Moral Hazard and Repeated Games with Many Agents*

Takuo Sugaya — Stanford GSB
Alexander Wolitzky — MIT

August 2, 2023

Abstract

Motivated by the problem of sustaining cooperation in large communities with limited information, we analyze the relationship between population size, monitoring, and incentive instruments in moral hazard problems and repeated games with individual-level noise. We identify the *per-capita channel capacity* of the monitoring structure as a key determinant of the possibility of cooperation. In repeated games, cooperation is impossible if per-capita channel capacity is much smaller than discounting. Conversely, for a class of monitoring structures, a folk theorem holds when per-capita channel capacity is much larger than discounting. If attention is restricted to *linear perfect public equilibria* (which model collective incentive-provision), cooperation is possible only under much more severe parameter restrictions. Analogous results hold for static moral hazard problems with many agents and bounded rewards.

Keywords: moral hazard, teams, repeated games, large populations, individual-level noise, χ^2-divergence, mutual information, channel capacity, folk theorem, random monitoring, linear equilibrium, tail test, large deviations

JEL codes: C72, C73

*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, even in the absence of external contractual enforcement (Ostrom, 1990; Ellickson, 1991; Seabright, 2004). Cooperation in large groups usually seems to rely on accurate monitoring of individual agents' actions, together with sanctions that narrowly target deviators. These are key features of the community resource management settings documented by Ostrom (1990), as well as the local public goods provision environment studied by Miguel and Gugerty (2005), who in a development context found that parents who fell behind on their school fees and other voluntary contributions faced social sanctions.[1] Large cartels appear to operate on similar principles. For example, the Federation of Quebec Maple Syrup Producers—a government-sanctioned cartel that organizes more than 7,000 producers, accounting for over 90% of Canadian maple syrup production—strictly monitors its members' sales, and producers who violate its rules regularly have their sugar shacks searched and their syrup impounded, and can also face fines, legal action, and ultimately the seizure of their farms (Kuitenbrouwer, 2016; Edmiston and Hamilton, 2018). In contrast, we are not aware of any evidence that individual maple syrup producers—or the parents studied by Miguel and Gugerty, or the farmers, fishers, and herders studied by Ostrom—are motivated by the fear of starting a price war or other general breakdown of cooperation.

The principle that large-group cooperation requires precise monitoring and personalized

[1] Similar effects have also been found in the context of group lending (Karlan, 2007; Feigenberg, Field, and Pande, 2013).

sanctions seems like common sense, but it is not reflected in current repeated game models. The standard analysis of repeated games with patient players (e.g., Fudenberg, Levine, and Maskin, 1994; henceforth FLM) fixes all parameters of the game except the discount factor δ and considers the limit as $\delta \to 1$. This approach does not capture situations where, while players are patient ($\delta \approx 1$), they are not necessarily patient in comparison to the population size N (so $(1-\delta)N$ may or may not be close to 0). In addition, since standard results are based on statistical identification conditions that hold generically regardless of the number of players, they also do not capture the possibility that more information may be required to support cooperation in larger groups. Finally, since there is typically a vast multiplicity of cooperative equilibria in the $\delta \to 1$ limit, standard results also say little about what kind of strategies must be used to support large-group cooperation: for example, whether it is better to rely on personalized sanctions (e.g., fines) or collective ones (e.g., price wars).

This paper extends the standard analysis of repeated games with imperfect public monitoring by letting the population size, discount factor, stage game, and monitoring structure all vary together. These aspects of the repeated game can vary in a flexible manner: we assume only a uniform upper bound on the magnitude of the players' stage-game payoffs and a uniform lower bound on the amount of "individual-level noise." Our main results provide necessary and sufficient conditions for cooperation as a function of N, δ, and a measure of the "informativeness" of the monitoring structure. We also show that cooperation is possible only under much more restrictive conditions if society exclusively relies on collective sanctions, such as price wars (or, a la Hume, "the abandoning of the whole project"). In sum, we show that large-group cooperation requires a lot of patience and/or a lot of information, and cannot be based on collective sanctions for reasonable parameter values.

We now preview our main ideas and results. We model individual-level noise by assuming that each player i's action a_i stochastically determines an *individual-level outcome* x_i, independently across players, and that the distribution of observed signals y (the *outcome monitoring structure*) depends on the action profile $a = (a_i)$ only through the outcome profile $x = (x_i)$. This setup follows earlier work by Fudenberg, Levine, and Pesendorfer (1998; henceforth FLP) and al-Najjar and Smorodinsky (2000, 2001; henceforth A-NS). We find that a useful measure of the informativeness of the outcome monitoring structure is its *channel capacity*, C. This is a standard measure in information theory, which in our context is defined as the maximum expected reduction in uncertainty (entropy) about the outcome profile x that results from observing the signal y. Channel capacity obeys the elementary inequality $C \leq \log |Y|$, where Y is the set of possible signal realizations. Due to this inequality, using channel capacity permits more general results as compared to measuring informativeness by the number of possible signal realizations (as FLP and A-NS do). At the same time, channel capacity is convenient to work with, as it lets us apply tools from information theory such as Pinsker's inequality and the chain rule for mutual information, which play key roles in our analysis.

We begin by considering static moral hazard problems with individual-level noise, many agents, and an exogenous bound on rewards, $\bar{w}$. This static analysis develops many of the ideas used in our repeated games results in a simple and canonical context. Our first result (Theorem 1) is that if $\bar{w}^2 C/N$—the product of the square of the reward bound $\bar{w}$ and the per-capita channel capacity C/N—is small, then cooperation is impossible: all implementable outcomes are consistent with approximately myopic play. This relationship among $\bar{w}$, C, and N is tight (Theorem 2). We then model collective incentive-provision by restricting all players to receive the same reward. This restriction makes a bang-bang reward structure optimal, so increasing C beyond $\log(2)$ (i.e., one bit) is no longer valuable, but the restriction does not have a large effect on the relationship between $\bar{w}$ and N required to support cooperation. However, if in addition the expected reward is required to be bounded independent of $\bar{w}$ and N, then cooperation is impossible if there exists $\rho > 0$ such that $\bar{w}/\exp(N^{1-\rho})$ is small (Theorem 3).[2] Since this condition holds even if $N \to \infty$ much slower than $\bar{w} \to \infty$, we interpret this result as a near-impossibility theorem for large-group cooperation based on collective incentives.

We then turn to repeated games. Here, our first result (Theorem 4) is that cooperation is impossible if $C/((1-\delta)N)$ is small. Thus, the condition for cooperation in a repeated game with discount factor δ is the same as that in a static game with reward bound $\bar{w} = (1-\delta)^{-1/2}$. This result builds on a general necessary condition for cooperation in repeated games that we establish in a companion paper (Sugaya and Wolitzky, 2023a; henceforth SW). Compared to that result, the key difference is that here we consider games with individual-level 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

[2] In contrast, requiring bounded expected rewards does not effect the conclusion of Theorem 1 or (when C is proportional to N) Theorem 2.

the channel capacity of the outcome monitoring structure.

Our next result (Theorem 5) provides a partial converse to Theorem 4 for a specific monitoring structure: *random monitoring*, where in each period a certain number M out of the N players are chosen at random and their outcomes are perfectly revealed, while nothing is learned about the other players' outcomes. Under random monitoring, channel capacity is proportional to the number of monitored players M, and we show that if $M/\left(\left(1-\delta\right)N\log\left(N\right)\right)$ is large then cooperation is possible: a large set of payoffs arise as perfect equilibria in the repeated game, i.e., a folk theorem holds. This result implies that the condition on δ, N, and C in Theorem 4 is tight up to $\log\left(N\right)$ slack. Moreover, while random monitoring is admittedly special, in Appendix D we generalize Theorem 5 to a similar result that holds for any *product-structure monitoring* (Theorem 7).

Our final result (Theorem 6) considers the implications of restricting society to collective incentives in repeated games. We formalize this restriction by focusing on *linear perfect public equilibria*, where all on-equilibrium-path continuation payoff vectors lie on a line in $\mathbb{R}^N$. When the stage game is symmetric and the line in question is the 45° line, linear 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). We show that if there exists $\rho > 0$ such that $\left(1-\delta\right)\exp\left(N^{1-\rho}\right)$ is large, then all equilibrium payoffs are consistent with approximately myopic play.[3] This result is analogous to the bounded-expected-rewards case of Theorem 3.

1.1 Related Literature

Prior research on repeated games 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. Relative to SW, the current paper introduces two features that are specific to large-population games: individual-level noise and the possibility that N varies together

[3] It is well-known that strongly symmetric equilibria are typically less efficient than general perfect public equilibria. Our result is instead that the relationship between N and δ required for *any* non-trivial incentive provision differs dramatically between strongly symmetric/linear equilibria and general ones.

with discounting and monitoring. Individual-level noise is crucial for our anti-folk theorems (Theorems 1 and 4), while letting N vary with discounting and monitoring is the key novelty in our folk theorems (Theorems 5 and 7).

The most relevant folk theorems are due to FLM, Kandori and Matsushima (1998), and SW. However, these papers fix the stage game while taking $\delta \to 1$ (and also letting monitoring vary, in the case of SW), and their proof approach does not easily extend to the case where N and δ vary together. Our proof of Theorems 5 and 7 takes a different approach, which is based on "block strategies" as in Matsushima (2004) and Hörner and Olszewski (2006), and involves a novel application of some large deviations bounds.

Other than that in SW, the most relevant anti-folk theorems are those of FLP, A-NS, Pai, Roth, and Ullman (2014), and Awaya and Krishna (2016, 2019). Following earlier work by Green (1980) and Sabourian (1990), these papers establish conditions under which play is approximately myopic as $N \to \infty$ for fixed δ.[4] 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. The key difference is that these results rely on bounds on the strength of players' incentives that have a worse order in $1-\delta$ 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.[5]

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, 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 anti-folk theorem for linear equilibrium, as we explain in Section 4.4.[6]

In Sugaya and Wolitzky (2021), we studied the relationship among N, δ, and monitoring in repeated random-matching games with private monitoring and incomplete information,

[4] 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 FLP/a-NS's individual noise assumption.

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

[6] Another somewhat related question is the rate of convergence of the equilibrium payoff set as $\delta \to 1$ (Hörner and Takahashi, 2016; Sugaya and Wolitzky, 2023b).

where each player is "bad" (i.e., a *Defect* commitment type) with some probability. In that model, society has enough information to determine which players are bad after a single period of play, but this information is disaggregated, and supporting cooperation requires sufficiently quick information diffusion. In contrast, the current paper has complete information and public monitoring, so the analysis concerns monitoring precision (the "amount" of information available to society) rather than the speed of information diffusion (the "distribution" of information). In general, whether the key obstacle to cooperation is that societal information is *insufficient* or *disaggregated* distinguishes "large-population repeated games," such as FLP, A-NS, and the current paper, from "community enforcement" models, such as Kandori (1992), Ellison (1994), and our earlier paper.

Finally, we also contribute to the literature on static moral hazard in teams (Alchian and Demsetz, 1972; Holmström, 1982), especially the branch that considers the implications of limited information 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 in terms of information or attention: if a manager can extract at most C bits of information about the performance of N subordinates, then, whenever there is some noise in the subordinates' performance, the maximum number of subordinates that the manager can incentivize with any contract with bounded rewards is of order C.

2 Moral Hazard in Large Teams

We begin by considering static games with individual-level noise, imperfect monitoring, and bounded rewards.

The Game. The game to be played involves a finite set of players $I = \{1, \ldots, N\}$, a finite product set of actions $A = \times_{i \in I} A_i$, and a payoff function $u_i : A \to \mathbb{R}$ for each $i \in I$. The interpretation is that $u_i(a)$ is player i's expected payoff at action profile a. We denote the range of player i's payoff function by $\bar{u}_i = \max_{a,a'} u_i(a) - u_i(a')$. We also assume that $|A_i| \geq 2$ for all i.

Noise. There is a finite product set of *individual outcomes* $X = \times_{i \in I} X_i$ and a rowstochastic *noise matrix* $\pi^i \in [0,1]^{A_i \times X_i}$ for each player i such that, when action profile $a \in A$ is played, outcome profile $x \in X$ is realized with probability $\pi_{a,x} = \prod_i \pi^i_{a_i,x_i}$. We call the pair (X, π) a *noise structure*. Let $\underline{\pi}^i = \min_{a_i,x_i} \pi^i_{a_i,x_i}$ and assume that $\min_i \underline{\pi}^i > 0$: we call this assumption *individual-level noise*. The point of this setup is that signals will depend on a only through x. Note that, since $|A_i| \geq 2$ for all i, we have $\underline{\pi} \leq 1/2$.

For a natural example of a noise structure, suppose that there is some 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, $X = A$, and $\pi_{a_i,a_i'}$ is the probability that player i "trembles" to a_i' when she intends to take a_i. We refer to this example as *noisy actions*.

For any $\bar{u} > 0$ and $\underline{\pi} > 0$, the game is $(\bar{u}, \underline{\pi})$*-bounded* if the range of payoffs is bounded above by $\bar{u}$ and individual-level noise is bounded below by $\underline{\pi}$: that is, if $\bar{u}_i \leq \bar{u}$ and $\underline{\pi}_i \geq \underline{\pi}$ for all i. We also call a game $\bar{u}$*-bounded* or $\underline{\pi}$*-bounded* when only one of these bounds is imposed. Note that $\underline{\pi}$-boundedness implies that $|X_i| \leq 1/\underline{\pi}$ for all i.

