Do a Few Bad Apples Spoil the Barrel?: Community Enforcement with Incomplete Information*

Takuo Sugaya and Alexander Wolitzky

Stanford GSB and MIT

January 16, 2020

Abstract

We study the repeated prisoner's dilemma with random matching when some players may be "bad types" who never cooperate. We establish an anti-folk theorem: with anonymous players, cooperation is impossible in large groups under a smoothness assumption on the distribution of the number of bad types. Communities may avoid this grim outcome by segregating themselves into smaller sub-groups, at the cost of forgoing some gains from trade. Making players' identities observable does not help much: cooperation remains impossible in groups whose size N is large relative to the discount factor δ, in that $(1-\delta)\sqrt{N} \to \infty$. However, allowing within-match cheap talk supports cooperation in much larger groups: those where $(1-\delta)\log N \to 0$. Thus, in contrast to the situation where all players are rational, communication is essential for supporting cooperation in large groups in the presence of a few bad apples.

Keywords: community enforcement, repeated games, incomplete information, communication

JEL codes: C72, C73, D83

*For helpful comments, we thank Daron Acemoglu, Glenn Ellison, Drew Fudenberg, Yuval Heller, Matt Jackson, George Mailath, Stephen Morris, Satoru Takahashi, Omer Tamuz, and seminar particiants at Princeton, Stanford, and Wharton.

1 Introduction

Economists have long asked whether large, decentralized groups of agents with limited information about individuals' past behavior can be expected to cooperate. Investigating this question may help us understand what types of information, communication, enforcement, and patterns of interaction are required to support trust and pro-social norms of behavior in different kinds of groups.

Early results by Kandori (1992) and Ellison (1994) suggest that groups can often support cooperation despite minimal information by relying on *contagion*, a form of *collective punishment*: whenever a player sees anyone defect, she starts defecting against everyone. While collective punishment works in groups where everyone is perfectly rational and forward-looking and no one ever makes a mistake, it is not clear whether this approach remains effective in more realistic situations. Kandori writes (p. 71),

> Even when cooperation is sustained, [contagion] has the unfortunate feature that innocent players will necessarily be punished. Also, it is fragile in that a little bit of noise ("trembling hands") causes complete breakdown of cooperation in the community. This may be the main reason why we do not observe such a norm very often. Given this point, we feel that a norm should be evaluated not only by its equilibrium payoffs, but also by its (suitably defined) "robustness."

Ellison challenges this conclusion by showing that a version of contagion strategies are robust to i.i.d. noise, where each player is forced to defect with probability ε every period. However, he also notes (p. 578)

> What is probably more important practically and harder to overcome is that the argument above deals only with trembles. If one player were "crazy" and always played D [defect]... the contagious strategies would not support cooperation. In large populations, the assumption that all players are rational and know their opponents' strategies may be both very important to the conclusions and fairly implausible.

In this paper, we probe the limits of community enforcement when some players are "bad types" who always defect. Our first result—Theorem 1—is a stark anti-folk theorem: with anonymous players, cooperation is impossible in large groups under a smoothness assumption on the distribution of the number of bad types, no matter what strategies players use or how patient they may be.

That is, for population size N and discount factor δ, cooperation is impossible when N is large, even if $(1-\delta)\,N$ is small.

The key assumption behind this result is not that the number of bad types is large, but that it is uncertain. When players are anonymous, cooperation is a public good: a player's decision to cooperate rather than defect in the current period is individually costly and benefits all her opponents equally. Therefore, a rational player can be deterred from pretending to be bad—and thus contributing less to the total "amount of cooperation" in society—only if the amount of cooperation is significantly lower when there is one extra bad type. Since the amount of cooperation is bounded, when the variance of the number of bad types goes to infinity, the expected amount of cooperation goes to zero. For example, we show that this is the case when each player is bad with independent probability ε, for any fixed $\varepsilon \in (0,1)$, and $N \to \infty$.

Theorem 1 provides a simple rationale for why individuals sometimes choose to interact in small rather than large groups: the larger the group, the more likely it is to contain bad apples, and bad apples spoil cooperation. We investigate optimal group size with anonymous agents, taking into account that restricting interaction to a smaller group reduces the available gains from trade. In the simple case where each player is bad with independent probability ε and marginal gains from trade are constant in group size, we find that optimal group size is on the order of $1/\varepsilon$ under both contagion strategies (where a single bad type completely destroys cooperation with patient players; see Proposition 1) and optimal ex post equilibrium strategies (where we show that each additional bad type reduces the amount of cooperation by a constant factor; see Proposition 2).[1]

We then move from the benchmark case of anonymous agents to the more realistic setting where players' identities are observable. In this environment, it is obvious that the presence of bad types does not destroy cooperation when players are sufficiently patient, as players can simply forgo community enforcement altogether and treat the repeated random matching game as a collection of asynchronous two-player games, in which case the fact that one partner may turn out to be bad has no bearing on interactions with different partners. However, this approach works only if players are very patient—in that $(1-\delta)\,N$ is small—while a central promise of community enforcement is to support cooperation for much lower, more realistic discount factors: for example, without bad types, contagion strategies form a Nash equilibrium whenever $(1-\delta)\log N$ is small. Our second main result—Theorem 2—shows that in the presence of bad types, even non-anonymous players

[1] Here we consider an upper bound on ex post equilibrium payoffs: we conjecture the bound is attainable, but we have not proved this.

cannot support any cooperation when $(1-\delta)\sqrt{N}$ is large. Thus, community enforcement can be effective only for the limited range of discount factors where $(1-\delta) \in \left(\frac{1}{cN}, \frac{c}{\sqrt{N}}\right)$ for some constant c.[2]

We also establish a version of Theorem 2 in the case where each player is forced to defect with probability ε every period, independently across players and periods (rather than each player being a bad type with probability ε, in which case she always defects). Here we show that cooperation is impossible when $(1-\delta) N^{1/4}$ is large. This implies that Ellison's result that contagion strategies are robust to i.i.d. noise requires a qualitatively higher discount factor than is needed without noise.

The intuition for Theorem 2 is that, to avoid the negative conclusion of Theorem 1, players must be able to identify and punish defectors individually. However, in the absence of explicit communication, it takes a long time for information about defectors' identities to percolate through the community. For instance, since a player observes one bit of information each period (i.e., whether her partner cooperated or defected), it is impossible for her to perfectly learn whether each of the other $N-1$ players is a bad type in fewer than $N-1$ periods. While a player can get a noisy signal of whether each of her opponents is a bad type with fewer than $N-1$ bits/periods, we show that at least $\sqrt{N}$ bits/periods are needed to provide enough information to incentivize cooperation. Hence, cooperation is possible only if $\delta^{\sqrt{N}}$ is not too small, and hence if $(1-\delta)\sqrt{N}$ is not too large.

The obvious way out of this impossibility result is to allow pre-play cheap talk communication between matched players. Our final main results—Theorems 3 and 4—show that this lets the group support cooperation whenever $(1-\delta)\log N$ is small. In this setting, cooperation can be achieved as an approximate Nash equilibrium using simple strategies where each player keeps track of a "blacklist" of opponents whom she believes have ever defected against a rational player, players share their blacklists with each other prior to taking actions, and each player defects against the opponents on her blacklist. Cooperation can be achieved as an exact sequential equilibrium using a more complicated, "block belief-free" construction.

Taken together, our results show that the possibility that some players may be "bad types" who never cooperate dramatically undermines cooperation both when players are anonymous (and

[2] In interpreting these results, it may be helpful to suppose that players match once every Δ units of real time with fixed discount rate $r > 0$, so $\delta = e^{-r\Delta}$, and hence $(1-\delta) N \approx r\Delta N$. Since each pair of players interacts $1/(\Delta \times (N-1))$ times per unit of real time on average, $(1-\delta) N$ small means that each pair of players interact frequently, while $(1-\delta) N$ large means that each of pair of players rarely interact.

can be arbitrarily patient) and when their identities are observable (but they are only "moderately patient"). However, this obstacle to cooperation can be mitigated by restricting cooperation to occur only within smaller sub-groups, and it can be eliminated entirely if players communicate actively, sharing information about defectors' identities.

Our paper relates to several branches of the literature. Most directly, we contribute to the literature on community enforcement in repeated games, originating with the work of Kandori (1992), Ellison (1994), Harrington (1995), and Okuno-Fujiwara and Postlewaite (1995). A few papers in this literature consider incomplete information, focusing like us on the case where players are either rational or are bad types who are committed to always defecting.[3] Ghosh and Ray (1996) show that the presence of bad types can help support cooperation in the prisoner's dilemma (PD) with voluntary separation, by making players reluctant to cheat their current partners and return to the matching pool. Dilmé (2016) and Heller and Mohlin (2018) show that the presence of bad types can make cooperative equilibria more robust in the PD with random matching and information about one's partner's past actions, by making the observation that the partner defected in the past informative of his being a bad type. Among the many differences with our work, perhaps the most fundamental is that these papers study ways in which the presence of bad types *helps* support cooperation (in more intricate models with voluntary separation or information about past play), while we ask when bad types undermine cooperation (in the canonical random-matching PD).[4]

We also contribute to the question of when the folk theorem holds in repeated games. Deb, Sugaya, and Wolitzky (2019) generalize Kandori and Ellison's positive results to establish the folk theorem for repeated games with anonymous random matching, when all players are rational. In contrast, Theorem 1 is an anti-folk theorem, which applies whenever each player has an arbitrarily small probability of being a bad type. Introducing a small amount of incomplete information in repeated games also leads to anti-folk theorems in "reputation" models (Mailath and Samuelson, 2006). Among other differences, reputation models typically feature a single patient player, and the folk theorem usually holds with multiple, equally patient players (Cripps and Thomas, 1997; Chan, 2000; Cripps, Dekel, and Pesendorfer, 2005).

Another related theoretical literature asks when introducing explicit communication opportunities expands the set of attainable payoffs, relative to what could be achieved with only "implicit

[3] However, as we discuss below, our Theorem 1 extends to much more general type spaces.

[4] One result closer in spirit to ours is Heller and Mohlin's Theorem 1, which shows that cooperation is impossible in the "offensive" (submodular) PD with bad types, while Takahashi (2010) showed that cooperative "belief-free" equilibria exist in this setting without bad types.

communication" via players' actions. When the discount factor δ is bounded away from 1, communication helps support cooperation by spreading the news that a player defected throughout the population more quickly (Ahn and Suominen, 2001; Dixit, 2003; Lippert and Spagnolo, 2011; Wolitzky, 2013; Ali and Miller, 2016; Balmaceda and Escobar, 2017). This effect may capture an important role of gossip in supporting cooperation (e.g., Sommerfeld, Krambeck, Semmann and Milinski, 2007). A more subtle question concerns the role of communication when δ is high. Here, Awaya and Krishna (2016, 2019) show that introducing explicit communication can expand the equilibrium payoff set by exploiting correlation between players' signals. Our Theorems 2 and 4 identify a novel role for communication: in large groups, when δ is high but less than $1 - 1/N$, communication supports cooperation by letting the group identify defectors in roughly $\log N$ rather than N periods, thus facilitating individual rather than collective punishment.

Our analysis also relates to the broader issue of whether the benefits from cooperation are greater in larger or smaller groups. Several authors have identified wide or generalized trust as an important factor of economic growth, while noting that many societies nonetheless feature narrower circles of trust within smaller, more cohesive groups (Coleman, 1988; Putnam, 1994; Fukuyama, 1995; Knack and Keefer, 1997). The standard model of "local trust" is that enforcement is too weak to support global cooperation: δ is too low to support cooperation between players who meet infrequently, so players cooperate only with local contacts (e.g., Kranton, 1996; Dixit, 2003). In contrast, our results (Propositions 1 and 2) show that, even if $\delta \approx 1$, players can benefit from restricting interaction to small groups, because larger groups are more likely to contain a few bad types, whose presence undermines cooperation. This alternative account of local trust generates new comparative statics: while the standard model predicts that the size of cooperative groups increases as δ increases, reaching global cooperation when $\delta \approx 1$, we predict that the size of cooperative groups increases as ε (the probability each player is bad) decreases, and that global cooperation requires not only $\delta \approx 1$ but also $\varepsilon \approx 0$.

2 Model

A set $I = \{1, \ldots, N\}$ of N players interact in discrete time, $t = 1, 2, \ldots$, with N even. Each period, the players randomly match in pairs to play the prisoner's dilemma (PD):

	C	D
C	$1,1$	$-L,1+G$
D	$1+G,-L$	$0,0$

where $G, L > 0$ and $G < 1 + L$, so (C, C) maximizes the sum of stage-game payoffs.

Each player is either *rational* or *bad*. Rational players maximize expected discounted payoffs with discount factor $\delta \in (0, 1)$. Bad players always play D. Bad players can thus be viewed as types for whom D is a dominant strategy in the repeated game, or as "commitment types" as in the literature on reputation in repeated games (Mailath and Samuelson, 2006).

For each set of players $S \subset I$, there is a commonly known prior probability $p(S)$ that the set of bad players is precisely S. Thus, p is a probability distribution on subsets of I. We assume that p symmetric, in that for any two sets $S, S' \subset I$ containing the same number of players, we have $p(S) = p(S')$. In particular, conditional on the event that there are exactly n bad players, the probability that any given player is bad equals n/N. We denote the probability that there are n bad players by p_n, and denote the probability that a given player i is bad by $\varepsilon = \sum_{S \ni i} p(S)$. Assume $\varepsilon < 1$, so rational types exist with positive probability.

Thus, for fixed payoff parameters G and L, the repeated game is parameterized by the triple (N, δ, p). We consider sequences $(N, \delta, p)_l$ indexed by $l \in \mathbb{N}$.

A leading special case of the model involves *independent types*, where each player is bad with independent probability ε. In this case we parameterize the game by (N, δ, ε).

We consider two information structures. In the *anonymous PD*, players take actions without observing their opponent's identity, and observe only their opponent's actions at the end of each period. That is, letting $\mu_t(i) \in I \backslash \{i\}$ denote player i's period-t partner, and letting $\omega_{i,t} = a_{\mu_t(i),t}$, player i's history at the beginning of period t is $h_i^t = (a_{i,\tau}, \omega_{i,\tau})_{\tau=1}^{t-1}$, with $h_i^1 = \emptyset$.

In the *non-anonymous PD*, players additionally observe their opponent's identity before taking actions. Thus, player i's history at the beginning of period t is $h_i^t = \Big((\mu_\tau(i), a_{i,\tau}, \omega_{i,\tau})_{\tau=1}^{t-1}, \mu_t(i) \Big)$, with $h_i^1 = \mu_1(i)$.

A *strategy* σ_i for player i maps histories h_i^t to $\Delta(\{C, D\})$, for each t. The interpretation is that player i plays $\sigma_i(h_i^t)$ at history h_i^t when rational; when bad, she always plays D. Given a strategy profile $\sigma = (\sigma_i)_i$, denote player i's expected discounted per-period payoff conditional on the event that the set of bad players is $S \not\ni i$ by $U_i(S)$, and denote player i's expected discounted per-period payoff when rational by

$$U_i = \frac{1}{1-\varepsilon} \sum_{S \not\ni i} p(S) U_i(S).$$

Our measure of average payoffs is

$$U = \frac{1}{N} \sum_i U_i = \frac{1}{(1-\varepsilon) N} \sum_{S \subset I} p(S) \sum_{i \notin S} U_i(S).$$

Thus, U equals the average over players of their expected payoffs when rational, or equivalently the expectation over sets of bad players S of the total payoff of the rational players $i \notin S$, divided by the expected number of rational players. With independent types, expected average payoffs ex ante (before players learn their types) are at most U, since a player's expected payoff when rational must exceed her expected payoff when bad (as a rational player has the option of always playing D). Also, since a player's minmax payoff is 0 and the maximum sum of stage-game payoffs is 2, in any Nash equilibrium we have $U \in [0, 1]$ (even if types are correlated).

In Section 5.3, we augment the game by allowing pre-play cheap talk communication. Until then, no explicit communication is allowed. Players do of course draw inferences from observing their opponents' actions, so in this sense "implicit communication" through actions is possible. Deb, Sugaya, and Wolitzky (2019) showed that such implicit communication is very powerful in complete-information random matching games with patient players. One of the themes of the current paper is that implicit communication is much less powerful in the presence of bad types.

3 No Cooperation in Large Anonymous Groups without Communication

Our benchmark result says that cooperation is impossible in a large anonymous population without pre-play communication, regardless of how patient players are (in particular, even if $(1-\delta) N$ is small). This result requires a smoothness assumption on the distribution of the number of bad types. To formulate this assumption, given (N, p), denote the conditional probability that n out of the first $N-1$ players are bad, given that player N is rational, by

$$q_n = \frac{\Pr(\text{player } N \text{ is rational} | n \text{ players are bad}) \times p_n}{\Pr(\text{player } N \text{ is rational})} = \frac{N-n}{N} \frac{p_n}{1-\varepsilon} \quad \text{for } n \in \{0, \dots, N-1\},$$

with $q_N = 0$ by convention. Similarly, denote the conditional probability that $n-1$ out of the first $N-1$ players are bad, given that player N is rational, by

$$q_n^- = q_{n-1} \text{ for } n \in \{1, \dots, N\},$$

with $q_0^- = 0$ by convention. Note that $q = (q_n)_{n=0}^N$ and $q^- = (q_n^-)_{n=0}^N$ are both probability distributions on $\{0, \dots, N\}$. Denote the total variation distance between q and q^- by

$$\Delta_{q,q^-} = \max_{\mathcal{N} \subset \{0,\dots,N\}} \left| \sum_{n \in \mathcal{N}} \left(q_n - q_n^- \right) \right|. \quad (1)$$

We say that a sequence $(N, p)_l$ has an *unpredictable number of bad types* if

$$\lim_{l \to \infty} \Delta_{q,q^-} = 0.$$

Similarly, a sequence $(N, \delta, p)_l$ has an *unpredictable number of bad types* if this true of $(N, p)_l$.

Theorem 1 *In the anonymous PD, for any parameters (N, δ, p) and any corresponding Nash equilibrium average payoff U, we have*

$$U \leq \frac{1+G}{\min\{G, L\}} \Delta_{q,q^-}. \quad (2)$$

Hence, for any sequence $(N, \delta, p)_l$ with an unpredictable number of bad types and any corresponding sequence of Nash equilibrium average payoffs $(U)_l$, we have $\lim_{l \to \infty} U_l = 0$.

To see the intuition, recall that when players are anonymous, cooperation is a public good: a player's decision to take C rather than D in the current period benefits all her opponents equally. We can therefore view each rational player as deciding whether to "contribute" to the public good by following her equilibrium strategy, or to "shirk" by playing D at every history (thus "mimicking" a bad type). We show that, for any given number of shirkers in the population, a shirker gets at least $\frac{1+G}{1+(G-L)_+}$ times the payoff of a contributor. A rational player is therefore willing to contribute only if the overall level of cooperation decreases substantially when the number of shirkers increases by 1. When there is enough variance in the number of shirkers (i.e., when the number of bad types is unpredictable), this implies that the expected level of cooperation is low.[5]

[5] For example, in the independent types case with fixed $\varepsilon \in (0, 1)$, the variance of the *fraction* of bad types in the population goes to 0 as $N \to \infty$, by the law of large numbers. What matters for this argument though is the variance

Before proving the theorem, we give some examples of sequences $(N,p)_l$ with unpredictable numbers of bad types.

Suppose p is log-concave: $\frac{p_n}{p_{n-1}} \geq \frac{p_{n+1}}{p_n}$ for all $n \in \{1,\ldots,N-1\}$. Then the maximum in (1) is attained by a set $\mathcal{N}$ that takes a "threshold" form $\mathcal{N} = \{n^*,\ldots,N\}$ for some threshold $n^* \in \{0,\ldots,N\}$. This yields

$$\Delta_{q,q^-} = q_{n^*-1} = \frac{N-n^*+1}{N}\frac{p_{n^*-1}}{1-\varepsilon}.$$

Therefore, when p is log-concave, the sequence $(N,\delta,p)_l$ has an unpredictable number of bad types if and only if $\max_{n \leq N-1} \frac{N-n}{N} p_n \to 0$.

This is a mild condition. For example, with independent types, p is log-concave,[6] and $\max_{n \leq N-1} \frac{N-n}{N} p_n \to 0$ whenever ε remains bounded away from 0 and 1 as $N \to \infty$.

For an example of a sequence of distributions where the number of bad types is *not* unpredictable, consider independent types with εN held constant at some $\bar{n} \in \mathbb{N}$ as $N \to \infty$, so the distribution of the number of bad types converges to a Poisson distribution with parameter $\bar{n}$. Then

$$\Delta_{q,q^-} = q_{\bar{n}} = \frac{N-\bar{n}}{N}\binom{N}{\bar{n}}\varepsilon^{\bar{n}}(1-\varepsilon)^{N-\bar{n}-1}.$$

For instance, if $\bar{n} = 1$—on average, there is exactly one bad player in the population—then

$$\Delta_{q,q^-} = (N-1)\,\varepsilon\,(1-\varepsilon)^{N-2} \sim \frac{1}{e}.$$ [7]

Thus, if $\frac{1+G}{\min\{G,L\}}$ is close to 1 then (2) bounds U significantly below 1 even if on average there is only a single bad player.[8] If instead $\bar{n}$ is sufficiently large, then Stirling's approximation gives

$$\Delta_{q,q^-} = \frac{N-\bar{n}}{N}\binom{N}{\bar{n}}\varepsilon^{\bar{n}}(1-\varepsilon)^{N-\bar{n}-1} \sim \frac{1}{\sqrt{2\pi\bar{n}}}.$$

Distributions for which $\Delta_{q,q^-} > 1/e$ for large N exist, but seem rather artificial. For example, if the number of bad types is known in advance to be even, then $\Delta_{q,q^-} = 1$.

of the *number* of bad types, which goes to ∞ as $N \to \infty$.

[6] Here $p_n = \binom{N}{n}\varepsilon^n(1-\varepsilon)^{N-n}$, and hence $\frac{p_n}{p_{n-1}} = \frac{N-n+1}{n}\frac{\varepsilon}{1-\varepsilon}$, which is decreasing in n.

[7] Throughout, $\sim$ denotes asymptotic equality.

[8] The bound is tight in this case: when $\bar{n} = 1$, under contagion strategies (discussed in Section 4) expected payoffs converge to $1/e$ as $\delta \to 1$, since $1/e$ is the probability that there are no bad types. This coincides with (2) when $\frac{1+G}{\min\{G,L\}} \approx 1$.

Proof. We allow players to access a public randomization device. This only expands the set of equilibrium payoffs.

Consider any equilibrium $\sigma = (\sigma_1, \ldots, \sigma_N)$. At the beginning of the game, use public randomization to draw a random permutation ϕ of I, and then let the players follow strategy profile $\tilde{\sigma} = \left(\sigma_{\phi(1)}, \ldots, \sigma_{\phi(N)}\right)$. This gives another equilibrium with the same average payoffs as σ, and viewed from the perspective of the beginning of the game (before ϕ is drawn), it is symmetric. It is thus without loss of generality to consider symmetric equilibria.

