwell, and Chris Sanchirico. "Collusion and Price Rigidity." *Review of Economic Studies* 71.2 (2004): 317-349.

[8] Awaya, Yu, and Vijay Krishna. "On Communication and Collusion." *American Economic Review* 106.2 (2016): 285-315.

[9] Awaya, Yu, and Vijay Krishna. "Communication and Cooperation in Repeated Games." *Theoretical Economics* 14.2 (2019): 513-553.

[10] Bowles, Samuel, and Herbert Gintis. *A Cooperative Species: Human Reciprocity and its Evolution.* Princeton University Press, 2011.

[11] Boyd, Robert, and Peter J. Richerson. "The Evolution of Reciprocity in Sizable Groups." *Journal of Theoretical Biology* 132.3 (1988): 337-356.

[12] Calvo, Guillermo A., and Stanislaw Wellisz. "Supervision, Loss of Control, and the Optimum Size of the Firm." *Journal of political Economy* 86.5 (1978): 943-952.

[13] Camera, Gabriele, and Marco Casari. "Cooperation among Strangers under the Shadow of the Future." *American Economic Review* 99.3 (2009): 979-1005.

[14] Camera, Gabriele, Marco Casari, and Maria Bigoni. "Money and Trust among Strangers." *Proceedings of the National Academy of Science*s 110.37 (2013): 14889-14893.

[15] Cover, Thomas M., and Joy A. Thomas. *Elements of Information Theory, 2nd ed.* Wiley (2006).

[16] Deb, Joyee, Takuo Sugaya, and Alexander Wolitzky. "The Folk Theorem in Repeated Games with Anonymous Random Matching." *Econometrica* 88.3 (2020): 917-964.

[17] Duffy, John, and Jack Ochs. "Cooperative Behavior and the Frequency of Social Interaction." *Games and Economic Behavior* 66.2 (2009): 785-812.

[18] Ekmekci, Mehmet, Olivier Gossner, and Andrea Wilson. "Impermanent Types and Permanent Reputations." *Journal of Economic Theory* 147.1 (2012): 162-178.

[19] Ellickson, Robert C. *Order without Law: How Neighbors Settle Disputes.* Harvard University Press, 1991.

[20] Ellison, Glenn. "Cooperation in the Prisoner's Dilemma with Anonymous Random Matching." *Review of Economic Studies* 61.3 (1994): 567-588.

[21] Faingold, Eduardo. "Reputation and the Flow of Information in Repeated Games." *Econometrica* 88.4 (2020): 1697-1723.

[22] Feigenberg, Benjamin, Erica Field, and Rohini Pande. "The Economic Returns to Social Interaction: Experimental Evidence from Microfinance." *Review of Economic Studies* 80.4 (2013): 1459-1483.

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

[24] Fudenberg, Drew, and David K. Levine. "Efficiency and Observability with Long-Run and Short-Run Players." J*ournal of Economic Theory* 62.1 (1994): 103-135.

[25] 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.

[26] Fudenberg, Drew, and David K. Levine. "Repeated Games with Frequent Signals." *Quarterly Journal of Economics* 124.1 (2009): 233-265.

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

[28] Fudenberg, Drew, David Levine, and Wolfgang Pesendorfer. "When are Nonanonymous Players Negligible?" *Journal of Economic Theory* 79.1 (1998): 46-71.

[29] Gossner, Olivier. "Simple Bounds on the Value of a Reputation." *Econometrica* 79.5 (2011): 1627-1641.

[30] Gossner, Olivier, Penelope Hernández, and Abraham Neyman. "Optimal Use of Communication Resources." *Econometrica* 74.6 (2006): 1603-1636.

[31] Green, Edward J. "Noncooperative Price Taking in Large Dynamic Markets." *Journal of Economic Theory* 22.2 (1980): 155-182.

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

[33] Hellman, Ziv, and Ron Peretz. "A survey on entropy and economic behaviour." *Entropy* 22.2 (2020): 157.

[34] Holmström, Bengt. "Moral Hazard in Teams." *Bell Journal of Economics* (1982): 324-340.

[35] Kandori, Michihiro. "Social Norms and Community Enforcement." *Review of Economic Studies* 59.1 (1992): 63-80.

[36] Karlan, Dean S. "Social Connections and Group Banking." *Economic Journal* 117.517 (2007): F52-F84.

[37] Mailath, George, and Larry Samuelon. *Repeated Games and Reputations: Long-Run Relationships.* Oxford University Press, 2006.

[38] McKay, Brendan. "On Littlewood's Estimate for the Binomial Distribution." *Advances in Applied Probability* 21.2 (1989), 475-478.

[39] Miguel, Edward, and Mary Kay Gugerty. "Ethnic Diversity, Social Sanctions, and Public Goods in Kenya." *Journal of Public Economics* 89.11-12 (2005): 2325-2368.

[40] 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).

[41] Neyman, Abraham, and Daijiro Okada. "Strategic Entropy and Complexity in Repeated Games." *Games and Economic Behavior* 29.1-2 (1999): 191-223.

[42] Neyman, Abraham, and Daijiro Okada. "Repeated Games with Bounded Entropy." *Games and Economic Behavior* 30.2 (2000): 228-247.

[43] Ostrom, Elinor. *Governing the Commons: The Evolution of Institutions for Collective Action.* Cambridge University Press, 1990.

[44] Pai, Mallesh M., Aaron Roth, and Jonathan Ullman. "An Antifolk Theorem for Large Repeated Games." *ACM Transactions on Economics and Computation (TEAC)* 5.2 (2016): 1-20.

[45] Qian, Yingyi. "Incentives and Loss of Control in an Optimal Hierarchy." *Review of Economic Studies* 61.3 (1994): 527-544.

[46] Sabourian, Hamid. "Anonymous Repeated Games with a Large Number of Players and Random Outcomes." *Journal of Economic Theory* 51.1 (1990): 92-110.

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

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

[49] Sason, Igal, and Sergio Verdú. "f-Divergence Inequalities." *IEEE Transactions on Information Theory* 62.11 (2016): 5973-6006.

[50] Seabright, Paul. *The Company of Strangers.* Princeton University Press, 2004.

[51] Sugaya, Takuo, and Alexander Wolitzky. "A Few Bad Apples Spoil the Barrel: An Anti-folk Theorem for Anonymous Repeated Games with Incomplete Information." *American Economic Review* 110.12 (2020): 3817-35.

[52] Sugaya, Takuo, and Alexander Wolitzky, "Communication and Community Enforcement." *Journal of Political Economy* 129.9 (2021): 2595-2628.

[53] Sugaya, Takuo, and Alexander Wolitzky, "Monitoring versus Discounting in Repeated Games." *Econometrica* 91.5 (2023): 1727-1761.

Online Appendix

This appendix establish a folk theorem for repeated games with public, product structure monitoring where the discount factor, monitoring structure, and stage game (including the number of players N) vary simultaneously. A consequence is Corollary 1 in the main text, which implies that the relationship among N, δ, and C in Theorem 3 is tight up to a $\log(N)$ factor.

A monitoring structure $(\mathcal{Y}, q)$ has a *product structure* if there exist sets $(\mathcal{Y}_i)_{i\in I}$ and a family of conditional distributions $(q_i(y_i|x_i))_{i,,y_{ii},x_i}$ such that $\mathcal{Y} = \prod_i \mathcal{Y}_i$ and $q(y|x) = \prod_i q_i(y_i|x_i)$ for all y, x. That is, the public signal y consists of conditionally independent signals of each player's individual outcome. Note that if $(\mathcal{Y}, q)$ has a product structure, then so does $(\mathcal{Y}, p)$, meaning that there exists a family of conditional distributions $(p_i(y_i|a_i))_{i,,y_{ii},a_i}$ (given by $p_i(y_i|a_i) = \sum_{x_i} \pi_i(x_i|a_i) q_i(y_i|x_i)$) such that $p(y|a) = \prod_i p_i(y_i|a_i)$ for all y, a.

Next, for any $\eta > 0$, we say that an action monitoring structure $(\mathcal{Y}, p)$ satisfies *η-individual identifiability* if

$$\sum_{y_i:p_i(y_i|a_i)\geq\eta} \frac{(p_i(y_i|a_i) - p_i(y_i|\alpha_i))^2}{p_i(y_i|a_i)} \geq \eta \quad \text{for all } i \in I, a_i \in \mathcal{A}_i, \alpha_i \in \Delta(\mathcal{A}_i \setminus \{a_i\}). \tag{12}$$

This condition is a variant of Fudenberg, Levine, and Maskin (1994)'s individual full rank condition and Kandori and Matsushima's (1998) assumption (A2"). It says that the influence on the signal distribution (measured by χ^2-divergence) of a deviation from a_i to any mixed action α_i supported on $\mathcal{A}_i \setminus \{a_i\}$ is at least η^2, ignoring signals that occur with probability less than η^2 under a_i. Intuitively, this requires that deviations from a_i are detectable, and that in addition detection does not rest on very rare signal realizations. This assumption will ensure that players can be incentivized through rewards whose variance and maximum absolute value are both of order $(1-\delta)/\eta$.[24]

Finally, denote the feasible payoff set by $F = \text{co}\left\{\{u(a)\}_{a\in\mathcal{A}}\right\} \subseteq \mathbb{R}^N$ (where co denotes convex hull). Let $F^* \subseteq F$ denote the set of payoff vectors that weakly Pareto-dominate a payoff vector which is a convex combination of static Nash payoffs: that is, $v \in F^*$ if $v \in F$ and there exists a collection of static Nash equilibria (α_n) and non-negative weights (β_n) such that $v \geq \sum_n \beta_n u(\alpha_n)$ and $\sum_n \beta_n = 1$. For each $v \in \mathbb{R}^N$ and $\varepsilon > 0$, let $B_v(\varepsilon) = \prod_i [v_i - \varepsilon, v_i + \varepsilon]$ and let $B(\varepsilon) = \left\{v \in \mathbb{R}^N : B_v(\varepsilon) \subseteq F^*\right\}$. That is, $B(\varepsilon)$ is the set of payoff vectors $v \in \mathbb{R}^N$ such that the cube with center v and side-length 2ε lies entirely within F^*. Let $E \subseteq \mathbb{R}^N$ denote the set of PPE payoff vectors .