Monitoring. An *outcome monitoring structure* (Y, q) consists of a finite set of possible signal realizations Y and a family of conditional probability distributions $q(y|x)$. The signal distribution thus depends only on the realized outcome profile. The outcome monitoring structure (Y, q) is a primitive object in our model: we are interested in properties of (Y, q) that (together with the other model primitives) are necessary or sufficient for supporting cooperative outcomes.

Given an outcome monitoring structure (Y, q), we denote the probability of signal profile y at action profile a by $p(y|a) = \sum_x \pi_{a,x} q(y|x)$. We refer to the pair (Y, p) as the *action monitoring structure* induced by (X, π, Y, q). The action monitoring structure (Y, p) is a derived object in our model: it plays an important role in our analysis, but we will avoid imposing assumptions directly on (Y, p), and instead consider the implications of properties of the noise structure (X, π) and the outcome monitoring structure (Y, q) for (Y, p).[7]

Without loss, we assume that for every $y \in Y$, there exists $x \in X$ such that $q(y|x) > 0$. Since $\underline{\pi}^i > 0$ for each i, this implies that p has full support: $p(y|a) > 0$ iff $y \in Y$.

Contracts. A *contract for player* i is a function $w_i : A \times Y \to \mathbb{R}$ specifying a *reward*

[7] In SW, we allow general monitoring structures and directly consider properties of the action monitoring structure (Y, p). The current paper imposes the additional structure that (Y, p) factors into a noise structure (X, π) and an outcome monitoring structure (Y, q). This additional structure lets us formulate the individual-level noise assumption.

$w_i(a, y)$ 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 non-negative and bounded from above: there exists $\bar{w} > 0$ such that $w_i(a, y) \in [0, \bar{w}]$ for all i, a, y. We say that 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 $s_i : A_i \to \Delta(A_i)$. The interpretation is that when player i is recommended a_i, she instead plays $s_i(a_i)$. A distribution over action profiles $\alpha \in \Delta(A)$ is a *correlated equilibrium* if there exists a contract w such that, for any player i and manipulation s_i,

$$\sum_{a,y} \alpha(a) \left(u_i(a) + p(y|a) w_i(a, y)\right) \geq \sum_{a,y} \alpha(a) \left(u_i(s_i(a_i), a_{-i}) + p(y|s_i(a_i), a_{-i}) w_i(a, y)\right). \quad (1)$$

[8]

We say that a correlated equilibrium is *public* if $\alpha \in \prod_i \Delta(A_i)$ and there exists a public contract satisfying (1) for all i, s_i.

Mutual Information and Channel Capacity. Given a distribution of outcomes $\xi \in \Delta(X)$, a standard measure of the informativeness of a signal y about the realized outcome x is the *mutual information* between x and y, defined as

$$\mathbf{I}(\xi) = \sum_{x \in X, y \in \bar{Y}} \xi(x) q(y|x) \log \left(\frac{q(y|x)}{\sum_{x' \in X} \xi(x') q(y|x')} \right).$$

[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 (X, π) by

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

[8] Here and throughout, u_i and p are linearly extended to mixed actions, as usual.

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

Finally, define the *channel capacity* of the tuple (X,π,Y,q) as

$$C=\max_{\xi\in\vartheta}\mathbf{I}\left(\xi\right).$$

Channel capacity is an exogenous measure of the informativeness of y about x, as it is defined as a function of only the noise structure (X,π) and the outcome monitoring structure (Y,q).[10] Note that C is no greater than the entropy of the signal y, which in turn is at most $\log|Y|$ (Theorem 2.6.3 of Cover and Thomas, 2006; henceforth CT). Channel capacity plays a central role in information theory, because it is the maximum rate at which information can be transmitted over a noisy channel (*Shannon's channel coding theorem*, CT 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 arises in our analysis because it obeys useful properties, in particular the *chain rule* (CT, Theorem 2.5.2) and *Pinsker's inequality* (CT, Lemma 11.6.1). These properties play key roles in our analysis.[11]

Almost-Myopic Play. Player i's *gain* from manipulation s_i at an action profile distribution $\alpha\in\Delta\left(A\right)$ is

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

Note that α is a correlated equilibrium iff there exists w such that

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

Player i's *maximum gain* at $\alpha\in\Delta\left(A\right)$ is $\bar{g}_i\left(\alpha\right)=\max_{s_i:A_i\to\Delta(A_i)}g_i\left(s_i,\alpha\right)$. For any $\varepsilon>0$,

[10] We define C as the maximum of $\mathbf{I}\left(\zeta\right)$ over $\zeta\in\vartheta$ rather than $\zeta\in\Delta\left(X\right)$, because only $\zeta\in\vartheta$ can ever arise. This definition makes our results stronger than they would be if we instead took the maximum over all $\zeta\in\Delta\left(X\right)$.

[11] We are not aware of prior papers that employ entropy methods in static moral hazard problems. In repeated games, these 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, our analysis is not very related to these papers'.

the set of ε-*myopic action distributions* is

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

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

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

That is, 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 small deviation gains. Otherwise, it allows a few players to have large gains. We discuss this point further in Section 5.1, following our results.

3 Conditions for (Non-)Cooperation in Moral Hazard

Our first result says that if per-capita channel capacity is much smaller than the ratio of noise and the (squared) maximum reward, then equilibrium play is almost-myopic. Cooperation in large groups thus requires a lot of information or large rewards.

Theorem 1 *Every correlated equilibrium in a $\underline{\pi}$-bounded game is ε-myopic, for*

$$\varepsilon = (1 - 2\underline{\pi}) \sqrt{\frac{2\bar{w}^2 C}{\underline{\pi} N}}.$$

Theorem 1 is similar to earlier results by FLP and A-NS. The main difference is measuring information by channel capacity rather than the number of possible signal realizations. The approach of these papers would yield $\log|Y|$ in place of $\sqrt{C}$ in Theorem 1, which gives a considerably weaker result as $\sqrt{C} \ll C \leq \log|Y|$. Moreover, in addition to yielding a stronger result, information theory also allows a shorter proof.[12]

[12] Theorem 1 is also closely related to Lemma 3 of Sugaya and Wolitzky (2021), which is a general result on the maximum average "influence" of N binary random variables on a signal with channel capacity C.

Proof. For any $a \in A$, $i \in I$, and $a_i' \in A_i$, we have

$$
\begin{aligned}
& \sum_y \left(p\left(y|a\right) - p\left(y|a_i', a_{-i}\right)\right)_+ \\
= & \sum_y \left(\sum_{x_i} \left(\pi_{a_i,x_i} - \pi_{a_i',x_i}\right) \Pr\left(y|x_i\right) \right)_+ \\
= & \sum_y \left(\sum_{x_i} \left(\pi_{a_i,x_i} - \pi_{a_i',x_i}\right) \left(\Pr\left(y|x_i\right) - \Pr\left(y\right)\right) \right)_+ \\
\le & \sum_y \sum_{x_i} \left(\pi_{a_i,x_i} - \pi_{a_i',x_i}\right)_+ \left(\Pr\left(y|x_i\right) - \Pr\left(y\right)\right)_+ + \sum_y \sum_{x_i} \left(\pi_{a_i,x_i} - \pi_{a_i',x_i}\right)_- \left(\Pr\left(y|x_i\right) - \Pr\left(y\right)\right)_- \\
= & \sum_{x_i} \left(\pi_{a_i,x_i} - \pi_{a_i',x_i}\right)_+ \sum_y \left(\Pr\left(y|x_i\right) - \Pr\left(y\right)\right)_+ + \sum_{x_i} \left(\pi_{a_i,x_i} - \pi_{a_i',x_i}\right)_- \sum_y \left(\Pr\left(y|x_i\right) - \Pr\left(y\right)\right)_- \\
\le & \sum_{x_i} \left(\pi_{a_i,x_i} - \pi_{a_i',x_i}\right)_+ \max_{x_i'} \sum_y \left(\Pr\left(y|x_i'\right) - \Pr\left(y\right)\right)_+ \\
& + \sum_{x_i} \left(\pi_{a_i,x_i} - \pi_{a_i',x_i}\right)_- \min_{x_i'} \sum_y \left(\Pr\left(y|x_i'\right) - \Pr\left(y\right)\right)_- \\
\le & \left(1 - 2\underline{\pi}\right) \left(\max_{x_i'} \sum_y \left(\Pr\left(y|x_i'\right) - \Pr\left(y\right)\right)_+ - \min_{x_i'} \sum_y \left(\Pr\left(y|x_i'\right) - \Pr\left(y\right)\right)_- \right).
\end{aligned}
$$

Next, for any $x_i' \in X_i$, we have

$$
\begin{aligned}
\sum_y \left(\Pr\left(y|x_i'\right) - \Pr\left(y\right)\right)_+ & \le \sqrt{\frac{1}{2} \sum_y \Pr\left(y|x_i'\right) \log\left(\frac{\Pr\left(y|x_i'\right)}{\Pr\left(y\right)}\right)} \\
& = \sqrt{\frac{1}{2\Pr\left(x_i'\right)} \sum_y \Pr\left(x_i', y\right) \log\left(\frac{\Pr\left(x_i', y\right)}{\Pr\left(x_i'\right)\Pr\left(y\right)}\right)} \\
& \le \sqrt{\frac{1}{2\underline{\pi}} \sum_{y,x_i} \Pr\left(x_i, y\right) \log\left(\frac{\Pr\left(x_i, y\right)}{\Pr\left(x_i\right)\Pr\left(y\right)}\right)} = \sqrt{\frac{\mathbf{I}^a\left(x_i; y\right)}{2\underline{\pi}}},
\end{aligned}
$$

where the first inequality is by Pinsker (CT, Lemma 11.6.1); the second inequality holds because, for each x_i, $\Pr\left(x_i\right) \ge \underline{\pi}$ and

$$
\sum_y \Pr\left(x_i, y\right) \log\left(\frac{\Pr\left(x_i, y\right)}{\Pr\left(x_i\right)\Pr\left(y\right)}\right) = \frac{1}{\Pr\left(x_i\right)} \sum_y \Pr\left(y|x_i\right) \log\left(\frac{\Pr\left(y|x_i\right)}{\Pr\left(y\right)}\right) \ge 0,
$$

by the information inequality (CT, Theorem 2.6.3); and the last equality is by the definition of mutual information (here $\mathbf{I}^a(\cdot,\cdot)$ denotes mutual information when action profile a is played). Similarly, $-\sum_y (\Pr(y|x_i') - \Pr(y))_- \leq \sqrt{\mathbf{I}^a(x_i;y)/2\underline{\pi}}$.

Now, for any correlated equilibrium α and accompanying contract w, and for every player i and manipulation s_i, we have

$$\begin{aligned} g_i(s_i,\alpha) &\leq \sum_{a,y} \alpha(a)\left(p(y|a) - p(y|a_i',a_{-i})\right) w_i(a,y) \\ &\leq \sum_a \alpha(a) \sum_y \left(p(y|a) - p(y|a_i',a_{-i})\right)_+ \bar{w} \leq (1-2\underline{\pi})\sqrt{\frac{2\bar{w}^2\mathbf{I}^a(x_i;y)}{\underline{\pi}}}. \end{aligned}$$

Hence,

$$\begin{aligned} \frac{1}{N}\sum_i \bar{g}_i(\alpha) &\leq \frac{1}{N}\sum_i (1-2\underline{\pi})\sqrt{\frac{2\bar{w}^2\mathbf{I}^a(x_i;y)}{\underline{\pi}}} \\ &\leq (1-2\underline{\pi})\sqrt{\frac{2\bar{w}^2\sum_i \mathbf{I}^a(x_i;y)}{\underline{\pi}N}} \\ &= (1-2\underline{\pi})\sqrt{\frac{2\bar{w}^2\mathbf{I}^a(x;y)}{\underline{\pi}N}} \leq (1-2\underline{\pi})\sqrt{\frac{2\bar{w}^2 C}{\underline{\pi}N}}, \end{aligned}$$

where the second inequality is by Jensen; the equality is by the chain rule for mutual information (CT, Theorem 2.5.2), because (x_i) are independent conditional on a; and the second inequality is by the definition of channel capacity, because the distribution of x given a lies in ϑ . ■

Without individual-level noise, detectability cannot be bounded in terms of channel capacity, and Theorem 1 fails. For example, suppose that the stage game is an N-player prisoner's dilemma with a binary public signal y, where $y = 0$ if *every* player cooperates, and $y = 1$ if *any* player defects. Mutual cooperation is then a correlated equilibrium for a fixed value of $\bar{w}$, independent of N: if $w_i(0) = \bar{w}$ and $w_i(1) = 0$ for all i, then each player's incentive is the same as in a single agent problem. However, in this example the channel capacity of (Y, q) is only $\log 2$ (i.e., one bit). Thus, without individual noise, a monitoring structure can support strong incentives even if it not very "informative" in terms of channel capacity. In contrast, Theorem 1 shows that with individual noise, only informative signals can support strong incentives.

Our next result is that Theorem 1 is tight up to a factor of $\sqrt{\pi/2}$. Thus, for any fixed noise level, the relationship between team size, information, and the maximum reward captured by Theorem 1 is tight up to a constant factor.

Theorem 2 *For any N, $C \in (0, (\log 2) N]$, $\bar{w}$, and $\underline{\pi} \in (0, 1/2)$, there exists a $\underline{\pi}$-bounded game and a correlated equilibrium that is not ε-myopic, for*

$$\varepsilon = (1 - 2\underline{\pi}) \sqrt{\frac{\bar{w}^2 C}{N}}.$$

Proof. Let q be the smaller solution to

$$\frac{C}{N} = q \log(2q) + (1 - q) \log(2(1 - q)).$$

Note that, since $C/N \in (0, \log 2]$, we have $q \in [0, 1/2)$.