Fix a symmetric equilibrium, and let u_n be the expected payoff of a rational player when there are n bad players in the population, with the convention that $u_N = 0$. We claim that, for any $n \in \{0, \ldots, N\}$, the expected payoff of a bad player when there are n bad players is at least

$$\frac{1+G}{1+(G-L)_+}(u_n)_+ .^{9} \tag{3}$$

To see this, let α_t be the ex ante probability with which each rational player plays C in period t (when there are n bad players). Since matching is uniformly random, α_t is also the ex ante probability with which a rational player plays C in period t, conditional on matching with another rational player (and, also, conditional on matching with a bad player). Hence, the average period-t payoff a rational player receives when matched with another rational player is upper-bounded by the value of the linear program

$$\max_{\alpha^{CC}, \alpha^{CD}, \alpha^{DC}, \alpha^{DD} \in [0,1]} \alpha^{CC}(1) + \alpha^{CD}(-L) + \alpha^{DC}(1+G) + \alpha^{DD}(0)$$
$$\text{subject to } \alpha^{CC} + \alpha^{CD} = \alpha^{CC} + \alpha^{DC} = \alpha_t, \ \alpha^{CC} + \alpha^{CD} + \alpha^{DC} + \alpha^{DD} = 1.$$

Substituting $\alpha^{CD} = \alpha^{DC} = \alpha_t - \alpha^{CC}$ into the objective, this program is equivalent to

$$\max_{\alpha^{CC} \in [0, \alpha_t]} \alpha^{CC}(L-G) + \alpha_t(1+G-L).$$

The solution is $\alpha^{CC} = \alpha_t$ if $L \geq G$ and $\alpha^{CC} = 0$ if $L < G$, which yields value $\alpha_t\left(1+(G-L)_+\right)$. Thus, the average period-t payoff a rational player receives when matched with another rational player is at most $\alpha_t\left(1+(G-L)_+\right)$, while rational players always receive non-positive payoffs when matched with bad players. In contrast, the average period-t payoff a bad player receives when matched with a rational player equals $\alpha_t(1+G)$, while bad players receive payoff 0 when matched

[9] For $x \in \mathbb{R}$, $(x)_+ = \max\{x, 0\}$.

with each other. Thus, since there are $(N-n)$ rational players, the average period-t payoff of a rational player is at most $\alpha_t\left(1+(G-L)_+\right)(N-n)/N$, while that of a bad player is at least $\alpha_t(1+G)(N-n)/N$. Moreover, the average per-period payoff of a bad player is always non-negative. This comparison holds for every t, establishing (3).

Note that, if the true number of bad players is $n \leq N-1$, a rational player who plays *Always Defect* (the strategy that plays D at every history) receives the same payoff as that obtained by a bad player when the true number of such players is $n+1$, which by (3) is at least

$$\frac{1+G}{1+(G-L)_+}\left(u_{n+1}\right)_+,$$

Since in any equilibrium each player must prefer her equilibrium strategy to playing *Always Defect*, we have

$$\sum_{n=0}^{N-1} q_n\left(u_n\right)_+ \geq \frac{1+G}{1+(G-L)_+}\sum_{n=0}^{N-1} q_n\left(u_{n+1}\right)_+.$$

Note that

$$\sum_{n=0}^{N-1} q_n\left(u_{n+1}\right)_+ = \sum_{n=1}^{N} q_{n-1}\left(u_n\right)_+ = \sum_{n=1}^{N} q_n^-\left(u_n\right)_+ = \sum_{n=0}^{N-1} q_n^-\left(u_n\right)_+,$$

where the last line uses $q_0^- = u_N = 0$. Hence, we have

$$\sum_{n=0}^{N-1} q_n\left(u_n\right)_+ \geq \frac{1+G}{1+(G-L)_+}\sum_{n=0}^{N-1} q_n^-\left(u_n\right)_+. \tag{4}$$

Subtracting $\sum_{n=0}^{N-1} q_n^-\left(u_n\right)_+$ from both sides and noting that $\left(u_n\right)_+ \in [0,1]$ for each n, we obtain

$$\frac{\min\{G,L\}}{1+(G-L)_+}\sum_{n=0}^{N-1} q_n^-\left(u_n\right)_+ \leq \sum_{n=0}^{N-1}\left(q_n - q_n^-\right)\left(u_n\right)_+ \leq \max_{\mathcal{N}\subset\{0,\ldots,N-1\}}\sum_{n\in\mathcal{N}}\left(q_n - q_n^-\right) \leq \Delta_{q,q^-}.$$

Hence, since $U = \sum_{n=0}^{N-1} q_n u_n$, we have

$$U \leq \sum_{n=0}^{N-1} q_n\left(u_n\right)_+ = \sum_{n=0}^{N-1} q_n^-\left(u_n\right)_+ + \sum_{n=0}^{N-1}\left(q_n - q_n^-\right)\left(u_n\right)_+ \leq \frac{1+(G-L)_+}{\min\{G,L\}}\Delta_{q,q^-} + \Delta_{q,q^-} = \frac{1+G}{\min\{G,L\}}\Delta_{q,q^-}.$$

■

Theorem 1 is related to several prior results with the flavor that players behave selfishly in games where their probability of being pivotal for others' decisions is small. In static models, Rob (1989) and Mailath and Postlewaite (1990) show that the probability of efficient public good provision goes to zero in large groups of agents with privately-known valuations for the good. In complete-information repeated games, Green (1980), Sabourian (1990), Fudenberg, Levine, and Pesendorfer (1998), Al-Najjar and Smorodinsky (2001), and Pai, Roth, and Ullmann (2017) show that only approximate stage-game Nash equilibria can be played when each player's action has only a small impact on other players' signals; however, in these papers "approximate" and "small" depend on δ, so they do not yield anti-folk theorems in the sense of Theorem 1 (i.e., inefficiency for all δ). In addition, the key step in the proof of Theorem 1—that when there are n bad types each of them receives payoff at least $\frac{1+G}{1+(G-L)_+}(u_n)_+$, and hence (4) holds since a rational player can pretend to be bad—does not seem especially similar to the proofs of any of these results.

Suppose that, in addition to possibly being committed to *Always Defect*, players could be committed to various other strategies with arbitrary probabilities. Continue to define U_i as the expected payoff of the rational type of player i, with $U = \frac{1}{N}\sum_i U_i$, continue to define p_n as the probability that there are n *Always Defect* types, and similarly leave the definitions of q_n and q_n^+ unchanged. Then Theorem 1 and its proof remain valid—in reading the proof for this more general model, one must only interpret u_n as the expected payoff of a rational player conditional on the event that there are n *Always Defect* types, without conditioning on whether the remaining $N-n-1$ players are rational or are committed to some other strategy. Our finding that cooperation is impossible in large anonymous games with incomplete information thus holds for very general type spaces, so long as *Always Defect* types are present with non-vanishing probability, and the distribution of the number of such types is smooth.

If instead the number of *Always Defect* types can be predicted with sufficient accuracy then we expect a folk theorem to hold, even if the predicted number of *Always Defect* types is large. For example, we anticipate that if it is known in advance that exactly n out of a fixed number of players N are bad (even if it is not known in advance which players are the bad ones), then for sufficiently high δ there is an equilibrium where all rational players receive payoffs approximately $1-\frac{n}{N-1}(1+L)$ (which is the payoff of a rational player when rational players always cooperate and bad players always defect), provided this payoff is non-negative. The intuition is that if a rational player pretended to be bad, this deviation would eventually be detected by her opponents (as now $n+1$ players would be defecting, rather than n) and could be punished.[10]

Theorem 1 generalizes beyond uniform random matching. Suppose that, given history profile $h^t = \left(h_i^t\right)_i$, players i and j meet with probability $\psi_{ij}\left(h^t\right)$. Thus, the matching process can be non-

[10] We believe this result could be proved by extending the proof of Deb, Sugaya, and Wolitzky (2019).

stationary and history-dependent. We do assume that the matching process is ex ante symmetric across players: for any permutation f on I, we have $\psi_{ij}\left(\left(h_i^t\right)_i\right) = \psi_{f(i)f(j)}\left(\left(h_{f(i)}^t\right)_i\right)$ for all $\left(i,j,t,h^t\right)$. The game is now parameterized by (N,δ,p,ψ). We also assume that the meeting probabilities between any two players cannot be too unequal: letting

$$R := \frac{\sup_{i,j,t,h^t} \psi_{ij}\left(h^t\right)}{\inf_{i,j,t,h^t} \psi_{ij}\left(h^t\right)},$$

we assume that

$$R^2 < \frac{1+G}{1+(G-L)_+}.$$

Theorem 1 can be extended to show that, for any Nash equilibrium average payoff U, we have

$$U \leq \frac{1+G}{1+G-R^2\left(1+(G-L)_+\right)}\Delta_{q,q^+}.$$

We prove this in Appendix A.1. Note that uniform random matching corresponds to $R=1$, which recovers Theorem 1.

We have established that, under a smoothness assumption on the distribution of the number of bad types (e.g., if each player is bad with independent probability bounded away from 0 and 1), cooperation is impossible under the following three conditions:

1. The group is large.

2. Players are anonymous.

3. Players cannot communicate.

The rest of the paper relaxes these assumptions in turn.

4 The Optimal Size of Groups of Anonymous, Patient Players

Assume independent types with fixed $\varepsilon \in (0,1)$. Theorem 1 implies that cooperation is impossible when $N \to \infty$. Hence, a natural way to support cooperation is to segregate the population into smaller groups.[11] This has the benefit of reducing the prevalence of bad types in each group; however, there is also a cost in terms of foregone gains from trade with individuals excluded from

[11] The question of what determines whether individuals cooperate in larger or smaller groups is a fundamental one. See, e.g., Seabright (2010) or Bowles and Gintis (2011) for broad perspectives on this issue.

the group. We model this cost in a very simple way: we assume that, in a group of N players, the players match to play the prisoner's dilemma $f(N)$ times per unit of real time, where f is an increasing and weakly concave function. Here, concavity of f reflects diminishing marginal gains from trade from increasing the size of the group—in the simple case where marginal gains from trade are constant, f is linear.

In this section, we derive some simple results on the efficient size of a group of anonymous, patient players, where "patient" here means that we consider the limit $\delta \to 1$ for fixed N.

A challenge is specifying which equilibrium we expect to be played among a population of N agents. We will consider two diametrically opposed approaches and show that they yield very similar results. Our results in this section are thus quite suggestive; still, they are intended more to point out some implications of our model for the questions of optimal group size and local vs. generalized trust, rather than providing a completely definitive solution.

The first approach is to consider the *contagion strategies* of Kandori (1992): each player takes C until she observes a single play of D, and subsequently takes D forever. Contagion strategies are the simplest and best-understood strategies for supporting cooperation in a population of anonymous players. For any fixed N, $\varepsilon \in \left(0, \frac{1}{1+L}\right)$, and $\eta > 0$, contagion strategies form an η-Nash equilibrium for sufficiently high δ.[12] Moreover, as $\delta \to 1$, per-period payoffs under contagion strategies converge to

$$(1-\varepsilon)^N f(N).$$

This follows because if all players are rational, cooperation endures forever and each player receives a per-period payoff of $f(N)$; if instead even one player is bad, almost surely defection eventually spreads throughout the population, and players' per-period payoffs converge to 0 as $\delta \to 1$. Treating N as a continuous variable for convenience, the optimal group size is given by the solution to the first-order condition

$$\frac{f'(N)}{f(N)} = -\log(1-\varepsilon).^{13}$$

Since $-\log(1-\varepsilon) = \varepsilon + O(\varepsilon)$, in the case of constant marginal gains from trade (i.e., f linear), we obtain the following simple result:

[12] Whether they form a sequential (hence, exact Nash) equilibrium depends on parameters. A partial characterization is provided by equations (1) and (2) of Kandori (1992). In the iterated limit where first $\delta \to 1$ and then $N \to \infty$, the right-hand side of Kandori's equation (2) can be shown to converge to a constant $c^* < 1$. A sufficient condition for contagion strategies to form a sequential equilibrium in this iterated limit is then $\frac{\min\{G,L\}}{1+G} \geq c^*$; a sufficient condition for them *not* to form a sequential equilibrium is $\frac{\max\{G,L\}}{1+G} < c^*$.

[13] The solution is unique since f is increasing and concave.

Proposition 1 *Under contagion strategies, with independent types and constant marginal gains from trade, the optimal size of a group of anonymous, patient players is approximately equal to* $\frac{1}{\varepsilon}$.

Furthermore, writing $f(N) = \alpha N$ *for some* $\alpha > 0$, *when* ε *is small per-period payoffs in a group of optimal size* $\frac{1}{\varepsilon}$ *can be approximated as*

$$(1-\varepsilon)^{\frac{1}{\varepsilon}} \frac{\alpha}{\varepsilon} \sim \frac{\alpha}{e\varepsilon}.$$

The above analysis under contagion strategies is unsatisfactory insofar as contagion strategies do not deliver the best possible equilibrium payoffs. An upper bound on the best payoffs that can be attained in an *ex post Nash equilibrium* (i.e., an equilibrium in which players' strategies are optimal regardless of the number of commitment types) in a population of N players is given by

$$\sum_{n=0}^{N} \binom{N}{n} \varepsilon^n (1-\varepsilon)^{N-n} \left(\frac{1+(G-L)_+}{1+G} \right)^n f(N). \tag{5}$$

To see this, note that, as in the proof of Theorem 1, in every equilibrium in which rational players obtain payoffs u_n when there are n bad types, bad types obtain payoffs at least $\frac{1+G}{1+(G-L)_+} u_n$ when there are n bad types. Since in an ex post equilibrium a rational player must not benefit from pretending to be bad no matter how many bad types are present, we have

$$u_n \geq \frac{1+G}{1+(G-L)_+} u_{n+1} \text{ for all } n \in \{0, \ldots, N-1\}. \tag{6}$$

Since $u_0 \leq f(N)$, this implies that $u_n \leq \left(\frac{1+(G-L)_+}{1+G} \right)^n f(N)$ for each n. Since $U = \sum_{n=0}^{N} p_n u_n$ and $p_n = \binom{N}{n} \varepsilon^n (1-\varepsilon)^{N-n}$, we obtain (5).

We believe (6) is an important robustness criterion: if this condition is violated, a rational player who somehow learns her opponents' types before the game begins will sometimes have an incentive to deviate. Moreover, we conjecture that, for any $\eta > 0$, payoffs arbitrarily close to the upper bound (5) can be attained in an η-Nash equilibrium in the iterated limit where first $\delta \to 1$ and then $N \to \infty$.[14] This motivates the problem of choosing N to maximize (5). We will see that doing so gives the same qualitative result as for contagion strategies.

[14] We believe this can be proved by combining the arguments of Deb, Sugaya, and Wolitzky (2019) and Sugaya and Yamamoto (2019).

Note that, by the binomial theorem,

$$\sum_{n=0}^{N}\binom{N}{n}\varepsilon^{n}(1-\varepsilon)^{N-n}\left(\frac{1+(G-L)_{+}}{1+G}\right)^{n}f(N)=\left(1-\varepsilon+\varepsilon\frac{1+(G-L)_{+}}{1+G}\right)^{N}f(N).$$

Hence, the optimal group size is given by the solution to the first-order condition

$$\frac{f'(N)}{f(N)}=-\log\left(1-\varepsilon+\varepsilon\frac{1+(G-L)_{+}}{1+G}\right)=\frac{\min\{G,L\}}{1+G}\varepsilon+O(\varepsilon).$$

In the case of constant marginal gains from trade, we obtain

Proposition 2 *In the conjectured optimal ex post η-Nash equilibrium, with independent types and constant marginal gains from trade, the optimal size of a group of anonymous, patient players is approximately equal to $\frac{1+G}{\min\{G,L\}}\frac{1}{\varepsilon}$.*

Furthermore, writing $f(N)=\alpha N$ for some $\alpha>0$, when ε is small per-period payoffs in a group of optimal size $\frac{1+G}{\min\{G,L\}}\frac{1}{\varepsilon}$ can be approximated as

$$\left(1-\frac{\min\{G,L\}}{1+G}\varepsilon\right)^{\frac{1+G}{\min\{G,L\}}\frac{1}{\varepsilon}}\frac{1+G}{\min\{G,L\}}\frac{\alpha}{\varepsilon}\sim\frac{1+G}{\min\{G,L\}}\frac{\alpha}{e\varepsilon}.$$

Thus, whether we consider contagion strategies or optimal ex post equilibrium strategies, both the optimal group size and expected per-period payoffs in an optimally sized group are on the order of $1/\varepsilon$. This gives a simple account of both the benefits of restricting interaction to smaller groups (reducing the prevalence of bad types), and the comparative statics of optimal group size (the more fear of bad types, ε, the smaller the optimal group size).

We end this section by noting an important contrast with Theorem 1. Suppose that, rather than considering contagion strategies or optimal ex post equilibrium strategies, we had assumed that the group could attain the upper bound for Nash equilibrium payoffs from Theorem 1. In the constant marginal gains from trade case where $f(N)=\alpha N$, the optimal group size N would then be chosen to maximize

$$\begin{aligned}\frac{1+G}{\min\{G,L\}}\Delta_{q,q^{+}}\alpha N &= \frac{1+G}{\min\{G,L\}}\binom{N}{\varepsilon N}\varepsilon^{\varepsilon N}(1-\varepsilon)^{(1-\varepsilon)N}\alpha N\\ &\sim \frac{1+G}{\min\{G,L\}}\frac{1}{\sqrt{2\pi\varepsilon N}}\alpha N,\end{aligned}$$

which yields $N=\infty$. However, we do not know if this upper bound is attainable (even for η-Nash equilibrium); even if it is attainable, such an equilibrium would necessarily violate ex post incentive compatibility, (5).

5 Communication is Essential with Moderately Patient Players

We now return to large-population games ($N \to \infty$), but relax the assumptions that players are anonymous and cannot communicate. We consider in turn the case where players are non-anonymous but still cannot communicate, and the case where players are non-anonymous and can communicate. We do *not* treat the case where players can communicate while retaining anonymity—the reason is that communication can typically be used as a "work-around" for anonymity, for example by having players establish "passwords" the first time they meet, and subsequently using these passwords to identify themselves before taking actions. See Deb (2019) for an analysis involving such cryptographic issues, which we wish to set aside in the current paper.

The main result of this section is that communication is essential for supporting cooperation when players are moderately patient. When players are not anonymous, cooperation is possible when players are very patient—in that $(1-\delta)N \to 0$—even in the absence of explicit communication. However, without communication this high degree of patience is "almost necessary": we show that cooperation is impossible if $(1-\delta)\sqrt{N} \to \infty$. In stark contrast, when pre-play communication is allowed, cooperation is possible whenever $(1-\delta)\log N \to 0$. Communication thus exponentially increases the maximum discount rate at which cooperation can be supported.

5.1 Cooperation with Very Patient, Non-Anonymous Players

We start with a simple folk theorem for non-anonymous players. Consider the overall repeated game as consisting of $N(N-1)/2$ bilateral relationships, one for each pair of distinct players. The result states that, if $(1-\delta)N \to 0$, then any profile of feasible and strictly individually rational payoffs among these bilateral relationships can be supported in equilibrium. The intuition is that such payoffs can be supported by grim trigger strategies within each bilateral relationship.

Let $F = \text{co}\{(0,0),(1,1),(1+G,-L),(-L,1+G)\}$ denote the convex hull of the feasible payoff set in the two-player PD. Let $F^\eta = \{(v_1,v_2) \in F : v_1, v_2 \geq \eta\}$ denote the set of feasible payoffs where each player receives payoff at least $\eta > 0$. Given a sequence of games $(N,\delta,p)_l$, let E_l denote the corresponding sequence of rational-player sequential equilibrium payoff vectors: that is, $(u_i)_i \in E$ indicates that there exists a sequential equilibrium in which each player i's expected payoff when rational equals u_i. Let $\varepsilon_l = \sum_{S \ni i} p_l(S)$ denote the probability that a given player i is bad.

Proposition 3 *In the non-anonymous PD, fix $\eta > 0$ and a sequence $(N, \delta, p)_l$ satisfying $\lim_{l\to\infty}(1-\delta_l) N_l = 0$ and $\limsup_{l\to\infty} \varepsilon_l < 1$. For each $l \in \mathbb{N}$ and each $i, j \in I_l$ with $i \neq j$, fix $(v_{i,j}, v_{j,i}) \in F^\eta$. There exists $\bar{l} > 0$ such that, for all $l > \bar{l}$, there exists a payoff vector $v \in E_l$ satisfying*

$$\left| \left(\frac{1}{N_l - 1} \sum_{j \in I_l : j \neq i} (1-\varepsilon_l) v_{i,j} \right) - v_i \right| < \varepsilon_l \eta \quad \text{for all } i \in I_l.$$

Proof. See Appendix A.2. ■

We conjecture that an even larger set of payoffs can be supported in equilibrium using more complex strategies. For example, player 1 is willing to accept a negative present value payoff in her relationship with player 2, so long as she is compensated for this by a positive payoff in her relationship with player 3. In principle, such payoffs vectors can be supported by having players occasionally communicate implicitly via actions.[15] We do not pursue such a result here, since Proposition 3 suffices to make the point that cooperation is easily sustained in random matching games with non-anonymous players when $(1-\delta) N \to 0$.

Proposition 3 is a useful baseline result, but it does not provide much reassurance about the scope for cooperation in large groups at realistic discount factors. Moreover, the strategies used to prove the result do not rely on "community enforcement" at all—the community interacts as a collection of pairs of agents, where player i's behavior towards player j has no effect on her treatment by any third party k. This is not an accurate model of large-group cooperation, and it does not really address whether community enforcement is possible with non-anonymous players.

We formalize the question of whether community enforcement is possible by asking if cooperation can be supported with less patient players: in particular, when $(1-\delta)$ goes to 0 at a rate more like $1/\log N$ than $1/N$. Our remaining results show that this is possible when players can communicate, but not otherwise.

5.2 No Cooperation with Moderately Patient Players who Cannot Communicate

Assume independent types with fixed $\varepsilon \in (0, 1)$. The game is thus parameterized by the pair (N, δ).

[15] Of course, such communication would have to be incentivized.

Theorem 2 *In the non-anonymous PD with independent types, for any sequence of parameters* $(N, \delta)_l$ *satisfying* $\lim_{l \to \infty} (1 - \delta_l) \sqrt{N_l} = \infty$ *and any corresponding sequence of Nash equilibrium average payoffs* $(U)_l$*, we have* $\lim_{l \to \infty} U_l = 0$*.*

For an intuition, recall that Theorem 1 shows that cooperation is impossible when the only available punishments are collective—everyone's continuation payoff is reduced when a player switches from the rational equilibrium strategy to *Always Defect*. Non-anonymity introduces the possibility of individual punishments—where only defecting players are punished—*if* players are able to communicate the defectors' identities quickly enough. Without explicit communication, each player can transmit only a single bit of information every period. We show that it takes at least $\sqrt{N}$ periods for such information to reflect the defectors' identities with sufficient accuracy to provide incentives for cooperation. Hence, cooperation is impossible if $(1 - \delta) \sqrt{N} \to \infty$.

Proof. We show that $U \to 0$ along any sequence of equilibria even when players observe not only their own partners' identities but the entire match realization, and even when the Nash equilibrium condition is relaxed to require only that players weakly prefer their equilibrium strategies to *Always Defect*. This implies the theorem, as in any equilibrium of the game where players observe only their own partners' identities, players prefer their equilibrium strategies to *Always Defect*, and we show that $U \to 0$ under all strategy profiles that satisfy this condition even for the more permissive notion of strategy that allows actions to depends on the entire match realization.

Let μ^t denote the first t periods of the match realization, and let $h_i^t = (a_{i,\tau}, \omega_{i,\tau})_{\tau=1}^{t-1}$ denote the history of player i's own actions and opponent's actions at the beginning of period t. Slightly abusing notation, let σ_i denote a strategy for player i in the game where she observes the match realization, where $\sigma_i \left(h_i^t, \mu^t\right)$ is the (possibly mixed) action taken by player i in period t at history $\left(h_i^t, \mu^t\right)$. (Note that μ^t includes the identity of i's period-t partner.) Let 0_i (resp., 1_i) denote the event that player i is rational (resp., bad). Finally, for $x_i \in \{0_i, 1_i\}$ and $x_j \in \{0_j, 1_j\}$, let $\Pr\left(h_i^t, h_j^t | x_i, x_j, \mu^t\right)$ denote the probability that, under strategy profile σ, h_i^t and h_j^t are the period-t histories of player i and player j, conditional on the event (x_i, x_j) and the event that the match realization is μ^t.