We establish the following folk theorem.

Theorem 5 *Fix any $\bar{u} > 0$. For any $\varepsilon > 0$, there exists $k > 0$ such that, for any repeated game with $\bar{u}$-bounded payoffs and public, product structure monitoring satisfying η-individual identifiability, where*

$$(1-\delta)\log(N) < k\eta, \tag{13}$$

[24] If (12) were weakened by taking the sum over all y_i (rather than only y_i such that $p_i(y_i|a_i) \geq \eta$), player i could be incentivized by rewards with variance $O((1-\delta)/\eta)$, but not necessarily with maximum absolute value $O((1-\delta)/\eta)$. Our analysis requires controlling both the variance and absolute value of players' rewards, so we need the stronger condition. We also note that the current definition of η-individual identifiability coincides with $\sqrt{\eta}$-individual identifiability in the terminology in SW.

we have $B(\varepsilon) \subseteq E$.

To prove Corollary 1 from Theorem 5, consider a game where $\mathcal{X} = \mathcal{A}$ with uniform noise, so that $\pi_i(a_i'|a_i) = \underline{\pi}$ for all $i, a_i, a_i' \neq a_i$, and assume that $\underline{\pi} < (\max_i |\mathcal{A}_i| + 1)^{-1}$. Suppose that the outcome monitoring structure $(\mathcal{Y}, q)$ is given by *η-random monitoring*, where in every period the public signal perfectly reveals each player's identity and realized individual outcome with probability η. That is, under η-random monitoring, $\mathcal{Y}_i = \mathcal{X}_i \cup \{\emptyset\}$ for all i, and

$$q_i(y_i|x_i) = \begin{cases} \eta & \text{if } y_i = x_i, \\ 0 & \text{if } y_i \in \mathcal{X}_i \backslash \{x_i\}, \\ 1-\eta & \text{if } y_i = \emptyset, \end{cases} \quad \text{so that} \quad p_i(y_i|a_i) = \begin{cases} \eta\pi_i(y_i|a_i) & \text{if } y_i \in \mathcal{X}_i, \\ 1-\eta & \text{if } y_i = \emptyset. \end{cases}$$

Note that the channel capacity of random monitoring is at most $\eta N \log(\max_i |\mathcal{A}_i|)$. In addition, random monitoring satisfies $\eta\underline{\pi}$-individual identifiability, because, any i, a_i, α_i, we have

$$\begin{aligned} \sum_{y_i: p_i(y_i|a_i) \geq \eta\underline{\pi}} \frac{(p_i(y_i|a_i) - p_i(y_i|\alpha_i))^2}{p_i(y_i|a_i)} &\geq \frac{\left(p_i(a_i|a_i) - \max_{a_i' \neq a_i} p_i(a_i|a_i')\right)^2}{p_i(a_i|a_i)} \\ &= \frac{(\eta(1 - (|\mathcal{A}_i| - 1)\underline{\pi}) - \eta\underline{\pi})^2}{\eta(1 - (|\mathcal{A}_i| - 1)\underline{\pi})} \\ &\geq \eta(1 - |\mathcal{A}_i|\underline{\pi}) \geq \eta\underline{\pi}, \end{aligned}$$

where the last inequality uses $\underline{\pi} < (\max_i |\mathcal{A}_i| + 1)^{-1}$. Thus, by Theorem 5, η-random monitoring is a monitoring structure with channel capacity at most $C = \eta N \log(\max_i |\mathcal{A}_i|)$ under which a folk theorem holds whenever $(1-\delta) N \log(N)/C \to 0$.

Theorem 5 is a folk theorem for PPE in repeated games with public monitoring.[25] The standard proof approach, following Fudenberg, Levine, and Maskin (1994) and Kandori and Matsushima (1998), relies on transferring continuation payoffs among the players along hyperplanes that are tangent to the boundary of the PPE payoff set. Unfortunately, this approach encounters difficulties when N and δ vary simultaneously. The problem is that when N is large, changing each player's continuation payoff by a small amount can result in a large overall movement in the continuation payoff vector. Mathematically, Fudenberg, Levine, and Maskin's proof relies on the equivalence of the L^1 norm and the Euclidean norm in $\mathbb{R}^N$. Since this equivalence is not uniform in N, their proof does not apply when N and δ vary simultaneously.[26] Our proof is instead based on the "block strategy" approach intro-

[25] Specifically, it is a "Nash threat" folk theorem, as F^* is the set of payoffs that Pareto-dominate a convex combination of static Nash equilibria. To extend this result to a "minmax threat" theorem, players must be made indifferent among all actions in the support of a mixed strategy that minmaxes an opponent. This requires a stronger identifiability condition, similar to Kandori and Matsushima's assumption (A1).

[26] With random monitoring of M players, the per-period movement in each player's continuation payoff required to provide incentives is of order $(1-\delta) N/M$, so the movement of the continuation payoff vector in $\mathbb{R}^N$ is $O\left((1-\delta) N^{3/2}/M\right)$. For any ball $B \subseteq F^*$, consider the problem of generating the point $v = \operatorname{argmax}_{w \in B} w_1$ using continuation payoffs drawn from B. To satisfy promise keeping, player 1's continuation payoff must be within distance $O(1-\delta)$ of v, so the largest possible movement along a translated tangent hyperplane is $O\left(\sqrt{1-\delta}\right)$. FLM's proof approach thus requires that $(1-\delta) N^{3/2}/M \ll \sqrt{1-\delta}$, or

duced by Matsushima (2004) and Hörner and Olszewski (2006) in the context of repeated games with private monitoring. We sketch the proof here, deferring the details to the next section.

Fix any $v \in B(\varepsilon)$. We show that, for sufficiently large δ, the cube $B_v(\varepsilon/2)$ is self-generating. Since $B(\varepsilon)$ is compact, this implies that, for sufficiently large δ, $B(\varepsilon)$ is self-generating, and hence $B(\varepsilon) \subseteq E$.

Since $B_v(\varepsilon/2)$ is a cube, for each extreme point $v^* \in B_v(\varepsilon/2)$, there exists $\zeta \in \{-1, 1\}^N$ such that $v_i^* \in \text{argmax}_{\omega \in B_v(\varepsilon/2)} \zeta_i \omega_i$ for all i. To self-generate $B_v(\varepsilon/2)$, it is sufficient that, for each $\zeta \in \{-1, 1\}^N$ and v^* satisfying $v_i^* \in \text{argmax}_{\omega \in B_v(\varepsilon/2)} \zeta_i \omega_i$ for all i, we can find a number $T \in \mathbb{N}$, a T-period strategy σ, and a history-contingent continuation payoff $\omega\left(h^{T+1}\right)$ such that the following three conditions hold:

Promise Keeping $v_i^* = (1-\delta) \sum_{t=1}^{T} \delta^{t-1} \mathbb{E}^{\sigma}\left[u_i(a_t)\right] + \delta^T \mathbb{E}^{\sigma}\left[\omega_i\left(h^{T+1}\right)\right]$ for all i.

Incentive Compatibility $\tilde{\sigma}_i = \sigma_i$ is optimal in the T-period repeated game with objective $\mathbb{E}^{\tilde{\sigma}_i, \sigma_{-i}}\left[(1-\delta) \sum_{t=1}^{T} \delta^{t-1} u_i(a) + \delta^T \omega_i\left(h^{T+1}\right)\right]$, for all i.

Self Generation $\omega\left(h^{T+1}\right) \in B_v(\varepsilon/2)$ for all h^{T+1}.

Since $B_v(\varepsilon/2)$ is the cube with center v and side-length ε, and $v_i^* \in \text{argmax}_{\omega \in B_v(\varepsilon/2)} \zeta_i \omega_i$ for all i, we have $\omega\left(h^{T+1}\right) \in B_v(\varepsilon/2)$ if and only if $\zeta_i\left(\omega_i\left(h^{T+1}\right) - v_i\right) \in [-\varepsilon, 0]$ for all i. Thus, defining $\psi_i\left(h^{T+1}\right) = \left(\delta^T/(1-\delta)\right)\left(\omega_i\left(h^{T+1}\right) - v_i^*\right)$, we can rewrite the above conditions as

Promise Keeping $v_i = \frac{1-\delta}{1-\delta^T} \mathbb{E}^{\sigma}\left[\sum_{t=1}^{T} \delta^{t-1} u(a_t) + \psi_i\left(h^{T+1}\right)\right]$ for all i.

Incentive Compatibility $\tilde{\sigma}_i = \sigma_i$ is optimal in the T-period repeated game with objective $\mathbb{E}^{\tilde{\sigma}_i, \sigma_{-i}}\left[\sum_{t=1}^{T} \delta^{t-1} u(a) + \psi_i\left(h^{T+1}\right) | \sigma_i', \sigma_{-i}\right]$, for all i.