Consider an N-player prisoner's dilemma, where $A_i = X_i = Y_i = \{c, d\}$; $\pi_{a_i, x_i} = 1 - \underline{\pi}$ if $a_i = x_i$ and $\underline{\pi}$ otherwise; and $q(y|x) = \prod_i q_i$, where $q_i = 1 - q$ if $y_i = x_i$, $q_i = q$ otherwise. Suppose that each player gains g by taking d rather than c, regardless of the opponents' actions.

Note that this monitoring structure does have channel capacity C. This follows because the mutual information is maximized when the x_i's are independent *Bernoulli* $(1/2)$ variables, which gives $C = N(q \log(2q) + (1 - q) \log(2(1 - q)))$, as desired.

Now consider the action distribution where everyone takes c with probability 1, together with the contract $w_i(a, y) = \mathbf{1}_{\{y_i = c\}} \bar{w}$. This is a correlated equilibrium iff

$$g \leq (1 - 2\pi)(1 - 2q)\bar{w}.$$

Thus, if g is such that this expression holds with equality, we have a correlated equilibrium that is not ε-myopic, for any $\varepsilon < (1 - 2\pi)(1 - 2q)\bar{w}$. It thus suffices to show that $1 - 2q > \sqrt{C/N}$, or equivalently

$$1 - 2q > \sqrt{q \log(2q) + (1 - q) \log(2(1 - q))}.$$

But it is straightforward to verify that this inequality holds for all $q \in [0, 1/2)$. ■

Our third result concerns a restricted class of equilibria, which model collective incentive provision, as in price wars a la Green and Porter (1984), or Hume's threat of "the abandoning of the whole project." We say that a public equilibrium is *linear* if the reward vectors lie on a line: for each player $i \neq 1$, there exists a constant $b_i \in \mathbb{R}$ such that, for all signals y, y', we have $w_i(y') - w_i(y) = b_i(w_1(y') - w_1(y))$. For any $c > 0$, we say that a linear equilibrium has *c-bounded expected rewards* if $\mathbb{E}[w_1(y)] \leq c$.

Cooperation in an arbitrary linear equilibrium is possible under conditions similar to those in Theorem 1 and 2. In contrast, cooperation in a linear equilibrium with bounded expected rewards is possible only if the maximum reward is extremely large relative to the population size. Later on, we will see that this near-impossibility result caries over to repeated games, where boundedness arises endogenously as an implication of self generation and promise keeping.

Theorem 3 *Fix any $\underline{\pi} > 0$, and $c > 0$.*

For any $\varepsilon > 0$, there exists $k > 0$ such that, in any $\underline{\pi}$-bounded game where

$$\frac{N}{\bar{w}^2} > k,$$

all linear equilibria are ε-myopic.

Moreover, for any $\varepsilon > 0$ and $\rho > 0$, there exists $k > 0$ such that, in any $\underline{\pi}$-bounded game where

$$\frac{\exp\left(N^{1-\rho}\right)}{\bar{w}} > k,$$

all linear equilibria with c-bounded expected rewards are ε-myopic.

Theorem 3 differs from Theorem 1 in the required relationship between N and $\bar{w}$ (under bounded expected rewards), and also in that Theorem 3 does not depend on C. Intuitively, the optimal linear equilibria take 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. Thus, Theorem 3 cannot be refined by incorporating C.

The proof of Theorem 3 is deferred to the appendix. To see the idea, suppose the game is symmetric and that $Y = X = A$ with binary actions and symmetric noise, so that $|A_i| = 2$, $\pi_{a_i,a_i} = 1 - \underline{\pi}$ and $\pi_{a_i,a_i'} = \underline{\pi}$ for all $a_i \neq a_i'$, and $q(y|x) = \mathbf{1}\{y = x\}$. Suppose we wish to enforce a symmetric pure action profile $\vec{a}_0 = (a_0, \ldots, a_0)$, where $\bar{g}_i(\vec{a}_0) = \nu$. By standard arguments, an optimal linear equilibrium takes the form of a "tail test," where $w_i(y) = \bar{w}$ for all i if the number n of players for whom $y_i = a_0$ exceeds a threshold n^*, and otherwise $w_i(y) = 0$ for all i.[13] Due to individual-level noise, when N is large the distribution of n is approximately normal, with mean $(1 - \underline{\pi})N$ and standard deviation $\sqrt{\underline{\pi}(1-\underline{\pi})N}$. Denote the threshold z-score of a tail test with threshold 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. Incentive compatibility then gives

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

This condition fails for any z^* when $N/\bar{w}^2$ is sufficiently large, which delivers the first part of the theorem. Now, if we also require that expected rewards are c-bounded, then

$$(1 - \Phi(z^*))\,\bar{w} \leq c. \tag{3}$$

Yet, the combination of (2) and (3) is extremely restrictive when N is large. Indeed, dividing (2) by (3), we obtain

$$\frac{\phi(z^*)}{1-\Phi(z^*)} \geq \frac{\nu}{c}\sqrt{\underline{\pi}(1-\underline{\pi})N}.$$

The left-hand side of this inequality is the standard normal Mills ratio, which is approximately equal to z^* when $z^* \gg 0$. Hence, z^* must increase at least linearly with $\sqrt{N}$. But since $\phi(z^*)$ decreases exponentially with z^*, and hence exponentially with N, the second part of the theorem now follows from (2).[14]

4 Repeated Games with Many Players

We now turn to repeated games with individual-level noise and imperfect monitoring.

[13] The analysis of tail tests as optimal incentive contracts under normal noise goes back to Mirrlees (1975). The logic of Theorem 3 shows 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.

[14] Conversely, if $\pi^i_{a_i,a_i}$ is sufficiently large for all i and a_i, and $\exp(N^{1+\rho})/\bar{w} \to 0$ for some $\rho > 0$, then any action profile a can be supported with bounded expected rewards by a tail test where $w_i(y) = \bar{w}$ only if $y_i = a_i$ for *every* player i.

4.1 Model

We consider stage games as in the previous sections, but replace contracts with equilibrium continuation play. Formally, a *repeated game with individual-level noise* $\Gamma = (I, A, u, X, \pi, Y, q, \delta)$ is described by a stage game (I, A, u), a noise structure (X, π), an outcome monitoring structure (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 according to $\pi_{a,x}$, and (iv) the signal y is drawn according to $q(y|x)$ and is publicly observed.[15] 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.[16] The set of PPE payoff vectors is denoted by $E \subseteq \mathbb{R}^N$. A *repeated game outcome* $\mu \in \Delta((A \times X \times Y)^\infty)$ (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 μ defines a marginal distribution over period t actions $\alpha_t^\mu \in \Delta(A)$, 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

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

[15] Since we will assume that players do not observe their own payoffs in addition to their signals, it is natural to require that players' realized payoffs are determined by their signals, and hence depend on a only through x. However, this assumption is not necessary for our analysis. In addition, our results are unchanged if each player i also observes her own individual outcome x_i. We discuss generalizations to privately observed signals y_i below.

[16] As usual, this definition allows players to consider deviations to arbitrary, non-public strategies; however, such deviations are irrelevant because, whenever a player's opponents use public strategies, she has a public strategy as a best response.

4.2 Conditions for Non-Cooperation in Repeated Games

The analogue of Theorem 1 for repeated games is as follows.

Theorem 4 *Any Nash equilibrium occupation measure in a $(\bar{u}, \bar{\pi})$-bounded repeated game is ε-myopic (and hence any Nash equilibrium payoff vector is ε-myopic), for*

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

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 almost myopic. This result is analogous to Theorem 1 with a maximum reward of $\bar{w} = (1-\delta)^{-1/2}$. This may be somewhat counterintuitive, as continuation payoffs in a repeated game are weighted by $(1-\delta)^{-1}$, not $(1-\delta)^{-1/2}$. However, under imperfect monitoring it is impossible that each period is solely responsible for determining continuation play, so an average incentive strength of $(1-\delta)^{-1}$ cannot be attained. This last point is formalized by Theorem 1 of SW, on which Theorem 4 relies.

Proof. For any player i, manipulation s_i, and action profile distribution α, define

$$\chi_i^2(s_i,\alpha) = \sum_{a,y}\alpha(a)\,p(y|a)\left(\frac{p(y|a)-p(y|s_i(a_i),a_{-i})}{p(y|a)}\right)^2.$$

(When $\alpha(a)=1$ for some action profile a, this is the χ^2-*divergence* of the manipulated signal distribution $p(\cdot|s_i(a_i),a_{-i})$ from the prescribed distribution $p(\cdot|a)$.) By Theorem 1 of SW, for any Nash equilibrium outcome μ, any player i, and any manipulation s_i, we have

$$g_i(s_i,\alpha^\mu) \le \bar{u}\sqrt{\frac{\delta}{1-\delta}\chi_i^2(s_i,\alpha^\mu)}.$$

Hence, by Jensen,

$$\frac{1}{N}\sum_i g_i(s_i,\alpha^\mu) \le \bar{u}\sqrt{\frac{\delta}{1-\delta}\frac{1}{N}\sum_i \chi_i^2(s_i,\alpha^\mu)}.$$

The proof is now completed by an application of the following lemma.

Lemma 1 *For any profile of manipulations (s_i) and any action profile distribution α, we* *have*

$$\sum_i \chi_i^2 (s_i, \alpha) \leq \frac{4C}{\underline{\pi}^2}.$$

The proof is a simple application of the Cauchy-Schwarz and Pinsker inequalities and the chain rule for mutual information, and is deferred to the appendix. ■

When N is large, the necessary condition for cooperation implied by Theorem 4—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 Y is fixed independently of N, then, since $C \leq \log |Y|$, the necessary condition implies that δ must converge to 1 at least as fast as $N \to \infty$, 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 pollution in a common water source, the output of team production, or some other aggregate statistic.

However, in other settings C scales linearly with N, so that $(1-\delta) N/C$ is small whenever players are patient (regardless of the population size). In repeated games with random matching (Kandori, 1992; Ellison, 1994; Deb, Sugaya, and Wolitzky, 2020), players match in pairs each period and $y_t^i = a_{m(i,t),t}$, where $m(i,t) \in I \backslash \{i\}$ denotes player i's period-t partner. In these games, $C = N \log |A_i|$, so per-capita channel capacity is independent of N. Intuitively, in random matching games each player gets a distinct signal of the overall action profile, so the total amount of information available to society is proportional to the population size. Channel capacity also 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 websites like eBay and AirBnB. In general, C/N may be constant in settings where players are monitored "separately," rather than being monitored jointly through an aggregate statistic.

Remark 1 *In applications like Cournot competition, pollution, or team production, the signal space may be modeled as a continuum, in which case the constraint* $C \leq \log |Y|$ *is vacuous. However, our results extend to the case where* Y *is a compact metric space and there exists another compact metric space* Z *and a function* $f^N : X^N \to Z$ *(which can vary with* N*) such that the signal distribution admits a conditional density of the form* $q_{Y|Z}(y|z)$*, where* Y*,* Z*, and* $q_{Y|Z}$ *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 term with variance fixed independent of N.) In this case,

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

which is bounded by

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

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

Remark 2 *Prior results by FLP, A-NS, and Pai, Roth, and Ullman (2014) establish anti-folk theorems as $N \to \infty$ for fixed δ. If we let N and δ vary together, the arguments in these papers could be used to show that cooperation is impossible if $(1-\delta)^2 N \to \infty$.*[17] *In contrast, Theorem 4 implies the stronger result that cooperation is impossible if $(1-\delta) N \to \infty$. The improvement comes from applying Theorem 1 of SW. Moreover, Theorem 5 will imply that the relationship between $1-\delta$ and N in Theorem 4 is tight up to* log *terms.*

Remark 3 *Theorem 4 can easily be generalized to allow private monitoring or correlation by a mediator. Indeed, the same result applies for the* blind game *where the signal y is observed only by a mediator, who privately recommends actions to the players.*[18]

4.3 Cooperation under Random Monitoring

We now give a partial converse to Theorem 4. The result we state here—Theorem 5—implies that the relationship among N, δ, and C in Theorem 4 is tight up to a $\log(N)$ term. However, Theorem 5 assumes "random monitoring," a particular monitoring structure. In Appendix D, we establish a much more general folk theorem (Theorem 7) that allows N, δ, and monitoring to vary simultaneously, and which implies Theorem 5 as a corollary.

We require some additional terminology. A monitoring structure (Y, q) has a *product structure* if there exist sets $(Y_i)_{i \in I}$ and a family of conditional distributions $(q_i(y_i|x_i))_{i,,y_{ii},x_i}$

[17] The same result obtains by applying Theorem 1 directly to a repeated game, since continuation payoffs are weighted by $(1-\delta)^{-1}$.

[18] See SW for more on the blind game.

such that $Y = \prod_i Y_i$ 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 (Y, q) has a product structure, then so does (Y, p), meaning that there exists a family of conditional distributions $(p_i(y_i|a_i))_{i,,y_ii,a_i}$ (given by $p_i(y_i|a_i) = \sum_{x_i} \pi^i_{a_i,x_i} q_i(y_i|x_i)$) such that $p(y|a) = \prod_i p_i(y_i|a_i)$ for all y, a. A particular product monitoring structure is *random monitoring*. Under random monitoring, at the end of every period a certain number M of players are selected uniformly at random, and the public signal perfectly reveals their identities and their realized individual outcomes. That is, under random monitoring of M players, $Y_i = X_i \cup \{\emptyset\}$ for all i, and

$$q_i(y_i|x_i) = \begin{cases} \frac{M}{N} & \text{if } y_i = x_i, \\ 0 & \text{if } y_i \in X_i \backslash \{x_i\}, \\ 1 - \frac{M}{N} & \text{if } y_i = \emptyset. \end{cases}$$

Note that the channel capacity of random monitoring is no more than $M \log(\max_i |X_i|)$.