When i's opponents play σ_{-i}, the outcome distribution when $x_i = 0_i$ but i deviates to *Always Defect* is the same as that when $x_i = 1_i$. Since in any equilibrium each player prefers her equilibrium strategy to *Always Defect*, we have

$$\begin{aligned}
&(1-\delta)\sum_t \delta^{t-1}\sum_{\mu^t}\Pr\left(\mu^t\right)\sum_{j\neq i}\frac{1}{N-1}\sum_{h_i^t,h_j^t}\left(\begin{array}{c}(1-\varepsilon)\Pr\left(h_i^t,h_j^t|0_i,0_j,\mu^t\right)u\left(\sigma_i(h_i^t,\mu^t),\sigma_j(h_j^t,\mu^t)\right)\\ +\varepsilon\Pr\left(h_i^t,h_j^t|0_i,1_j,\mu^t\right)u\left(\sigma_i(h_i^t,\mu^t),D\right)\end{array}\right)\\
\geq\ &(1-\delta)\sum_t \delta^{t-1}\sum_{\mu^t}\Pr\left(\mu^t\right)\sum_{j\neq i}\frac{1}{N-1}\sum_{h_i^t,h_j^t}\left(\begin{array}{c}(1-\varepsilon)\Pr\left(h_i^t,h_j^t|1_i,0_j,\mu^t\right)u\left(D,\sigma_j(h_j^t,\mu^t)\right)\\ +\varepsilon\Pr\left(h_i^t,h_j^t|1_i,1_j,\mu^t\right)u\left(D,D\right)\end{array}\right), \qquad (7)
\end{aligned}$$

where $u(\cdot,\cdot)$ is the stage-game payoff function, extended to mixed actions in the usual manner.

The first step of the proof puts this “incentive compatibility” constraint in a more convenient form and sums it over players $i \in I$. (Proofs of lemmas are deferred to the appendix.)

Lemma 1 *If (7) holds for all $i \in I$ then*

$$\begin{aligned}
&(1-\delta)\sum_t \delta^{t-1}\sum_{\mu^t}\Pr\left(\mu^t\right)\sum_i\frac{1}{N}\sum_{h_i^t}\Pr\left(h_i^t|0_i,\mu^t\right)\Pr\left(\sigma_i(h_i^t,\mu^t)=C\right)\min\{G,L\}\\
\leq\ &(1-\varepsilon)(1-\delta)\sum_t \delta^{t-1}\sum_{\mu^t}\Pr\left(\mu^t\right)\sum_{i,j\neq i}\frac{1}{N(N-1)}\sum_{h_j^t}\left(\begin{array}{c}\Pr\left(h_j^t|0_i,0_j,\mu^t\right)\\ -\Pr\left(h_j^t|1_i,0_j,\mu^t\right)\end{array}\right)_+(1+G) \qquad (8)
\end{aligned}$$

The heart of the proof of Theorem 2 consists of showing that the right-hand side of (8) goes to 0 as $l \to \infty$. Intuitively, this amounts to showing that the (expected, discounted, average) impact of player i's type on the signals of players $j \neq i$ is small.

Lemma 2 *If $(1-\delta)\sqrt{N} \to \infty$, then the right-hand side of (8) goes to 0.*

To see that Lemmas 1 and 2 imply the theorem, note that

$$U \leq 2(1-\varepsilon)(1-\delta)\sum_t \delta^{t-1}\sum_{\mu^t}\Pr\left(\mu^t\right)\sum_i\frac{1}{N}\sum_{h_i^t}\Pr\left(h_i^t|0_i,\mu^t\right)\Pr\left(\sigma_i(h_i^t,\mu^t)=C\right),$$

as this bound would result if total within-match payoffs equalled 2 whenever either partner cooperated and equalled 0 otherwise, which gives an upper bound on total payoffs since $1+G-L<2$. Since Lemmas 1 and 2 imply that

$$(1-\delta)\sum_t \delta^{t-1}\sum_{\mu^t}\Pr\left(\mu^t\right)\sum_i\frac{1}{N}\sum_{h_i^t}\Pr\left(h_i^t|0_i,\mu^t\right)\Pr\left(\sigma_i(h_i^t,\mu^t)=C\right)\to 0,$$

we have $U \to 0$ as well.

The proof of Lemma 2 relies on two further lemmas, which are pure probability theory results.[16]

Lemma 3 *Let $X_1, X_2, \ldots, X_N$ be i.i.d. binary random variables with $\Pr(X_i = 1) = \varepsilon$, and let S be a binary random variable defined on the same probability space. Then*

$$\sum_{i=1}^{N} \sum_{s\in\{0,1\}} \left(\Pr(s|0_i) - \Pr(s|1_i)\right)_+ \leq Np(N),$$

where

$$p(N) = \max_{n\leq N-1} \begin{pmatrix} N-1 \\ n \end{pmatrix} \varepsilon^n (1-\varepsilon)^{N-1-n}.$$

Moreover,

$$p(N) \sim \sqrt{\frac{1}{2N\pi\varepsilon(1-\varepsilon)}}. \tag{9}$$

Lemma 4 *Let $X_1, X_2, \ldots, X_N$ be i.i.d. binary random variables with $\Pr(X_i = 1) = \varepsilon$, and let S be a k-dimensional binary random variable defined on the same probability space. Then*

$$\sum_{i=1}^{N} \sum_{s\in\{0,1\}^k} \left(\Pr(s|0_i) - \Pr(s|1_i)\right)_+ \leq kNp(N).$$

We show how these lemmas imply Lemma 2. Since h_j^t is a $2(t-1)$-dimensional binary random variable whose distribution, conditional on $x_j = 0_j$ and μ^t, depends on the $N-1$ binary random variables $(X_i)_{i\neq j}$, we have

$$\begin{aligned}
& (1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}\sum_{\mu^t}\Pr\left(\mu^t\right)\sum_{i,j\neq i}\frac{1}{N(N-1)}\sum_{h_j^t}\left(\Pr\left(h_j^t|0_i,0_j,\mu^t\right)-\Pr\left(h_j^t|1_i,0_j,\mu^t\right)\right)_+ \\
\leq\ & (1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}\frac{1}{N(N-1)}2(t-1)N(N-1)p(N-1) \qquad \text{(by Lemma 4)} \\
=\ & 2\frac{\delta}{1-\delta}p(N-1) \\
\sim\ & 2\frac{\delta}{1-\delta}\sqrt{\frac{1}{2(N-1)\pi\varepsilon(1-\varepsilon)}} \qquad \text{(by (9))}.
\end{aligned}$$

Hence, if $(1-\delta)\sqrt{N} \to \infty$, then (8) goes to 0. ■

[16] We are not aware of a reference for these results. Similar results include Theorems 1 and 2 of Al-Najjar and Smorodinsky (2000) and Lemma 1 of Mossel et al. (2019).

Theorem 2 assumes each player is bad with probability ε, in which case she takes D in every period. We can establish a similar result in the related model with i.i.d. noise, where each player is forced to play D with independent probability ε in every period.

Theorem 2' *In the non-anonymous PD with i.i.d. noise, for any sequence of parameters* $(N, \delta)_l$ *satisfying* $\lim_{l\to\infty} (1-\delta_l)(N_l)^{1/4} = \infty$ *and any corresponding sequence of Nash equilibrium average payoffs* $(U)_l$*, we have* $\lim_{l\to\infty} U_l = 0$.

Proof. See Appendix A.6. ■

Comparing Theorems 2 and 2', we see that $(1-\delta)\sqrt{N} \to \infty$ is enough to guarantee that cooperation is impossible with perfectly persistent noise, while (for our proof approach) the stronger condition $(1-\delta) N^{1/4} \to \infty$ is required to rule out cooperation with i.i.d. noise.

To see the rough intuition, recall that Theorem 2 comes from considering the period 1 incentive constraint that a player prefers her equilibrium strategy to deviating to *Always Defect*. When noise is not perfectly persistent, we must consider not only the period 1 incentive constraint, but incentive constraints in every period. At a key step in the proof of Theorem 2', we take a discounted sum over incentive constraints, weighting the period t constraint by δ^t. This introduces an extra $(1-\delta)$ term in the denominator of the resulting aggregated constraint so that, where we have a $(1-\delta)\sqrt{N}$ term in the proof of Theorem 2, we now have a $(1-\delta)^2\sqrt{N}$ term.

Theorem 2' suggests that the early community enforcement literature was too optimistic about the possibility of enforcing cooperation through the threat of collective punishment. Ellison (1994) emphasizes that contagion strategies can support cooperation for relatively low discount factors in the absence of noise: implicitly, the required discount factor is approximately $1 - 1/\log N$. In stark contrast, Theorem 2 shows that, for arbitrarily small positive noise, the discount factor required to support cooperation is at least $1 - 1/N^{1/4}$. The presence of noise thus exponentially decreases the maximum discount rate at which cooperation can be supported.[17]

5.3 Cooperation with Moderately Patient Players who Can Communicate

Our final result shows that, if players can send cheap talk messages to their partners before taking actions, cooperation is possible whenever $(1-\delta)\log N \to 0$. We assume that all players (both

[17] However, the quantitative implications of this theoretical result may emerge only for very large groups. For example, Ellison computes the minimum discount factor required for cooperation for $N = 2$, 4, 10, 100, and 1000. Since $\log 1000 > 1000^{1/4}$, Theorem 2' does not directly imply that Ellison's calculations would be greatly affected by introducing noise. One could of course try to sharpen Theorem 2' to give more bite for smaller values of N. Here we have chosen to set aside small-N considerations and instead present the simplest and cleanest versions of our results.

rational types and bad types) communicate strategically to maximize their expected utility,[18] and we assume that the set of possible messages is an arbitrarily large finite set. Throughout this subsection, we fix a sequence $(N, \delta, p)_l$ and assume there exists $\bar{\alpha} \in (0,1)$ such that

$$\lim_{l \to \infty} \sum_{S \subset I : |S| < (1-\bar{\alpha})N} p(S) = 1. \tag{10}$$

That is, asymptotically almost surely, at least fraction $\bar{\alpha}$ of the population is rational. For example, with independent types and fixed $\varepsilon \in (0,1)$, (10) holds for any $\bar{\alpha} < 1 - \varepsilon$. (In general, we do not assume independent types in this subsection.)

We prove two versions of our result. First, we show that for any $\eta > 0$ there exists an η-Nash equilibrium in which rational types always cooperate with each other on path.

Theorem 3 *Fix a sequence $(N, \delta, p)_l$ satisfying (10) and $\lim_{l \to \infty} (1-\delta) \log N = 0$. In the non-anonymous PD with cheap talk, for every $\eta > 0$, there exists $\bar{l} > 0$ such that, for each $l \geq \bar{l}$, there exists an η-Nash equilibrium in which rational players always cooperate with each other along the equilibrium path of play.*

This is a simple result, and its proof (in Appendix A.7) involves strategies that seem quite realistic. Each player keeps track of a "blacklist" of players whom she believes have previously played D against a rational opponent at some point in the past. Every period, all players communicate their blacklists to their partners before taking actions. Players take C against opponents who are not on their blacklists, and take D against opponents on their blacklists.

To see that these strategies form an η-Nash whenever (10) is satisfied and $(1-\delta) \log N \to 0$, first note that, if a player defects against a rational opponent, she is added to his blacklist, and her blacklisted status then spreads through the population "exponentially quickly," regardless of her own future behavior. Formally, we rely on the following lemma.

Lemma 5 *Consider uniform random matching among N agents. Suppose that agent 1 knows a "rumor" in period 1, and in every period all agents other than agent 2 who know the rumor share it with their partners; agent 2, meanwhile, never shares the rumor. Then, letting $T = Z \log_2 N$,*

[18] Here we interpret bad types as having the same preferences as rational types, while being constrained to take D in every period. Our results would be easier to prove if we could instead freely specify the communication strategy of bad types. On the other hand, our results would be harder to prove if bad types were assumed to choose their communication strategies "adversarially."

there exist constants $c > 0$ and $\bar{Z} > 0$ (independent of N) such that, for all $Z > \bar{Z}$, the probability that everyone knows the rumor at time T is at least $1 - \exp(-cZ)$.

Moreover, suppose there are N different rumors, where initially agent i knows rumor i, and agent $i + 1$ shares all rumors except rumor i. Then, letting $T = Z \log_2 N$, there exist constants $c > 0$ and $\bar{Z} > 0$ (independent of N) such that, for all $Z > \bar{Z}$, the probability that everyone knows all *N rumors at time T is at least $1 - \exp(-cZ)$.*

Proof. Frieze and Grimmett (1985) prove a similar result in the related model where, every period, each informed player shares the rumor with a receiver selected uniformly at random from the population—rather than having players meet in pairs, as in the current model.[19] Since pairwise matching yields a different stochastic process for the number of informed players, we provide a complete proof in Online Appendix B.1.

The basic idea, though, is the same as in Frieze and Grimmett. So long as most players are uninformed, informed players are unlikely to meet each other, so the number of informed players grows exponentially. Then, once most players are informed, *un*informed players are unlikely to meet each other, so the number of uninformed players shrinks exponentially. ■

By Lemma 5, a player who takes D against a rational opponent is very likely to find herself completely excluded from cooperation within $O(\log N)$ periods. Hence, if $(1 - \delta) \log N \approx 0$, deviating to D against a rational opponent is unprofitable.

However, other deviations from this strategy profile may be (slightly) profitable—this is why it is only an η-Nash equilibrium. First, a standard problem is that a player who punishes a deviant rational opponent gets blacklisted herself, so (off path) players do not have incentives to punish deviators. This problem, though, is easily addressed by modifying the criterion for getting on the blacklist. One simple fix here would be specifying that a player gets blacklisted only if she plays D against an opponent who simultaneously plays C against her.

A much more serious problem arises in the low-probability event that a player learns that the fraction of rational players in the population is actually much smaller than $\bar{\alpha}$. In the extreme, suppose player 1 witnesses (and/or is told about) a large number of defecting players, and eventually comes to believe that player 2 is the only other rational player in the population. Then, when player 1 meets player 2, if $(1 - \delta) N \approx \infty$ she should play D against him even if he is not on her blacklist—this follows because players 1 and 2 now effectively find themselves in a two-player repeated game

[19] Frieze and Grimmett also do not consider the possibility that a single agent refuses to spread the rumor. While we need to take this feature into account (since we cannot rely on a deviant player to "self-incriminate"), it has little effect on the proof of Lemma 5.

with discount factor $\delta/(N-1)$ (since they meet on average once every $N-1$ periods). Moreover, this problem cannot easily be avoided by specifying that players take D if they learn that there are few other rational players: when facing such strategies a player should rely on her higher-order beliefs about the event that there are few rational players, and the equilibrium could easily unravel.

We therefore need a more sophisticated approach to construct an exact Nash equilibrium (indeed, a sequential equilibrium). The basic idea is to concede the impossibility of cooperation in the rare event that there are few rational types, while preventing unraveling by adjusting players' continuation payoffs to make them indifferent to cooperating when the number of rational types is close to the cutoff. Not only does this proof approach let us construct an exact sequential equilibrium, it also lets us also support a wider range of payoffs. On the other hand, not surprisingly, the strategies used in the proof (in Online Appendix B.2) are considerably more complicated than those used to prove Theorem 3.

To state this more general theorem, first fix (N, δ, p), and denote the (random) set of rational players by $\theta^* \subset I$. For each θ^*, let $F(\theta^*)$ denote the set of feasible payoff profiles where players outside θ^* always play D. That is, letting $\mathbf{a}_i : \{-i\} \to \{C, D\}$ specify an action for player i as a function of her opponent's identity, player i's expected payoff as a function of $\mathbf{a} = (\mathbf{a}_j)_{j \in I}$ equals $\hat{u}_i(\mathbf{a}) = \frac{1}{N-1} \sum_j u_i(\mathbf{a}_i(j), \mathbf{a}_j(i))$. We define $F(\theta^*) = \mathrm{co}(\{\hat{u}(\mathbf{a})\}_{\mathbf{a} \in \mathbf{A}(\theta^*)}) \subset \mathbb{R}^N$, where $\mathbf{A}(\theta^*) = \{\mathbf{a} : \mathbf{a}_j(k) = D \ \forall j \notin \theta^*, k \neq j\}$. Let $F^*(\theta^*) = F(\theta^*) \cap \mathbb{R}^N_+$ denote the set of feasible and individually rational payoffs. Note that $F^*(\theta^*)$ implicitly depends on N (but not on δ or p).

Now fix a sequence $(N, \delta, p)_l$ satisfying (10). For any $\alpha \in (0, \bar{\alpha})$ and $\eta \in (0,1)$, we define $F^{\alpha,\eta} \subset \mathbb{R}^N_+$ as the set of payoff profiles $v \in \mathbb{R}^N_+$ such that there exists $\boldsymbol{v} \in \mathbb{R}^{N|\Theta|}_+$ satisfying the following three conditions:

1. Letting $v^{\theta^*} \in \mathbb{R}^N_+$ denote the θ^*-component of $\boldsymbol{v}$, we have $v = \sum_{\theta^*} p(I \backslash \theta^*) v^{\theta^*}$.

2. For each θ^* satisfying $|\theta^*| \geq \alpha N$, we have $B^\eta\left(v^{\theta^*}\right) \subset F^*(\theta^*)$, where B^η denotes the ball of radius η.

   In contrast, for each θ^* satisfying $|\theta^*| < \alpha N$, we have $v^{\theta^*} = 0$.

3. For each $i \in I$ and each $\theta^*, \theta^{*\prime}$ satisfying (i) $|\theta^*|, \left|\theta^{*\prime}\right| \geq \alpha N$, (ii) $\theta^* \ni i$, and (iii) $\theta^{*\prime} \not\ni i$, we have $v_i^{\theta^*} - v_i^{\theta^{*\prime}} \geq \eta$.

Intuitively, $F^{\alpha,\eta}$ is the set of feasible and strictly individually rational expected payoffs such that no cooperation occurs when $|\theta^*| < \alpha N$ and each player's expected payoff is strictly greater when she is rational than when she is bad, where all strict constraints hold with η slack.[20] Note that $F^{\alpha,\eta}$ implicitly depends on N and p (but not δ). In addition, if $\alpha = 0$ then $F^{\alpha,\eta} \supset F^{\eta}$, where F^{η} is the payoff set obtained in Proposition 3. Together with continuity in α, this implies that $F^{\alpha,\eta}$ is non-empty for all sufficiently small α and all $\eta \in (0,1)$. Let E^* denote the set of sequential equilibrium payoff profiles, which implicitly depends on N, δ, and p.[21]

Theorem 4 *Fix a sequence $(N,\delta,p)_l$ satisfying (10), and fix any $\alpha \in (0,\bar{\alpha})$ and $\eta \in (0,1)$. In the non-anonymous PD with cheap talk, if $\lim_l (1-\delta_l)\log N_l = 0$ then $F^{\alpha,\eta} \subseteq E^*$ for sufficiently large l.*

The $(1-\delta)\log N \to 0$ sufficient condition in Theorem 4 is nearly the best possible: The maximum number of players who could learn about a deviation by player i within t periods is 2^t. Thus, for each $\eta \in (0,1)$, the "cost" to player i from deviating is at most

$$\begin{aligned}\sum_{t=1}^{\infty} \delta^t \min\left\{\frac{2^t}{N}, 1\right\}(1+G+L) &\leq \left(\sum_{t=1}^{\eta \log N} \frac{2^t}{N} + \delta^{\eta \log N}\right)(1+G+L) \\ &\leq \left(\frac{\eta \log N \times N^{\eta}}{N} + \exp\left(-\eta(1-\delta)\log N\right)\right)(1+G+L),\end{aligned}$$

which goes to 0 as $l \to \infty$ whenever $(1-\delta)\log N \to \infty$. Thus, if $(1-\delta)\log N \to \infty$ then for sufficiently large l the unique Nash equilibrium is *Always Defect*.

The proof of Theorem 4 proceeds by constructing a *block belief-free equilibrium*. Block belief-free equilibria were introduced by Hörner and Olszewski (2006) in the context of repeated games with almost-perfect monitoring, and were extended to community enforcement games by Deb, Sugaya, and Wolitzky (2019), and to ex post equilibria in games with incomplete information by Sugaya and Yamamoto (2019). The current proof combines elements from these three papers. The main novelty is that, since cooperation is impossible in the rare event that there are very few normal types, we must keep track of players' beliefs about the number of normal types. In particular, the equilibrium cannot be ex post with respect to the set of normal types. On the other hand, the availability of cheap talk makes it much easier to provide incentives for truthful communication, relative to the case where communication can be implemented only through payoff-relevant actions.

[20] Strict individual rationality implies that for each θ^* all players receive strictly positive expected payoffs, including bad players. One might instead want to require bad players to receive payoff 0. In this case, the definition of $F(\theta^*)$ can be modified by assuming that rational players always take D when matched with committed players. The proof of Theorem 4 then goes through as written, except that equation (26) in Online Appendix B.2 is imposed only for players $i \in \theta$.

[21] The set E^* differs from the set E defined in Section 5.1, in that the latter set contains vectors of players' expected payoffs conditional on being rational, while E^* contains vectors of unconditional expected payoffs.

We end this section by clarifying that the divergent conclusions of Theorems 2 and 4 depend on the assumption in Theorem 4 that the set of possible cheap talk messages is arbitrarily large, and in particular is large enough to include all subsets of I (the set of the players' "names"). Suppose the set of possible messages each player can send to her partner is required to be a finite set M with cardinality $|M| = \kappa$. Assume independent types with $\varepsilon \in (0,1)$, so the game is now parameterized by (N, δ, κ).

Theorem 2" *In the non-anonymous PD with independent types and cheap talk from message sets of cardinality* κ*, for any sequence of parameters* $(N, \delta, \kappa)_l$ *satisfying* $\lim_{l\to\infty} (1-\delta_l)\sqrt{N_l}/\log\kappa = \infty$ *and any corresponding sequence of Nash equilibrium average payoffs* $(U)_l$*, we have* $\lim_{l\to\infty} U_l = 0$*.*

Thus, for cheap talk to support cooperation for "much" lower discount factors in the sense of Theorem 4, the cardinality of the message set must increase exponentially in N.

Proof. The proof is the same as that of Theorem 2, except for two steps. First, instead of considering a deviation from the equilibrium strategy to *Always Defect*, we must consider a deviation to playing *Always Defect* together with sending cheap talk messages as if one were truly a bad type. This yields equation (7). Second, player j's history h_j^t may now be viewed as a vector of $(2 + \lceil \log_2 \kappa \rceil)(t-1)$ binary random variables, rather than $2(t-1)$ as in the model without cheap talk. Replacing $2(t-1)$ with $(2 + \lceil \log_2 \kappa \rceil)(t-1)$ in the last step of the proof of Theorem 2 implies that a sufficient condition for $U \to 0$ is

$$(2 + \lceil \log_2 \kappa \rceil) \frac{\delta}{1-\delta} \sqrt{\frac{1}{2(N-1)\pi\varepsilon(1-\varepsilon)}} \to 0.$$

This holds whenever $(1-\delta)\sqrt{N}/\log\kappa \to \infty$ ■

6 Conclusion

This paper has analyzed community enforcement in the presence of "bad types" who never cooperate. We established three main results. First, when players are anonymous, cooperation cannot occur in large groups under a smoothness condition on the distribution of the number of bad types, no matter how patient players might be. This anti-folk theorem stands in sharp contrast to the case where all players are known to be rational. Second, making players' identities observable does not allow cooperation to arise, unless players are patient relative to the size of the group, in that $(1-\delta)\sqrt{N}$ is not too large. Third, if communication is introduced in the model with observable identities, cooperation becomes possible for much more plausible discount factors: here, $(1-\delta)\log N \to 0$ suffices to support cooperation. We also consider the optimal size of a group of anonymous players, showing that it is approximately equal to $1/\varepsilon$ under both contagion strategies and optimal ex post equilibrium strategies.