Self Generation $-\frac{\delta^T}{1-\delta}\varepsilon \leq \zeta_i \psi_i\left(h^{T+1}\right) \leq 0$ for all i, h^{T+1}. Moreover, since $\lim_{\delta \to 1} -\frac{\delta^T}{1-\delta}\varepsilon = -\infty$, it suffices to require that $\zeta_i \psi_i\left(h^{T+1}\right) \leq 0$ for all i, h^{T+1}.

Fix ζ and v^*, and take $T = O\left((1-\delta)^{-1}\right)$. We construct a T-period strategy σ and a "reward function" $\psi_i\left(h^{T+1}\right)$ that satisfy the above conditions.

By (12), for each recommendation r_i, there exists $f_{i,r_i}(y_i)$ such that (i) augmenting player i's utility by $f_{i,r_i}(y_i)$ incentivizes her to take r_i, (ii) the expectation of $f_{i,r_i}(y_i)$ when player i takes r_i equals 0, and (iii) the variance of $f_{i,r_i}(y_i)$ is of order $1/\eta$. Indeed, these properties are achieved by taking $f_{i,r_i}(y_i)$ proportional to the likelihood ratio difference $\min_{\alpha_i \in \Delta(\mathcal{A}_i \setminus \{a_i\})} \left(p_i(y_i|a_i) - p_i(y_i|\alpha_i)\right)/p_i(y_i|a_i)$. (See Lemma 5.)

equivalently $(1-\delta) N^3/M^2 \ll 1$, while we assume only $(1-\delta) N \log(N)/M \ll 1$. Hence, while the conditions for Theorem 5 are tight up to $\log(N)$ slack, Fudenberg, Levine, and Maskin's approach would instead require slack $N^2/M \geq N$. On the other hand, in SW, we extend Fudenberg, Levine, and Maskin's proof to give a folk theorem where discounting and monitoring vary simultaneously for a fixed stage game. There, this approach works because N is fixed.

Since $v \in B(\varepsilon)$ and $v^* \in B_v(\varepsilon/2)$, there exists $\bar{\alpha} \in \Delta(\mathcal{A})$ such that $\zeta_i(u_i(\bar{\alpha}) - v_i^*) = \varepsilon/2$. Suppose that the recommendation profile r is drawn according to $\bar{\alpha}$ by public randomization (and players follow their recommendations), and define the reward function $\tilde{\psi}_i(h^{T+1}) = \sum_t \delta^{t-1} f_{i,r_{i,t}}(y_{i,t}) - \zeta_i \frac{1-\delta^T}{1-\delta}\frac{\varepsilon}{2}$. We call $\tilde{\psi}_i(h^{T+1})$ the "base reward." We show that this strategy and reward function satisfy promise keeping and incentive compatibility, and also satisfy self generation with high probability. We then show how to modify the strategy and reward function to ensure that self generation is always satisfied.

Since $f_{i,r_i}(y_i)$ has 0 mean, promise keeping is immediate:

$$v_i = \frac{1-\delta}{1-\delta^T}\mathbb{E}^{\sigma}\left[\sum_{t=1}^{T}\delta^{t-1}u_i(a) + \tilde{\psi}_i(h^{T+1})\right] = u_i(\bar{\alpha}) - \zeta_i\frac{\varepsilon}{2} = v_i^*.$$

Next, incentive compatibility holds because

$$\mathbb{E}^{\tilde{\sigma}_i,\sigma_{-i}}\left[\sum_{t=1}^{T}\delta^{t-1}u_i(a_t) + \tilde{\psi}_i(h^{T+1})\right] = \mathbb{E}^{\tilde{\sigma}_i,\sigma_{-i}}\left[\sum_{t=1}^{T}\delta^{t-1}\left(u_i(a_t) + f_{i,r_{i,t}}(y_{i,t})\right)\right] - \frac{1-\delta^T}{1-\delta}\zeta_i\frac{\varepsilon}{2},$$

so the augmented per-period payoff is $u_i(a) + f_{i,r_{i,t}}(y_{i,t})$. Moreover, since the variance of f_{i,r_i} is $O(1/\eta)$ and T is $O\left((1-\delta)^{-1}\right)$, by a standard concentration inequality, the self generation constraint $\zeta_i\tilde{\psi}_i(h^{T+1}) \leq 0$ holds for all i with probability at least

$$N\exp\left(-\frac{\frac{1-\delta^T}{1-\delta}\zeta_i\frac{\varepsilon}{2}}{\sqrt{T\eta}}\right) \approx \exp\left(-\sqrt{\frac{(1-\delta)\log N}{\eta}}\right).$$

Therefore, by (13), self generation holds with high probability when k is small. (Lemmas 6 and 8.)

We now modify the strategy and reward to satisfy self-generation at every history. To this end, define a stopping time as the first period τ such that

$$\zeta_i\sum_{t=1}^{\tau}\delta^{t-1}f_{i,r_{i,t}}(y_{i,t}) > \bar{f}, \tag{14}$$

where $\bar{f}$ is a positive constant less than $\left(\left(1-\delta^T\right)/(1-\delta)\right)\varepsilon/2$. That is, in (the random) period τ, for a player, the base reward $\tilde{\psi}_i(h^{T+1})$ becomes abnormal. If no such period arises, define $\tau = T$. By the same concentration argument as above, no player's base reward is abnormal (that is, $\tau = T$) with high probability: in particular,

$$\Pr(\tau < T) \approx \exp\left(-\sqrt{\frac{(1-\delta)\log N}{\eta}}\right). \tag{15}$$

We now define the modified strategy.

If $\tau = T$, then in every period r is drawn according to $\bar{\alpha}$ and the reward equals $\tilde{\psi}_i(h^{T+1})$.

If $\tau < T$, then let I^* be the set of players whose base reward satisfies (14). For each

$i \in I^*$, we add or subtract a constant from the rewards of players $-i$ to satisfy self generation. Since monitoring has a product structure, players $-i$ cannot control the realization of player i's reward. Thus, this addition or subtraction does not affect incentives.

If I^* is a singleton, $I^* = \{i\}$, then player i starts taking a static best response. Meanwhile, players $-i$ take r_{-i} drawn from $\bar{\alpha}$ if $\zeta_i = 1$, and take static Nash actions $\left(\alpha_j^{NE}\right)_{j \neq i}$ if $\zeta_i = -1$. Let $u_i(\zeta_i)$ be player i's resulting instantaneous payoff. Since $v^* \in F^*$, we have $\zeta_i (u_i(\zeta_i) - u_i(\bar{\alpha})) \geq 0$. Hence, if player i's period t reward is fixed at $u_i(\bar{\alpha}) - u_i(\zeta_i)$, self generation is satisfied, and player i's period t augmented payoff equals $u_i(\bar{\alpha})$. If instead $|I^*| \geq 2$, then all players' subsequent rewards equal 0.

Since $\tau = T$ with high probability by (15), expected payoffs under the modified strategy and reward are close to v. Further adjusting the rewards by a small constant thus achieves promise keeping. Moreover, self generation now holds by construction. Finally, for any period $t > \tau$, incentive compatibility holds, because either a player's reward is fixed and she is supposed to take a static best response, or she is incentivized by the base reward function.

To complete the proof, it remains to establish incentive compatibility for periods $t \leq \tau$. For $t \leq \tau$, player i's augmented period t payoff is $u_i(a_i, r_{-i}) + f_{i,r_i}(y_i))$. Thus, to show that it is optimal for player i to follow her recommendation, it suffices to show that she cannot gain by manipulating the stopping time τ.

Since monitoring has a product structure, player i cannot influence others' rewards. Player i also cannot improve her augmented period t payoff by manipulating her own reward, because both $u_i(r) + f_{i,r_{i,t}}(y_{i,t}))$ and $u_i(\zeta_i) + u_i(\bar{\alpha}) - u_i(\zeta_i))$ equal $u(\bar{\alpha})$ regardless of whether $t \leq \tau$ or $t > \tau$. However, there is one possible benefit from manipulation: once τ realizes with $I^* = \{i\}$, the chance of a constant being added or subtracted from player i's reward vanishes, but if τ first realizes with $I^* \neq \{i\}$, this addition or subtraction occurs. To prevent this adjustment from affecting player i's incentive, a "fictitious" recommendation $\tilde{r}_t$ is drawn according to $\bar{\alpha}$, and a fictitious signal $\tilde{y}$ is drawn according to $p(\tilde{y}|\tilde{r})$, and the base rewards are updated according to the fictitious recommendations and signals even when $t > \tau$. (See (30) for the definition of the fictitious recommendations and signals.) If player $j \neq i$'s fictitious base reward satisfies (14), we add or subtract a constant from player i's reward. (See (31) for the definition of the event that induces this addition or subtraction. Note also that this fictitious update of player j's base reward is used solely to satisfy player i's incentives and does not affect player j's reward.) Given this modification, player i does not have an incentive to manipulate her own reward to manipulate the distribution of τ (Lemma 7), and hence incentive compatibility holds (Lemma 9).

E Proof of Theorem 5

E.1 Preliminaries

Fix any $\varepsilon > 0$. If $\varepsilon \geq \bar{u}/2$ then $B(\varepsilon) = \emptyset$ and the conclusion of the theorem is trivial, so assume without loss that $\varepsilon < \bar{u}/2$. We begin with two preliminary lemmas. First, for each $i \in I$ and $r_i \in \mathcal{A}_i$, we define a function $f_{i,r_i} : \mathcal{Y}_i \to \mathbb{R}$ that will later be used to specify player i's continuation payoff as a function of y_i.