We require that individual-level noise is not too extreme. Specifically, define the *maximum detectability* of a noise structure (X, π) as

$$\Delta = \sup \left\{ \tilde{\Delta} : \sum_{x_i : \pi_{a_i,x_i} \geq \tilde{\Delta}} \pi_{a_i,x_i} \left(\frac{\pi_{a_i,x_i} - \pi_{a'_i,x_i}}{\pi_{a_i,x_i}} \right)^2 \geq \tilde{\Delta} \quad \text{for all } i \in I, a_i \neq a'_i \in A_i \right\}.$$

This quantity is equal to the maximum detectability $\max_{i,s_i,\alpha} \chi^2_i(s_i, \alpha)$ (as defined in the proof of Theorem 4) of the action monitoring structure (Y, p) induced by the noise structure (X, π) together with perfect monitoring of outcomes (i.e., $q(y|x) = \mathbf{1}\{y = x\}$), when we ignore outcomes that occur with probability less than Δ. For example, with noisy actions (i.e., $X = A$), maximum detectability satisfies

$$\Delta > \min_{i, a_i \neq a'_i} \pi_{a_i,a_i} - 2\pi_{a'_i,a_i},$$

and is thus close to 1 when the "tremble probability" $1 - \pi_{a_i,a_i}$ is close to 0 for all i and a_i.[19]

Finally, denote the feasible payoff set by $F = \text{co}\left\{\{u(a)\}_{a \in A}\right\} \subseteq \mathbb{R}^N$ (where co denotes convex hull). Let $F^* \subseteq F$ denote the set of payoff vectors that weakly Pareto-dominate a

[19] This follows because $\sum_{x_i:\pi_{a_i,x_i} \geq \pi_{a_i,a_i} - 2\pi_{a'_i,a_i}} \pi_{a_i,x_i} \left(\frac{\pi_{a_i,x_i} - \pi_{a'_i,x_i}}{\pi_{a_i,x_i}} \right)^2 \geq \frac{\left(\pi_{a_i,a_i} - \pi_{a'_i,a_i}\right)^2}{\pi_{a_i,a_i}} \geq \pi_{a_i,a_i} - 2\pi_{a'_i,a_i}$.

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^*. For example, in Appendix C we consider a canonical public-goods game where each player chooses *Contribute* or *Don't Contribute*, and a player's payoff is the fraction of players who contribute less a constant $c \in (0,1)$ (independent of N) if she contributes herself. In this game, we show that for every $v \in (0, 1-c)$ there exists $\varepsilon > 0$ such that the symmetric payoff vector where all players receive payoff v lies in $B(\varepsilon)$, for all N.

Our folk theorem for random monitoring is as follows.

Theorem 5 *Fix any $\bar{u} > 0$ and $\Delta > 0$. For any $\varepsilon > 0$, there exists $k > 0$ such that, in any $\bar{u}$-bounded repeated game with random monitoring of $M \leq N$ players and a noise structure with maximum detectability Δ, where*

$$\frac{(1-\delta) N \log(N)}{M\Delta} < k,$$

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

Theorem 5 implies that the relationship among N, δ, and C in Theorem 4 is tight up to a $\log(N)$ term. To see this, note that in a $(\bar{u}, \underline{\pi})$-bounded game, random monitoring of M players has a channel capacity of at most $M \log(1/\underline{\pi})$. Thus, under random monitoring of M players with a noise structure with any fixed maximum detectability $\Delta > 0$, Theorem 4 implies that all Nash equilibrium payoff vectors are consistent with approximately myopic play if $(1-\delta) N/M \to \infty$, while Theorem 5 implies that a perfect folk theorem holds if $(1-\delta) N \log(N)/M \to 0$.

In Appendix D, we generalize Theorem 5 from random monitoring to arbitrary product-structure monitoring. This more general result (Theorem 7) is no harder to prove than Theorem 5, but it is less tightly connected to Theorem 4 because it relies on statistical conditions which are imposed directly on the action monitoring structure (Y, p). For this reason, we defer Theorem 7 to the appendix.

Theorem 5 (as well as its generalized version, Theorem 7) is a folk theorem for PPE in repeated games with public monitoring.[20] The standard proof approach, following FLM 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, FLM'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.[21]

Our proof (which is sketched in Appendix E, with details deferred to the online appendix) is instead based on the "block strategy" approach introduced by Matsushima (2004) and Hörner and Olszewski (2006) in the context of repeated games with private monitoring. We view the repeated game as a sequence of T-period blocks, where T is a number proportional to $1/(1-\delta)$. At the beginning of each block, a target payoff vector is determined by public randomization, and with high probability the players take actions throughout the block that deliver the target payoff. Players accrue promised continuation payoff adjustments whenever they are monitored, and the distribution of target payoffs in the next block is set to deliver the promised adjustments. To provide incentives, the required payoff adjustment when a player is monitored is of order N/M, the inverse of the monitoring probability. By the law of large numbers, when $T \gg N/M$, with high probability the total adjustment that a given player accrues over a T-period block is much smaller than $T \sim 1/(1-\delta)$, and is thus small enough that it can be delivered by appropriately specifying the distribution of target payoffs at the start of the next block.

The main difficulty in the proof is caused by the low-probability event that a player

[20] 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).

[21] 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 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 7 are tight up to $\log(N)$ slack, FLM's approach would instead require slack $N^2/M \geq N$. On the other hand, in SW, we extend FLM's proof to give a folk theorem where discounting and monitoring vary simultaneously for a fixed stage game. There, FLM's approach works because N is fixed.

accrues an unusually large total adjustment over a block, so that at some point there is no room to provide additional incentives. In this case, the player can no longer be incentivized to take a non-myopic best response, and all players' behavior in the current block must change. Hence, if any player's payoff adjustment is "abnormal," all players' payoffs in the block may be far from the target equilibrium payoffs.

The proof ensures that rare payoff-adjustment abnormalities do not compromise either ex ante efficiency or the players' incentives. Efficiency is preserved if the block-length T is large enough that the probability that *any* player's payoff adjustment is abnormal is small. Since the per-period payoff adjustment for each player is $O(N/M)$ and the length of a block is $O(1/(1-\delta))$, standard concentration bounds imply that the probability that a given player's payoff adjustment is abnormal is $\exp(-O(M/((1-\delta)N)))$. Hence, by the union bound, the probability that any player's adjustment is abnormal is at most $N\exp(-O(M/((1-\delta)N)))$, which converges to 0 when $(1-\delta)N\log(N)/M \to 0$. This step accounts for the $\log(N)$ gap between Theorem 4 and 5.

Finally, since all players' payoffs are affected whenever any player's payoff adjustment becomes abnormal, incentives would be threatened if a player's action influenced the probability that other players' adjustments become abnormal. We avoid this problem by letting each player's adjustment depend only on the public signals of her own actions. Such a separation of payoff adjustments across players is possible under product structure monitoring. We do not know if Theorems 5 and 7 can be extended to non-product structure monitoring without introducing qualitatively larger (i.e., polynomial) slack.

4.4 Non-Cooperation under Collective Sanctions

We now consider an arbitrary public monitoring structure. We say that a PPE is *linear* if all continuation payoff vectors lie on a line: 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 $w_i(h') - w_i(h) = b_i(w_1(h') - w_1(h))$, where $w_i(h)$ denotes player i's equilibrium continuation payoff at history h. Relabeling the players if necessary, we can take $|b_i| \leq 1$ for all i without loss. Note that if $b_i \geq 0$ for all i then the players' preferences over histories are all aligned; while if $b_i < 0$ for some i then the players can be divided into two groups with opposite preferences. This notion of linear equilibrium generalizes *strongly symmetric equilibrium* (SSE) in symmetric games, where $b_i = 1$ for all i.

Our result for linear PPE is as follows.

Theorem 6 *Fix any $\bar{u} > 0$ and $\underline{\pi} > 0$. For any $\varepsilon > 0$ and $\rho > 0$, there exists $k > 0$ such that, in any $(\bar{u}, \underline{\pi})$-bounded repeated game with public monitoring where*

$$(1 - \delta) \exp \left(N^{1-\rho}\right) > k,$$

all linear PPE payoff vectors are ε-myopic.

Theorem 6 is analogous to the bounded-expected-reward case of Theorem 3. The main difference is that, in the current result, bounded expected rewards arise endogenously as a consequence of self generation and promise keeping.

Theorem 6 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 $(1-\delta)^{-1}$.[22] As a tail test is optimal in their setting, the logic of Theorems 3 and 6 implies that 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. Theorems 3 and 6 suggest that their anti-folk theorem extends whenever variance asymptotically dominates $\left(\log (1-\delta)^{-1}\right)^{1/(1-\rho)}$ for some $\rho > 0$, while their folk theorem extends whenever variance is asymptotically dominated by $\left(\log (1-\delta)^{-1}\right)^{1/(1+\rho)}$ for some $\rho > 0$.

5 Discussion

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

Recall that Theorem 4 gives conditions under which all equilibrium payoffs lie in the set

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

[22] 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.

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" is that *all* players' deviations gains are less than ε. The corresponding payoff vectors are the *static ε-correlated equilibrium payoffs*, given by

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

We now 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 Theorem 4). We then give a condition under which maximum per-capita utilitarian welfare $\sum_i v_i/N$ is "similar" in $V(\varepsilon)$ and $CE(c\sqrt{\varepsilon})$, 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 $CE(c\sqrt{\varepsilon})$ are similar if each player's action has 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 $X = A$ and $\underline{\pi}^i = \underline{\pi} \in (0, 1/3)$ for all i. Note that this game is $(3, \underline{\pi})$-bounded.

In this game, for any $\varepsilon > 0$, when N is sufficiently large, we have $(H, B, \ldots, B) \in A(\varepsilon)$, 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, Theorem 4 does 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., $Y = \{H, L\}$, $q(y|x) = \mathbf{1}\{y = x_1\}$) has channel capacity $\log 2$ and supports the payoff vector

$$\left(\frac{1-3\underline{\pi}}{1-2\underline{\pi}}, \frac{1-3\underline{\pi}}{1-2\underline{\pi}}, \ldots, \frac{1-3\underline{\pi}}{1-2\underline{\pi}}\right), \quad \text{for all } \delta > \frac{1}{2-3\underline{\pi}} \text{ and all } N \geq 2.$$[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 as a repeated game equilibrium with "not very informative" monitoring. The reason is that only one player (the seller) is tempted to deviate at the efficient action profile, so monitoring one player suffices to support this action profile regardless of the population size (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, maximum per-capita utilitarian welfare in $V(\varepsilon)$ and $CE\left(\sqrt{8d\varepsilon}\right)$ are "similar."

Proposition 1 *Assume that per-capita externalities are bounded by d. Then, for any $\varepsilon \in (0, d)$ and any $v \in V(\varepsilon)$, there exists $v' \in CE\left(\sqrt{8d\varepsilon}\right)$ such that*

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

5.2 Conclusion

This paper has developed a theory of large-group cooperation based on moral hazard problems and repeated games with individual-level noise where the population size, discount factor, stage game, and monitoring structure all vary together in a flexible manner. Our

[23] This is a standard calcuation, which results from considering "forgiving 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 $\phi = (1-\delta)/(\delta - 3\delta\underline{\pi})$, and substituting this into the value recursion $v = (1-\delta)(1) + \delta(1 - \pi\phi)v$ yields $v = (1-3\underline{\pi})/(1-2\underline{\pi})$.

main results establish necessary and sufficient conditions for cooperation, which identify the ratio of the discount rate and the *per-capita channel capacity* of the outcome monitoring structure as a key statistic. For a class of monitoring structures, our necessary and sufficient conditions coincide up to $\log(N)$ slack. We also show that cooperation in a *linear equilibrium* is possible only under much more stringent conditions. This last result demonstrates a sense in which large-group cooperation must rely on personalized sanctions.

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 monitoring being decentralized (as in community enforcement models) and the possibility that some players may be irrational or fail to understand the equilibrium being played (as in our earlier work, Sugaya and Wolitzky, 2020, 2021). Combining these features may help develop a richer and more realistic perspective on the prospects for large-group cooperation. We have also only scratched the surface of the implications of limited monitoring precision for organizational design, for example the design of large hierarchies. This seems like another promising direction for future work.

As for applied work, more systematic empirical or experimental evidence on the determinants of cooperation in large-population repeated games would be valuable.[24] In particular, our results predict that, while either personalized or collective sanctions can work in small groups, personalized sanctions are much more effective in large groups. It would be interesting to test this prediction.

[24] 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. As explained in the introduction, 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 (without individual-level noise), where they find that larger groups cooperate less.

Appendix

A Proof of Theorem 3

Fix a linear equilibrium α 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

$$\begin{aligned} g_i(s_i, \alpha) &\leq \sum_r \alpha(r) \left(\mathbb{E}^r [b_i w(y)] - \mathbb{E}^{(s_i(r_i), r_{-i})} [b_i w(y)]\right) \\ &\leq \sum_r \alpha(r) \max_{a_i \in A_i} \left(\mathbb{E}^r [w(y)] - \mathbb{E}^{(a_i, r_{-i})} [w(y)]\right). \end{aligned}$$

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 and $\bar{w}$:

$$\max_{(Y,p),r,a,w} \frac{2}{N} \sum_i \left(\mathbb{E}^r [w(y)] - \mathbb{E}^{(a_i, r_{-i})} [w(y)]\right) \quad \text{s.t.}$$
$$w(y) \in [0, \bar{w}] \quad \text{for all } y.$$

To prove the first statement in the theorem, it suffices to show that the value of this program converges to 0 along any sequence $(N, \bar{w})$ where $N/\bar{w}^2 \to \infty$.