Introducing incomplete information into repeated games with random matching raises several interesting questions. For instance, what happens in games other than the prisoner's dilemma, or for type spaces that do not include "bad" (i.e., *Always Defect*) types? What is the role of voluntary separation, assortative matching, or other "homophilic" interaction structures in such models? (Our analysis of optimal group size in Section 4 is a simple first step in this direction.) Finally, our analysis has emphasized the necessity of individual rather than collective punishment for supporting cooperation. This issue seems understudied in the repeated games literature relative to its importance in narrative accounts of group cooperation, such as Ostrom (1990).

A Appendix: Omitted Proofs

A.1 Non-Uniform Matching in Theorem 1

Since the matching process is assumed symmetric across players, as in the proof of Theorem 1 it is without loss to restrict attention to symmetric equilibria. Fix $n \in \{0, \ldots, N-1\}$, and let α_t be the ex ante probability with which each rational player plays C in period t, when there are n bad players: for each $i \in I$,

$$\alpha_t = \Pr\left(a_{i,t} = C | i \notin S, |S| = n\right),$$

where S is the set of bad players. Note that, for any set of players $\mathcal{S} \subset I$ and any h^t-measurable events $\mathcal{E}$ and $\mathcal{E}'$, we have

$$1/R \leq \frac{\Pr\left(\mu_i\left(t\right) \in \mathcal{S} | \mathcal{E}\right)}{\Pr\left(\mu_i\left(t\right) \in \mathcal{S} | \mathcal{E}'\right)} \leq R.$$

Hence, by Bayes' rule,

$$\begin{aligned}
\Pr\left(a_{i,t} = C | i \notin S, |S| = n, \mu_i\left(t\right) \notin S\right) &= \frac{\Pr\left(a_{i,t} = C | i \notin S, |S| = n\right) \Pr\left(\mu_i\left(t\right) \notin S | a_{i,t} = C, i \notin S, |S| = n\right)}{\Pr\left(\mu_i\left(t\right) \notin S | i \notin S, |S| = n\right)} \\
&\leq R\alpha_t,
\end{aligned}$$

and similarly

$$\Pr\left(a_{i,t}=C|i\notin S,|S|=n,\mu_i\left(t\right)\in S\right)\geq\alpha_t/R.$$

Now, as in the proof of Theorem 1, the average period-t payoff a rational player receives when matched with another rational player is at most $R\alpha_t\left(1+(G-L)_+\right)$, while the average period-t payoff a bad player receives when matched with a rational player is at least $\alpha_t\left(1+G\right)/R$. Hence, letting u_n be the expected payoff of a rational player when there are n bad players, the expected payoff of a bad player is at least

$$\frac{1}{R^2}\frac{1+G}{1+(G-L)_+}\left(u_n\right)_+. \tag{11}$$

Following the rest of the proof of Theorem 1 with (3) replaced by (11), we obtain

$$\sum_{n=0}^{N-1}q_n\left(u_n\right)_+\geq\frac{1}{R^2}\frac{1+G}{1+(G-L)_+}\sum_{n=0}^{N-1}q_n^+\left(u_n\right)_+,$$

and hence

$$\frac{1+G-R^2\left(1+(G-L)_+\right)}{R^2\left(1+(G-L)_+\right)}\sum_{n=0}^{N-1}q_n^+\left(u_n\right)_+\leq\sum_{n=0}^{N-1}\left(q_n-q_n^+\right)\left(u_n\right)_+\leq\Delta_{q,q^+}$$

and

$$\begin{aligned}U &\leq \sum_{n=0}^{N-1}q_n^+\left(u_n\right)_++\sum_{n=0}^{N-1}\left(q_n-q_n^+\right)\left(u_n\right)_+\leq\frac{R^2\left(1+(G-L)_+\right)}{1+G-R^2\left(1+(G-L)_+\right)}\Delta_{q,q^+}+\Delta_{q,q^+}\\ &= \frac{1+G}{1+G-R^2\left(1+(G-L)_+\right)}\Delta_{q,q^+}.\end{aligned}$$

A.2 Proof of Proposition 3

By Lemma 2 of Fudenberg and Maskin (1991), there exists $\bar{\delta}<1$ such that, for all $(v_{i,j},v_{j,i})\in F^\eta$, there exists a sequence of pure action profiles whose discounted average payoffs equal $(v_{i,j},v_{j,i})$ and whose continuation payoffs starting from any time t are within $\eta/2$ of $(v_{i,j},v_{j,i})$. Call this action path $\left(a_t^{i,j}\right)_{i\neq j,t\in\mathbb{N}}$.

Suppose each player i conditions her behavior against j on the history of outcomes in past (i,j) matches, and in particular follows $\left(a_t^{i,j}\right)_{t\in\mathbb{N}}$ if this path has been followed so far in the (i,j) matches, and otherwise reverts to D in these matches forever. By construction, this strategy profile is a sequential equilibrium if, for all $i \neq j$, we have

$$(1-\delta)\max\{G, L\} \leq \frac{\delta}{N-1}(1-\varepsilon_l)\left(v_{i,j} - \frac{\eta}{2}\right).$$

Since $v_{i,j} - \frac{\eta}{2} \geq \frac{\eta}{2}$ for all $i \neq j$ by hypothesis, a sufficient condition for this profile to be a sequential equilibrium is $\delta \geq \frac{1}{2}$ and

$$(1-\delta)N \leq \frac{\eta(1-\varepsilon_l)}{4\max\{G, L\}}.$$

If $\lim_l (1-\delta)N = 0$ and $\limsup_l \varepsilon_l < 1$, there exists $\bar{l} > 0$ such that this inequality is satisfied for all $l > \bar{l}$.

In the resulting sequential equilibrium, player i (when rational) obtains payoff $v_{i,j}$ against player j when player j is rational. When player j is bad, i obtains the payoff from action path $\left(a_t^{i,j}\right)_{t\in\mathbb{N}}$ until j deviates from this path, and then obtains payoff 0 forever. Suppose the first deviation by j from action path $\left(a_t^{i,j}\right)_{t\in\mathbb{N}}$ occurs in period t. Then i's payoff against j is at least $\left(1-\delta^t\right)u_{i,j}^{<t} + \delta^t(1-\delta)(-L)$, where $u_{i,j}^{<t}$ is i's average payoff from the first $t-1$ periods of action path $\left(a_t^{i,j}\right)_{t\in\mathbb{N}}$. Note that $u_{i,j}^{<t}$ satisfies

$$\left(1-\delta^t\right)u_{i,j}^{<t} + \delta^t u_{i,j}^{\geq t} = v_{i,j},$$

where $u_{i,j}^{\geq t}$ is i's average payoff starting from period t under action path $\left(a_t^{i,j}\right)_{t\in\mathbb{N}}$, and $u_{i,j}^{\geq t} \leq v_{i,j} + \eta/2$. Hence,

$$\left(1-\delta^t\right)u_{i,j}^{<t} \geq \left(1-\delta^t\right)v_{i,j} - \delta^t\frac{\eta}{2} \geq -\frac{\eta}{2}.$$

Therefore, for δ sufficiently high that $(1-\delta)L \leq \eta/2$, i's payoff against j is at least

$$\left(1-\delta^t\right)u_{i,j}^{<t} + \delta^t(1-\delta)(-L) \geq -\frac{\eta}{2} - \frac{\eta}{2} = -\eta.$$

Moreover, i's payoff against j is non-positive, since j always defects. Hence, i's ex ante expected payoff $\tilde{w}_{i,j}^l$ satisfies

$$v_i \in \left[\frac{1}{N_l - 1}\sum_{j\in I_l, j\neq i}\left((1-\varepsilon_l)v_{i,j} - \varepsilon_l\eta\right), \frac{1}{N_l-1}\sum_{j\in I_l, j\neq i}(1-\varepsilon_l)v_{i,j}\right],$$

as desired.

A.3 Proof of Lemma 1

Since $u(D,D)=0$ and $u(C,D)=-L$, (7) is equivalent to

$$
\begin{aligned}
&(1-\delta)\sum_t \delta^{t-1}\sum_{\mu^t}\Pr\left(\mu^t\right)\sum_{j\neq i}\frac{1}{N-1}\sum_{h_i^t,h_j^t}\left(\begin{array}{c}\Pr\left(h_i^t,h_j^t|0_i,0_j,\mu^t\right)u\left(\sigma_i(h_i^t,\mu^t),\sigma_j(h_j^t,\mu^t)\right)\\ -\frac{\varepsilon}{1-\varepsilon}\Pr\left(h_i^t|0_i,1_j,\mu^t\right)\Pr\left(\sigma_i(h_i^t,\mu^t)=C\right)L\end{array}\right)\\
\geq\ &(1-\delta)\sum_t \delta^{t-1}\sum_{\mu^t}\Pr\left(\mu^t\right)\sum_{j\neq i}\frac{1}{N-1}\sum_{h_i^t,h_j^t}\Pr\left(h_i^t,h_j^t|1_i,0_j,\mu^t\right)u\left(D,\sigma_j(h_j^t,\mu^t)\right).
\end{aligned}
$$

Subtracting a like term from both sides, this necessary condition may be rewritten as

$$
\begin{aligned}
&(1-\delta)\sum_t \delta^{t-1}\sum_{\mu^t}\Pr\left(\mu^t\right)\sum_{j\neq i}\frac{1}{N-1}\sum_{h_i^t,h_j^t}\left(\begin{array}{c}\Pr\left(h_i^t,h_j^t|0_i,0_j,\mu^t\right)\left(\begin{array}{c}u\left(\sigma_i(h_i^t,\mu^t),\sigma_j(h_j^t,\mu^t)\right)\\ -u\left(D,\sigma_j(h_j^t,\mu^t)\right)\end{array}\right)\\ -\frac{\varepsilon}{1-\varepsilon}\Pr\left(h_i^t|0_i,1_j,\mu^t\right)\Pr\left(\sigma_i(h_i^t,\mu^t)=C\right)L\end{array}\right)\\
\geq\ &(1-\delta)\sum_t \delta^{t-1}\sum_{\mu^t}\Pr\left(\mu^t\right)\sum_{j\neq i}\frac{1}{N-1}\sum_{h_i^t,h_j^t}\left(\Pr\left(h_i^t,h_j^t|1_i,0_j,\mu^t\right)-\Pr\left(h_i^t,h_j^t|0_i,0_j,\mu^t\right)\right)u\left(D,\sigma_j(h_j^t,\mu^t)\right).
\end{aligned}
$$

Since $u(C,a)-u(D,a)\leq-\min\{G,L\}$ and $u(D,a)\in\{0,1+G\}$ for each $a\in\{C,D\}$, a weaker necessary condition is

$$
\begin{aligned}
&(1-\delta)\sum_t \delta^{t-1}\sum_{\mu^t}\Pr\left(\mu^t\right)\sum_{j\neq i}\frac{1}{N-1}\sum_{h_i^t,h_j^t}\left(\begin{array}{c}\Pr\left(h_i^t,h_j^t|0_i,0_j,\mu^t\right)\Pr\left(\sigma_i(h_i^t,\mu^t)=C\right)\\ +\frac{\varepsilon}{1-\varepsilon}\Pr\left(h_i^t|0_i,1_j,\mu^t\right)\Pr\left(\sigma_i(h_i^t,\mu^t)=C\right)\end{array}\right)\min\{G,L\}\\
\leq\ &(1-\delta)\sum_t \delta^{t-1}\sum_{\mu^t}\Pr\left(\mu^t\right)\sum_{j\neq i}\frac{1}{N-1}\sum_{h_i^t,h_j^t}\left(\Pr\left(h_i^t,h_j^t|0_i,0_j,\mu^t\right)-\Pr\left(h_i^t,h_j^t|1_i,0_j,\mu^t\right)\right)_+(1+G),
\end{aligned}
$$

or equivalently

$$
\begin{aligned}
&(1-\delta)\sum_t \delta^{t-1}\sum_{\mu^t}\Pr\left(\mu^t\right)\sum_{h_i^t}\Pr\left(h_i^t|0_i,\mu^t\right)\Pr\left(\sigma_i(h_i^t,\mu^t)=C\right)\min\{G,L\}\\
\leq\ &(1-\varepsilon)(1-\delta)\sum_t \delta^{t-1}\sum_{\mu^t}\Pr\left(\mu^t\right)\sum_{j\neq i}\frac{1}{N-1}\sum_{h_j^t}\left(\Pr\left(h_j^t|0_i,0_j,\mu^t\right)-\Pr\left(h_j^t|1_i,0_j,\mu^t\right)\right)_+(1+G).
\end{aligned}
$$

Summing this necessary condition over i and dividing by N yields (8).

A.4 Proof of Lemma 3

Let us first suppose that S takes a threshold form, where $s=1$ if $\sum_i \zeta_i x_i \geq C$ for some constants $\zeta_i \in \{-1,+1\}$ for each $i \in I$ and some threshold $C \in \mathbb{Z}$, and $s=0$ otherwise. Then

$$\sum_i \sum_{s\in\{0,1\}} \left(\Pr(s|0_i) - \Pr(s|1_i)\right)_+ = \sum_i \Pr\left(\sum_{j\neq i} \zeta_j x_j = C - \frac{1}{2}(1-\zeta_i)\right).$$

For each i,

$$\Pr\left(\sum_{j\neq i} \zeta_j x_j = C - \frac{1}{2}(1-\zeta_i)\right) \leq \max_{n\leq N-1} \begin{pmatrix} N-1 \\ n \end{pmatrix} \varepsilon^n (1-\varepsilon)^{N-1-n} = p(N).$$

Hence,

$$\sum_i \Pr\left(\sum_{j\neq i} \zeta_j x_j = C - \frac{1}{2}(1-\zeta_i)\right) \leq Np(N).$$

Moreover, Stirling's approximation gives (9).

Thus, to prove the lemma, it suffices to show that $\sum_i \sum_{s\in\{0,1\}} \left(\Pr(s|0_i) - \Pr(s|1_i)\right)_+$ is maximized over all binary random variables S by an S with a threshold form.

Note that, for each $s \in \{0,1\}$ and each i,

$$\begin{aligned}
\Pr(s|0_i) - \Pr(s|1_i) &= \frac{\Pr(s,0_i)}{\Pr(0_i)} - \frac{\Pr(s) - \Pr(s,0_i)}{\Pr(1_i)} \\
&= \frac{1}{\varepsilon(1-\varepsilon)}\left(\Pr(s,0_i)\Pr(1_i) - \Pr(s)\Pr(0_i) + \Pr(s,0_i)\Pr(0_i)\right) \\
&= \frac{1}{\varepsilon(1-\varepsilon)}\left(\Pr(s,0_i) - \Pr(s)\Pr(0_i)\right).
\end{aligned}$$

Moreover,

$$\Pr(1,0_i) - \Pr(1)\Pr(0_i) = \Pr(0_i) - \Pr(0,0_i) - (1-\Pr(0))\Pr(0_i) = -\left(\Pr(0,0_i) - \Pr(0)\Pr(0_i)\right).$$

Hence, letting $\zeta_i \in \{-1,+1\}$ denote the sign of $\Pr(1,0_i) - \Pr(1)\Pr(0_i)$, we have

$$\sum_i \sum_{s\in\{0,1\}} \left(\Pr(s|0_i) - \Pr(s|1_i)\right)_+ = \frac{1}{\varepsilon(1-\varepsilon)} \sum_i \zeta_i \left(\Pr(1,0_i) - \Pr(1)\Pr(0_i)\right).$$

Denoting the underlying state space by Ω, we have

$$\sum_i \zeta_i \left(\Pr(1,0_i) - \Pr(1)\Pr(0_i)\right) = \sum_{\omega\in\Omega} \Pr(\omega) S(\omega) \left(\sum_i \zeta_i \left(1 - X_i(\omega) - \varepsilon\right)\right).$$

Clearly, this expression is maximized over signals $S : \Omega \to \{0,1\}$ by setting $S(\omega) = 1$ if $\sum_i \zeta_i X_i(\omega) \leq \lfloor \sum_i \zeta_i (1-\varepsilon) \rfloor$ and $S(\omega) = 0$ otherwise. Thus, $\sum_i \sum_{s\in\{0,1\}} \left(\Pr(s|0_i) - \Pr(s|1_i)\right)_+$ is maximized by a threshold signal.

A.5 Proof of Lemma 4

We argue by induction on k. The $k = 1$ case is Lemma 3.

Suppose the lemma holds for $k-1$. Let $s = (s_1, \ldots, s_k) \in \{0,1\}^k$. Then,

$$\begin{aligned}
& \sum_i \sum_{s=\left(s^{k-1},s_k\right)} \left(\Pr(s|0_i) - \Pr(s|1_i)\right)_+ \\
= & \sum_i \sum_{\left(s^{k-1},s_k\right)} \left(\Pr\left(s^{k-1}|0_i\right)\Pr\left(s_k|0_i,s^{k-1}\right) - \Pr\left(s^{k-1}|1_i\right)\Pr\left(s_k|1_i,s^{k-1}\right)\right)_+ \\
= & \sum_i \sum_{\left(s^{k-1},s_k\right)} \begin{pmatrix} \Pr\left(s_k|0_i,s^{k-1}\right)\left(\Pr\left(s^{k-1}|0_i\right) - \Pr\left(s^{k-1}\right)\right) \\ +\Pr\left(s^{k-1}\right)\left(\Pr\left(s_k|0_i,s^{k-1}\right) - \Pr\left(s_k|1_i,s^{k-1}\right)\right) \\ +\Pr\left(s_k|1_i,s^{k-1}\right)\left(\Pr\left(s^{k-1}\right) - \Pr\left(s^{k-1}|1_i\right)\right) \end{pmatrix}_+ \\
\leq & \begin{array}{l} \sum_i \sum_{s^{k-1}} \left(\Pr\left(s^{k-1}|0_i\right) - \Pr\left(s^{k-1}\right)\right)_+ \\ +\sum_{s^{k-1}} \Pr\left(s^{k-1}\right) \sum_i \sum_{s_k} \left(\Pr\left(s_k|0_i,s^{k-1}\right) - \Pr\left(s_k|1_i,s^{k-1}\right)\right)_+ \\ +\sum_i \sum_{s^{k-1}} \left(\Pr\left(s^{k-1}\right) - \Pr\left(s^{k-1}|1_i\right)\right)_+ \end{array}.
\end{aligned} \tag{12}$$

For the first term of (12), we have

$$\begin{aligned}
& \sum_i \sum_{s^{k-1}} \left(\Pr\left(s^{k-1}|0_i\right) - \Pr\left(s^{k-1}\right)\right)_+ \\
= & \sum_i \sum_{s^{k-1}} \left(\Pr\left(s^{k-1}|0_i\right) - (1-\varepsilon)\Pr\left(s^{k-1}|0_i\right) - \varepsilon\Pr\left(s^{k-1}|1_i\right)\right)_+ \\
= & \varepsilon \sum_i \sum_{s^{k-1}} \left(\Pr\left(s^{k-1}|0_i\right) - \Pr\left(s^{k-1}|1_i\right)\right)_+ \leq \varepsilon (k-1) Np(N)
\end{aligned}$$

by the inductive hypothesis. Similarly, for the last term, we have

$$\begin{aligned}\sum_i \sum_{s^{k-1}} \left(\Pr\left(s^{k-1}\right) - \Pr\left(s^{k-1}|1_i\right)\right)_+ &= (1-\varepsilon) \sum_i \sum_{s^{k-1}} \left(\Pr\left(s^{k-1}|0_i\right) - \Pr\left(s^{k-1}|1_i\right)\right)_+ \\ &\leq (1-\varepsilon)(k-1) Np(N).\end{aligned}$$

Finally, for the second term, since conditional on each s^{k-1}, s_k is a binary signal that depends on $X_1, \ldots, X_N$, we have

$$\sum_{s^{k-1}} \Pr\left(s^{k-1}\right) \sum_i \sum_{s_k} \left(\Pr\left(s_k|0_i, s^{k-1}\right) - \Pr\left(s_k|1_i, s^{k-1}\right)\right)_+ \leq \sum_{s^{k-1}} \Pr\left(s^{k-1}\right) Np(N) = Np(N),$$

by Lemma 3. In total, (12) is bounded by $kNp(N)$, as desired.

A.6 Proof of Theorem 2'

Let $1_{i,t}$ denote the event that player i is hit by noise (i.e., forced to play D) in period t, and let $X_{i,t}$ denote the indicator for this event. We assume that $\Pr(1_{i,t}) = \varepsilon$ for each i and t, independently across i and t. Thus, the difference between the current theorem and Theorem 2 is that there the variables $(X_{i,t})_{t=1}^{\infty}$ were assumed to be perfectly correlated, while here we assume independence.