Lemma 5 *Under η-individual identifiability, for each $i \in I$ and $r_i \in \mathcal{A}_i$ there exists a function $f_{i,r_i} : \mathcal{Y}_i \to \mathbb{R}$ such that*

$$\begin{aligned}
\mathbb{E}\left[f_{i,r_i}(y_i)|r_i\right] - \mathbb{E}\left[f_{i,r_i}(y_i)|a_i\right] &\geq \bar{u} \quad \textit{for all } a_i \neq r_i, && (16)\\
\mathbb{E}\left[f_{i,r_i}(y_i)|r_i\right] &= 0, && (17)\\
\mathrm{Var}\left(f_{i,r_i}(y_i)|r_i\right) &\leq \bar{u}^2/\eta, \quad \textit{and} && (18)\\
\left|f_{i,r_i}(y_i)\right| &\leq 2\bar{u}/\eta \quad \textit{for all } y_i. && (19)
\end{aligned}$$

Proof. Fix i and r_i. Let $\mathcal{Y}_i^* = \{y_i : p_i(y_i, r_i) \geq \eta\}$, and let

$$p_i(r_i) = \left(\sqrt{p_i(y_i|r_i)}\right)_{y_i \in \mathcal{Y}_i^*} \quad \text{and} \quad P_i(r_i) = \bigcup_{a_i \neq r_i} \left(\frac{p_i(y_i|a_i)}{\sqrt{p_i(y_i|r_i)}}\right)_{y_i \in \mathcal{Y}_i^*}.$$

Note that (12) is equivalent to $d\left(p_i(r_i), \mathrm{co}\left(P_i(r_i)\right)\right) \geq \sqrt{\eta}$ for all $i \in I, r_i \in \mathcal{A}_i$, where $d(\cdot,\cdot)$ denotes Euclidean distance in $\mathbb{R}^{|\mathcal{Y}_i^*|}$. Hence, by the separating hyperplane theorem, there exists $x = (x(y_i))_{y_i \in \mathcal{Y}_i^*} \in \mathbb{R}^{|\mathcal{Y}_i^*|}$ such that $\|x\| = 1$ and $(p_i(r_i) - p) \cdot x \geq \sqrt{\eta}$ for all $p \in P_i(r_i)$. By definition of p_i and P_i, this implies that $\sum_{y_i \in \mathcal{Y}_i^*} \left(p_i(y_i|r_i) - p_i(y_i|a_i)\right) x(y_i) \geq \sqrt{\eta p_i(y_i|r_i)}$ for all $a_i \neq r_i$. Now define

$$\begin{aligned}
f_{i,r_i}(y_i) &= \frac{\bar{u}}{\sqrt{\eta}} \left(\frac{x(y_i)}{\sqrt{p_i(y_i|r_i)}} - \sum_{\tilde{y}_i \in \mathcal{Y}_i} \frac{p(\tilde{y}_i|r_i)}{\sqrt{p_i(\tilde{y}_i|r_i)}} x_i(\tilde{y}_i)\right) \quad \text{for all } y_i \in \mathcal{Y}_i^*, \quad \text{and} \\
f_{i,r_i}(y_i) &= 0 \quad \text{for all } y_i \notin \mathcal{Y}_i^*.
\end{aligned}$$

Clearly, conditions (16) and (17) hold. Moreover, since $\mathbb{E}\left[f_{i,r_i}(y_i)|r_i\right] = 0$ and the term $\sum_{\tilde{y}_i \in \mathcal{Y}_i} \sqrt{p(\tilde{y}_i|r_i)} x_i(\tilde{y}_i)$ is independent of y_i, we have

$$\mathrm{Var}\left(f_{i,r_i}(y_i)|r_i\right) = \mathbb{E}\left[\frac{\bar{u}^2 x(y_i)^2}{\eta p_i(y_i|r_i)}\right] - \mathbb{E}\left[\frac{\bar{u} x_i(y_i)}{\sqrt{\eta p_i(y_i|r_i)}}\right]^2 \leq \frac{\bar{u}^2}{\eta} \sum_{y_i \in \mathcal{Y}_i^*} x(y_i)^2 \leq \frac{\bar{u}^2}{\eta},$$

and hence (18) holds. Finally, (19) holds since, for each $y_i \in \mathcal{Y}_i^*$,

$$\left|f_{i,r_i}(y_i)\right| \leq \left(\frac{|x(y_i)| + \sum_{\tilde{y}_i \in \mathcal{Y}_i^*} p(\tilde{y}_i|r_i)\, |x_i(\tilde{y}_i)|}{\sqrt{\eta p_i(y_i|r_i)}}\right) \bar{u} \leq \left(1 + \sum_{\tilde{y}_i \in \mathcal{Y}_i^*} p(\tilde{y}_i|r_i)\right) \frac{\bar{u}}{\eta} \leq \frac{2\bar{u}}{\eta}.$$

■

Now fix $i \in I$ and $r_i \in \mathcal{A}_i$, and suppose that $y_{i,t} \sim p_i(\cdot|r_i)$ for each period $t \in \mathbb{N}$, independently across periods (which would be the case in the repeated game if r_i were taken in every period). By (18), for any $T \in \mathbb{N}$, we have

$$\mathrm{Var}\left(\sum_{t=1}^{T} \delta^{t-1} f_{i,r_i}(y_{i,t})\right) = \sum_{t=1}^{T} \delta^{2(t-1)} \mathrm{Var}\left(f_{i,r_i}(y_{i,t})\right) \leq \frac{1-\delta^{2T}}{1-\delta^2} \frac{\bar{u}^2}{\eta}.$$

Together with (17) and (19), Bernstein's inequality now implies that, for any $T \in \mathbb{N}$ and $\bar{f} \in \mathbb{R}_+$, we have

$$\Pr\left(\sum_{t=1}^{T} \delta^{t-1} f_{i,r_i}(y_{i,t}) \geq \bar{f}\right) \leq \exp\left(-\frac{\bar{f}^2 \eta}{2\left(\frac{1-\delta^{2T}}{1-\delta^2}\bar{u}^2 + \frac{2}{3}\bar{f}\bar{u}\right)}\right). \tag{20}$$

Our second lemma fixes T and $\bar{f}$ so that the bound in (20) is sufficiently small, and some other conditions used in the proof also hold.

Lemma 6 *There exists $k > 0$ such that, whenever $(1-\delta)\log(N)/\eta < k$, there exist $T \in \mathbb{N}$ and $\bar{f} \in \mathbb{R}$ that satisfy the following three inequalities:*

$$60\bar{u}N \exp\left(-\frac{\left(\frac{\bar{f}}{3}\right)^2 \eta}{2\left(\frac{1-\delta^{2T}}{1-\delta^2}\bar{u}^2 + \frac{2}{3}\frac{\bar{f}}{3}\bar{u}\right)}\right) \leq \varepsilon, \tag{21}$$

$$8\frac{1-\delta}{1-\delta^T}\left(\bar{f} + \frac{2\bar{u}}{\eta}\right) \leq \varepsilon, \tag{22}$$

$$4\bar{u}\frac{1-\delta^T}{\delta^T} + \frac{1-\delta}{\delta^T}\left(\bar{f} + \frac{2\bar{u}}{\eta}\right) \leq \varepsilon. \tag{23}$$

Proof. Let T be the largest integer such that $8\bar{u}\left(1-\delta^T\right)/\delta^T \leq \varepsilon$, and let

$$\bar{f} = \sqrt{36\log\left(\frac{60\bar{u}}{\varepsilon}\right)\log(N)\frac{1-\delta^T}{1-\delta}\frac{\bar{u}^2}{\eta}}.$$

Note that if $(1-\delta)\log(N)/\eta \to 0$ then $1-\delta^T \to \varepsilon/(\varepsilon + 8\bar{u})$, and hence $(1-\delta)\log(N)/\left(\eta\left(1-\delta^T\right)\right) \to 0$. Therefore, there exists $k > 0$ such that, whenever $(1-\delta)\log(N)/\eta < k$, we have

$$\frac{4}{9}\sqrt{36\log\left(\frac{60\bar{u}}{\varepsilon}\right)\log(N)\frac{1-\delta}{1-\delta^T}\frac{1}{\eta}} \leq 1 \quad \text{and} \tag{24}$$

$$8\bar{u}\left(\sqrt{36\log\left(\frac{60\bar{u}}{\varepsilon}\right)\log(N)\frac{1-\delta}{1-\delta^T}\frac{1}{\eta}} + \frac{1-\delta}{1-\delta^T}\frac{2}{\eta}\right) \leq \varepsilon. \tag{25}$$

It now follows from straightforward algebra (provided in Appendix E.4) that (21)–(23) hold for every $k \geq \bar{k}$. ■

E.2 Equilibrium Construction

Fix any k, T, and $\bar{f}$ that satisfy (21)–(23), as well any $v \in B(\varepsilon)$. For each extreme point v^* of $B_v(\varepsilon/2)$, we construct a PPE in a T-period, finitely repeated game augmented with continuation values drawn from $B_v(\varepsilon/2)$ that generates payoff vector v^*. By standard arguments, this implies that $B_v(\varepsilon/2) \subseteq E(\Gamma)$, and hence that $v \in E(\Gamma)$.[27] Since $v \in B(\varepsilon)$ was chosen arbitrarily, it follows that $B(\varepsilon) \subseteq E(\Gamma)$.