Since $\pi^i_{a_i, x_i} \geq \underline{\pi}$ for all i, the solution to this program involves taking $Y = X = A$, $q(y|x) = \mathbf{1}\{y = x\}$ for all y, x, $r_i \neq a_i$ for all i, $\pi^i_{r_i, r_i} = 1 - \underline{\pi}$, $\pi^i_{a_i, r_i} = \underline{\pi}$, and

$$w(y) = \begin{cases} \bar{w} & \text{if } \{i : y_i = r_i\} \geq n^*, \\ 0 & \text{if } \{i : y_i = r_i\} < n^*, \end{cases}$$

for some $n^* \in \{0, 1, \ldots, N\}$.[25] The value of the program is then

$$\max_{n^* \in \{0,1,\ldots,N\}} 2\bar{w}(1 - 2\underline{\pi}) \binom{N-1}{n^*-1} (1 - \underline{\pi})^{n^*-1} \underline{\pi}^{N-n^*}.$$

That is, the value equals $\bar{w}(1 - 2\underline{\pi})$ times the maximum of the *Binomial* $(N, \underline{\pi})$ probability mass function. The first statement in the theorem follows as the latter quantity is propor-

[25] A proof is contained in an earlier version of the paper, available on request.

tional to $N^{-(1/2)}$, by the De Moivre-Laplace theorem.

For the second statement, we impose the additional constraint that $\mathbb{E}[w(y)] \leq c$. The solution takes the same form as above, except that now we have

$$w(y) = \begin{cases} \bar{w} & \text{if } \{i : y_i = r_i\} > n^*, \\ \beta\bar{w} & \text{if } \{i : y_i = r_i\} = n^*, \\ 0 & \text{if } \{i : y_i = r_i\} < n^*, \end{cases}$$

for some $n^* \in \{0, 1, \ldots, N\}$ and $\beta \in [0,1]$, as it may be optimal to randomize the reward at the cutoff to satisfy $\mathbb{E}[w(y)] \leq c$ with equality. Letting $n = |\{i : y_i = r_i\}|$ and $n_{-i} = |\{j \neq i : y_j = r_j\}|$, the program becomes

$$\max_{n^* \in \{0,1,\ldots,N\}, \beta \in [0,1]} 2\bar{w}(1 - 2\underline{\pi})(\beta \Pr(n_{-i} = n^* - 1 | r_{-i}) + (1-\beta)\Pr(n_{-i} = n^* | r_{-i})) \tag{4}$$

$$\text{s.t.} \quad \beta \Pr(n = n^* | r) + \Pr(n \geq n^* + 1 | r) \leq \frac{c}{\bar{w}}, \tag{5}$$

where

$$\Pr(n_{-i} = n^* | r_{-i}) = \binom{N-1}{n^*} \underline{\pi}^{n^*} (1 - \underline{\pi})^{N-1-n^*} \qquad \text{and} \qquad \Pr(n = n^* | r) = \binom{N}{n^*} \underline{\pi}^{n^*} (1 - \underline{\pi})^{N-n^*}.$$

Now fix $\rho > 0$ and a sequence, indexed by k, of pairs $(N, \bar{w})$ satisfying $\exp(N^{1-\rho})/\bar{w} > k$ and pairs (n^*, β) satisfying (5). Fix $\varepsilon > 0$, and suppose toward a contradiction that, for every $\bar{k}$, there is some $k \geq \bar{k}$ such that the value of (4) exceeds ε. Taking a subsequence and relabeling $\bar{k}$ if necessary, this implies that there exists $\bar{k}$ such that, for every $k \geq \bar{k}$, the value of (4) exceeds ε.

We consider two cases and derive a contradiction in each of them.

First, suppose that there exists $d > 0$ such that, for every $\tilde{k}$, there is some $k \geq \tilde{k}$ satisfying $|\underline{\pi} - (n^* - 1)/(N-1)| > d$. By Hoeffding's inequality (Boucheron, Lugosi, and Massart, 2013, Theorem 2.8),

$$\Pr(n_{-i} \geq n^* - 1 | r_{-i}) \leq \exp\left(-2\left(\underline{\pi} - \frac{n^* - 1}{N - 1}\right)^2 (N-1)\right).$$

Hence, for every $\tilde{k}$, there is some $k \geq \tilde{k}$ such that the value of (4) is at most

$$2\bar{w}\left(1-2\underline{\pi}\right)\exp\left(-2\left(\underline{\pi}-\frac{n^*-1}{N-1}\right)^2\left(N-1\right)\right) \leq 2\bar{w}\left(1-2\underline{\pi}\right)\exp\left(-2d^2\left(N-1\right)\right).$$

Since $\exp\left(N^{1-\rho}\right)/\bar{w} \to \infty$, we have $\bar{w}\exp\left(-2d^2\left(N-1\right)\right) \to 0$ for all $d > 0$, and hence (4) is less than ε for sufficiently large k, a contradiction.

Second, suppose that for any $d > 0$, there exists $\tilde{k}$ such that, for every $k \geq \tilde{k}$, we have

$$\left|\underline{\pi}-\frac{n^*-1}{N-1}\right| \leq d. \tag{6}$$

For this case, we rely on the following lemma.

Lemma 2 *For any $m \in \mathbb{N}$ and any $\gamma > 0$, there exists $\tilde{k}$ such that, for every $k \geq \tilde{k}$, we have*

$$\frac{\beta \Pr\left(n = n^*|r\right) + \Pr\left(n \geq n^*+1|r\right)}{\beta \Pr\left(n_{-i} = n^*-1|r_{-i}\right) + \left(1-\beta\right)\Pr\left(n_{-i} = n^*|r_{-i}\right)} \geq m\left(1-\gamma\right). \tag{7}$$

Proof. Fix $d > 0$ and take k sufficiently large that (6) holds. For any $m \in \mathbb{N}$, we have

$$\begin{aligned}
\frac{\Pr\left(n \geq n^*+1|r\right)}{\Pr\left(n_{-i} = n^*|r_{-i}\right)} &= \sum_{n=n^*+1}^{N} \frac{N\left(1-\underline{\pi}\right)}{N-n^*}\frac{\left(N-n^*\right)!n^*!}{\left(N-n\right)!n!}\left(\frac{\underline{\pi}}{1-\underline{\pi}}\right)^{n-n^*} \\
&\geq \sum_{n=n^*+1}^{N} \frac{N\left(1-d\right)}{N-1}\left(\frac{N-n^*}{n}\right)^{n-n^*}\left(\frac{n^*-1-d\left(N-1\right)}{N-n^*+d\left(N-1\right)}\right)^{n-n^*} \\
&\geq \sum_{n=n^*+1}^{n^*+m} \left(1-d\right)\left(\frac{N-n^*}{n^*+m}\times\frac{n^*-1-d\left(N-1\right)}{N-n^*+d\left(N-1\right)}\right)^{m} \\
&= m\left(1-d\right)\left(\frac{N-n^*}{n^*+m}\times\frac{n^*-1-d\left(N-1\right)}{N-n^*+d\left(N-1\right)}\right)^{m}.
\end{aligned}$$

By (6), for any $\gamma' > 0$, for sufficiently large k we have $\left(n^*-1\right)/\left(n^*+m\right) \geq 1-\gamma'$, and hence

$$\begin{aligned}
\frac{N-n^*}{n^*+m}\times\frac{n^*-1-d\left(N-1\right)}{N-n^*+d\left(N-1\right)} &\geq \left(1-\gamma'\right)\frac{N-n^*}{n^*-1}\times\frac{n^*-1-d\left(N-1\right)}{N-n^*+d\left(N-1\right)} \\
&= \left(1-\gamma'\right)\frac{1-d\frac{N-1}{n^*-1}}{1+d\frac{N-1}{N-n^*}} \\
&\geq \left(1-\gamma'\right)\frac{1-\frac{d}{\underline{\pi}-d}}{1+\frac{d}{1-\underline{\pi}-d}} = \frac{\left(1-\gamma'\right)\left(\underline{\pi}-2d\right)\left(1-\underline{\pi}-d\right)}{\left(\underline{\pi}-d\right)\left(1-\underline{\pi}\right)},
\end{aligned}$$

which converges to $1-\gamma'$ as $d \to 0$. Hence, for any $\gamma > 0$, there exists $\tilde{k}$ sufficiently large such that, for every $k \geq \tilde{k}$,

$$\frac{\Pr(n \geq n^* + 1 | r)}{\Pr(n_{-i} = n^* | r_{-i})} \geq m(1-d)\left(\frac{(1-\gamma')(\underline{\pi} - 2d)(1 - \underline{\pi} - d)}{(\underline{\pi} - d)(1 - \underline{\pi})}\right)^m \geq m(1-\gamma).$$

We therefore have

$$\frac{\beta \Pr(n = n^* | r) + \Pr(n \geq n^* + 1 | r)}{\Pr(n_{-i} = n^* | r_{-i})} \geq \frac{\Pr(n \geq n^* + 1 | r)}{\Pr(n_{-i} = n^* | r_{-i})} \geq m(1-\gamma).$$

Similarly, for any m and $\gamma > 0$, there exists $\tilde{k}$ such that, for every $k \geq \tilde{k}$, we have

$$\frac{\beta \Pr(n = n^* | r) + \Pr(n \geq n^* + 1 | r)}{\Pr(n_{-i} = n^* - 1 | r_{-i})} \geq m(1-\gamma).$$

Together, these inequalities imply that, for any m and $\gamma > 0$, there exists $\tilde{k}$ such that, for every $k \geq \tilde{k}$, (7) holds. ■

Thus, for any $m \in \mathbb{N}$ and any $\gamma > 0$, there exists $\tilde{k}$ such that, for every $k \geq \tilde{k}$, the value of (4) satisfies

$$\begin{aligned}
& 2\bar{w}(1-2\underline{\pi})(\beta \Pr(n_{-i} = n^* - 1 | r_{-i}) + (1-\beta)\Pr(n_{-i} = n^* | r_{-i})) \\
\leq \; & 2c(1-2\underline{\pi})\frac{\beta \Pr(n_{-i} = n^* - 1 | r_{-i}) + (1-\beta)\Pr(n_{-i} = n^* | r_{-i})}{\beta \Pr(n = n^* | r) + \Pr(n \geq n^* + 1 | r)} \quad \text{(by (5))} \\
\leq \; & \frac{2c(1-2\underline{\pi})}{m(1-\gamma)} \quad \text{(by (7))}.
\end{aligned}$$

Taking m and γ such that $2c(1-2\underline{\pi})/(m(1-\gamma)) < \varepsilon$ gives the desired contradiction.

B Proof of Lemma 1

We establish an intermediate lemma.

Lemma 3 *For any player i, any manipulation s_i, and any action profile a, we have*

$$\chi_i^2(s_i, a) \leq \frac{4\mathbf{I}^a(x_i; y)}{\underline{\pi}^2},$$

where $\mathbf{I}^a(\cdot;\cdot)$ denotes mutual information when action profile a is played.

Proof. Let Pr denote the probability distribution over (X, Y) when a is played. For any $x_i \in X_i$ and $y \in Y$, we have $\Pr(x_i, y) = \pi_{a_i,x_i} \Pr(y|x_i) = p(y|a)\Pr(x_i|y)$. Hence, since $\pi_{a_i,x_i} \geq \underline{\pi}$, we have

$$(\Pr(y|x_i) - p(y|a))^2 = \left(\frac{p(y|a)}{\pi_{a_i,x_i}}(\Pr(x_i|y) - \pi_{a_i,x_i})\right)^2 \leq \left(\frac{p(y|a)}{\underline{\pi}}(\Pr(x_i|y) - \pi_{a_i,x_i})\right)^2. \tag{8}$$

For any player i, manipulation s_i, and action profile a, we thus have

$$\begin{aligned}
\chi_i^2(s_i, a) &= \sum_{y(\in \bar{Y})} \frac{(p(y|a) - p(y|s_i(a_i), a_{-i}))^2}{p(y|a)} = \sum_y \frac{\left(\sum_{x_i}\left(\pi_{a_i,x_i} - \pi_{s_i(a_i),x_i}\right)\Pr(y|x_i)\right)^2}{p(y|a)} \\
&= \sum_y \frac{\left(\sum_{x_i}\left(\pi_{a_i,x_i} - \pi_{s_i(a_i),x_i}\right)(\Pr(y|x_i) - p(y|a))\right)^2}{p(y|a)} \\
&\leq \sum_{x_i}\left(\pi_{a_i,x_i} - \pi_{s_i(a_i),x_i}\right)^2 \sum_y \frac{\sum_{x_i}(\Pr(y|x_i) - p(y|a))^2}{p(y|a)} \\
&\leq \frac{2}{\underline{\pi}^2}\sum_y p(y|a)\sum_{x_i}(\Pr(x_i|y) - \pi_{a_i,x_i})^2 \leq \frac{2}{\underline{\pi}^2}\sum_y p(y|a)\left(\sum_{x_i}|\Pr(x_i|y) - \pi_{a_i,x_i}|\right)^2 \\
&\leq \frac{4}{\underline{\pi}^2}\sum_y p(y|a)\sum_{x_i}\Pr(x_i|y)\log\left(\frac{\Pr(x_i|y)}{\pi_{a_i,x_i}}\right) = \frac{4\mathbf{I}^a(x_i; y)}{\underline{\pi}^2},
\end{aligned}$$

where the first inequality follows by Cauchy-Schwarz, the second follows by (8) and $\sum_{x_i}\left(\pi_{a_i,x_i} - \pi_{a_i',x_i}\right)^2 \leq 2$, the third is immediate, and the fourth follows by Pinsker. ■

We thus have

$$\sum_i \chi_i^2(s_i, a) \leq \frac{4}{\underline{\pi}^2}\sum_i \mathbf{I}^a(x_i; y) = \frac{4}{\underline{\pi}^2}\mathbf{I}^a(x; y) \leq \frac{4}{\underline{\pi}^2}C,$$

where the equality follows by the chain rule for mutual information, and the second inequality follows by the definition of channel capacity.