As in the proof of Theorem 2, we consider the relaxed environment where players can observe the entire match realization. That is, player i can condition her period-t action on $h_i^t = \left((a_{i,\tau}, \omega_{i,\tau}, x_{i,\tau})_{\tau=1}^{t-1}, x_{i,t}\right)$ as well as μ^t. For each t_0, the ex ante expectation of player i's equilibrium continuation payoff starting from period t_0 equals

$$\sum_{t=t_0}^{\infty} \delta^{t-t_0} \sum_{\mu^t} \Pr\left(\mu^t\right) \sum_{j\neq i} \frac{1}{N-1} \sum_{h_i^t, h_j^t} \Pr\left(h_i^t, h_j^t|\mu^t\right) u_i\left(\sigma_i(h_i^t, \mu^t), \sigma_j(h_j^t, \mu^t)\right).$$

Suppose player i deviates to the strategy that follows the equilibrium strategy up to period t_0, always plays D in period t_0, and subsequently follows the equilibrium strategy. This strategy yields the same payoff up to period t_0, and yields expected continuation payoff starting from period t_0

equal to

$$\sum_{\mu^{t_0}} \Pr\left(\mu^{t_0}\right) \sum_{j\neq i} \frac{1}{N-1} \sum_{h_i^{t_0}, h_j^{t_0}} \Pr\left(h_i^{t_0}, h_j^{t_0} | \mu^{t_0}\right) u_i\left(D, \sigma_j(h_j^{t_0}, \mu^{t_0})\right)$$
$$+ \sum_{t=t_0}^{\infty} \delta^{t-t_0} \sum_{\mu^t} \Pr\left(\mu^t\right) \sum_{j\neq i} \frac{1}{N-1} \sum_{h_i^t, h_j^t} \Pr\left(h_i^t, h_j^t | 1_{i,t_0}, \mu^t\right) u_i\left(\sigma_i(h_i^t, \mu^t), \sigma_j(h_j^t, \mu^t)\right).$$

The equilibrium continuation payoff must exceed the deviant continuation payoff. This implies that

$$\begin{aligned}
& \sum_{\mu^{t_0}} \Pr\left(\mu^{t_0}\right) \sum_{h_i^{t_0}} \Pr\left(h_i^{t_0} | \mu^{t_0}\right) \Pr\left(\sigma_i(h_i^{t_0}, \mu^{t_0}) = C_i\right) \min\{G, L\} \\
\leq \; & \sum_{t=t_0}^{\infty} \delta^{t-t_0} \sum_{\mu^t} \Pr\left(\mu^t\right) \sum_{j\neq i} \frac{1}{N-1} \sum_{h_i^t, h_j^t} \left(\Pr\left(h_i^t, h_j^t | \mu^t\right) - \Pr\left(h_i^t, h_j^t | 1_{i,t_0}, \mu^t\right)\right) u_i\left(\sigma_i(h_i^t, \mu^t), \sigma_j(h_j^t, \mu^t)\right) \\
\leq \; & \sum_{t=t_0}^{\infty} \delta^{t-t_0} \sum_{\mu^t} \Pr\left(\mu^t\right) \sum_{j\neq i} \frac{1}{N-1} \sum_{h_j^t} \left|\Pr\left(h_j^t | \mu^t\right) - \Pr\left(h_j^t | 1_{i,t_0}, \mu^t\right)\right| (1+G+L) \\
= \; & (1-\varepsilon) \sum_{t=t_0}^{\infty} \delta^{t-t_0} \sum_{\mu^t} \Pr\left(\mu^t\right) \sum_{j\neq i} \frac{1}{N-1} \sum_{h_j^t} \left|\Pr\left(h_j^t | 0_{i,t_0}, \mu^t\right) - \Pr\left(h_j^t | 1_{i,t_0}, \mu^t\right)\right| (1+G+L). \qquad (13)
\end{aligned}$$

Note that

$$\sum_{h_j^t} \left|\Pr\left(h_j^t | 0_{i,t_0}, \mu^t\right) - \Pr\left(h_j^t | 1_{i,t_0}, \mu^t\right)\right| = 2 \sum_{h_j^t} \left(\Pr\left(h_j^t | 0_{i,t_0}, \mu^t\right) - \Pr\left(h_j^t | 1_{i,t_0}, \mu^t\right)\right)_+ .$$

Since $h_j^t = \left(\left(a_{i,\tau}, \omega_{i,\tau}, x_{i,\tau}\right)_{\tau=1}^{t-1}, x_{i,t}\right) \in \{0,1\}^{3(t-1)+1}$ is a binary signal of dimension $3(t-1)+1 < 3t$, Lemma 4 implies that

$$\sum_{j\neq i} \sum_{h_j^t} \left(\Pr\left(h_j^t | 0_{i,t_0}, \mu^t\right) - \Pr\left(h_j^t | 1_{i,t_0}, \mu^t\right)\right)_+ \leq 3t\,(N-1)\,p\,(N-1).$$

Therefore, the last line of (13) is at most

$$6\,(1-\varepsilon)\,(1+G+L)\,p\,(N-1) \sum_{t=t_0}^{\infty} \delta^{t-t_0} t = 6\,(1-\varepsilon)\,(1+G+L)\,p\,(N-1)\,\frac{(1-\delta)\,t_0 + \delta}{(1-\delta)^2}.$$

Hence, for each $t \geq 1$, we have

$$\sum_{\mu^t} \Pr\left(\mu^t\right) \sum_{h_i^t} \Pr\left(h_i^t|\mu^t\right) \Pr\left(\sigma_i(h_i^t,\mu^t) = C_i\right) \leq 6\left(1-\varepsilon\right) \frac{1+G+L}{\min\left\{G,L\right\}} p\left(N-1\right) \frac{\left(1-\delta\right)t+\delta}{\left(1-\delta\right)^2}. \quad (14)$$

Now, as in the proof of Theorem 2, we have

$$U \leq 2\left(1-\delta\right) \sum_t \delta^{t-1} \sum_{\mu^t} \Pr\left(\mu^t\right) \sum_i \frac{1}{N} \sum_{h_i^t} \Pr\left(h_i^t|\mu^t\right) \Pr\left(\sigma_i(h_i^t,\mu^t) = C_i\right).$$

Hence, summing (14) over $t = 1, ..., \infty$, with the period t inequality multiplied by $\left(1-\delta\right)\delta^{t-1}$, we obtain

$$\begin{aligned}
U &\leq 12\left(1-\varepsilon\right) \frac{1+G+L}{\min\left\{G,L\right\}} p\left(N-1\right) \sum_{t=1}^{\infty} \frac{\delta^{t-1}\left(\left(1-\delta\right)t+\delta\right)}{1-\delta} \\
&= 12\left(1-\varepsilon\right) \frac{1+G+L}{\min\left\{G,L\right\}} p\left(N-1\right) \frac{1+\delta}{\left(1-\delta\right)^2} \\
&\sim 12\left(1-\varepsilon\right) \frac{1+G+L}{\min\left\{G,L\right\}} \sqrt{\frac{1}{2N\pi\varepsilon\left(1-\varepsilon\right)}} \frac{1+\delta}{\left(1-\delta\right)^2} \qquad \text{(by Lemma 3).}
\end{aligned}$$

Therefore, if $\left(1-\delta\right)N^{1/4} \to \infty$, we have $U \to 0$.

A.7 Proof of Theorem 3

Equilibrium strategy. Each player i enters each period t with a "blacklist" $I_{i,t}^D \subset I$. Let $I_{i,1}^D = \emptyset$ for each i.

In period t, player i truthfully reports $I_{i,t}$ to her period-t opponent $\mu_t(i)$ (whether or not i is rational). When rational, i then takes action C if $\mu_t(i) \notin I_{i,t}^D$, and takes D otherwise. Committed players take D.

Denote the report of player i's opponent by $\hat{I}_{\mu_t(i),t}^D$. At the end of period t, i's blacklist updates to

$$I_{i,t+1}^D = \begin{cases} I_{i,t}^D \cup \hat{I}_{\mu_t(i),t}^D & \text{if } \mu_t(i) \text{ played } C \text{ or } i \text{ is bad,} \\ I_{i,t}^D \cup \hat{I}_{\mu_t(i),t}^D \cup \left\{\mu_t(i)\right\} & \text{if } \mu_t(i) \text{ played } D \text{ and } i \text{ is rational.} \end{cases}$$

We prove that this strategy gives an η-Nash equilibrium by (i) computing lower bounds on the equilibrium payoffs of rational and bad types, (ii) computing upper bounds on the payoffs of rational and bad types from any unilateral deviation, and (iii) showing that the latter cannot exceed the former by more than η.

Rational type equilibrium payoff. Suppose i is rational, let S be the set of bad players, and suppose that $|S| = n$. Fix any $T, Z \in \mathbb{N}$. The probability that every bad player meets a rational player at least once by period T is at least $1 - n\left(\frac{n-1}{N-1}\right)^T$. Conditional on this event, by Lemma 5, $I^D_{i,T+Z\log_2 N} = S$ with probability at least $1 - \exp(-cZ)$. Hence, with probability at least $1 - n\left(\frac{N-n}{N-1}\right)^T - \exp(-cZ)$, starting from period $T + Z\log_2 N$ player i obtains payoff 1 when she meets a rational type and obtains payoff 0 when she meets a bad type, for an expected payoff of $\frac{N-1-n}{N-1}$. For the first $T + Z\log_2 N$ periods, and with probability at most $n\left(\frac{N-n}{N-1}\right)^T + \exp(-cZ)$ for the rest of the game, player i's payoff is at least $-L$. In total, rational player i's equilibrium expected payoff, conditional on the event $|S| = n$, is at least

$$\frac{N-1-n}{N-1} - \min_{T,Z\in\mathbb{N}}\left\{\left(1-\delta^{T+Z\log_2 N}\right) + n\left(\frac{n-1}{N-1}\right)^T + \exp(-cZ)\right\}(1+L).$$

Taking the expectation with respect to n, rational player i's equilibrium unconditional expected payoff is at least

$$\sum_n p_n \frac{N-1-n}{N-1} - \sum_n p_n \min_{T,Z\in\mathbb{N}}\left\{\left(1-\delta^{T+Z\log_2 N}\right) + n\left(\frac{n-1}{N-1}\right)^T + \exp(-cZ)\right\}(1+L).$$

Now fix some $\hat{\alpha} \in (0, \bar{\alpha})$, and let $T, Z \in \mathbb{N}$ be the smallest integers such that

$$\max\left\{N(1-\hat{\alpha})^T, \exp(-cZ)\right\} \leq \frac{\eta}{4(1+\max\{G,L\})}.$$

For all $n \leq (1-\bar{\alpha})N$, we have $n\left(\frac{n-1}{N-1}\right)^T \leq N\left(\frac{(1-\bar{\alpha})N-1}{N-1}\right)^T$. For sufficiently large l, we have $\left(\frac{(1-\bar{\alpha})N-1}{N-1}\right)^T \leq (1-\hat{\alpha})^T$, and hence $n\left(\frac{N-n}{N-1}\right)^T \leq N(1-\hat{\alpha})^T \leq \frac{\eta}{4(1+\max\{G,L\})}$. Finally, we have

$$1-\delta^{T+Z\log_2 N} \leq (1-\delta)(T+Z\log_2 N) \leq (1-\delta)\hat{c}\log_2 N \text{ for some constant } \hat{c} > 0,$$

and $(1-\delta)\log_2 N$ converges to 0 as $l \to \infty$ by hypothesis. Hence, for sufficiently large l, for all $n \leq (1-\bar{\alpha})N$ we have

$$\min_{T,Z\in\mathbb{N}}\left\{\left(1-\delta^{T+Z\log_2 N}\right) + n\left(\frac{N-n}{N-1}\right)^T + \exp(-cZ)\right\}(1+L) \leq \frac{\eta}{4} + \frac{\eta}{4} + \frac{\eta}{12} = \frac{7}{12}\eta.$$

By (10), $\Pr(n \leq (1-\bar{\alpha})N) \to 1$ as $l \to \infty$. Hence, for sufficiently large l, player i's equilibrium unconditional expected payoff is at least

$$\sum_n p_n \frac{N-1-n}{N-1} - \frac{7}{12}\eta - \frac{1}{12}\eta = \sum_n p_n \frac{N-1-n}{N-1} - \frac{2}{3}\eta. \tag{15}$$

Bad type equilibrium payoff. Here we take the trivial bound that, when player i is bad, her equilibrium payoff is non-negative.

Rational type deviation payoff. We derive an upper bound for player i's payoff under any unilateral deviation. To this end, suppose that player i can observe whether her opponent is rational or bad before acting, and always takes D against bad opponents. Moreover, suppose player i's opponents blacklist her if they learn that she took D against a rational player through a path of players that excludes player i herself: that is, if player i played D against a rational opponent in period τ, then a rational player j takes D against i in period $t > \tau$ if there exists a sequence of players $(j_\tau, j_{\tau+1}, \ldots, j_{t-1})$ such that $j_\tau = \mu_\tau(i)$, $j \in \{j_\tau, \ldots, j_{t-1}\}$, $i \notin \{j_\tau, \ldots, j_{t-1}\}$, and $j_{t'+1} = \mu_{t'+1}(j_{t'})$ for each $t' \in \{\tau, \ldots, t-2\}$. By Lemma 5, if player i takes D against a rational player in period τ, then with probability $1 - \exp(-cZ)$ everyone takes D against player i starting from period $Z + \log_2 N$. Hence, player i's expected payoff at most

$$\sum_{n=0}^{N-1} p_n \frac{N-1-n}{N-1} + \left(\left(1 - \delta^{Z \log_2 N}\right) + \exp(-cZ)\right)(1+G).$$

Since $\exp(-cZ) \leq \frac{\eta}{4(1+\max\{G,L\})}$ and $(1-\delta)\log_2 N \to 0$, for sufficiently large l this is at most

$$\sum_{n=0}^{N-1} p_n \frac{N-1-n}{N-1} + \frac{1}{3}\eta.$$

Comparing this upper bound with the lower bound (15), the equilibrium strategy is η-optimal.

Bad type deviation payoff. Since player i always takes D when bad, if she meets a rational player for the first time in period τ, her continuation payoff starting from period $\tau + Z\log_2 N$ is 0 with probability at least $1 - \exp(-cZ)$. (As in the case where player i is rational, this holds regardless of player i's own behavior following period τ.) Since player i's continuation payoff against bad opponents is non-positive, her payoff under any unilateral deviation is at most

$$\delta^\tau \left(\left(1 - \delta^{Z \log_2 N}\right) + \exp(-cZ)\right)(1+G).$$

Again, for sufficiently large l, this is at most $\eta/3$. Hence, the equilibrium strategy is η-optimal.

References

[1] Ahn, Illtae, and Matti Suominen. "Word-Of-Mouth Communication and Community Enforcement." *International Economic Review* 42.2 (2001): 399-415.

[2] Ali, S. Nageeb, and David A. Miller. "Ostracism and Forgiveness." *American Economic Review* 106.8 (2016): 2329-48.

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

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

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

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

[7] Balmaceda, Felipe, and Juan F. Escobar. "Trust in Cohesive Communities." *Journal of Economic Theory* 170 (2017): 289-318.

[8] Bowles, Samuel, and Herbert Gintis. *A Cooperative Species: Human Reciprocity and its Evolution.* Princeton University Press, 2013.

[9] Chan, Jimmy. "On the Non-existence of Reputation Effects in Two-person Infinitely-repeated Games." *Working Paper* (2000).

[10] Cripps, Martin W., and Jonathan P. Thomas. "Reputation and Perfection in Repeated Common Interest Games." *Games and Economic Behavior* 18.2 (1997): 141-158.

[11] Cripps, Martin W., Eddie Dekel, and Wolfgang Pesendorfer. "Reputation with Equal Discounting in Repeated Games with Strictly Conflicting Interests." *Journal of Economic Theory* 121.2 (2005): 259-272.

[12] Coleman, James S. "Social Capital in the Creation of Human Capital." *American Journal of Sociology* 94 (1988): S95-S120.

[13] Deb, Joyee. "Cooperation and Community Responsibility," *Journal of Political Economy*, Forthcoming.

[14] Deb, Joyee, Takuo Sugaya and Alex Wolitzky. "The Folk Theorem in Repeated Games with Anonymous Random Matching." *Econometrica*, Forthcoming.

[15] Dilmé, Francesc. "Helping Behavior in Large Societies." *International Economic Review* 57.4 (2016): 1261-1278.

[16] Dixit, Avinash. "Trade Expansion and Contract Enforcement." *Journal of Political Economy* 111.6 (2003): 1293-1317.

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

[18] Frieze, Alan M., and Geoffrey R. Grimmett. "The Shortest-path Problem for Graphs with Random Arc-lengths." *Discrete Applied Mathematics* 10.1 (1985): 57-77.

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

[20] Fukuyama, Francis. *Trust: The Social Virtues and the Creation of Prosperity. Vol. 99.* New York, NY: Free Press (1995).

[21] Ghosh, Parikshit, and Debraj Ray. "Cooperation in Community Interaction without Information Flows." *Review of Economic Studies* 63.3 (1996): 491-519.

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

[23] Harrington Jr, Joseph E. "Cooperation in a One-shot Prisoners' Dilemma." *Games and Economic Behavior* 8.2 (1995): 364-377.

[24] Heller, Yuval, and Erik Mohlin. "Observations on Cooperation." *Review of Economic Studies* 85.4 (2017): 2253-2282.

[25] Hörner, Johannes, and Wojciech Olszewski. "The Folk Theorem for Games with Private Almost-perfect Monitoring." *Econometrica* 74.6 (2006): 1499-1544.

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

[27] Knack, Stephen, and Philip Keefer. "Does Social Capital have an Economic Payoff? A Cross-Country Investigation," *Quarterly Journal of Economics* 112.4 (1997): 1251-1288.

[28] Kranton, Rachel E. "Reciprocal Exchange: a Self-sustaining System." *American Economic Review* 86.4 (1996): 830-851.

[29] Lippert, Steffen, and Giancarlo Spagnolo. "Networks of Relations and Word-of-mouth Communication." *Games and Economic Behavior* 72.1 (2011): 202-217.

[30] Mailath, George J., and Andrew Postlewaite. "Asymmetric Information Bargaining Problems with Many Agents." *Review of Economic Studies* 57.3 (1990): 351-367.

[31] Mailath, George J., and Larry Samuelson. *Repeated Games and Reputations: Long-run Relationships.* Oxford University Press, 2006.

[32] Mossel, Elchanan, Manuel Mueller-Frank, Allan Sly, and Omer Tamuz. "Social Learning Equilibria." *Working Paper* (2019).

[33] Okuno-Fujiwara, Masahiro, and Andrew Postlewaite. "Social Norms and Random Matching Games." *Games and Economic Behavior* 9.1 (1995): 79-109.

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

[35] Pai, Mallesh M., Aaron Roth, and Jonathan Ullman. "An Antifolk Theorem for Large Repeated Games." *ACM Transactions on Economics and Computation (TEAC)* 5.2 (2017): 10.

[36] Putnam, Robert D., Robert Leonardi, and Raffaella Y. Nanetti. *Making Democracy Work: Civic Traditions in Modern Italy.* Princeton University Press, 1994.

[37] Rob, Rafael. "Pollution Claim Settlements under Private Information." *Journal of Economic Theory* 47.2 (1989): 307-333.

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

[39] Seabright, Paul. *The Company of Strangers: A Natural History of Economic Life.* Princeton University Press, 2010.

[40] Sommerfeld, Ralf D., Hans-Jürgen Krambeck, Dirk Semmann, and Manfred Milinski. "Gossip as an Alternative for Direct Observation in Games of Indirect Reciprocity." *Proceedings of the National Academy of Sciences* 104.44 (2007): 17435-17440.

[41] Sugaya, Takuo, and Yuichi Yamamoto. "Common Learning and Cooperation in Repeated Games." *Working Paper* (2019).

[42] Takahashi, Satoru. "Community Enforcement when Players Observe Partners' Past Play," *Journal of Economic Theory* 145.1 (2010): 42-62.

[43] Wolitzky, Alexander. "Cooperation with Network Monitoring." *Review of Economic Studies* 80.1 (2012): 395-427.

B Online Appendix

B.1 Proof of Lemma 5

For each $i \neq j$, let $N_t^{-i}(j)$ denote the (random) number of players in the set $-i = I \backslash \{i\}$ who, by period t, have met a player in $-i$ who met a player in $-i$ who... met player j. We wish to show that there exists a constant $c > 0$ such that $\Pr\left(N_T^{-i}(j) = N - 1 \; \forall i, j\right) \geq 1 - \exp(-cZ)$. The idea of the proof is to show that, with high probability, $\min_{i,j} N_t^{-i}(j)$ grows exponentially in t until it reaches a constant fraction of N, and that subsequently $N - \min_{i,j} N_t^{-i}(j)$ shrinks exponentially.

We first show that $\min_{i,j} N_t^{-i}(j)$ grows exponentially until it reaches $\frac{2}{3}N$.

Lemma 6 *There exists $\bar{\alpha} \in \left(0, \frac{1}{2}\right]$ such that, for every N and $n \leq \frac{2}{3}N$,*

$$
\begin{aligned}
& \Pr\left(\min_{i,j} N_{t+1}^{-i}(j) \leq (1 + \bar{\alpha}) \min_{i,j} N_t^{-i}(j) | \min_{i,j} N_t^{-i}(j) = n\right) \\
\leq \quad & N(N-1) \frac{e}{2\pi \bar{\alpha}^{\frac{1}{2}} (1 - \bar{\alpha})^{\frac{1}{2}}} \left(\frac{1}{2}\right)^{\frac{\bar{\alpha} n}{2}}.
\end{aligned}
$$

Proof. By monotonicity in the number of informed players and symmetry, it suffices to prove that, for each particular $i \neq j$,

$$
\Pr\left(N_{t+1}^{-i}(j) \leq (1 + \bar{\alpha}) N_t^{-i}(j) | N_t^{-i}(j) = n\right) \leq \frac{e}{2\pi \bar{\alpha}^{\frac{1}{2}} (1 - \bar{\alpha})^{\frac{1}{2}}} \left(\frac{1}{2}\right)^{\frac{\bar{\alpha} n}{2}}.
$$

Fixing $i \neq j$, and suppressing i and j in the notation, let I_t be the set of players who received player j's message through a path excluding i by period t: thus, $|I_t| = n$. Note that, for each number $n' \leq N - n$ with the same parity as n, $\Pr(N_{t+1} = n + n' | N_t = n)$ is at most

$$
\begin{aligned}
& \underbrace{\binom{n}{n'}}_{\substack{\text{who in } I_t \text{ meets} \\ \text{players in } I \backslash (I_t \cup \{i\})}} \times \underbrace{\frac{n-1}{N-1}}_{\substack{\text{“first” player in } I_t \\ \text{meets someone in } I_t}} \times \underbrace{\frac{n-3}{N-3}}_{\substack{\text{“second” (remaining) player in } I_t \\ \text{meets some (remaining) player in } I_t}} \times \cdots \times \frac{n'+1}{N-n+n'+1} \\
= \; & \binom{n}{n'} \prod_{k=1}^{\frac{n-n'}{2}} \frac{n-2k+1}{N-2k+1}.
\end{aligned}
$$

(This is an upper bound, as we neglect the probability that the players in I_t who are selected to meet someone in $I \backslash (I_t \cup \{i\})$ actually do so.) Similarly, for each n' with the opposite parity as n, $\Pr(N_{t+1} = n + n' | N_t = n)$ is at most

$$
\begin{aligned}
& \underbrace{n}_{\text{who in } I_t \text{ meets } i} \times \underbrace{\frac{1}{N-1}}_{\text{prob. of meeting } i} \times \underbrace{\binom{n-1}{n'}}_{\substack{\text{who in } I_t \text{ meets} \\ \text{players in } I \backslash (I_t \cup \{i\})}} \times \underbrace{\frac{n-2}{N-3} \times \ldots \times \frac{n'+1}{N-n+n'}}_{\substack{\text{remaining players in } I_t \\ \text{match with each other}}} \\
= \; & \binom{n-1}{n'} \prod_{k=1}^{\frac{n-n'+1}{2}} \frac{n-2k+2}{N-2k+1} \leq \binom{n}{n'} \prod_{k=1}^{\frac{n-n'+1}{2}} \frac{n-2k+2}{N-2k+1}.
\end{aligned}
$$

For any $\alpha \in (0, \frac{1}{2}]$, if $n' \leq \alpha n$, Stirling's formula gives

$$\binom{n}{n'} \leq \frac{en^{n+\frac{1}{2}}e^{-n}}{2\pi\left(\alpha n\right)^{\alpha n+\frac{1}{2}}e^{-\alpha n}\left(\left(1-\alpha\right)n\right)^{(1-\alpha)n+\frac{1}{2}}e^{-(1-\alpha)n}} \leq \frac{e}{2\pi\left(\alpha\right)^{\alpha n+\frac{1}{2}}\left(1-\alpha\right)^{(1-\alpha)n+\frac{1}{2}}}.$$

We also have

$$\prod_{k=1}^{\frac{n-n'}{2}} \frac{n-2k+1}{N-2k+1} \leq \left(\frac{n-1}{N-1}\right)^{\frac{n-n'}{2}} \leq \left(\frac{n}{N-1}\right)^{\frac{(1-\alpha)n}{2}}, \text{ and}$$

$$\prod_{k=1}^{\frac{n-n'+1}{2}} \frac{n-2k+2}{N-2k+1} \leq \left(\frac{n}{N-1}\right)^{\frac{n-n'+1}{2}} \leq \left(\frac{n}{N-1}\right)^{\frac{(1-\alpha)n}{2}}.$$

Therefore, for any $\alpha \in (0, \frac{1}{2}]$ and $n' \leq \alpha n$, we have

$$\Pr\left(N_{t+1} = n + n' | N_t = n\right) \leq \frac{e}{2\pi\left(\alpha\right)^{\alpha n+\frac{1}{2}}\left(1-\alpha\right)^{(1-\alpha)n+\frac{1}{2}}}\left(\frac{n}{N-1}\right)^{\frac{(1-\alpha)n}{2}},$$

and hence

$$\begin{aligned}
\Pr\left(N_{t+1} \leq n + \alpha n | N_t = n\right) &\leq \frac{e\left(\alpha n+1\right)}{2\pi\left(\alpha\right)^{\alpha n+\frac{1}{2}}\left(1-\alpha\right)^{(1-\alpha)n+\frac{1}{2}}}\left(\frac{n}{N-1}\right)^{\frac{(1-\alpha)n}{2}} \\
&= \frac{e}{2\pi\alpha^{\frac{1}{2}}\left(1-\alpha\right)^{\frac{1}{2}}}\left(\frac{\left(\alpha n+1\right)^{\frac{2}{\alpha n}}}{\alpha^2\left(1-\alpha\right)^{2\frac{1-\alpha}{\alpha}}}\left(\frac{n}{N-1}\right)^{\frac{1-\alpha}{\alpha}}\right)^{\frac{\alpha n}{2}} \\
&\leq \frac{e}{2\pi\left(\alpha\right)^{\frac{1}{2}}\left(1-\alpha\right)^{\frac{1}{2}}}\left(\frac{e^2}{\alpha^2\left(1-\alpha\right)^{2\frac{1-\alpha}{\alpha}}}\left(\frac{2}{3}\frac{N}{N-1}\right)^{\frac{1-\alpha}{\alpha}}\right)^{\frac{\alpha n}{2}}. \quad (16)
\end{aligned}$$

Fix $\bar{\alpha} \in \left(0, \frac{1}{2}\right]$ such that

$$\frac{e^2}{\bar{\alpha}^2\left(1-\bar{\alpha}\right)^{2\frac{1-\alpha}{\alpha}}}\left(\frac{8}{9}\right)^{\frac{1-\bar{\alpha}}{\bar{\alpha}}} < \frac{1}{2}.$$

Such an $\bar{\alpha}$ exists as the left-hand side of this inequality goes to 0 as $\bar{\alpha} \to 0$. Since $N \geq 4$, we have $\frac{2}{3}\frac{N}{N-1} \leq \frac{8}{9}$. Hence, substituting $\alpha = \bar{\alpha}$ in (16), we have, for every N and $n \leq \frac{2}{3}N$,

$$\Pr\left(N_{t+1} \leq n + \bar{\alpha} n | N_t = n\right) \leq \frac{e}{2\pi\left(\bar{\alpha}\right)^{\frac{1}{2}}\left(1-\bar{\alpha}\right)^{\frac{1}{2}}}\left(\frac{1}{2}\right)^{\frac{\bar{\alpha} n}{2}},$$

as desired. ■

Fix $\bar{\alpha}$ satisfying the conditions of Lemma 6. Let $n^*(N)$ satisfy

$$N\left(N-1\right)\frac{e}{2\pi\bar{\alpha}^{\frac{1}{2}}\left(1-\bar{\alpha}\right)^{\frac{1}{2}}}\left(\frac{1}{2}\right)^{\frac{\bar{\alpha} n^*(N)}{2}} = \frac{1}{4}.$$

Note that

$$n^*(N) = \hat{c}\left(\log_2 N + \log_2 (N-1)\right),$$

where $\hat{c} > 0$ is a constant independent of N. The following lemma is an immediate consequence of Lemma 6.