Specifically, for each $\zeta \in \{-1,1\}^N$ and $v^* = \operatorname{argmax}_{v \in B_v(\varepsilon/2)} \zeta \cdot v$, we construct a public strategy profile σ in a T-period, finitely repeated game (which we call a *block strategy profile*) together with a continuation value function $\omega : H^{T+1} \to \mathbb{R}^N$ such that, letting $\psi_i\left(h^{T+1}\right) = \frac{\delta^T}{1-\delta}\left(\omega_i\left(h^{T+1}\right) - v_i^*\right)$, we have

$$\textit{Promise Keeping:} \quad v_i^* = \frac{1-\delta}{1-\delta^T}\mathbb{E}^{\sigma}\left[\sum_{t=1}^{T}\delta^{t-1}u_{i,t} + \psi_i\left(h^{T+1}\right)\right] \quad \text{for all } i, \tag{26}$$

$$\textit{Incentive Compatibility:} \quad \sigma_i \in \operatorname*{argmax}_{\tilde{\sigma}_i} \mathbb{E}^{\tilde{\sigma}_i,\sigma_{-i}}\left[\sum_{t=1}^{T}\delta^{t-1}u_{i,t} + \psi_i\left(h^{T+1}\right)\right] \quad \text{for all } i, \tag{27}$$

$$\textit{Self Generation:} \quad \zeta_i\psi_i\left(h^{T+1}\right) \in \left[-\frac{\delta^T}{1-\delta}\varepsilon, 0\right] \quad \text{for all } i \text{ and } h^{T+1}. \tag{28}$$

Fix $\zeta \in \{-1,1\}^N$ and $v^* = \operatorname{argmax}_{v \in B_v(\varepsilon/2)} \zeta \cdot v$. We construct a block strategy profile σ and continuation value function ψ which, in the next subsection, we show satisfy these three conditions. This will complete the proof of the theorem.

First, fix a correlated action profile $\bar{\alpha} \in \Delta(\mathcal{A})$ such that

$$u_i(\bar{\alpha}) = v_i^* + \zeta_i\varepsilon/2 \quad \text{for all } i, \tag{29}$$

and fix a probability distribution over static Nash equilibria $\alpha^{NE} \in \Delta\left(\prod_i \Delta(\mathcal{A}_i)\right)$ such that $u_i\left(\alpha^{NE}\right) \leq v_i^* - \varepsilon/2$ for all i. Such $\bar{\alpha}$ and α^{NE} exist because $v^* \in B_v(\varepsilon/2)$ and $B_v(\varepsilon) \subseteq F^*$.

We now construct the block strategy profile σ. For each player $i \in I$ and period $t \in \{1,\ldots,T\}$, we define a *state* $\theta_{i,t} \in \{0,1\}$ for player i in period t. The states are determined by the public history, and so are common knowledge among the players. We first specify players' prescribed actions as a function of the state, and then specify the state as a function of the public history.

Prescribed Equilibrium Actions: For each period t, let $r_t \in \mathcal{A}$ be a pure action profile which is drawn by public randomization at the start of period t from the distribution $\bar{\alpha} \in \Delta(\mathcal{A})$ fixed in (29), and let $\varrho_t^{NE} \in \prod_i \Delta(\mathcal{A}_i)$ be a mixed action profile which is drawn by public randomization at the start of period t from the distribution α^{NE}. The prescribed equilibrium actions are defined as follows.

1. If $\theta_{i,t} = 0$ for all $i \in I$, the players take $a_t = r_t$.

2. If there is a unique player i such that $\theta_{i,t} = 1$, the players take $a_t = (r_i', r_{-i,t})$ for some $r_i' \in BR_i(r_{-i,t})$ if $\zeta_i = 1$, and they take ϱ_t^{NE} if $\zeta_i = -1$, where $BR_i(r_{-i}) = \operatorname{argmax}_{a_i \in \mathcal{A}_i} u_i(a_i, r_{-i})$ is the set of i's best responses to r_{-i}.

3. If there is more than one player i such that $\theta_{i,t} = 1$, the players take ϱ_t^{NE}.

[27] Specifically, at each history h^{T+1} that marks the end of a block, public randomization can be used to select an extreme point v^* to be targeted in the following block, with probabilities chosen so that the expected payoff $\mathbb{E}[v^*]$ equals the promised continuation value $w\left(h^{T+1}\right)$.

Let $\alpha_t^* \in \prod_i \Delta(\mathcal{A}_i)$ denote the distribution of prescribed equilibrium actions, prior to public randomization z_t.

(It may be helpful to informally summarize the prescribed actions. So long as $\theta_{i,t} = 0$ for all players, the players take actions drawn from the target action distribution $\bar{\alpha}$. If $\theta_{i,t} = 1$ for multiple players, the inefficient Nash equilibrium distribution α^{NE} is played. If $\theta_{i,t} = 1$ for a unique player i, player i starts taking static best responses; moreover, if $\zeta_i = -1$ then α^{NE} is played.)

It will be useful to introduce the following additional state variable $S_{i,t}$, which summarizes player i's prescribed action as a function of $(\theta_{j,t})_{j\in I}$:

1. $S_{i,t} = 0$ if $\theta_{j,t} = 0$ for all $j \in I$, or if there exists a unique player $j \neq i$ such that $\theta_{j,t} = 1$, and for this player we have $\zeta_j = 1$. In this case, player i is prescribed to take $a_{i,t} = r_{i,t}$.

2. $S_{i,t} = NE$ if $\theta_{i,t} = 0$ and either (i) there exists a unique player j such that $\theta_{j,t} = 1$, and for this player we have $\zeta_j = -1$, or (ii) there are two distinct players j, j' such that $\theta_{j,t} = \theta_{j',t} = 1$. In this case, player i is prescribed to take $\varrho_{i,t}^{NE}$.

3. $S_{i,t} = BR$ if $\theta_{i,t} = 1$. In this case, player i is prescribed to best respond to her opponents' actions (which equal either $r_{-i,t}$ or $\varrho_{-i,t}^{NE}$, depending on ζ_i and $(\theta_{j,t})_{j\neq i}$.)

States: At the start of each period t, conditional on the public randomization draw of $r_t \in \mathcal{A}$ described above, an additional ("fictitious") random variable $\tilde{y}_t \in \mathcal{Y}$ is also drawn by public randomization, with distribution $p(\tilde{y}_t|r_t)$. That is, the distribution of the public randomization draw $\tilde{y}_t$ conditional on the draw r_t is the same as the distribution of the realized public signal profile $\tilde{y}_t$ at action profile r_t; however, the distribution of $\tilde{y}_t$ depends only on the public randomization draw r_t and not on the players' actions. For each player i and period t, let $f_{i,r_{i,t}} : \mathcal{Y}_i \to \mathbb{R}$ be defined as in Lemma 5, and let

$$f_{i,t} = \begin{cases} f_{i,r_{i,t}}(y_{i,t}) & \text{if } S_{i,t} = 0, \\ f_{i,r_{i,t}}(\tilde{y}_{i,t}) & \text{if } S_{i,t} = NE, \\ 0 & \text{if } S_{i,t} = BR. \end{cases} \tag{30}$$

Thus, the value of $f_{i,t}$ depends on the state $(\theta_{n,t})_{n\in I}$, the target action profile r_t (which is drawn from distribution $\bar{\alpha}$ as described above), the public signal y_t, and the additional variable $\tilde{y}_t$.[28] Later in the proof, $f_{i,t}$ will be a component of the "reward" earned by player i in period t, which will be reflected in player i's end-of-block continuation payoff function $\psi : H^{T+1} \to \mathbb{R}$.

We can finally define $\theta_{i,t}$ as

$$\theta_{i,t} = \mathbf{1}\left\{\exists t' \le t : \left|\sum_{t''=1}^{t'-1} \delta^{t''-1} f_{i,t''}\right| \ge \bar{f}\right\}. \tag{31}$$

[28] Intuitively, introducing the variable $\tilde{y}_t$, rather than simply using $y_{i,t}$ everywhere in (30), ensures that the distribution of $f_{i,t}$ does not depend on player i's opponents' strategies.

That is, $\theta_{i,t}$ is the indicator function for the event that the magnitude of the component of player i's reward captured by $(f_{i,t''})_{t''=1}^{t'-1}$ exceeds $\bar{f}$ at any time $t' \leq t$.

This completes the definition of the equilibrium block strategy profile σ. Before proceeding further, we note that a unilateral deviation from σ by any player i does not affect the distribution of the state vector $\left((\theta_{j,t})_{j\neq i}\right)_{t=1}^{T}$. (However, such a deviation does affect the distribution of $(\theta_{i,t})_{t=1}^{T}$.)

Lemma 7 *For any player i and block strategy $\tilde{\sigma}_i$, the distribution of the random vector $\left((\theta_{j,t})_{j\neq i}\right)_{t=1}^{T}$ is the same under block strategy profile $(\tilde{\sigma}_i, \sigma_{-i})$ as under block strategy profile σ.*

Proof. Since $\theta_{j,t} = 1$ implies $\theta_{j,t+1} = 1$, it suffices to show that, for each t, each $J \subseteq I \backslash \{i\}$, each h^t such that $J = \{j \in I \backslash \{i\} : \theta_{j,t} = 0\}$, and each z_t, the probability $\Pr\left((\theta_{j,t+1})_{j\in J} | h^t, z_t, a_{i,t}\right)$ is independent of $a_{i,t}$. Since $\theta_{j,t+1}$ is determined by h^t and $f_{j,t}$, it is enough to show that $\Pr\left((f_{j,t})_{j\in J} | h^t, z_t, a_{i,t}\right)$ is independent of $a_{i,t}$.