C The Set $B(\varepsilon)$ in A Public-Goods Game

Consider the public-goods game where each player chooses *Contribute* or *Don't Contribute*, and a player's payoff is the fraction of players who contribute less a constant $c \in (0, 1)$ (independent of N) if she contributes herself. Fix any $v \in (0, 1 - c)$, let $v = (v, \ldots, v) \in \mathbb{R}^N$, and let $\varepsilon = cv(1-c-v)/4 > 0$. We show that $B_v(\varepsilon) \subseteq F$ for all N. Since no one contributing is a Nash equilibrium with 0 payoffs, this implies that $B_v(\varepsilon) \subseteq F^*$, and hence $v \in B(\varepsilon)$, for all N.

Fix any N. Since the game is symmetric, to show that $B_v(\varepsilon) \subseteq F$ it suffices to show that, for any number $n \in \{0, \ldots, N\}$, there exists a feasible payoff vector where n "favored" players receive payoffs no less than $v + \varepsilon$, and the remaining $N - n$ "disfavored" players receive payoffs no more than $v - \varepsilon$. First, consider the mixed action profile α^1 where favored players contribute with probability $\frac{v+\varepsilon}{1-c}$ and disfavored players always contribute. At this profile, favored players receive payoff $f(n) := \frac{n}{N}\frac{v+\varepsilon}{1-c} + \left(1 - \frac{n}{N}\right)(1) - c\frac{v+\varepsilon}{1-c}$, while disfavored players receive payoff $g(n) := \frac{n}{N}\frac{v+\varepsilon}{1-c} + \left(1 - \frac{n}{N}\right)(1) - c$. Now consider the mixed action profile α^2 where favored players contribute with probability $\frac{(v+\varepsilon)^2}{(1-c)f(n)}$ and disfavored players contribute with probability $\frac{v+\varepsilon}{f(n)}$. Note that each player's payoff at profile α^2 equals her payoff at profile α^1 multiplied by $\frac{v+\varepsilon}{f(n)}$. Therefore, at profile α^2, favored players receive payoff $f(n)\frac{v+\varepsilon}{f(n)} = v + \varepsilon$, while disfavored players receive payoff

$$\begin{aligned} g(n)\frac{v+\varepsilon}{f(n)} &= \left(f(n) - \left(1 - \frac{v+\varepsilon}{1-c}\right)c\right)\frac{v+\varepsilon}{f(n)} \\ &\leq v + \varepsilon - \left(1 - \frac{v+\varepsilon}{1-c}\right)c(v+\varepsilon) \quad \text{(since } f(n) \leq 1\text{)} \\ &\leq v - \varepsilon \quad \text{(since } \varepsilon = cv(1-c-v)/4\text{)}. \end{aligned}$$

D A More General Folk Theorem

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

$$\sum_{y_i : p_i(y_i|a_i) \geq \eta^2} p_i(y_i|a_i)\left(\frac{p_i(y_i|a_i) - p_i(y_i|\alpha_i)}{p_i(y_i|a_i)}\right)^2 \geq \eta^2 \quad \text{for all } i \in I, a_i \in A_i, \alpha_i \in \Delta(A_i \backslash \{a_i\}). \tag{9}$$

This condition is a variant of FLM's individual full rank condition and Kandori and Matsushima's (1998) assumption (A2"). It says that the detectability of a deviation from a_i to any mixed action α_i supported on $A_i \backslash \{a_i\}$ is at least η^2, from the perspective of an observer who ignores 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 $(1-\delta)/\eta^2$.[26]

Our general folk theorem is as follows.

Theorem 7 *Fix any $\bar{u} > 0$. For any $\varepsilon > 0$, there exists $k > 0$ such that, for any $\bar{u}$-bounded repeated game with product structure monitoring satisfying η-individual identifiability and*

$$\frac{(1-\delta)\log(N)}{\eta^2} < k, \tag{10}$$

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

To prove Theorem 5 from Theorem 7, it suffices to show that random monitoring of M players with a noise structure with Δ detectability satisfies $\sqrt{\Delta M/N}$-individual identifiability. To see this, note that, under random monitoring of M players, we have

$$p_i(y_i|a_i) = \begin{cases} \frac{M}{N}\pi_{a_i,y_i} & \text{if } y_i \in X_i, \\ 1-\frac{M}{N} & \text{if } y_i = \emptyset. \end{cases}$$

We then have

$$\sum_{y_i: p_i(y_i|a_i) \geq \Delta M/N} p_i(y_i|a_i)\left(\frac{p_i(y_i|\alpha_i) - p_i(y_i|a_i)}{p_i(y_i|a_i)}\right)^2 = \frac{M}{N}\sum_{x_i:\pi_{a_i,x_i}\geq\Delta}\pi_{a_i,x_i}\left(\frac{\pi_{a_i,x_i}-\pi_{a_i',x_i}}{\pi_{a_i,x_i}}\right)^2 \geq \frac{\Delta M}{N}.$$

Hence, random monitoring of M players with a noise structure with Δ detectability satisfies $\sqrt{\Delta M/N}$-individual identifiability.

E Sketch of the Proof of Theorem 7

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 $B(\varepsilon) \subseteq E$.

[26] If (9) were weakened by taking the sum over all y_i (rather than only y_i such that $p_i(y_i|a_i) \geq \eta^2$), player i could be incentivized by rewards with variance $O\left((1-\delta)/\eta^2\right)$, but not necessarily with maximum absolute value $O\left((1-\delta)/\eta^2\right)$. Our analysis requires controlling both the variance and absolute value of players' rewards, so we need the stronger condition.

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

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

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

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

Since $B_v(\varepsilon/2)$ is the cube with center v and side-length ε, and $v_i^* \in \operatorname{argmax}_{w \in B_v(\varepsilon/2)} \zeta_i w_i$ for all i, we have $w\left(h^{T+1}\right) \in B_v(\varepsilon/2)$ iff $\zeta_i\left(w_i\left(h^{T+1}\right) - v_i\right) \in [-\varepsilon, 0]$ for all i. Thus, defining $\psi_i\left(h^{T+1}\right) = \left(\delta^T/(1-\delta)\right)\left(w_i\left(h^{T+1}\right) - v_i^*\right)$, 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 (9), 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 η^2. Indeed, these properties are achieved by taking $f_{i,r_i}(y_i)$ proportional to the likelihood ratio difference $\min_{\alpha_i \in \Delta(A_i\setminus\{a_i\})}\left(p_i(y_i|a_i) - p_i(y_i|\alpha_i)\right)/p_i(y_i|a_i)$. (See Lemma 5.)

Since $v \in B(\varepsilon)$ and $v^* \in B_v(\varepsilon/2)$, there exists $\bar{\alpha} \in \Delta(A)$ such that $\zeta_i\left(u_i(\bar{\alpha}) - v_i^*\right) = \varepsilon/2$. 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 $\tilde{\psi}_i\left(h^{T+1}\right) = \sum_t \delta^{t-1} f_{i,r_{i,t}}(y_{i,t}) - \zeta_i\frac{1-\delta^T}{1-\delta}\frac{\varepsilon}{2}$. We call $\tilde{\psi}_i\left(h^{T+1}\right)$ 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:

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

Next, incentive compatibility holds because

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

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 $O(\eta^2)$ and T is $O\left((1-\delta)^{-1}\right)$, by a standard concentration inequality, the self generation constraint $\zeta_i\tilde{\psi}_i\left(h^{T+1}\right) \leq 0$ holds for all i with probability at least

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

Therefore, by (10), self generation holds with high probability when k is small. (See 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

$$\zeta_i\sum_{t=1}^{\tau}\delta^{t-1}f_{i,r_{i,t}}(y_{i,t}) > \bar{f}, \tag{11}$$

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 $\tilde{\psi}_i\left(h^{T+1}\right)$ becomes abnormal. If no such period arises, define $\tau = T$. By the same concentration argument as above, abnormality does not happen to any player's base reward (that is, $\tau = T$) with high probability: in particular,

$$\Pr(\tau < T) \approx \exp\left(-\sqrt{\frac{(1-\delta)\log N}{\eta^2}}\right). \tag{12}$$

We now define the modified strategy.

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

If $\tau < T$, then let I^* be the set of players whose base reward satisfies (11). 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 $\zeta_i \left(u_i(\zeta_i) - u_i(\bar{\alpha})\right) \geq 0$. 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 (12), 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 potential 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 (27) for the definition of the fictitious recommendations and signals.) If player $j \neq i$'s fictitious base reward satisfies (11), we add or subtract a constant from player i's reward. (See (28) 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 (see Lemma 7), and hence incentive compatibility holds (Lemma 9).

F Proof of Theorem 6

Fix a linear PPE 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

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

Since $V(\varepsilon)$ is convex, it suffices to show that $\underline{v}, \bar{v} \in V(\varepsilon)$.

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

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

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

Moreover, if the constraint $\omega(r, y) \in \left[0, \frac{\delta}{1-\delta}\bar{u}\right]$ is replaced with $\omega(r, y) \in \left[-\frac{\delta}{1-\delta}\bar{u}, 0\right]$, 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 $\alpha \in \Delta(A)$ and $w : A \times Y \to E$ such that

$$\begin{aligned} v &= \mathbb{E}^{\alpha}[u(r) + \delta w(r, y)] \quad \text{and} \\ \mathbb{E}^{\alpha}[u_i(r) + \delta w_i(r, y) | r_i = a_i] &\geq \mathbb{E}^{\alpha, a_i'}[u_i(a_i', r_{-i}) + \delta w_i(r, y) | r_i = a_i] \quad \text{for all } i, a_i \in \operatorname{supp} \alpha_i, a_i' \in A_i. \end{aligned}$$

Since $v \in E$ and $w(r,y) \in E$ for all r, y, we have $v_i - w_i(r,y) = b_i(v_1 - w_1(r,y))$ for all i, r, y. Since $\bar{v}_1 \geq v_1$ for all $v \in E$, if $v = \bar{v}$ then $w_1(r,y) \leq v_1$ for all r, y. Hence, taking $v = \bar{v} = (1-\delta)u(\alpha) + \delta b \mathbb{E}[w(r,y)|\alpha]$ and defining $\omega(r,y) = \frac{\delta}{1-\delta}(\bar{v}_1 - w_1(r,y)) \in \left[0, \frac{\delta}{1-\delta}\bar{u}\right]$ for all r, y, and letting $\mathbb{E}[\cdot]$ denote expectation where $y \sim p(\cdot|a)$, we have, for all a, r,

$$\begin{aligned} u(a) - b\mathbb{E}[\omega(r,y)] &= u(a) - b\mathbb{E}\left[\frac{\delta}{1-\delta}(\bar{v}_1 - w_1(r,y))\right] \\ &= u(a) - \mathbb{E}\left[\frac{\delta}{1-\delta}(\bar{v} - w(r,y))\right] = (1-\delta)u(a) + \delta\mathbb{E}[w(r,y)], \end{aligned}$$

and the result follows. Similarly, if $v = \underline{v}$ then $w_1(r,y) \geq v_1$ 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

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

where the last line holds because $\mathbb{E}^r[\omega(y)] = u_1(r) - \bar{v}_1 \leq \bar{u}$. This is identical to the program in the bounded-expected-reward case of Theorem 3, with $\bar{w} = (\delta/(1-\delta))\bar{u}$ and $c = \bar{u}$. The result therefore follows from Theorem 3.

G Proof of Proposition 1

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 $\left|\sum_{i\in I}(v_i - v_i')\right|/N \leq 4\bar{u}/c$. (The stated conclusion follows by taking $c = \sqrt{8d/\varepsilon}$.) Fix $\varepsilon \in (0,d)$ and $\alpha \in A(\varepsilon)$. Let $J = \{i : \bar{g}_i(\alpha) > c\varepsilon/2\}$, and note that $|J| \leq 2N/c$. Let $\tilde{\alpha} \in \Delta(A)$ be an action distribution that has the same marginal on $A_{I\setminus J}$ 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 $I\setminus J$ is held fixed. Since $\left|u_i(a_j', a_{-j}) - u_i(a)\right| \leq d/N$ 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 $\bar{g}_i(\alpha) \le c\varepsilon/2$ (as $i \in I \backslash J$) and $4d/c \le c\varepsilon/2$ (as $c \ge \sqrt{8d/\varepsilon}$), we have $\bar{g}_i(\tilde{\alpha}) \le c\varepsilon$. Since we assumed that $\bar{g}_i(\tilde{\alpha}) \le c\varepsilon$ for all $i \in J$, we have $\bar{g}_i(\tilde{\alpha}) \le c\varepsilon$ 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 $|u_i(\tilde{\alpha}) - u_i(\alpha)| \le 2d/c \le 2\bar{u}/c$ for all $i \in I \backslash J$, and by definition of $\bar{u}$ we have $|u_i(\tilde{\alpha}) - u_i(\alpha)| \le \bar{u}$ for all $i \in J$. Since $c > 2$ and $|J| \le 2N/c$, we have $\left|\sum_{i \in I}(u_i(\tilde{\alpha}) - u_i(\alpha))\right| \le (N - 2N/c)\, 2\bar{u}/c + (2N/c)\,\bar{u} \le 4N\bar{u}/c$.