Lemma 7 *For every n satisfying $n^*(N) \leq n \leq \frac{2}{3}N$,*

$$\Pr\left(\min_{i,j} N_{t+1}^{-i}(j) \leq (1+\bar{\alpha}) \min_{i,j} N_t^{-i}(j) \,|\, \min_{i,j} N_t^{-i}(j) = n\right) \leq \frac{1}{4}.$$

We now consider the case where $n \leq n^*(N)$, considering first the subcase where $n \geq 12$.

Lemma 8 *There exists $\bar{N}_1$ such that, for every $N \geq \bar{N}_1$ and n satisfying $12 \leq n \leq n^*(N)$,*

$$\Pr\left(\min_{i,j} N_{t+1}^{-i}(j) \leq \frac{3}{2} \min_{i,j} N_t^{-i}(j) \,|\, \min_{i,j} N_t^{-i}(j) = n\right) \leq \frac{1}{4}.$$

Proof. Taking $\alpha = \frac{1}{2}$ in (16), we have

$$\begin{aligned}
& \Pr\left(\min_{i,j} N_{t+1}^{-i}(j) \leq \frac{3}{2} \min_{i,j} N_{t+1}^{-i} | \min_{i,j} N_t^{-i}(j) = n\right) \\
\leq \quad & N(N-1) \frac{e}{2\pi\alpha^{\frac{1}{2}}(1-\alpha)^{\frac{1}{2}}} \left(\frac{e^2}{\alpha^2 (1-\alpha)^{2\frac{1-\alpha}{\alpha}}} \left(\frac{n}{N-1}\right)^{\frac{1-\alpha}{\alpha}}\right)^{\frac{\alpha n}{2}} \\
= \quad & N(N-1) \frac{e}{2\pi \left(\frac{1}{2}\right)^{\frac{1}{2}} \left(\frac{1}{2}\right)^{\frac{1}{2}}} \left(\frac{e^2}{\left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^2} \frac{n}{N-1}\right)^{\frac{n}{4}}.
\end{aligned}$$

Since $12 \leq n \leq \hat{c}\left(\log_2 N + \log_2 (N-1)\right)$, this is at most

$$N(N-1) \frac{e}{\pi} \left(16e^2 \frac{\hat{c}\left(\log_2 N + \log_2 (N-1)\right)}{N-1}\right)^3,$$

which is less than $\frac{1}{4}$ for sufficiently large N. ■

The next lemma addresses the subcase with fewer than 12 informed players.

Lemma 9 *There exists $\bar{N}_2$ such that, for every $N \geq \bar{N}_2$,*

$$\Pr\left(\min_{i,j} N_{t+6}^{-i}(j) \leq 12 | \min_{i,j} N_t^{-i}(j) \geq 1\right) \leq \frac{1}{4}.$$

Proof. Fix $i \neq j$, and suppose $N_{t+6}^{-i}(j) \leq 12$. Since $\min_{i,j} N_t^{-i}(j) \geq 1$ and $12 < 2^4$, this is possible only if $N_{t'+1}^{-i}(j) = 2N_{t'}^{-i}(j)$ for at most 3 out of the 6 periods $t' \in \{t+1, \ldots, t+6\}$. That is, in at least 3 out of these 6 periods, some player in $I_{t'}^{-i}(j)$ must meet someone in $I_{t'}^{-i}(j) \cup \{i\}$. Since by hypothesis $N_{t'}^{-i}(j) \leq 12$ for each such period t', the probability of this event is at most $\begin{pmatrix} 6 \\ 3 \end{pmatrix} \times 12 \times \left(\frac{12}{N-1}\right)^3$. Hence,

$$\Pr\left(\min_{i,j} N_{t+6}^{-i}(j) \leq 12 | \min_{i,j} N_t^{-i}(j) \geq 1\right) \leq N(N-1) \frac{20 \times 12^4}{(N-1)^3},$$

which is less than $\frac{1}{4}$ for sufficiently large N. ■

In total, since $\bar{\alpha} \le \frac{1}{2}$, we have the following lemma:

Lemma 10 *For every $N \ge \max\left\{\bar{N}_1, \bar{N}_2\right\}$,*

1. *For any $\left(N_t^{-i}(j)\right)_{i,j}$ such that $\min_{i,j} N_t^{-i}(j) \ge 1$, we have*

$$\Pr\left(\min_{i,j} N_{t+6}^{-i}(j) \le 12 | \left(N_t^{-i}(j)\right)_{i,j}\right) \le \frac{1}{4}.$$

2. *For any $\left(N_t^{-i}(j)\right)_{i,j}$ such that $\min_{i,j} N_t^{-i}(j) = n$ satisfies $12 \le n \le \frac{2}{3}N$, we have*

$$\Pr\left(\min_{i,j} N_{t+1}^{-i}(j) \le (1+\bar{\alpha}) \min_{i,j} N_t^{-i} | \left(N_t^{-i}(j)\right)_{i,j}\right) \le \frac{1}{4}.$$

We now provide a symmetric bound for the case where $N_t^{-i}(j)$ is "large" for each $i \ne j$. Let $M_t^{-i}(j) = N - 1 - N_t^{-i}(j)$ be the number of players $-i$ who have not yet received player j's message through a path excluding i; and let $J_t^{-i}(j)$ be the set of such players.

Lemma 11 *There exists $\bar{N}_3$ such that, for each $N \ge \bar{N}_3$,*

1. *For any $\left(M_t^{-i}(j)\right)_{i,j}$ such that $\max_{i,j} M_t^{-i}(j) \le 12$, we have*

$$\Pr\left(\max_{i,j} M_{t+6}^{-i}(j) > 0 | \left(M_t^{-i}(j)\right)_{i,j}\right) \le \frac{1}{4}.$$

2. *For any $\left(M_t^{-i}(j)\right)_{i,j}$ such that $\max_{i,j} M_t^{-i}(j) = n$ satisfies $12 \le n \le \frac{1}{3}N$, we have*

$$\Pr\left(\max_{i,j} M_{t+1}^{-i}(j) \ge (1-\bar{\alpha}) \max_{i,j} M_t^{-i}(j) \,|\, \max_{i,j} M_t^{-i}(j) = n\right) \le \frac{1}{4}.$$

Proof. Lemmas 6–10 provide an upper bound for the probability that fraction $\bar{\alpha}$ of players in $I_t^{-i}(j)$ do not meet players outside of $I_t^{-i}(j) \cup \{i\}$. The current lemma provides an upper bound for the probability that fraction $\bar{\alpha}$ of players in $J_t^{-i}(j)$ do not meet players outside of $J_t^{-i}(j) \cup \{i\}$. The argument is symmetric. ■

We now combine Lemmas 10 and 11 to prove Lemma 5. We first assume $N \ge \max\{\bar{N}_1, \bar{N}_2, \bar{N}_3\}$.

We have the following properties. First, if $\min_{i,j} N_t^{-i}(j) < 12$, then $\min_{i,j} N_{t+6}^{-i}(j) \ge 12$ with probability at least $\frac{3}{4}$. Second, if $12 \le \min_{i,j} N_t^{-i}(j) \le \frac{2}{3}N$, then $\min_{i,j} N_{t+1}^{-i}(j) \ge (1+\bar{\alpha}) \min_{i,j} N_t^{-i}(j)$ with probability at least $\frac{3}{4}$. (And note that $\log_{(1+\bar{\alpha})} \frac{2}{3}N$ "increases" by a factor of $(1+\bar{\alpha})$ suffice to raise $\min_{i,j} N_t^{-i}(j)$ to $\frac{2}{3}N$.) Third, if $\frac{2}{3}N \le \min_{i,j} N_t^{-i}(j) \le N - 13$—or equivalently $12 \le \max_{i,j} M_t^{-i}(j) \le \frac{1}{3}N$—then $\max_{i,j} M_{t+1}^{-i}(j) \le (1-\bar{\alpha}) \max_{i,j} M_t^{-i}(j)$ with probability at least $\frac{3}{4}$. (Note that $\log_{(1-\bar{\alpha})} 3\frac{1}{N}$ "decreases" suffice to reduce $\max_{i,j} M_t^{-i}(j)$ to 12.) Finally, if $\max_{i,j} M_t^{-i}(j) \le 12$, then $\min_{i,j} N_{t+6}^{-i}(j) = N - 1$ (equivalently $\max_{i,j} M_t^{-i}(j) = 0$) with probability at least $\frac{3}{4}$.

Combining these properties, we see that $\Pr\left(\min_{i,j} N_T^{-i}(j) = N - 1\right)$ is lower-bounded by the probability that, out of $T/6$ Bernoulli random variables with parameter $\frac{3}{4}$, the realizations of at least $2+\log_{(1+\bar{\alpha})}\frac{2}{3}N+\log_{(1-\bar{\alpha})}3\frac{1}{N}$ of them equal 1. By Hoeffding's inequality, this probability is at least

$$1-\exp\left(-2\left(\frac{3}{4}-\frac{2+\log_{(1+\bar{\alpha})}\frac{2}{3}N+\log_{(1-\bar{\alpha})}3\frac{1}{N}}{\frac{T}{6}}\right)^2\frac{T}{6}\right).$$

If $T=Z\log_2 N$, then

$$\begin{aligned}\frac{2+\log_{(1+\bar{\alpha})}\frac{2}{3}N+\log_{(1-\bar{\alpha})}3\frac{1}{N}}{\frac{Z\log_2 N}{6}} &< \frac{2+(\log_2 N)\left(\frac{1}{\log_2(1+\bar{\alpha})}-\frac{1}{\log_2(1-\bar{\alpha})}\right)}{\frac{Z\log_2 N}{6}}\\ &< \frac{6}{Z}\left(2+\frac{1}{\log_2(1+\bar{\alpha})}-\frac{1}{\log_2(1-\bar{\alpha})}\right).\end{aligned}$$

Hence, there exists $\bar{Z}_1>0$ such that if $Z>\bar{Z}_1$ then, for all $N\geq\max\{\bar{N}_1,\bar{N}_2,\bar{N}_3\}$, we have

$$\frac{2+\log_{(1+\bar{\alpha})}\frac{2}{3}N+\log_{(1-\bar{\alpha})}3\frac{1}{N}}{\frac{Z\log_2 N}{6}}<\frac{1}{4},$$

and hence

$$\Pr\left(\min_{i,j}N_T^{-i}(j)=N-1\right)\geq 1-\exp\left(-2\left(\frac{1}{2}\right)^2\frac{Z\log_2 N}{6}\right)\geq 1-\exp\left(-\frac{1}{12}Z\right).$$

Finally, for the case $N<\max\{\bar{N}_1,\bar{N}_2,\bar{N}_3\}$, Hoeffding's inequality implies

$$\Pr\left(\min_{i,j}N_T^{-i}(j)=N-1\right)\geq 1-N(N-1)\exp\left(-2\left(\frac{1}{N-1}\right)^2T\right).$$

Hence, there exist $c_1>0$ and $\bar{T}>0$ such that, for all $N<\max\{\bar{N}_1,\bar{N}_2,\bar{N}_3\}$ and $T>\bar{T}$, we have

$$\Pr\left(\min_{i,j}N_T^{-i}(j)=N-1\right)\geq 1-\exp(-c_1T).$$

Taking $c=\min\left\{\frac{1}{12},c_1\right\}$ and $\bar{Z}=\max\left\{\bar{Z}_1,\bar{T}\right\}$ completes the proof.

B.2 Proof of Theorem 4

We first prove the following theorem:

Theorem 5 *Fix a sequence $(N,\delta,p)_l$ satisfying (10), and fix any $\alpha\in(0,\bar{\alpha})$ and $\eta\in(0,1)$. In the non-anonymous PD with cheap talk, if $\lim_l(1-\delta_l)\log N_l=0$ then for any $v\in F^{\alpha,\eta}$, we have $v\in E^*$ for sufficiently large l.*

In Section B.2.10, we extend this result to show that $F^{\alpha,\eta}\subseteq E^*$ for sufficiently large l.

To prove Theorem 5, we first describe a protocol for the community to circulate messages. This protocol has the feature that, with high probability, the number of periods it takes for everyone to learn the message is on the order of $\log N$; moreover, no single player can stop the rest of the community from learning. We then use this protocol as a building block in the construction of a block belief-free equilibrium.

B.2.1 Protocol for Players to Circulate Message m

Suppose each player i wishes to disseminate a message m_i throughout the community, where each m_i is an element of some finite set M_i. We say that *players circulate message* $m = (m_i)_i$ *for* T *periods* if players obey the following protocol for the next T periods:

In each period $t \in \{1, \ldots, T\}$, all players take action D, while sending cheap-talk messages. Each player j has a "state"

$$\left(\zeta_{j,t}^{I,-i}, \zeta_{j,t}^{M,-i}\right)_{i \neq j} \subset \times_{i \neq j} \left(I \times \left(\times_{k \neq i} M_k\right)\right).$$

Intuitively, $\zeta_{j,t}^{I,-i}$ is the set of players k whose message player j has heard (directly or indirectly) via a path that excludes i, and $\zeta_{j,t}^{M,-i}|_k \subset M_k$ is the set of messages reported to j as having been sent by k.[22] Formally, for each player j and $i \neq j$, $\left(\zeta_{j,1}^{I,-i}, \zeta_{j,1}^{M,-i}\right) = (\{j\}, (\emptyset, \ldots, \emptyset, m_j, \emptyset, \ldots, \emptyset))$. In each period t, given $\left(\zeta_{j,1}^{I,-i}, \zeta_{j,1}^{M,-i}\right)_{i \neq j}$, if player j meets player k, player j sends message $\left(\zeta_{j,1}^{I,-i}, \zeta_{j,1}^{M,-i}\right)_{i \notin \{j,k\}}$. That is, player j passes all of his information to player k, except for the "$-k$" information being circulated by players $-k$. Given her opponent's message $\left(\hat{\zeta}_{k,t}^{I,-i}, \hat{\zeta}_{k,t}^{M,-i}\right)_{i \notin \{j,k\}}$, for each $i \notin \{j, k\}$, player j's next-period state is given by $\zeta_{j,t+1}^{I,-i} = \zeta_{j,t}^{I,-i} \cup \hat{\zeta}_{k,t}^{I,-i}$ and $\zeta_{j,t+1}^{M,-i}|_n = \zeta_{j,t}^{M,-i}|_n \cup \hat{\zeta}_{k,t}^{M,-i}|_n$ for all $n \neq i$. For $i \in \{j, k\}$, let $\left(\zeta_{j,t+1}^{I,-i}, \zeta_{j,t+1}^{M,-i}\right) = \left(\zeta_{j,t}^{I,-i}, \zeta_{j,t}^{M,-i}\right)$. That is, for each player $n \neq i$, player j adds $\hat{\zeta}_{k,t}^{M,-i}|_n$ to the set of messages reported to her as having been sent by n. (Throughout, we use hatted variables to denote messages.)

At the end of period T, for each $i \neq j$, if $\zeta_{j,T}^{I,-i} = -i$ and $\left|\zeta_{j,T}^{M,-i}|_n\right| = 1$ for each $n \neq i$, we say player j *infers* message $m_{-i}(j) \in \times_{k \neq i} M_k$, where $m_{-i}(j)|_n$ is equal to the unique element of $\zeta_{j,T}^{M,-i}|_n$, for each n. Otherwise, we say player j infers $m_{-i}(j) = \texttt{error}$. We also say the match realization is *erroneous* if there exists disjoint players $i \neq j \neq k \neq i$ such that, by period T, player i has not met a player in $-k$ who met a player in $-k$ who… met player j. Otherwise, the match is *regular*.

Note that, if all players follow the protocol, then at the end of period T either the match is erroneous or $m_{-i}(j) = m_{-i}$ for all $i \neq j$. Moreover, if $T = Z \log_2 N$, by Lemma 5 the probability that the match is erroneous decreases exponentially in Z. We thus have

Lemma 12 *Let $T = Z \log_2 N$. There exist $c > 0$ and $\bar{Z} > 0$ such that, for all $Z > \bar{Z}$ and all l, we have*

$$\Pr\left(m_{-i}(j) = m_{-i} \ \forall i \neq j\right) \geq 1 - \exp(-cZ).$$

Note also that whether or not the event $\{m_{-i}(j) = m_{-i}\}$ obtains is independent of player i's behavior.

B.2.2 Period 1

The very first period of the repeated game plays a special role in our construction. We denote this period by 1^* rather than 1, to clarify that this is the first period of the infinitely repeated game, rather than the first period of a block. In period 1^*, every normal player is supposed to play C.

[22] For a vector $x \in X^{N-1}$ and $k \in \{1, \ldots, N-1\}$, we denote the k^{th} coordinate of x by $x|_k$.

Given the outcome of period 1^*, let θ denote the set of players who took $a_{i,1^*} = C$ as prescribed. (Note that $\theta \subset \theta^*$, as all committed players take D, and some rational players may also take D as the result of a deviation.) In our construction, only players in θ will cooperate with each other. The strategies we construct will take θ as a persistent "state variable," and we denote the set of possible states θ by $\Theta = 2^I$. Note that each player i's period-1^* history, $h_{i,1^*} = (\mu_{1^*}(i), a_{i,1^*}, \omega_{i,1^*})$, is directly informative of θ; for this reason, players' period-1^* histories will play a distinguished role in our construction.[23]

B.2.3 Block Belief-Free Structure

We now describe the general structure of our construction (following period 1^*) and present the corresponding equilibrium conditions.

Block Strategies. We view the repeated game from period 2 on as an infinite sequence of T^{**}-period blocks, where T^{**} is a number to be specified. At the beginning of every block, each player i selects a "strategy state" $x_i^\theta \in \{G, B\}$ for each $\theta \in \Theta$ from a full support probability distribution. Given the vector $\boldsymbol{x}_i = (x_i^\theta)_{\theta \in \Theta}$ and player i's period-1^* history $h_{i,1^*}$, player i plays a behavior strategy $\sigma_i^* (\boldsymbol{x}_i, h_{i,1^*})$ (her *block strategy*) within the block. That is, in every period $t = 1, \ldots, T^{**}$ of the block, $\sigma_i^* (\boldsymbol{x}_i, h_{i,1^*})$ specifies a probability distribution over cheap talk messages and actions as a function of player i's *block history* $h_i^t = ((\mu_\tau(i), m_{i,\tau}, m_{\mu_\tau(i),\tau}, a_{i,\tau}, \omega_{i,\tau})_{\tau=1}^{t-1}, \mu_t(i))$. Denote player i's strategy set in the T^{**}-period game by Σ_i.

Players are prescribed to play C in period 1^* and subsequently use the same strategy in each block. Thus, a player's entire repeated-game strategy can be summarized by a single block strategy, together with a policy for selecting the strategy state $\boldsymbol{x}_i$ at the start of each block.

Continuation Payoffs. Conditional on the persistent state being equal to θ, player i's equilibrium continuation payoff at the end of a block is a function only of player $(i-1)$'s state x_{i-1}^θ and history $h_{i-1}^{T^{**}}$ in the previous block. (Adopt here the convention that player-names are mod N, so player $(1-1)$ is player N.) Denote this continuation payoff by $w_i^\theta(x_{i-1}^\theta, h_{i-1}^{T^{**}})$.

Thus, player $(i-1)$ is the "arbiter" of player i's payoff, in that player $(i-1)$'s choice of her strategy state $\boldsymbol{x}_{i-1}$ determines player i's equilibrium continuation payoff in each state θ. This feature is typical of block belief-free constructions, such as those in Hörner and Olszewski (2006), Deb, Sugaya, and Wolitzky (2019), and Sugaya and Yamamoto (2019).

Beliefs. Players' belief systems $(\beta_i)_{i \in I}$ are specified as a function of the block strategy profile σ. Intuitively, players believe that trembles in the current block are much less likely than trembles in previous blocks, but that, within the current block, trembles in later periods are much more likely than trembles in earlier periods. This has two important implications. First, if a player reaches a history that can be explained by some past opponents' play that does not involve any deviations within the current block, she believes with probability 1 that no one deviated within the current block. Second, if a player reaches a history that cannot be explained without appealing to deviations within the current block, but can be explained by supposing that the only within-block deviation was made by her current opponent in the current period, then she believes with probability 1 that this is indeed what occurred.

To construct the belief system, first note that N and T^{**} determine the number of possible block history profiles $(h_i^t)_{i \in I, t \leq T^{**}}$.[24] Denote this number by $\tilde{c}$. Beliefs are derived from Bayes' rule along a sequence of completely mixed strategy profiles $(\sigma^l)_{l \in \mathbb{N}}$, in which each player i "trembles"

[23] We omit messages $(m_{i,1^*}, m_{\mu_{1^*}(i),1^*})$ in the description of $h_{i,1^*}$, as there is no communication in period 1 in our construction.

[24] The size of the message sets $|M_{i,t}|$ used in the construction will be explicitly determined as a function of N and T^{**} in the course of the proof.

uniformly over all messages and actions with probability $(1/l)^{\tilde{c}(T^{**}b-t)}$ in period $t \in \{1, \ldots, T^{**}\}$ of block b. As $l \to \infty$, the resulting beliefs display the properties discussed above.