Recall that $S_{j,t}$ is determined by h^t, and that if $j \in J$ (that is, $\theta_{j,t} = 0$) then $S_{j,t} \in \{0, NE\}$. If $S_{j,t} = 0$ then player j takes $r_{j,t}$, which is determined by z_t, $y_{j,t}$ is distributed according to $p_j(y_{j,t}|r_{j,t})$, and $f_{j,t}$ is determined by $y_{j,t}$, independently across players conditional on z_t. If $S_{j,t} = NE$ then $\tilde{y}_{j,t}$ is distributed according to $p_j(\tilde{y}_{j,t}|r_{j,t})$, where $r_{j,t}$ is determined by z_t, and $f_{j,t}$ is determined by $\tilde{y}_{j,t}$, independently across players conditional on z_t. Thus, $\Pr\left((f_{j,t})_{j\in J} | h^t, z_t, a_{i,t}\right) = \prod_{j\neq i} \Pr(f_{j,t}|S_{j,t}, r_{j,t})$, which is independent of $a_{i,t}$ as desired. ■

Continuation Value Function: We now construct the continuation value function $\psi : H^{T+1} \to \mathbb{R}^N$. For each player i and end-of-block history h^{T+1}, player i's continuation value $\psi_i\left(h^{T+1}\right)$ will be defined as the sum of T "rewards" $\psi_{i,t}$, where $t = 1, \ldots, T$, and a constant term c_i that does not depend on h^{T+1}.

The rewards $\psi_{i,t}$ are defined as follows:

1. If $\theta_{j,t} = 0$ for all $j \in I$, then

$$\psi_{i,t} = \delta^{t-1} f_{i,r_{i,t}}(y_{i,t}). \tag{32}$$

2. If $\theta_{i,t} = 1$ and $\theta_{j,t} = 0$ for all $j \neq i$, then

$$\psi_{i,t} = \delta^{t-1}\left(u_i(\bar{\alpha}) - u_i(\alpha_t^*)\right). \tag{33}$$

3. Otherwise,

$$\psi_{i,t} = \delta^{t-1}\left(-\zeta_i \bar{u} - u_i(\alpha_t^*) + \mathbf{1}\{S_{i,t} = 0\} f_{i,r_{i,t}}(y_{i,t})\right). \tag{34}$$

The constant c_i is defined as

$$c_i = -\mathbb{E}\left[\sum_{t=1}^{T} \delta^{t-1}\left(\mathbf{1}\left\{\max_{j\neq i} \theta_{j,t} = 0\right\} u_i(\bar{\alpha}) - \mathbf{1}\left\{\max_{j\neq i} \theta_{j,t} = 1\right\} \zeta_i \bar{u}\right)\right] + \frac{1-\delta^T}{1-\delta} v_i^*. \tag{35}$$

Note that, since $u_i(\bar{\alpha})$ and v_i^* are both feasible payoffs, we have

$$|c_i| \leq 2\bar{u}\frac{1-\delta^T}{1-\delta}. \tag{36}$$

Finally, for each i and h^{T+1}, player i's continuation value at end-of-block history h^{T+1} is defined as

$$\psi_i\left(h^{T+1}\right) = c_i + \sum_{t=1}^{T} \psi_{i,t}. \tag{37}$$

E.3 Verification of the Equilibrium Conditions

We now verify that σ and ψ satisfy promise keeping, incentive compatibility, and self generation. We first show that $\theta_{i,t} = 0$ for all i and t with high probability, and then verify the three desired conditions in turn.

Lemma 8 *We have*

$$\Pr\left(\max_{i\in I, t\in\{1,\ldots,T\}} \theta_{i,t} = 0\right) \geq 1 - \frac{\varepsilon}{20\bar{u}}. \tag{38}$$

Proof. By union bound, it suffices to show that, for each i, $\Pr\left(\max_{t\in\{1,\ldots,T\}} \theta_{i,t} = 1\right) \leq \varepsilon/20\bar{u}N$, or equivalently

$$\Pr\left(\max_{t\in\{1,\ldots,T\}} \left|\sum_{t'=1}^{t} \delta^{t-1} f_{i,t'}\right| \geq \bar{f}\right) \leq \frac{\varepsilon}{20\bar{u}N}. \tag{39}$$

To see this, let $\tilde{f}_{i,t} = f_{i,r_{i,t}}(\tilde{y}_{i,t})$. Note that the variables $\left(\tilde{f}_{i,t}\right)_{t=1}^{T}$ are independent (unlike the variables $(f_{i,t})_{t=1}^{T}$). Since $\left(\tilde{f}_{i,t'}\right)_{t'=1}^{t}$ and $(f_{i,t'})_{t'=1}^{t}$ have the same distribution if $S_{i,t} \neq BR$, while $f_{i,t} = 0$ if $S_{i,t} = BR$, we have

$$\Pr\left(\max_{t\in\{1,\ldots,T\}} \left|\sum_{t'=1}^{t} \delta^{t-1} f_{i,t'}\right| \geq \bar{f}\right) \leq \Pr\left(\max_{t\in\{1,\ldots,T\}} \left|\sum_{t'=1}^{t} \delta^{t-1} \tilde{f}_{i,t'}\right| \geq \bar{f}\right). \tag{40}$$

Since $\left(\tilde{f}_{i,t}\right)_{t=1}^{T}$ are independent, Etemadi's inequality implies that

$$\Pr\left(\max_{t\in\{1,\ldots,T\}} \left|\sum_{t'=1}^{t} \delta^{t-1} \tilde{f}_{i,t'}\right| \geq \bar{f}\right) \leq 3 \max_{t\in\{1,\ldots,T\}} \Pr\left(\left|\sum_{t'=1}^{t} \delta^{t-1} \tilde{f}_{i,t'}\right| \geq \frac{\bar{f}}{3}\right). \tag{41}$$

Letting $x_{i,t} = \delta^{t-1}\tilde{f}_{i,t}$, note that $|x_{i,t}| \leq 2\bar{u}/\eta$ with probability 1 by (19), $\mathbb{E}[x_{i,t}] = 0$ by (17), and

$$\operatorname{Var}\left(\sum_{t'=1}^{t} x_{i,t'}\right) = \sum_{t'=1}^{t} \operatorname{Var}(x_{i,t'}) \leq \sum_{t'=1}^{T} \operatorname{Var}(x_{i,t'}) = \frac{1-\delta^T}{1-\delta}\frac{\bar{u}^2}{\eta} \quad \text{by (18).}$$

Therefore, by Bernstein's inequality ((20), which again applies because $\left(\tilde{f}_{i,t}\right)_{t=1}^{T}$ are independent) and (21), we have, for each $t \leq T$,

$$\Pr\left(\left|\sum_{t'=1}^{t} \delta^{t'-1} \tilde{f}_{i,t'}\right| \geq \frac{\bar{f}}{3}\right) \leq \frac{\varepsilon}{60\bar{u}N}. \tag{42}$$

Finally, (40), (41), and (42) together imply (39). ■

Incentive Compatibility: We use the following lemma (proof in Appendix E.5).

Lemma 9 *For each player i and block strategy profile σ, incentive compatibility holds (i.e., (27) is satisfied) if and only if*

$$\operatorname{supp} \sigma_i\left(h^t\right) \subseteq \underset{a_{i,t} \in \mathcal{A}_i}{\operatorname{argmax}}\, \mathbb{E}^{\sigma_{-i}}\left[\delta^{t-1} u_{i,t} + \psi_{i,t} | h^t, a_{i,t}\right] \quad \text{for all } t \text{ and } h^t. \tag{43}$$

In addition, for all $t \leq t'$ and h^t, we have

$$\mathbb{E}^{\sigma}\left[\delta^{t'-1} u_{i,t} + \psi_{i,t'} | h^t\right] = \mathbb{E}^{\sigma}\left[\delta^{t'-1}\left(\mathbf{1}\left\{\max_{j \neq i} \theta_{j,t'} = 0\right\} u_i(\bar{\alpha}) - \mathbf{1}\left\{\max_{j \neq i} \theta_{j,t'} = 1\right\} \zeta_i \bar{u}\right) | h^t\right]. \tag{44}$$

We now verify that (43) holds. Fix a player i, period t, and history h^t. We consider several cases, which parallel the definition of the reward $\psi_{i,t}$.

1. If $\theta_{j,t} = 0$ for all $j \in I$, recall that the equilibrium action profile is the r_t that is prescribed by public randomization z_t. For each action $a_i \neq r_{i,t}$, by (16) and (32), and recalling that $\bar{u} \geq \max_a u_i(a) - \min_a u_i(a)$, we have

$$\begin{aligned} & \mathbb{E}^{\sigma_{-i}}\left[\delta^{t-1} u_{i,t} + \psi_{i,t} | h^t, z_t, a_{i,t} = r_{i,t}\right] - \mathbb{E}^{\sigma_{-i}}\left[\delta^{t-1} u_{i,t} + \psi_{i,t} | h^t, z_t, a_{i,t} = a_i\right] \\ = \; & \delta^{t-1}\left(\mathbb{E}\left[u_i(r_t) + f_{i,r_{i,t}}(y_{i,t}) | a_{i,t} = r_{i,t}\right] - \mathbb{E}\left[u_i(a_i, r_{-i,t}) + f_{i,r_{i,t}}(y_{i,t}) | a_{i,t} = a_i\right]\right) \\ \leq \; & 0, \quad \text{so (43) holds.} \end{aligned}$$

2. If $\theta_{i,t} = 1$ and $\theta_{j,t} = 0$ for all $j \neq i$, then the reward $\psi_{i,t}$ specified by (33) does not depend on $y_{i,t}$. Hence, (43) reduces to the condition that every action in $\operatorname{supp} \sigma_i(h^t)$ is a static best responses to $\sigma_{-i}(h^t)$. This conditions holds for the prescribed action profile, $(r_i' \in BR_i(r_{-i,t}), r_{-i,t})$ or $\varrho_{i,t}^{NE}$.