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] Billingsley, Patrick. *Probability and Measure, 3rd ed.* Wiley (1995).

[11] Boucheron, Stéphane, Gábor Lugosi, and Pascal Massart. *Concentration Inequalities: A Nonasymptotic Theory of Independence.* Oxford University Press, 2013.

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

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

[14] 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.

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

[16] 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.

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

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

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

[20] Edmiston, Jake, and Graeme Hamilton, "The Last Days of Quebec's Maple Syrup Rebellion," *National Post*, April 6, 2018.

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

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

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

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

[25] 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.

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

[27] 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.

[28] 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.

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

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

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

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

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

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

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

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

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

[38] Hörner, Johannes, and Wojciech Olszewski. "The Folk Theorem for Games with Private Almost-Perfect Monitoring" *Econometrica* 74.7 (2006): 1499-1544.

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

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

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

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

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

[44] Kuitenbrouwer, "'It's Crazy, Isn't It': Quebec's Maple Syrup Rebels Face Ruin as Cartel Crushes Dissent," *Financial Post*, December 5, 2016.

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

[46] 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.

[47] 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).

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

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

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

[51] 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.

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

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

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

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

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

[57] 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.

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

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

[60] Sugaya, Takuo, and Alexander Wolitzky, "Performance Feedback in Long-Run Relationships: A Rate of Convergence Approach." *Working Paper* (2023b).

Online Appendix

H Proof of Theorem 7

H.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 $r_i \in A_i$, we define a function $f_{i,r_i} : Y_i \to \mathbb{R}$ 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 $r_i \in A_i$ there exists a function $f_{i,r_i} : Y_i \to \mathbb{R}$ such that*

$$\begin{aligned}
\mathbb{E}\left[f_{i,r_i}(y_i)|r_i\right] - \mathbb{E}\left[f_{i,r_i}(y_i)|a_i\right] &\geq \bar{u} \quad \textit{for all } a_i \neq r_i, && (13)\\
\mathbb{E}\left[f_{i,r_i}(y_i)|r_i\right] &= 0, && (14)\\
\mathrm{Var}\left(f_{i,r_i}(y_i)|r_i\right) &\leq \bar{u}^2/\eta^2, \quad \textit{and} && (15)\\
\left|f_{i,r_i}(y_i)\right| &\leq 2\bar{u}/\eta^2 \quad \textit{for all } y_i. && (16)
\end{aligned}$$

Proof. Fix i and r_i. Let $Y_i^* = \{y_i : p_i(y_i, r_i) \geq \eta^2\}$, and let

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

Note that (9) is equivalent to $d(p_i(r_i), \mathrm{co}(P_i(r_i))) \geq \eta$ for all $i \in I, r_i \in A_i$, where $d(\cdot,\cdot)$ denotes Euclidean distance in $\mathbb{R}^{|Y_i^*|}$. Hence, by the separating hyperplane theorem, there exists $x = (x(y_i))_{y_i \in Y_i^*} \in \mathbb{R}^{|Y_i^*|}$ such that $\|x\| = 1$ and $(p_i(r_i) - p) \cdot x \geq \eta$ for all $p \in P_i(r_i)$. By definition of p_i and P_i, this implies that $\sum_{y_i \in Y_i^*} (p_i(y_i|r_i) - p_i(y_i|a_i))\, x(y_i) \geq \eta \sqrt{p_i(y_i|r_i)}$ for all $a_i \neq r_i$. Now define

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

Clearly, conditions (13) and (14) hold. Moreover, since $\mathbb{E}\left[f_{i,r_i}(y_i)|r_i\right] = 0$ and the term $\sum_{\tilde{y}_i \in Y_i} \sqrt{p(\tilde{y}_i|r_i)}\, x_i(\tilde{y}_i)$ is independent of y_i, we have

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

and hence (15) holds. Finally, (16) holds since, for each $y_i \in Y_i^*$,

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

■

Now fix $i \in I$ and $r_i \in A_i$, 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 (15), for any $T \in \mathbb{N}$, we have

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

Together with (14) and (16), Bernstein's inequality (Boucheron, Lugosi, and Massart, 2013, Theorem 2.10) now implies that, for any $T \in \mathbb{N}$ and $\bar{f} \in \mathbb{R}_+$, we have

$$\Pr\left(\sum_{t=1}^{T} \delta^{t-1} f_{i,r_i}(y_{i,t}) \geq \bar{f}\right) \leq \exp\left(-\frac{\bar{f}^2\eta^2}{2\left(\frac{1-\delta^{2T}}{1-\delta^2}\bar{u}^2 + \frac{2}{3}\bar{f}\bar{u}\right)}\right). \tag{17}$$

Our second lemma fixes T and $\bar{f}$ so that the bound in (17) is sufficiently small, and some other conditions used in the proof also hold.

Lemma 6 *There exists $k > 0$ such that, whenever $(1-\delta)\log(N)/\eta^2 < k$, there exist $T \in \mathbb{N}$ and $\bar{f} \in \mathbb{R}$ that satisfy the following three inequalities:*

$$60\bar{u}N\exp\left(-\frac{\left(\frac{\bar{f}}{3}\right)^2\eta^2}{2\left(\frac{1-\delta^{2T}}{1-\delta^2}\bar{u}^2 + \frac{2}{3}\frac{\bar{f}}{3}\bar{u}\right)}\right) \leq \varepsilon, \tag{18}$$

$$8\frac{1-\delta}{1-\delta^T}\left(\bar{f} + \frac{2\bar{u}}{\eta^2}\right) \leq \varepsilon, \tag{19}$$

$$4\bar{u}\frac{1-\delta^T}{\delta^T} + \frac{1-\delta}{\delta^T}\left(\bar{f} + \frac{2\bar{u}}{\eta^2}\right) \leq \varepsilon. \tag{20}$$

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

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

Note that if $(1-\delta)\log(N)/\eta^2 \to 0$ then $1-\delta^T \to \varepsilon/(\varepsilon+8\bar{u})$, and hence $(1-\delta)\log(N)/\left(\eta^2\left(1-\delta^T\right)\right) \to 0$. Therefore, there exists $k>0$ such that, whenever $(1-\delta)\log(N)/\eta^2 < k$, we have

$$\frac{4}{9}\sqrt{36\log\left(\frac{60\bar{u}}{\varepsilon}\right)\log(N)\frac{1-\delta}{1-\delta^T}\frac{1}{\eta^2}} \quad \leq \quad 1 \quad \text{and} \tag{21}$$

$$8\bar{u}\left(\sqrt{36\log\left(\frac{60\bar{u}}{\varepsilon}\right)\log(N)\frac{1-\delta}{1-\delta^T}\frac{1}{\eta^2}}+\frac{1-\delta}{1-\delta^T}\frac{2}{\eta^2}\right) \quad \leq \quad \varepsilon. \tag{22}$$

It now follows from straightforward algebra (provided in Appendix H.4) that (18)–(20) hold for every $k \geq \bar{k}$. ■

H.2 Equilibrium Construction

Fix any k, T, and $\bar{f}$ that satisfy (18)–(20), 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 $w : H^{T+1} \to \mathbb{R}^N$ such that, letting $\psi_i\left(h^{T+1}\right) = \frac{\delta^T}{1-\delta}\left(w_i\left(h^{T+1}\right)-v_i^*\right)$, we have

$$\textit{Promise Keeping:} \qquad v_i^* = \frac{1-\delta}{1-\delta^T}\mathbb{E}^{\sigma}\left[\sum_{t=1}^{T}\delta^{t-1}u_{i,t}+\psi_i\left(h^{T+1}\right)\right] \quad \text{for all } i, \tag{23}$$

$$\textit{Incentive Compatibility:} \qquad \sigma_i \in \operatorname*{argmax}_{\tilde{\sigma}_i}\mathbb{E}^{\tilde{\sigma}_i,\sigma_{-i}}\left[\sum_{t=1}^{T}\delta^{t-1}u_{i,t}+\psi_i\left(h^{T+1}\right)\right] \quad \text{for all } i, \tag{24}$$

$$\textit{Self Generation:} \qquad \zeta_i\psi_i\left(h^{T+1}\right) \in \left[-\frac{\delta^T}{1-\delta}\varepsilon, 0\right] \quad \text{for all } i \text{ and } h^{T+1}. \tag{25}$$

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 $\bar{\alpha} \in \Delta(A)$ such that

$$u_i(\bar{\alpha}) = v_i^* + \zeta_i\varepsilon/2 \quad \text{for all } i, \tag{26}$$

and fix a probability distribution over static Nash equilibria $\alpha^{NE} \in \Delta\left(\prod_i \Delta(A_i)\right)$ 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 $B_v(\varepsilon) \subseteq F^*$.

[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 $w\left(h^{T+1}\right)$.

We now construct the block strategy profile σ. For each player $i \in I$ and period $t \in \{1,\dots,T\}$, 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 $r_t \in A$ be a pure action profile which is drawn by public randomization at the start of period t from the distribution $\bar{\alpha} \in \Delta(A)$ fixed in (26), and let $\varrho_t^{NE} \in \prod_i \Delta(A_i)$ 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 $BR_i(r_{-i}) = \text{argmax}_{a_i \in A_i} u_i(a_i, r_{-i})$ 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}.

Let $\alpha_t^* \in \prod_i \Delta(A_i)$ 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 $r_t \in A$ described above, an additional ("fictitious") random variable $\tilde{y}_t \in Y$ 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 $f_{i,r_{i,t}} : Y_i \to \mathbb{R}$ be defined as in Lemma 5, and let

$$f_{i,t} = \begin{cases} f_{i,r_{i,t}}\left(y_{i,t}\right) & \text{if } S_{i,t} = 0, \\ f_{i,r_{i,t}}\left(\tilde{y}_{i,t}\right) & \text{if } S_{i,t} = NE, \\ 0 & \text{if } S_{i,t} = BR. \end{cases} \tag{27}$$

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

$$\theta_{i,t} = \mathbf{1}\left\{\exists t' \leq t : \left|\sum_{t''=1}^{t'-1} \delta^{t''-1} f_{i,t''}\right| \geq \bar{f}\right\}. \tag{28}$$

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 $p_j\left(y_{j,t}|r_{j,t}\right)$, 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 $p_j\left(\tilde{y}_{j,t}|r_{j,t}\right)$, 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, $\Pr\left((f_{j,t})_{j\in J} | h^t, z_t, a_{i,t}\right) = \prod_{j\neq i} \Pr\left(f_{j,t}|S_{j,t}, r_{j,t}\right)$, 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

[28] Intuitively, introducing the variable $\tilde{y}_t$, rather than simply using $y_{i,t}$ everywhere in (27), ensures that the distribution of $f_{i,t}$ does not depend on player i's opponents' strategies.

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

$$\psi_{i,t} = \delta^{t-1} f_{i,r_{i,t}}\left(y_{i,t}\right). \tag{29}$$

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

$$\psi_{i,t} = \delta^{t-1}\left(u_i\left(\bar{\alpha}\right) - u_i\left(\alpha_t^*\right)\right). \tag{30}$$

3. Otherwise,

$$\psi_{i,t} = \delta^{t-1}\left(-\zeta_i \bar{u} - u_i\left(\alpha_t^*\right) + \mathbf{1}\left\{S_{i,t} = 0\right\} f_{i,r_{i,t}}\left(y_{i,t}\right)\right). \tag{31}$$

The constant c_i is defined as

$$c_i = -\mathbb{E}\left[\sum_{t=1}^{T} \delta^{t-1}\left(\mathbf{1}\left\{\max_{j \neq i} \theta_{j,t} = 0\right\} u_i\left(\bar{\alpha}\right) - \mathbf{1}\left\{\max_{j \neq i} \theta_{j,t} = 1\right\} \zeta_i \bar{u}\right)\right] + \frac{1-\delta^T}{1-\delta} v_i^*. \tag{32}$$

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

$$\left|c_i\right| \leq 2\bar{u}\frac{1-\delta^T}{1-\delta}. \tag{33}$$

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

$$\psi_i\left(h^{T+1}\right) = c_i + \sum_{t=1}^{T} \psi_{i,t}. \tag{34}$$

H.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*

$$\Pr\left(\max_{i \in I, t \in \{1,\ldots,T\}} \theta_{i,t} = 0\right) \geq 1 - \frac{\varepsilon}{20\bar{u}}. \tag{35}$$

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

$$\Pr\left(\max_{t \in \{1,\ldots,T\}}\left|\sum_{t'=1}^{t} \delta^{t-1} f_{i,t'}\right| \geq \bar{f}\right) \leq \frac{\varepsilon}{20\bar{u}N}. \tag{36}$$

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

$$\Pr\left(\max_{t\in\{1,\ldots,T\}}\left|\sum_{t'=1}^{t}\delta^{t-1}f_{i,t'}\right| \geq \bar{f}\right) \leq \Pr\left(\max_{t\in\{1,\ldots,T\}}\left|\sum_{t'=1}^{t}\delta^{t-1}\tilde{f}_{i,t'}\right| \geq \bar{f}\right). \tag{37}$$

Since $\left(\tilde{f}_{i,t}\right)_{t=1}^{T}$ are independent, Etemadi's inequality (Billingsley, 1995; Theorem 22.5) implies that

$$\Pr\left(\max_{t\in\{1,\ldots,T\}}\left|\sum_{t'=1}^{t}\delta^{t-1}\tilde{f}_{i,t'}\right| \geq \bar{f}\right) \leq 3\max_{t\in\{1,\ldots,T\}}\Pr\left(\left|\sum_{t'=1}^{t}\delta^{t-1}\tilde{f}_{i,t'}\right| \geq \frac{\bar{f}}{3}\right). \tag{38}$$