Equilibrium Conditions. Fix $\alpha \in (0, \bar{\alpha})$, $\eta \in (0,1)$, and a target payoff $\tilde{v} \in F^{\alpha,\eta}$. Let $\tilde{v}^{\theta^*}$ be the associated value given θ^*. Let p^0 and p^1 denote, respectively, the probability that a given pair of players are both rational, and the probability that exactly one of them is rational. Define $\left(v^{\theta^*}\right)_{\theta^*}$ such that, for each i,

$$v_i^{\theta^*} = \begin{cases} \tilde{v}_i^{\theta^*} - \frac{1-\delta}{\delta} \frac{p^0 + p^1\left(\frac{1+G-L}{2}\right)}{\Pr(|\theta^*| \geq \alpha N)} & \text{if } |\theta^*| \geq \alpha N, \\ 0 & \text{otherwise,} \end{cases} \tag{17}$$

and let $v = \sum_{\theta^*} p\left(I \setminus \theta^*\right) v^{\theta^*}$. In order to show $\tilde{v} \in E^*$, it suffices to show that, for sufficiently large l,

$$\left((1-\delta)\left(p^0 + p^1\left(\frac{1+G-L}{2}\right)\right) + \delta v_i \right)_i \in E^*. \tag{18}$$

(Note that the left-hand side of this expression is player i's expected payoff when rational players play C in period 1 and receive continuation payoff $\left(v_i^{\theta^*}\right)_{\theta^*}$ starting in period 2.) This follows because

$$\begin{aligned} & (1-\delta)\left(p^0 + p^1\left(\frac{1+G-L}{2}\right)\right) + \delta v_i \\ = \; & (1-\delta)\left(p^0 + p^1\left(\frac{1+G-L}{2}\right)\right) \\ & + \delta \left(\sum_{\theta^* : |\theta^*| \geq \alpha N} p\left(I \setminus \theta^*\right) \left(\tilde{v}_i^{\theta^*} - \frac{1-\delta}{\delta} \frac{p^0 + p^1\left(\frac{1+G-L}{2}\right)}{\Pr\left(|\theta^*| \geq \alpha N\right)} \right) + \Pr\left(|\theta^*| < \alpha N\right)(0) \right) \\ = \; & \tilde{v}. \end{aligned}$$

Suppose that the index l for $(N, \delta, p)_l$ is large enough so that

$$\left| \frac{1-\delta}{\delta} \frac{p^0 + p^1\left(\frac{1+G-L}{2}\right)}{\Pr\left(|\theta^*| \geq \alpha N\right)} \right| \leq \frac{\eta}{2}.$$

(This holds for large l, since $\delta \to 1$ and $\Pr\left(|\theta^*| \geq \alpha N\right) \to 1$ as $l \to \infty$.) Then $\tilde{v} \in F^{\alpha,\eta}$ implies that v satisfies the following conditions:

1. For each θ^* satisfying $|\theta^*| \geq \alpha N$, we have $B^{\frac{\eta}{2}}\left(v^{\theta^*}\right) \subset F^*\left(\theta^*\right)$. In contrast, for each θ^* satisfying $|\theta^*| < \alpha N$, we have $v^{\theta^*} = 0$.

2. For each $i \in I$ and each $\theta^*, \theta^{*\prime}$ satisfying (i) $|\theta^*|, \left|\theta^{*\prime}\right| \geq \alpha N$, (ii) $\theta^* \ni i$, and (iii) $\theta^{*\prime} \not\ni i$, we have
$$v_i^{\theta^*} - v_i^{\theta^{*\prime}} \geq \frac{\eta}{2}. \tag{19}$$

We now provide a sufficient condition to establish (18).

Fix a block length $T^{**} \in \mathbb{N}$, a block strategy profile $\left(\sigma_i^*\left(\boldsymbol{x}_i, h_i^{1^*}\right)\right)_{i \in I, \boldsymbol{x}_i \in \{G,B\}^{|\Theta|}, h_i^{1^*} \in H_i^{1^*}}$, target payoffs (conditional on both θ and $\boldsymbol{x}$) $\left(v_i^{\theta}(x_{i-1}^{\theta})\right)_{i \in I, \theta \in \Theta, x_{i-1}^{\theta} \in \{G,B\}}$, and continuation payoffs $\left(w_i^{\theta}(x_{i-1}^{\theta}, h_{i-1}^{T^{**}+1})\right)_{i \in I, \theta \in \Theta, x_{i-1}^{\theta} \in \{G,B\}, h_{i-1}^{T^{**}+1} \in H_{i-1}^{T^{**}+1}}$. We will show that the following set of conditions is sufficient for $\sum_{\theta^*} p\left(I \setminus \theta^*\right) v^{\theta^*} \in E^*$. In what follows, $\mathbb{E}^{\sigma}\left[\cdot|\cdot\right]$ denotes conditional expectation under block strategy profile σ, with the corresponding belief system defined above given σ. We also write $\tilde{h}_i^{b,t} \in \tilde{H}_i^{b,t}$ for a generic history in period t of block b of the infinitely repeated game, and write $h_i^t \in H_i^t$ for a generic block history in period t of a block. (Thus, $\tilde{h}_i^{b,t}$ records the outcomes of $(b-1)T^{**} + t - 1$ periods of play, while h_i^t records the outcomes of $t-1$ periods.) Finally, we write $\tilde{h}_i^{b,0} \in \tilde{H}_i^{b,0}$ for a generic repeated game history at the beginning of block b, before the determination of the first match in the block.

1. *[Sequential Rationality]* For each $\boldsymbol{x} \in \{G,B\}^{N|\Theta|}$, each $i \in I$, each $h_i^{1^*} \in H_i^{1^*}$, each $t \in \{1, \dots, T^{**}\}$, each $b \in \mathbb{N}$, and each $\tilde{h}_i^{b,t} \in \tilde{H}_i^{b,t}$, $\sigma_i^*\left(\boldsymbol{x}_i, h_i^{1^*}\right)$ is a maximizer (over $\sigma_i \in \Sigma_i$) of

$$\sum_{h_{-i}^{1^*} \in H_{-i}^{1^*}} \beta_i\left(h_{-i}^{1^*} | \boldsymbol{x}_{-i}, \tilde{h}_i^{b,t}\right) \mathbb{E}^{\left(\sigma_i, \sigma_{-i}^*\left(\boldsymbol{x}_{-i}, h_{-i}^{1^*}\right)\right)}\left[\begin{array}{c} (1-\delta)\sum_{\tau=1}^{T^{**}} \delta^{\tau-1}\hat{u}_i(\mathbf{a}_\tau) \\ +\delta^{T^{**}} w_i^\theta(x_{i-1}^\theta, h_{i-1}^{T^{**}+1}) \end{array} \middle| \boldsymbol{x}_{-i}, \tilde{h}_i^{b,t}\right].$$

   (Here, the sum $\sum_{\tau=1}^{T^{**}}$ is taken over all periods in the current block b, where the current period $t \in \{(b-1)T^{**} + 2, \dots bT^{**} + 1\}$ is some period in block b. Note also that sequential rationality is imposed "ex post" over vectors $\boldsymbol{x}_{-i} \in \{G,B\}^{(N-1)|\Theta|}$. This is the defining feature of a block belief-free construction. However, optimality with respect to $h_{-i}^{1^*}$ is demanded only in expectation, not ex post.)

2. *[Promise Keeping]* For each $\theta \in \Theta$, $i \in I$, $x_{i-1}^\theta \in \{G,B\}^N$, $b \in \mathbb{N}$, and $\tilde{h}^{b,0} \in \tilde{H}^{b,0}$,

$$v_i^\theta(x_{i-1}^\theta) = \mathbb{E}^{\sigma^*\left(\boldsymbol{x}, h^{1^*}\right)}\left[(1-\delta)\sum_{t=1}^{T^{**}} \delta^{t-1}\hat{u}_i(\mathbf{a}_t) + \delta^{T^{**}} w_i^\theta(x_{i-1}^\theta, h_{i-1}^{T^{**}+1}) | \tilde{h}^{b,0}, \theta\right].$$

   (Note that player i's continuation payoff $v_i^\theta(x_{i-1}^\theta)$ is allowed to depend on $\tilde{h}^b$ only through θ.)

3. *[Self-Generation]* For each $\theta \in \Theta$, $i \in I$, we have either (i) $w_i^\theta(x_{i-1}^\theta, h_{i-1}^{T^{**}+1}) \in (v_i^\theta(B), v_i^\theta(G))$ for each $x_{i-1}^\theta \in \{G,B\}$ and $h_{i-1}^{T^{**}+1} \in H_{i-1}^{T^{**}+1}$, or (ii) $w_i^\theta(x_{i-1}^\theta, h_{i-1}^{T^{**}+1}) = v_i^\theta(B) = v_i^\theta(G)$ for each $x_{i-1}^\theta \in \{G,B\}$ and $h_{i-1}^{T^{**}+1} \in H_{i-1}^{T^{**}+1}$.

4. *[Feasibility]* For each $\theta \in \Theta$ and $i \in I$, we have either $v_i^\theta \in (v_i^\theta(B), v_i^\theta(G))$ or $v_i^\theta = v_i^\theta(B) = v_i^\theta(G)$.

   (This implies that, by appropriately randomizing her strategy state x_{i-1}^θ in the first block, player $(i-1)$ can deliver the target payoff v_i^θ to player i. Moreover, this randomization has full support.)

5. *[Incentive to take C in period 1^*]* For each $i \in I$,

$$\delta\sum_{\theta\in\Theta} \Pr(\theta | a_{i,1^*} = C, i \in \theta^*)\, v_i^\theta > (1-\delta)\max\{G, L\} + \delta\sum_{\theta\in\Theta} \Pr(\theta | a_{i,1^*} = D, i \in \theta^*)\, v_i^\theta.$$

Defining $\pi_i^\theta(x_{i-1}^\theta, h_{i-1}^{T^{**}+1}) := \frac{\delta^{T^{**}}}{1-\delta}\left(w_i^\theta(x_{i-1}^\theta, h_{i-1}^{T^{**}+1}) - v_i^\theta(x_{i-1}^\theta)\right)$, we can rewrite these conditions as follows:

1. *[Sequential Rationality]* For each $\boldsymbol{x} \in \{G,B\}^{N|\Theta|}$, $i \in I$, $h_i^{1^*} \in H_i^{1^*}$, $t \in \{1, \dots, T^{**}\}$, $b \in \mathbb{N}$,

and $\tilde{h}_i^{b,t} \in \tilde{H}_i^{b,t}$, $\sigma_i^*\left(\boldsymbol{x}_i, h_i^{1^*}\right)$ maximizes

$$\sum_{h_{-i}^{1^*} \in H_{-i}^{1^*}} \beta_i\left(h_{-i}^{1^*} | \boldsymbol{x}_{-i}, \tilde{h}_i^{b,t}\right) \mathbb{E}^{\left(\sigma_i, \sigma_{-i}^*\left(\boldsymbol{x}_{-i}, h_{-i}^{1^*}\right)\right)}\left[\sum_{\tau=1}^{T^{**}} \delta^{\tau-1} \hat{u}_i\left(\mathbf{a}_\tau\right)+\pi_i^\theta(x_{i-1}^\theta, h_{i-1}^{T^{**}+1}) | \boldsymbol{x}_{-i}, \tilde{h}_i^{b,t}\right]. \tag{20}$$

2. *[Promise Keeping]* For each $\theta \in \Theta$, $i \in I$, $x_{i-1}^\theta \in \{G, B\}^N$, $b \in \mathbb{N}$, and $\tilde{h}^{b,0} \in \tilde{H}^{b,0}$,

$$v_i^\theta(x_{i-1}^\theta)=\mathbb{E}^{\sigma^*(\boldsymbol{x})}\left[\frac{1-\delta}{1-\delta^{T^{**}}} \sum_{t=1}^{T^{**}} \delta^{t-1} \hat{u}_i\left(\mathbf{a}_t\right)+\pi_i^\theta(x_{i-1}^\theta, h_{i-1}^{T^{**}+1}) | \tilde{h}^{b,0}, \theta\right]. \tag{21}$$

3. *[Self-Generation]* For each $\theta \in \Theta$ and $i \in I$, either (i)

$$\operatorname{sign}\left(x_{i-1}^\theta\right) \pi_i^\theta(x_{i-1}^\theta, h_{i-1}^{T^{**}+1})>0 \text{ and } \left|\frac{1-\delta}{\delta^{T^{**}}} \pi_i^\theta(x_{i-1}^\theta, h_{i-1}^{T^{**}+1})\right|<v_i^\theta(G)-v_i^\theta(B) \tag{22}$$

for each $x_{i-1}^\theta \in \{G, B\}$ and $h_{i-1}^{T^{**}+1} \in H_{i-1}^{T^{**}+1}$, or (ii) $\pi_i^\theta(x_{i-1}^\theta, h_{i-1}^{T^{**}+1})=0$ and $v_i^\theta(G)=v_i^\theta(B)$ for each $x_{i-1}^\theta \in \{G, B\}$ and $h_{i-1}^{T^{**}+1} \in H_{i-1}^{T^{**}+1}$, where $\operatorname{sign}\left(x_{i-1}^\theta\right):=1_{\{x_{i-1}^\theta=B\}}-1_{\{x_{i-1}^\theta=G\}}$.

4. *[Feasibility]* For each $\theta \in \Theta$ and $i \in I$,

$$v_i^\theta \in (v_i^\theta(B), v_i^\theta(G)) \text{ or } v_i^\theta=v_i^\theta(B)=v_i^\theta(G). \tag{23}$$

5. *[Incentive to take C in period 1^*]* For each $i \in I$,

$$\delta \sum_{\theta \in \Theta} \operatorname{Pr}\left(\theta | a_{i, 1^*}=C, i \in \theta^*\right) v_i^\theta>(1-\delta) \max \{G, L\}+\delta \sum_{\theta \in \Theta} \operatorname{Pr}\left(\theta | a_{i, 1^*}=D, i \in \theta^*\right) v_i^\theta. \tag{24}$$

Lemma 13 *For all $\boldsymbol{v} \in \mathbb{R}^{N|\Theta|}$ and $\delta \in [0,1)$, if there exist $T^{**} \in \mathbb{N}$, $\left(\sigma_i^*\left(\boldsymbol{x}_i, h_i^{1^*}\right)\right)_{i \in I, \boldsymbol{x}_i \in \{G,B\}^{|\Theta|}, h_i^{1^*} \in H_i^{1^*}}$, $\left(v_i^\theta(x_{i-1}^\theta)\right)_{i \in I, \theta \in \Theta, x_{i-1} \in \{G,B\}}$, and $\left(\pi_i^\theta(x_{i-1}^\theta, h_{i-1}^{T^{**}+1})\right)_{i \in I, \theta \in \Theta, x_{i-1}^\theta \in \{G,B\}, h_{i-1}^{T^{**}+1} \in H_{i-1}^{T^{**}+1}}$ such that Conditions (20)–(24) are satisfied, then*

$$\left((1-\delta)\left(p^0+p^1\left(\frac{1+G-L}{2}\right)\right)+\delta \sum_{\theta^*} p\left(I \setminus \theta^*\right) v_i^{\theta^*}\right)_i \in E^*. \tag{25}$$

Proof. Conditions (21) and (22) imply that payoffs $\left(v_i^\theta(x_{i-1}^\theta)\right)_{i \in I, \theta \in \Theta, x_{i-1} \in \{G,B\}}$ can be delivered at the beginning of each block with full support state transition probabilities, and Condition (23) then implies that, by appropriately randomizing over $\left(x_{i-1}^\theta\right)_{i \in I, \theta \in \Theta}$ before the first block (i.e., before period 2 of the repeated game), the target expected payoff vector $\boldsymbol{v}$ can be delivered. This is as in, for example, Hörner and Olszewski (2006). Condition (20) is then a more stringent version of the resulting sequential rationality constraint, as it imposes sequential rationality for each realization of $\boldsymbol{x}_{-i}$, rather than only in expectation. Thus, Conditions (20)–(23) imply that the strategies $\left(\sigma_i^*\left(\boldsymbol{x}_i, h_i^{1^*}\right)\right)_{i \in I, \boldsymbol{x}_i \in \{G,B\}^{|\Theta|}, h_i^{1^*} \in H_i^{1^*}}$ are sequentially rational and deliver continuation payoffs $\boldsymbol{v}$ starting from the second period of the repeated game. Given this, (24) implies that it is optimal for rational players to take C in period 1^*. Finally, the resulting ex ante expected payoffs are given by (25). ■

To prove Theorem 4, it thus suffices to show that, for any $\tilde{v} \in F^{\alpha,\eta}$ and v defined by (17), for sufficiently large l there exist $T^{**} \in \mathbb{N}$, $\left(\sigma_i^*\left(\boldsymbol{x}_i, h_i^{1^*}\right)\right)_{i\in I, \boldsymbol{x}_i \in \{G,B\}^{|\Theta|}, h_i^{1^*}\in H_i^{1^*}}$, $\left(v_i^\theta(x_{i-1}^\theta)\right)_{i\in I,\theta\in\Theta,x_{i-1}\in\{G,B\}}$, and $\left(\pi_i^\theta(x_{i-1}^\theta, h_{i-1}^{T^{**}+1})\right)_{i\in I,\theta\in\Theta,x_{i-1}^\theta\in\{G,B\},h_{i-1}^{T^{**}+1}\in H_{i-1}^{T^{**}+1}}$ such that Conditions (20)–(24) are satisfied.

Condition (24). It is immediate that Condition (24) is satisfied for sufficiently large l. For, taking C gives player i payoff at least

$$\begin{aligned}&(1-\delta)\left(\Pr\left(a_{\mu_{1^*}(i)} = C | i \in \theta^*\right)(1) + \Pr\left(a_{\mu_{1^*}(i)} = D | i \in \theta^*\right)(-L)\right)\\ &+\delta\left(\Pr\left(|\theta| \geq \alpha N | i \in \theta\right) \min_{\theta:|\theta|\geq\alpha N, \theta\ni i} v_i^\theta + \left(1 - \Pr\left(|\theta|\geq \alpha N | i \in \theta\right)\right)(0)\right),\end{aligned}$$

while taking D gives player i payoff at most

$$(1-\delta)\left(\Pr\left(a_{\mu_{1^*}(i)} = C | i \in \theta^*\right)(1+G) + \Pr\left(a_{\mu_{1^*}(i)} = D | i \in \theta^*\right)(0)\right) + \delta \max_{\theta:|\theta|\geq\alpha N,\theta\not\ni i} v_i^\theta.$$

Since $\lim_l \Pr\left(|\theta| \geq \alpha N | i \in \theta\right) = 1$, (19) implies (24) for sufficiently large l.

B.2.4 Target Actions

We now define a target (opponent identity-contingent) action profile $\mathbf{a}^{x^\theta}$ for each state $\theta \subset I$.

For θ satisfying $|\theta| < \alpha N$, we define $\mathbf{a}_i^{x^\theta}(j) = D$ for all $x^\theta \in \{G, B\}^N$ and $i \neq j$. That is, all players are prescribed defection. In this case, we define $v_i^\theta(G) = v_i^\theta(B) = 0$. Note that, to satisfy (22), this requires $\pi_i^\theta\left(x_{i-1}^\theta, h_{i-1}^{T^{**}+1}\right) = 0$ for all $x_{i-1}^\theta \in \{G, B\}$ and $h_{i-1}^{T^{**}+1} \in H_{i-1}^{T^{**}+1}$.

For θ satisfying $|\theta| \geq \alpha N$, for each $x^\theta \in \{G,B\}^N$ we define $\mathbf{a}^{x^\theta}$ such that, for each $i \in I$, $u_i\left(\mathbf{a}^{x^\theta}\right) > v_i^\theta$ if $x_{i-1}^\theta = G$, and $u_i\left(\mathbf{a}^{x^\theta}\right) < v_i^\theta$ if $x_{i-1}^\theta = B$.[25] Define $v_i^\theta(G)$, $v_i^\theta(B)$, and $\bar{\varepsilon} > 0$ such that

$$\left(\max_{x^\theta : x_{i-1}^\theta = B} u_i\left(\mathbf{a}^{x^\theta}\right)\right)_+ \leq v_i^\theta(B) + 4\bar{\varepsilon} < v_i^\theta < v_i^\theta(G) - 4\bar{\varepsilon} \leq \min_{x^\theta:x_{i-1}^\theta = G} u_i\left(\mathbf{a}^{x^\theta}\right). \tag{26}$$

Note that such $\bar{\varepsilon}$ exists since $B^{\frac{\eta}{2}}\left(v^\theta\right) \subset F^*(\theta)$. With these definitions, (23) is satisfied.

B.2.5 Structure of the Block

Each block consists of the following sub-blocks: Let

$$K := \left\lceil \frac{\max\{G, L\}}{\bar{\varepsilon}} \right\rceil. \tag{27}$$

Now fix Z sufficiently large such that $Z \geq \bar{Z}$ (with c and $\bar{Z}$ given in Lemma 12) and

$$(K+3)\left(\frac{1}{Z} + 2\exp(-cZ)\right)\bar{u} \leq \bar{\varepsilon}, \tag{28}$$

[25] As in Hörner and Olszewski (2006) and several subsequent papers, it may actually be necessary for players to cycle through a sequence of distinct action profiles $\mathbf{a}^{x^\theta}$ to achieve average payoffs $u_i > v_i^\theta$ (resp., $u_i < v_i^\theta$) for i such that $x_{i-1}^\theta = G$ (resp., $x_{i-1}^\theta = B$). Accommodating this possibility poses no difficulty for the proof, so we follow Hörner and Olszewski (and others) in assuming that a single action profile suffices.

where $\bar{u} = 2\max\{L, 1+G\}$. Let $T = Z\log_2 N$. In what follows, recall that players always take action D while circulating information.

1. 1^***-communication sub-block** (the first T periods of the block): Players circulate information about h^{1^*}.
2. $\boldsymbol{x}$**-communication sub-block** (the next T periods): Players circulate information about $\boldsymbol{x}$.
3. **Supplemental round** 0 (the next T periods): Players circulate information about the first two sub-blocks.
4. **Main sub-block** k (there are K main sub-blocks, each lasting for $(1+Z)T$ periods, and each divided into the following two rounds):
   (a) **Main round** k (the first ZT periods of the sub-block): Players take the target actions (and do not send cheap talk messages).
   (b) **Supplemental round** k (the next T periods of the sub-block): Players circulate information about the history up to the end of main round k.

Recall that T^{**} denotes the length of the block, or equivalently the last period of supplemental round K. Let T^* denote the last period of main round K. Note that $T^{**} = (3 + K(1+Z))Z\log_2 N$. Since $(1-\delta)\log N \to 0$, we have

$$\liminf_{l\to\infty}(1-\delta)T^* \leq \liminf_{l\to\infty}(1-\delta)T^{**} = 0. \tag{29}$$

B.2.6 Reduction Lemma

We now show that, by communicating their histories during supplemental round K (the last such round in the block) and adjusting continuation payoffs appropriately, the players can effectively cancel the effects of discounting while letting continuation payoffs depend on $\left(\boldsymbol{x}_{-i}, h^{1^*}, h^{T^*+1}\right)$ rather than $\left(x_{i-1}, h_{i-1}^{T^{**}+1}\right)$ (when $|\theta| \geq \alpha N$).[26]

Let $\Sigma_i^{T^*}$ denote the set of i's block strategies up to period T^*. We show that the following conditions are sufficient for (25).