3. Otherwise: (a) If $S_{i,t} = 0$, then (43) holds because it holds in Case 1 above and (32) and (34) differ only by a constant independent of $y_{i,t}$. (b) If $S_{i,t} \neq 0$, then either $\theta_{j,t} = \theta_{j',t} = 1$ for distinct players j, j', or there exists a unique player $j \neq i$ with $\theta_{j,t} = 1$, and for this player we have $\zeta_j = -1$. In both cases, ϱ_t^{NE} is prescribed. Since the reward $\psi_{i,t}$ specified by (34) does not depend on $y_{i,t}$, (43) reduces to the condition that every action in $\operatorname{supp} \sigma_i(h^t)$ is a static best responses to $\sigma_{-i}(h^t)$, which holds for the prescribed action profile ϱ_t^{NE}.

Promise Keeping: This essentially holds by construction: we have

$$\begin{aligned}
& \frac{1-\delta}{1-\delta^T}\mathbb{E}^{\sigma}\left[\sum_{t=1}^{T}\delta^{t-1}u_{i,t}+\psi_i\left(h^{T+1}\right)\right] \\
= & \frac{1-\delta}{1-\delta^T}\left(\mathbb{E}^{\sigma}\left[\sum_{t=1}^{T}\left(\delta^{t-1}u_{i,t}+\psi_{i,t}\right)\right]+c_i\right) \quad \text{(by (37))} \\
= & \frac{1-\delta}{1-\delta^T}\mathbb{E}^{\sigma}\left[\sum_{t=1}^{T}\delta^{t-1}\left(\mathbf{1}\left\{\max_{j\neq i}\theta_{j,t}=0\right\}u_i\left(\bar{\alpha}\right)-\mathbf{1}\left\{\max_{j\neq i}\theta_{j,t}=1\right\}\zeta_i\bar{u}\right)+c_i\right] \quad \text{(by (44))} \\
= & v_i^* \quad \text{(by (35)), so (26) holds.}
\end{aligned}$$

Self Generation: We use the following lemma (proof in Appendix E.6).

Lemma 10 *For every end-of-block history h^{T+1}, we have*

$$\begin{aligned}
\zeta_i\sum_{t=1}^{T}\psi_{i,t} &\leq \bar{f}+\frac{2\bar{u}}{\eta} \quad \textit{and} && (45)\\
\left|\sum_{t=1}^{T}\psi_{i,t}\right| &\leq \bar{f}+\frac{2\bar{u}}{\eta}+2\bar{u}\frac{1-\delta^T}{1-\delta}. && (46)
\end{aligned}$$

In addition,

$$\zeta_i c_i \leq -\frac{1-\delta^T}{1-\delta}\frac{\varepsilon}{8}. \tag{47}$$

To establish self generation ((28)), it suffices to show that, for each h^{T+1}, $\zeta_i\psi_i\left(h^{T+1}\right)\leq 0$ and $\left|\psi_i\left(h^{T+1}\right)\right|\leq\left(\delta^T/\left(1-\delta\right)\right)\varepsilon$. This now follows because

$$\begin{aligned}
\zeta_i\psi_i\left(h^{T+1}\right) &= \zeta_i\left(c_i+\sum_{t=1}^{T}\psi_{i,t}\right)\leq-\frac{1-\delta^T}{1-\delta}\frac{\varepsilon}{8}+\bar{f}+2\bar{u}/\eta \quad \text{(by (45) and (47))} \\
&\leq \frac{1-\delta^T}{8\left(1-\delta\right)}\left(-\varepsilon+8\left(\frac{1-\delta}{1-\delta^T}\right)\left(\bar{f}+2\bar{u}/\eta\right)\right)\leq 0 \quad \text{(by (22)),} \quad \text{and} \\
\left|\psi_i\left(h^{T+1}\right)\right| &\leq \left|c_i\right|+\left|\sum_{t=1}^{T}\psi_{i,t}\right| \\
&\leq 4\bar{u}\frac{1-\delta^T}{1-\delta}+\bar{f}+2\bar{u}/\eta \quad \text{(by (36) and (46))} \\
&= \frac{1-\delta^T}{1-\delta}4\bar{u}+\bar{f}+2\bar{u}/\eta\leq\frac{\delta^T}{1-\delta}\varepsilon \quad \text{(by (23)),}
\end{aligned}$$

which completes the proof.

E.4 Omitted Details for the Proof of Lemma 6

We show that, with the stated definitions of T and $\bar{f}$, (24) and (25) imply (21)–(23). First, note that

$$\frac{1-\delta^2}{1-\delta^{2T}} = \frac{(1+\delta)(1-\delta)}{\left(1+\delta^T\right)\left(1-\delta^T\right)} < 2\frac{1-\delta}{1-\delta^T}.$$

Hence,

$$\begin{aligned}
\frac{2\bar{f}\left(1-\delta^2\right)}{9\bar{u}\left(1-\delta^{2T}\right)} &< \frac{4}{9\bar{u}}\frac{1-\delta}{1-\delta^T}\sqrt{36\log\left(\frac{60\bar{u}}{\varepsilon}\right)\log(N)\frac{1-\delta^T}{1-\delta}\frac{\bar{u}^2}{\eta}} \\
&= \frac{4}{9}\sqrt{36\log\left(\frac{60\bar{u}}{\varepsilon}\right)\log(N)\frac{1-\delta}{1-\delta^T}\frac{1}{\eta}} \leq 1 \quad \text{(by (24))}.
\end{aligned}$$

Therefore,

$$60\bar{u}N\exp\left(\frac{-\left(\frac{\bar{f}}{3}\right)^2\eta}{2\left(\frac{1-\delta^{2T}}{1-\delta^2}\bar{u}^2+\frac{2}{3}\frac{\bar{f}}{3}\bar{u}\right)}\right) \leq 60\bar{u}N\exp\left(\frac{-\left(\frac{\bar{f}}{3}\right)^2\eta}{2\left(\frac{1-\delta^{2T}}{1-\delta^2}\bar{u}^2+\frac{1-\delta^{2T}}{1-\delta^2}\bar{u}^2\right)}\right) = 60\bar{u}N\exp\left(\frac{-\bar{f}^2\eta}{36\frac{1-\delta^{2T}}{1-\delta^2}\bar{u}^2}\right).$$

Moreover,

$$\frac{\bar{f}^2\eta}{36\frac{1-\delta^{2T}}{1-\delta^2}\bar{u}^2} = \frac{36\log\left(\frac{60\bar{u}}{\varepsilon}\right)\log(N)\frac{1-\delta^T}{1-\delta}}{36\frac{1-\delta^{2T}}{1-\delta^2}} = \frac{1+\delta}{1+\delta^T}\log\left(\frac{60\bar{u}}{\varepsilon}\right)\log(N) \geq \log\left(\frac{60\bar{u}}{\varepsilon}\right)\log(N).$$

Hence, we have

$$60\bar{u}N\exp\left(\frac{-\left(\frac{\bar{f}}{3}\right)^2\eta}{2\left(\frac{1-\delta^{2T}}{1-\delta^2}\bar{u}^2+\frac{2}{3}\frac{\bar{f}}{3}\bar{u}\right)}\right) \leq 60\bar{u}N\exp\left(-\log\left(\frac{60\bar{u}}{\varepsilon}\right)\log(N)\right) = \varepsilon.$$

This establishes (21).

Next, we have

$$8\frac{1-\delta}{1-\delta^T}\left(\bar{f}+\frac{2\bar{u}}{\eta}\right) = 8\bar{u}\left(\sqrt{36\log\left(\frac{60\bar{u}}{\varepsilon}\right)\log(N)\frac{1-\delta}{1-\delta^T}\frac{1}{\eta}}+\frac{1-\delta}{1-\delta^T}\frac{2}{\eta}\right) \leq \varepsilon \quad \text{(by (25))}. \tag{48}$$

This establishes (22).

Finally, by (48) and $8\bar{u}\left(1-\delta^T\right)/\delta^T \leq \varepsilon$, we have

$$4\bar{u}\frac{1-\delta^T}{\delta^T}+\frac{1-\delta}{\delta^T}\left(\bar{f}+\frac{2\bar{u}}{\eta}\right) = 4\bar{u}\frac{1-\delta^T}{\delta^T}+\frac{1-\delta^T}{\delta^T}\frac{1-\delta}{1-\delta^T}\left(\bar{f}+\frac{2\bar{u}}{\eta}\right) \leq 4\frac{\varepsilon}{8}+\frac{\varepsilon}{8}\frac{\varepsilon}{8} \leq \varepsilon.$$

This establishes (23).

E.5 Proof of Lemma 9

We show that player i has a profitable one-shot deviation from σ_i at some history h^t if and only if (43) is violated at h^t. To see this, we first calculate player i's continuation payoff under σ from period $t+1$ onward (net of the constant c_i and the rewards already accrued $\sum_{t'=1}^{t}\psi_{i,t'}$). For each $t' \geq t+1$, there are several cases to consider.