Letting $x_{i,t} = \delta^{t-1}\tilde{f}_{i,t}$, note that $|x_{i,t}| \leq 2\bar{u}/\eta^2$ with probability 1 by (16), $\mathbb{E}[x_{i,t}] = 0$ by (14), and

$$\operatorname{Var}\left(\sum_{t'=1}^{t}x_{i,t'}\right) = \sum_{t'=1}^{t}\operatorname{Var}(x_{i,t'}) \leq \sum_{t'=1}^{T}\operatorname{Var}(x_{i,t'}) = \frac{1-\delta^T}{1-\delta}\frac{\bar{u}^2}{\eta^2} \quad \text{by (15).}$$

Therefore, by Bernstein's inequality ((17), which again applies because $\left(\tilde{f}_{i,t}\right)_{t=1}^{T}$ are independent) and (18), we have, for each $t \leq T$,

$$\Pr\left(\left|\sum_{t'=1}^{t}\delta^{t'-1}\tilde{f}_{i,t'}\right| \geq \frac{\bar{f}}{3}\right) \leq \frac{\varepsilon}{60\bar{u}N}. \tag{39}$$

Finally, (37), (38), and (39) together imply (36). ■

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

Lemma 9 *For each player i and block strategy profile σ, incentive compatibility holds (i.e., (24) is satisfied) if and only if*

$$\operatorname{supp}\sigma_i\left(h^t\right) \subseteq \operatorname*{argmax}_{a_{i,t}\in A_i}\mathbb{E}^{\sigma_{-i}}\left[\delta^{t-1}u_{i,t} + \psi_{i,t}|h^t, a_{i,t}\right] \quad \textit{for all } t \textit{ and } h^t. \tag{40}$$

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

$$\mathbb{E}^{\sigma}\left[\delta^{t'-1}u_{i,t} + \psi_{i,t'}|h^t\right] = \mathbb{E}^{\sigma}\left[\delta^{t'-1}\left(\mathbf{1}\left\{\max_{j\neq i}\theta_{j,t'} = 0\right\}u_i(\bar{\alpha}) - \mathbf{1}\left\{\max_{j\neq i}\theta_{j,t'} = 1\right\}\zeta_i\bar{u}\right)|h^t\right]. \tag{41}$$

We now verify that (40) 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 (13) and (29), and

recalling that $\bar{u} \geq \max_a u_i(a) - \min_a u_i(a)$, we have

$$\begin{aligned}
& \mathbb{E}^{\sigma_{-i}}\left[\delta^{t-1}u_{i,t} + \psi_{i,t}|h^t, z_t, a_{i,t} = r_{i,t}\right] - \mathbb{E}^{\sigma_{-i}}\left[\delta^{t-1}u_{i,t} + \psi_{i,t}|h^t, z_t, a_{i,t} = a_i\right] \\
= \; & \delta^{t-1}\left(\mathbb{E}\left[u_i(r_t) + f_{i,r_{i,t}}(y_{i,t})\,|a_{i,t} = r_{i,t}\right] - \mathbb{E}\left[u_i(a_i, r_{-i,t}) + f_{i,r_{i,t}}(y_{i,t})\,|a_{i,t} = a_i\right]\right) \\
\leq \; & 0, \quad \text{so (40) holds.}
\end{aligned}$$

2. If $\theta_{i,t} = 1$ and $\theta_{j,t} = 0$ for all $j \neq i$, then the reward $\psi_{i,t}$ specified by (30) does not depend on $y_{i,t}$. Hence, (40) 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 (40) holds because it holds in Case 1 above and (29) and (31) 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 (31) does not depend on $y_{i,t}$, (40) 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

$$\begin{aligned}
& \frac{1-\delta}{1-\delta^T}\mathbb{E}^{\sigma}\left[\sum_{t=1}^{T}\delta^{t-1}u_{i,t} + \psi_i\left(h^{T+1}\right)\right] \\
= \; & \frac{1-\delta}{1-\delta^T}\left(\mathbb{E}^{\sigma}\left[\sum_{t=1}^{T}\left(\delta^{t-1}u_{i,t} + \psi_{i,t}\right)\right] + c_i\right) \quad \text{(by (34))} \\
= \; & \frac{1-\delta}{1-\delta^T}\mathbb{E}^{\sigma}\left[\sum_{t=1}^{T}\delta^{t-1}\left(\mathbf{1}\left\{\max_{j\neq i}\theta_{j,t} = 0\right\}u_i(\bar{\alpha}) - \mathbf{1}\left\{\max_{j\neq i}\theta_{j,t} = 1\right\}\zeta_i\bar{u}\right) + c_i\right] \quad \text{(by (41))} \\
= \; & v_i^* \quad \text{(by (32)), so (23) holds.}
\end{aligned}$$

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

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

$$\zeta_i\sum_{t=1}^{T}\psi_{i,t} \leq \bar{f} + \frac{2\bar{u}}{\eta^2} \quad \text{and} \tag{42}$$

$$\left|\sum_{t=1}^{T}\psi_{i,t}\right| \leq \bar{f} + \frac{2\bar{u}}{\eta^2} + 2\bar{u}\frac{1-\delta^T}{1-\delta}. \tag{43}$$

In addition,

$$\zeta_i c_i \leq -\frac{1-\delta^T}{1-\delta}\frac{\varepsilon}{8}. \tag{44}$$

To establish self generation ((25)), 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

$$\begin{aligned}
\zeta_i \psi_i\left(h^{T+1}\right) &= \zeta_i\left(c_i + \sum_{t=1}^{T} \psi_{i,t}\right) \leq -\frac{1-\delta^T}{1-\delta}\frac{\varepsilon}{8} + \bar{f} + 2\bar{u}/\eta^2 \quad \text{(by (42) and (44))} \\
&\leq \frac{1-\delta^T}{8\left(1-\delta\right)}\left(-\varepsilon + 8\left(\frac{1-\delta}{1-\delta^T}\right)\left(\bar{f} + 2\bar{u}/\eta^2\right)\right) \leq 0 \quad \text{(by (19)),} \quad \text{and} \\
\left|\psi_i\left(h^{T+1}\right)\right| &\leq |c_i| + \left|\sum_{t=1}^{T} \psi_{i,t}\right| \\
&\leq 4\bar{u}\frac{1-\delta^T}{1-\delta} + \bar{f} + 2\bar{u}/\eta^2 \quad \text{(by (33) and (43))} \\
&= \frac{1-\delta^T}{1-\delta}4\bar{u} + \bar{f} + 2\bar{u}/\eta^2 \leq \frac{\delta^T}{1-\delta}\varepsilon \quad \text{(by (20)),}
\end{aligned}$$

which completes the proof.

H.4 Omitted Details for the Proof of Lemma 6

We show that, with the stated definitions of T and $\bar{f}$, (21) and (22) imply (18)–(20). First, note that

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

Hence,

$$\begin{aligned}
\frac{2\bar{f}\left(1-\delta^2\right)}{9\bar{u}\left(1-\delta^{2T}\right)} &< \frac{4}{9\bar{u}}\frac{1-\delta}{1-\delta^T}\sqrt{36\log\left(\frac{60\bar{u}}{\varepsilon}\right)\log\left(N\right)\frac{1-\delta^T}{1-\delta}\frac{\bar{u}^2}{\eta^2}} \\
&= \frac{4}{9}\sqrt{36\log\left(\frac{60\bar{u}}{\varepsilon}\right)\log\left(N\right)\frac{1-\delta}{1-\delta^T}\frac{1}{\eta^2}} \leq 1 \quad \text{(by (21)).}
\end{aligned}$$

Therefore,

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

Moreover,

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

Hence, we have

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

This establishes (18).

Next, we have

$$8\frac{1-\delta}{1-\delta^T}\left(\bar{f}+\frac{2\bar{u}}{\eta^2}\right)=8\bar{u}\left(\sqrt{36\log\left(\frac{60\bar{u}}{\varepsilon}\right)\log(N)\frac{1-\delta}{1-\delta^T}\frac{1}{\eta^2}}+\frac{1-\delta}{1-\delta^T}\frac{2}{\eta^2}\right)\le\varepsilon\quad\text{(by (22))}. \tag{45}$$

This establishes (19).

Finally, by (45) and $8\bar{u}\left(1-\delta^T\right)/\delta^T\le\varepsilon$, we have

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

This establishes (20).

H.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 (40) 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 $t'\ge t+1$, there are several cases to consider.

1. If $\theta_{j,t'}=0$ for all j, then by (14) and (29) 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 $j\ne i$, then by (30) 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 (14) and (31) (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 $S_{i,t'}\ne 0$, then by (31) 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, (41) holds, and player i's net continuation payoff under σ from period $t+1$ onward equals

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

By Lemma 7, the distribution of $\left((\theta_{n,t'})_{n\neq i}\right)_{t'=t+1}^{T}$ 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 $\mathbb{E}^{\sigma_{-i}}[\delta^{t-1}u_{i,t}+\psi_{i,t}|h^t,a_{i,t}]$.

H.6 Proof of Lemma 10

Define

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

Note that, by (29)–(31), we can write $\psi_{i,t}=\psi_{i,t}^{v}+\psi_{i,t}^{f}$. (Note that, if $\theta_{n,t}=0$ for all $n\in I$, we have $\alpha_t^*=\bar{\alpha}$ and hence $\psi_{i,t}^{v}+\psi_{i,t}^{f}=\delta^{t-1}f_{i,a_{i,t}}\left(y_{i,t}\right)$, as specified in (29).) We show that, for every end-of-block history h^{T+1}, we have

$$
\zeta_i\sum_{t=1}^{T}\psi_{i,t}^{v} \in \left[-2\bar{u}\frac{1-\delta^T}{1-\delta},0\right] \quad \text{and} \tag{46}
$$

$$
\left|\zeta_i\sum_{t=1}^{T}\psi_{i,t}^{f}\right| \leq \bar{f}+\frac{2\bar{u}}{\eta^2}. \tag{47}
$$

Since $\psi_{i,t}=\psi_{i,t}^{v}+\psi_{i,t}^{f}$, (46) and (47) imply (42) and (43), which proves the first part of the lemma.

For (46), 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 $u_i\left(\alpha_t^*\right)\geq\sum_a\bar{\alpha}\left(a\right)\min\left\{u_i\left(a\right),\max_{a_i'}u_i\left(a_i',a_{-i}\right)\right\}\geq u_i\left(\bar{\alpha}\right)$; and (ii) if $\zeta_i=-1$, we have $u_i\left(\alpha_t^*\right)\leq\max\left\{u_i\left(\bar{\alpha}\right),u_i\left(\alpha^{NE}\right)\right\}=u_i\left(\bar{\alpha}\right)$. In total, we have $\zeta_i\left(u_i\left(\bar{\alpha}\right)-u_i\left(\alpha_t^*\right)\right)\leq 0$. Since obviously $\zeta_i\left(u_i\left(\bar{\alpha}\right)-u_i\left(\alpha_t^*\right)\right)\geq-2\bar{u}$ and $-\bar{u}-\zeta_iu_i\left(\alpha_t^*\right)\geq-2\bar{u}$, we have

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

For (47), note that $S_{i,t}=0$ implies $\theta_{i,t}=0$, and hence

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

Since $\theta_{i,t+1}=1$ whenever $\left|\sum_{t'=1,..,t}\delta^{t-1}f_{i,a_{i,t}}\left(y_{i,t}\right)\right|\geq\bar{f}$, and in addition $\left|f_{i,a_{i,t}}\left(y_{i,t}\right)\right|\leq 2\bar{u}/\eta^2$ by (16), this inequality implies (47).

For the second part of the lemma, by (32), we have

$$
\begin{aligned}
\zeta_i c_i &= \zeta_i \left(-\mathbb{E}\left[\sum_{t=1}^{T} \delta^{t-1} \left(\mathbf{1}\left\{ \max_{j\neq i} \theta_{j,t} = 0 \right\} u_i\left(\bar{\alpha}\right) - \mathbf{1}\left\{ \max_{j\neq i} \theta_{j,t} = 1 \right\} \zeta_i \bar{u} \right) \right] + \frac{1-\delta^T}{1-\delta} v_i^* \right) \\
&= \mathbb{E}\left[\sum_{t=1}^{T} \delta^{t-1} \left(\mathbf{1}\left\{ \max_{j\neq i} \theta_{j,t} = 0 \right\} \zeta_i \left(v_i^* - u_i\left(\bar{\alpha}\right)\right) + \mathbf{1}\left\{ \max_{j\neq i} \theta_{j,t} = 1 \right\} \underbrace{\left(\bar{u} + \zeta_i v_i^*\right)}_{\in[0,2\bar{u}]} \right) \right] \\
&\leq \mathbb{E}\left[\sum_{t=1}^{T} \delta^{t-1} \left(\mathbf{1}\left\{ \max_{j\neq i} \theta_{j,t} = 0 \right\} \left(\frac{-\varepsilon}{2} \right) + \mathbf{1}\left\{ \max_{j\neq i} \theta_{j,t} = 1 \right\} 2\bar{u} \right) \right] \text{ by (26)} \\
&\leq -\frac{1-\delta^T}{1-\delta} \left(\left(1 - \frac{\varepsilon}{20\bar{u}} \right) \frac{\varepsilon}{2} + \left(\frac{\varepsilon}{20\bar{u}} \right) 2\bar{u} \right) \quad \text{(by (35))} \\
&\leq -\frac{1-\delta^T}{1-\delta} \frac{\varepsilon}{8} \quad \text{(as } \varepsilon < \bar{u}/2\text{).}
\end{aligned}
$$