1. If $\theta_{j,t'} = 0$ for all j, then by (17) and (32) we have
$\mathbb{E}^{\sigma}\left[\delta^{t'-1}u_{i,t'} + \psi_{i,t'}|h^{t'}\right] = \delta^{t'-1}\left(u_i\left(\alpha^*_{t'}\right) + \mathbb{E}\left[f_{i,r_{i,t'}}\left(y_{i,t'}\right)|r_{i,t'}\right]\right) = \delta^{t'-1}u_i\left(\bar{\alpha}\right)$.

2. If $\theta_{i,t'} = 1$ and $\theta_{j,t'} = 0$ for all $j \neq i$, then by (33) we have
$\mathbb{E}^{\sigma}\left[\delta^{t'-1}u_{i,t'} + \psi_{i,t'}|h^{t'}\right] = \delta^{t'-1}\left(u_i\left(\alpha^*_{t'}\right) + u_i\left(\bar{\alpha}\right) - u_i\left(\alpha^*_{t'}\right)\right) = \delta^{t'-1}u_i\left(\bar{\alpha}\right)$.

3. Otherwise: (a) If $S_{i,t'} = 0$, then by (17) and (34) (and recalling that player i's equilibrium action is $r_{i,t'}$ when $S_{i,t'} = 0$) we have
$\mathbb{E}^{\sigma}\left[\delta^{t'-1}u_{i,t'} + \psi_{i,t'}|h^{t'}\right] = \delta^{t'-1}\left(u_i\left(\alpha^*_{t'}\right) - \zeta_i\bar{u} - u\left(\alpha^*_{t'}\right) + \mathbb{E}\left[f_{i,r_{i,t'}}\left(y_{i,t'}\right)|r_{i,t'}\right]\right) = \delta^{t'-1}\left(-\zeta_i\bar{u}\right)$.
(b) If $S_{i,t'} \neq 0$, then by (34) we have
$\mathbb{E}^{\sigma}\left[\delta^{t'-1}u_{i,t'} + \psi_{i,t'}|h^{t'}\right] = \delta^{t'-1}\left(u_i\left(\alpha^*_{t'}\right) - \zeta_i\bar{u} - u\left(\alpha^*_{t'}\right)\right) = \delta^{t'-1}\left(-\zeta_i\bar{u}\right)$.

In total, (44) holds, and player i's net continuation payoff under σ from period $t+1$ onward equals

$$\mathbb{E}^{\sigma}\left[\sum_{t'=t+1}^{T}\delta^{t'-1}\left(\mathbf{1}\left\{\max_{j\neq i}\theta_{j,t'} = 0\right\}u_i\left(\bar{\alpha}\right) - \mathbf{1}\left\{\max_{j\neq i}\theta_{j,t'} = 1\right\}\zeta_i\bar{u}\right)|h^t\right].$$

By Lemma 7, the distribution of $\left(\left(\theta_{n,t'}\right)_{n\neq i}\right)_{t'=t+1}^{T}$ does not depend on player i's period-t action, and hence neither does player i's net continuation payoff under σ from period $t+1$ onward. Therefore, player i's period-t action $a_{i,t}$ maximizes her continuation payoff from period t onward if and only if it maximizes $\mathbb{E}^{\sigma_{-i}}[\delta^{t-1}u_{i,t} + \psi_{i,t}|h^t, a_{i,t}]$.

E.6 Proof of Lemma 10

Define

$$\begin{aligned}
\psi^v_{i,t} &= \begin{cases}\delta^{t-1}\left(u_i\left(\bar{\alpha}\right) - u_i\left(\alpha^*_t\right)\right) & \text{if } \theta_{j,t} = 0 \text{ for all } j \neq i,\\ \delta^{t-1}\left(-\zeta_i\bar{u} - u_i\left(\alpha^*_t\right)\right) & \text{otherwise,}\end{cases} \qquad \text{and}\\
\psi^f_{i,t} &= \begin{cases}\delta^{t-1}f_{i,a_{i,t}}\left(y_{i,t}\right) & \text{if either } \theta_{j,t} = 0 \text{ for all } j \text{ or } S_{i,t} = 0,\\ 0 & \text{otherwise.}\end{cases}
\end{aligned}$$

Note that, by (32)–(34), we can write $\psi_{i,t} = \psi^v_{i,t} + \psi^f_{i,t}$. (Note that, if $\theta_{n,t} = 0$ for all $n \in I$, we have $\alpha^*_t = \bar{\alpha}$ and hence $\psi^v_{i,t} + \psi^f_{i,t} = \delta^{t-1}f_{i,a_{i,t}}\left(y_{i,t}\right)$, as specified in (32).) We show that, for every end-of-block history h^{T+1}, we have

$$\zeta_i \sum_{t=1}^{T} \psi_{i,t}^{v} \in \left[-2\bar{u}\frac{1-\delta^T}{1-\delta}, 0\right] \quad \text{and} \tag{49}$$

$$\left|\zeta_i \sum_{t=1}^{T} \psi_{i,t}^{f}\right| \leq \bar{f} + \frac{2\bar{u}}{\eta}. \tag{50}$$

Since $\psi_{i,t} = \psi_{i,t}^{v} + \psi_{i,t}^{f}$, (49) and (50) imply (45) and (46), which proves the first part of the lemma.

For (49), note that, by definition of the prescribed equilibrium actions, if $\theta_{j,t} = 0$ for all $j \neq i$, then (i) if $\zeta_i = 1$, we have $u_i(\alpha_t^*) \geq \sum_a \bar{\alpha}(a) \min\left\{u_i(a), \max_{a_i'} u_i(a_i', a_{-i})\right\} \geq u_i(\bar{\alpha})$; and (ii) if $\zeta_i = -1$, we have $u_i(\alpha_t^*) \leq \max\left\{u_i(\bar{\alpha}), u_i\left(\alpha^{NE}\right)\right\} = u_i(\bar{\alpha})$. In total, we have $\zeta_i(u_i(\bar{\alpha}) - u_i(\alpha_t^*)) \leq 0$. Since obviously $\zeta_i(u_i(\bar{\alpha}) - u_i(\alpha_t^*)) \geq -2\bar{u}$ and $-\bar{u} - \zeta_i u_i(\alpha_t^*) \geq -2\bar{u}$, we have

$$\zeta_i \psi_{i,t}^{v} = \begin{cases} \delta^{t-1}\zeta_i(u_i(\bar{\alpha}) - u_i(\alpha_t^*)) & \text{if } \theta_{j,t} = 0 \text{ for all } j \neq i, \\ \delta^{t-1}(-\bar{u} - \zeta_i u_i(\alpha_t^*)) & \text{otherwise} \end{cases} \in \left[-2\bar{u}\delta^{t-1}, 0\right].$$

For (50), note that $S_{i,t} = 0$ implies $\theta_{i,t} = 0$, and hence

$$\left|\zeta_i \sum_{t=1}^{T} \psi_{i,t}^{f}\right| \leq \left|\zeta_i \sum_{t=1}^{T} \mathbf{1}\{\theta_{i,t} = 0\}\delta^{t-1} f_{i,a_{i,t}}(y_{i,t})\right|.$$

Since $\theta_{i,t+1} = 1$ whenever $\left|\sum_{t'=1,..,t} \delta^{t-1} f_{i,a_{i,t}}(y_{i,t})\right| \geq \bar{f}$, and in addition $\left|f_{i,a_{i,t}}(y_{i,t})\right| \leq 2\bar{u}/\eta$ by (19), this inequality implies (50).

For the second part of the lemma, by (35), we have

$$\begin{aligned}
\zeta_i c_i &= \zeta_i \left(-\mathbb{E}\left[\sum_{t=1}^{T} \delta^{t-1}\left(\mathbf{1}\left\{\max_{j\neq i} \theta_{j,t} = 0\right\} u_i(\bar{\alpha}) - \mathbf{1}\left\{\max_{j\neq i} \theta_{j,t} = 1\right\}\zeta_i \bar{u}\right)\right] + \frac{1-\delta^T}{1-\delta} v_i^*\right) \\
&= \mathbb{E}\left[\sum_{t=1}^{T} \delta^{t-1}\left(\mathbf{1}\left\{\max_{j\neq i} \theta_{j,t} = 0\right\}\zeta_i(v_i^* - u_i(\bar{\alpha})) + \mathbf{1}\left\{\max_{j\neq i} \theta_{j,t} = 1\right\}\underbrace{(\bar{u} + \zeta_i v_i^*)}_{\in[0,2\bar{u}]}\right)\right] \\
&\leq \mathbb{E}\left[\sum_{t=1}^{T} \delta^{t-1}\left(\mathbf{1}\left\{\max_{j\neq i} \theta_{j,t} = 0\right\}\left(\frac{-\varepsilon}{2}\right) + \mathbf{1}\left\{\max_{j\neq i} \theta_{j,t} = 1\right\}2\bar{u}\right)\right] \quad \text{by (29)} \\
&\leq -\frac{1-\delta^T}{1-\delta}\left(\left(1 - \frac{\varepsilon}{20\bar{u}}\right)\frac{\varepsilon}{2} + \left(\frac{\varepsilon}{20\bar{u}}\right)2\bar{u}\right) \quad \text{(by (38))} \\
&\leq -\frac{1-\delta^T}{1-\delta}\frac{\varepsilon}{8} \quad (\text{as } \varepsilon < \bar{u}/2).
\end{aligned}$$