1. *[Sequential Rationality]* For each $\boldsymbol{x} \in \{G,B\}^{N|\Theta|}$, $i \in I$, $h_i^{1^*} \in H_i^{1^*}$, $t \in \{1,\ldots,T^*\}$, $b \in \mathbb{N}$, and $\tilde{h}_i^{b,t} \in \tilde{H}_i^{b,t}$, $\sigma_i^{T^*}\left(\boldsymbol{x}_i, h_i^{1^*}\right)$ maximizes (over $\sigma_i \in \Sigma_i^{T^*}$)

$$\sum_{h_{-i}^{1^*}\in H_{-i}^{1^*}} \beta_i\left(h_{-i}^{1^*}|\boldsymbol{x}_{-i}, \tilde{h}_i^{b,t}\right)\mathbb{E}^{\left(\sigma_i, \sigma_{-i}^*\left(\boldsymbol{x}_{-i}, h_{-i}^{1^*}\right)\right)}\left[\begin{array}{c} 1_{\{\theta:|\theta|\geq\alpha N\}}\left(\sum_{\tau=1}^{T^*}\hat{u}_i(\mathbf{a}_\tau) + \pi_i^\theta(x_{-i}^\theta, h^{1^*}, h^{T^*+1})\right) \\ +1_{\{\theta:|\theta|<\alpha N\}}\sum_{\tau=1}^{T^*}\delta^{\tau-1}\hat{u}_i(\mathbf{a}_\tau) \\ |\boldsymbol{x}_{-i}, \tilde{h}_i^{b,t} \end{array}\right]. \tag{30}$$

2. *[Promise Keeping]* For each $\theta \in \Theta$, $\boldsymbol{x} \in \{G,B\}^{N|\Theta|}$, $i \in I$, $b \in \mathbb{N}$, and $\tilde{h}^{b,0} \in \tilde{H}^{b,0}$,

$$v_i^\theta(x_{i-1}^\theta) = \mathbb{E}^{\sigma^*\left(\boldsymbol{x}, h^{1^*}\right)}\left[\begin{array}{c} 1_{\{\theta:|\theta|\geq\alpha N\}}\frac{1}{T^*}\left(\sum_{\tau=1}^{T^*}\hat{u}_i(\mathbf{a}_\tau) + \pi_i^\theta(x_{-i}^\theta, h^{1^*}, h^{T^*+1})\right) \\ +1_{\{\theta:|\theta|<\alpha N\}}\frac{1-\delta}{1-\delta^{T^*}}\sum_{\tau=1}^{T^*}\delta^{\tau-1}\hat{u}_i(\mathbf{a}_\tau) \end{array} |\theta, \tilde{h}^{b,0}\right]. \tag{31}$$

[26] A similar but more complicated argument appears in Deb, Sugaya, and Wolitzky (2019).

3. *[Self-Generation]* For each $\theta \in \Theta$, $i \in I$, $\boldsymbol{x}_{i-1} \in \{G, B\}^{|\Theta|}$, $h^{1^*} \in H^{1^*}$, and $h^{T^*+1} \in H^{T^*+1}$,

$$\text{sign}\left(x_{i-1}^{\theta}\right) \pi_i^{\theta}(x_{-i}^{\theta}, h^{1^*}, h^{T^*+1}) \begin{cases} > 0 & \text{for } \theta \text{ satisfying } |\theta| \geq \alpha N, \\ = 0 & \text{for } \theta \text{ satisfying } |\theta| < \alpha N. \end{cases} \tag{32}$$

and

$$\left|\pi_i^{\theta}(x_{-i}^{\theta}, h^{1^*}, h^{T^*+1})\right| \begin{cases} \leq 2\bar{u}T^* & \text{for } \theta \text{ satisfying } |\theta| \geq \alpha N, \\ = 0 & \text{for } \theta \text{ satisfying } |\theta| < \alpha N. \end{cases} \tag{33}$$

Lemma 14 *For any sequence $(N, \delta, p)_l$ such that (10) holds and $(1-\delta)\log N \to 0$, suppose there exists $\bar{l}$ such that, for each $l \geq \bar{l}$ and corresponding $(N, \delta, p)_l$, there exist $\left(\sigma_i^{T^*}\left(\boldsymbol{x}_i, h_i^{1^*}\right)\right)_{i \in I, \boldsymbol{x}_i \in \{G,B\}^{|\Theta|}, h_i^{1^*} \in H_i^{1^*}}$ and $\left(\pi_i^{\theta}(x_{-i}^{\theta}, h^{1^*}, h^{T^*+1})\right)_{i \in I, x_{i-1} \in \{G,B\}, h^{1^*}, h^{T^*+1}}$ such that Conditions (30)–(33) are satisfied. Then, for sufficiently large l, (25) holds.*

Proof. We have fixed $v_i^{\theta}(x_{i-1}^{\theta})$. Fix $\sigma_i^{T^*}$ and π_i^{θ} satisfying (30)–(33). We will construct $\tilde{\sigma}_i^*$, $\tilde{\pi}_i^{\theta}$, and $\tilde{v}_i^{\theta}(x_{i-1}^{\theta})$ that satisfy (20)–(23).

We extend strategy $\sigma_i^{T^*} \in \Sigma_i^{T^*}$ to a strategy $\tilde{\sigma}_i^* \in \Sigma_i$ by specifying that players circulate message $m = (m_i)_i = (\boldsymbol{x}_i, h_i^{1^*}, h_i^{T^*+1})_i$ in supplemental round K.

Given player $i-1$'s history in supplemental round K, we define $\tilde{\pi}_i^{\theta}(x_{i-1}, h_{i-1}^{T^{**}+1})$ as follows. (i) If $m_{-i}(i-1) = \texttt{error}$, then $\tilde{\pi}_i^{\theta}(x_{i-1}, h_{i-1}^{T^{**}+1}) = 0$ for each $x_{i-1}, h_{i-1}^{T^{**}+1}$. (ii) Otherwise, player $i-1$ infers $(h_{-i}^{1^*}(i-1), h_{-i}^{T^*+1}(i-1))$. Since matching is pairwise, there exists a unique $h^{1^*}(i-1), h^{T^*+1}(i-1)$ that is consistent with $(h_{-i}^{1^*}(i-1), h_{-i}^{T^*+1}(i-1))$. Given $h^{T^*+1}(i-1)$, let $\mathbf{a}_t(i-1)$ be the action in period t. We define

$$\tilde{\pi}_i^{\theta}(x_{i-1}, h_{i-1}^{T^{**}+1}) = \begin{cases} \begin{array}{c} \frac{T^*\left(1-\delta^{T^*}\right)\text{sign}\left(x_{i-1}^{\theta}\right)\bar{u} + \sum_{t=1}^{T^*}\left(1-\delta^{t-1}\right)\hat{u}_i(\mathbf{a}_t(i-1))}{\Pr(m_{-i}(i-1) \neq \text{error})} \\ + \frac{\pi_i^{\theta}(\boldsymbol{x}_{-i}(i-1), h^{1^*}(i-1), h^{T^*+1}(i-1))}{\Pr(m_{-i}(i-1) \neq \text{error})} \end{array} & \text{if } |\theta| \geq \alpha N, \\ 0 & \text{if } |\theta| < \alpha N. \end{cases}$$

Finally, we define

$$\tilde{v}_i^{\theta}(x_{i-1}^{\theta}) = \begin{cases} \frac{1-\delta}{1-\delta^{T^{**}}} T^* v_i^{\theta}(x_{i-1}^{\theta}) + (1-\delta) T^* \text{sign}\left(x_{i-1}^{\theta}\right)\bar{u} & \text{if } |\theta| \geq \alpha N, \\ v_i^{\theta}(x_{i-1}^{\theta}) & \text{if } |\theta| < \alpha N. \end{cases}$$

As $l \to \infty$, since $\frac{1-\delta}{1-\delta^{T^{**}}}T^* \to 1$ and $(1-\delta)T^* \to 0$, we have $\tilde{v}_i^{\theta}(x_{i-1}^{\theta}) \to v_i^{\theta}(x_{i-1}^{\theta})$ uniformly for each θ. It remains to show that $\tilde{\sigma}_i^*\left(\boldsymbol{x}_i, h_i^{1^*}\right)$, $\tilde{\beta}^*$, $\tilde{\pi}_i^{\theta}$, and $\tilde{v}_i^{\theta}(x_{i-1}^{\theta})$ satisfy (20)–(23).

Note that player i's payoff depends on the outcome of play in supplemental round K only through her stage game payoffs (which are maximized by taking D) and $m_{-i}(i-1)$. Since player i cannot affect the distribution of $m_{-i}(i-1)$, following $\sigma_i^*\left(\boldsymbol{x}_i, h_i^{1^*}\right)|_{h_i^{T^*+1}}$ is optimal. Given this, by the law of iterated expectation, in period $t \leq T^*$, the expected value of

$$\sum_{\tau=t}^{T^*} \delta^{t-1}\hat{u}_i\left(\mathbf{a}_{\tau}\right) + \tilde{\pi}_i^{\theta}(x_{i-1}, h_{i-1}^{T^{**}})$$

given θ, $\boldsymbol{x}$, h^{1^*}, and $\tilde{h}^t$ is equal to

$$1_{\{\theta:|\theta|\geq\alpha N\}}\left(\left(1-\delta^{T^*}\right)T^*\text{sign}\left(x_{i-1}^{\theta}\right)\bar{u} + \sum_{\tau=t}^{T^*}\hat{u}_i\left(\mathbf{a}_{\tau}\right) + \pi_i^{\theta}(x_{-i}^{\theta}, h^{1^*}, h^{T^*+1})\right) + 1_{\{\theta:|\theta|<\alpha N\}}\sum_{\tau=t}^{T^*}\delta^{\tau-1}\hat{u}_i\left(\mathbf{a}_{\tau}\right).$$

Ignoring the constant $\left(1-\delta^{T^*}\right) T^* \text{sign}\left(x_{i-1}^{\theta}\right)\bar{u}$, (30) implies (20).

For $|\theta| < \alpha N$, (21)–(23) hold since all the payoffs and rewards are zero regardless of x_{i-1} and $h_{i-1}^{T^{**}+1}$. For $|\theta| \geq \alpha N$, (21) follows from (31) given the definition of $\tilde{v}_i^{\theta}(x_{i-1}^{\theta})$. In addition, (32) and (33) imply (22) for sufficiently large l since $\delta \to 1$ and $v_i^{\theta}(G) - v_i^{\theta}(B) \geq 8\bar{\varepsilon}$ by (26). Finally, (26) implies that there is $4\bar{\varepsilon}$ slack between $v_i^{\theta}(x_{i-1}^{\theta})$ and v_i^{θ} for each θ satisfying $|\theta| \geq \alpha N$. Hence, (23) holds with $\tilde{v}_i^{\theta}(x_{i-1}^{\theta})$ for sufficiently large l. ■

B.2.7 Equilibrium Strategies

We now complete the description of the equilibrium strategies.

It will be useful to define the notion of a "detectable deviation" by player i. As we will see, given player i's period 1^* history $h_i^{1^*}$ and her strategy state $\boldsymbol{x}_i$, her block strategy is pure along the equilibrium path of play. Given $h_i^{1^*}$ and an on-path period t block history h_i^t, we say that a period t message $m_{i,t}$ is a *detectable deviation* if there does not exist a strategy state $\hat{\boldsymbol{x}}_i$ such that $\left(h_i^t, m_{i,t}\right)$ occurs with positive probability given $\left(\hat{\boldsymbol{x}}_i, h_i^{1^*}\right)$; similarly, given a pair $\left(h_i^t, m_{i,t}\right)$, an action $a_{i,t}$ is a *detectable deviation* if there does not exist a strategy state $\hat{\boldsymbol{x}}_i$ such that $\left(h_i^t, m_{i,t}, a_{i,t}\right)$ occurs with positive probability given $\left(\hat{\boldsymbol{x}}_i, h_i^{1^*}\right)$. We say a player *detectably deviates* if she plays a detectable deviation.

1^*-Communication Sub-Block Players circulate message $m = (m_i)_i$, where m_i is the set of players whom player i knows to have taken C in period 1^*: that is, $m_i = \{i, \mu_{1^*}(i)\} \cap \theta$.

Let h_i^{T+1} be player i's history at the end of the sub-block. We define $\theta\left(h_i^{T+1}\right) = \emptyset$ if, for some $j \neq i$, either $\zeta_{i,T}^{I,-j} \neq -j$ (i.e., i does not receive each player's message through a path excluding j) or $m_{-j}(i) = \texttt{error}$ (i.e., i receives inconsistent messages through a path excluding j). We also define $\theta\left(h_i^{T+1}\right) = \emptyset$ if there exist $j \neq j' \neq k \neq j$ such that $m_{-j}(i)|_k \neq m_{-j'}(i)|_k$. Otherwise, we define $\theta\left(h_i^{T+1}\right) = \bigcup_{j\neq i}\bigcup_{k\neq j} m_{-j}(i)|_k$ (i.e., $\theta\left(h_i^{T+1}\right)$ is the set of players who i has been told took C in period 1^*).

Lemma 12 immediately implies the following result.

Lemma 15 *Suppose all players follow the protocol. There exist $c > 0$ and $\bar{Z} > 0$ such that, for all $Z > \bar{Z}$ and all l, we have*

$$\Pr\left(\theta(h_i^{T+1}) = \theta \ \forall i\right) \geq 1 - \exp\left(-cK\right).$$

We record two key properties of player i's beliefs about θ. Suppose the current block is block b.

First, for each $t \geq T+1$, $\tilde{h}_i^{b,t} \in \tilde{H}_i^{b,t}$, and $\boldsymbol{x}_{-i} \in \{G, B\}^{N-1}$, player i believes that $\theta \supseteq \theta(h_i^{T+1})$:

$$\sum_{\theta \supseteq \theta(h_i^{T+1})} \beta_i\left(\theta | \boldsymbol{x}_{-i}, \tilde{h}_i^{b,t}\right) = 1. \tag{34}$$

This is trivial if $\theta_i(h_i^{T+1}) = \emptyset$. Otherwise, since trembles in earlier blocks are more likely, $\beta_i\left(\theta = \theta(h_i^{T+1}) | \boldsymbol{x}_{-i}, \tilde{h}_i^{b,T+1}\right) = 1$ (i.e., player i believes that $\theta = \theta\left(h_i^{T+1}\right)$ at the end of the 1^*-communication sub-block). Moreover, since trembles are more likely in later periods within the block, player i continues to believe that $\theta = \theta\left(h_i^{T+1}\right)$ for the duration of the block.

Second, for each $t \geq T+1$, $\tilde{h}_i^{b,t} \in \tilde{H}_i^{b,t}$, and $\boldsymbol{x}_{-i} \in \{G,B\}^{N-1}$, player i believes that $\theta \supseteq \theta(h_j^{T+1})$ for each $j \neq i$:

$$\sum_{\theta \supseteq \theta(h_j^{T+1})} \beta_{i,t}\left(\theta | \boldsymbol{x}_{-i}, \tilde{h}_i^{b,t}\right) = 1. \tag{35}$$

This holds by similar reasoning. Note that, if player i deviates in the 1*-communication sub-block, this can only switch $\theta(h_j^{T+1})$ from θ to $\emptyset$, and hence cannot affect the probability that $\theta \supseteq \theta(h_j^{T+1})$.

$\boldsymbol{x}$-Communication Sub-Block Players circulate message $m = (m_i)_i = (\boldsymbol{x}_i)_i$. Slightly abusing notation, let $h_i^{\leq 0}$ denote player i's history at the end of the $\boldsymbol{x}$-communication sub-block.

If player i infers $m_{-j}(i) = \texttt{error}$ for some $j \neq i$, we define $\boldsymbol{x}(i) = \boldsymbol{B}$, where $\boldsymbol{B} \in \{G,B\}^{N|\Theta|}$ denotes the vector with B in every component. If instead i infers some $m_{-j}(i) \in \times_{n \neq i} M_n$ for each $j \neq i$, then:

1. If there exists $\hat{\boldsymbol{x}}_{-i} \in \{G,B\}^{(N-1)|\Theta|}$ such that $m_{-j}(i)|_n = \hat{\boldsymbol{x}}_{-i}|_n$ for all $j \neq i \neq n \neq j$, we define $\boldsymbol{x}(i) = (\boldsymbol{x}_i, \hat{\boldsymbol{x}}_{-i})$.
2. Otherwise, we define $\boldsymbol{x}(i) = \boldsymbol{B}$.

Finally, we define $x(i) = \boldsymbol{x}(i)^{\theta\left(h_i^{T+1}\right)}$.

Supplemental Round 0 Players circulate message $m = (m_i)_i = \left(h_i^{\leq 0}\right)_i$. Let $h_i^{<1}$ denote player i's history at the end of supplemental round 0.

We define $I^D\left(h_i^{<1}\right) = 1$ if any of the following hold:

1. $\left|(\theta(h_i^{T+1}))\right| < \alpha N$.
2. Player i detectably deviates in either the $\boldsymbol{x}$-communication sub-block or supplemental round 0.
3. $m_{-j}(i) = \texttt{error}$ for some $j \neq i$ in either the $\boldsymbol{x}$-communication sub-block or supplemental round 0.

Otherwise, for each $j \neq i$, $m_{-j}(i) = \times_{n \neq i} h_n^{\leq 0}$ for some $\times_{n \neq i} h_n^{\leq 0} \in \times_{n \neq i} H_n^{\leq 0}$. If there exists a player $j \neq i$ such that, according to history $\left(h_i^{\leq 0}, m_{-j}(i)\right)$, player j detectably deviated in the 1*-communication sub-block or the $\boldsymbol{x}$-communication sub-block, then we define $I^D\left(h_i^{<1}\right) = 1$. Otherwise, we define $I^D\left(h_i^{<1}\right) = 0$.

Main Sub-Block k, $k \in \{1,\ldots,K\}$ For each $k \in \{1,\ldots,K\}$, each player i enters sub-block k with state variables $x(i) \in \{G,B\}^N$ and $I^D\left(h_i^{<k}\right) \in \{0,1\}$. The state variable $x(i)$ was determined at the end of the $\boldsymbol{x}$-communication sub-block, and remains constant throughout the main sub-blocks. The state variable $I^D\left(h_i^{<1}\right)$ was determined at the end of supplemental round 0; the state variable $I^D\left(h_i^{<k}\right)$ may switch from 0 to 1 during some main sub-block, in which case it remains equal to 1 for the duration of the block.

We now define player i's strategy in main sub-block k as a function of $x(i)$ and $I^D\left(h_i^{<k}\right)$, and then specify how $I^D\left(h_i^{<k+1}\right)$ evolves.

Main round actions as a function of $x(i)$ *and* $I^D\left(h_i^{<k}\right)$*:* If $I^D\left(h_i^{<k}\right)=1$, then player i takes D throughout the round. If $I^D\left(h_i^{<k}\right)=0$, then player i takes $\mathbf{a}_i^{x(i)}$ throughout the round, unless she herself deviates from $\mathbf{a}_i^{x(i)}$ during the round. If such a deviation occurs, she takes D for the rest of the round. Let $h_i^{\leq k}$ denote player i's history at the end of main round k.

Supplemental round communication as a function of $h_i^{\leq k}$: Players circulate message $m = (m_i)_i = \left(h_i^{\leq k}\right)_i$.

Determination of $I^D\left(h_i^{<k+1}\right)$*:* Set $I^D\left(h_i^{<k+1}\right)=1$ if any of the following hold:

1. $I^D\left(h_i^{<k}\right)=1$.
2. Player i detectably deviated during main sub-block k.
3. $m_{-j}\left(i\right)=\texttt{error}$ for some $j\neq i$ during supplemental round k.
4. For each $j\neq i$, $m_{-j}\left(i\right)=\times_{n\neq i}h_n^{\leq k}$ for some $\times_{n\neq i}h_n^{\leq k}\in\times_{n\neq i}H_n^{\leq k}$, and there exists a player $j\neq i$ such that, according to history $\left(h_i^{\leq k}, m_{-j}\left(i\right)\right)$, player j detectably deviated during main round k.

Otherwise, set $I^D\left(h_i^{<k+1}\right)=0$.

B.2.8 Reward Function

Given the above block strategy profile, we now define the reward function $\pi_i^\theta(x_{-i}^\theta, h^{1^*}, h^{T^*+1})$. For θ satisfying $|\theta|<\alpha N$, define $\pi_i^\theta(x_{-i}^\theta, h^{1^*}, h^{T^*+1})=0$ for all $(x_{-i}^\theta, h^{1^*}, h^{T^*+1})$. This satisfies Conditions (31)–(33); we verify Condition (30) (sequential rationality) in the next subsection. For the remainder of this section, assume $|\theta|\geq\alpha N$.

Given $\left(h^{1^*}, h^{T^*+1}\right)$, we define $\chi_i(h^{1^*}, h^{T^*+1})=1$ if there exists a player $j\neq i$ who detectably deviated from the prescribed block strategy (according to h^{T^*+1}) or if the match realization was erroneous in any round in the current block (again, according to h^{T^*+1}). We define $\chi_i(h^{1^*}, h^{T^*+1})=0$ otherwise. Lemma 12 immediately implies the following result.

Lemma 16 *Suppose player i's opponents follow the prescribed strategy. For all l and all θ (and regardless of player i's own strategy), we have*

$$\Pr\left(\chi_i(h^{1^*}, h^{T^*+1})=1|\theta\right)\leq(3+K)\exp\left(-cZ\right).$$

Next, define $I_i^D(h^{1^*}, h^{T^*+1})=1$ if player i detectably deviated from the prescribed strategy, and define $I_i^D(h^{1^*}, h^{T^*+1})=0$ otherwise. Finally, given $\left(h^{1^*}, h^{T^*+1}\right)$ satisfying $I_i^D(h^{1^*}, h^{T^*+1})=0$, define $\hat{\boldsymbol{x}}_i(h^{1^*}, h^{T^*+1})$ to be that value of $\hat{\boldsymbol{x}}_i$ for which the history $\left(h_i^{1^*}, h_i^{T^*+1}\right)$ is consistent with player i taking strategy $\sigma_i^{T^*}(\hat{\boldsymbol{x}}_i, h_i^{1^*})$. Such $\hat{\boldsymbol{x}}_i$ is uniquely determined since player i communicates $\hat{\boldsymbol{x}}_i$ in the $\boldsymbol{x}$-communication sub-block.

Given a profile of actions and observations for player i's opponents, (a_{-i}, ω_{-i}), let a_i denote the unique action for player i consistent with these observations, and let $\boldsymbol{a}=(a_i, a_{-i})$.[27] Define the

[27] By carrying the extra notation $\left(\mu\left(j\right)\right)_{j\neq i}$ in the vector $\left(a_{-i}, \omega_{-i}, \left(\mu\left(j\right)\right)_{j\neq i}\right)$ (i.e., information about who matched with whom), we can simply specify $a_i=\omega_j$, for j satisfying $\mu\left(j\right)=i$. Even without this extra information, a_i is uniquely identified from (a_{-i}, ω_{-i}); see Lemma 2 of Deb, Sugaya, and Wolitzky (2019).

function $\pi_i^{\text{cancel}}(x_{i-1}, a_{-i}, \omega_{-i}) : \{G, B\} \times A^{N-1} \times A^{N-1} \to [-\bar{u}, \bar{u}]$ such that, for each $\boldsymbol{a} \in A^N$, we have

$$\begin{cases} \hat{u}_i(\boldsymbol{a}) + \pi_i^{\text{cancel}}(x_{i-1}, a_{-i}, \omega_{-i}) = \operatorname{sign}(x_{i-1}) \frac{1}{2}\bar{u} \\ \operatorname{sign}(x_{i-1}) \pi_i^{\text{cancel}}(x_{i-1}, a_{-i}, \omega_{-i}) \geq 0 \end{cases} \tag{36}$$

Thus, the function $\pi_i^{\text{cancel}}(x_{i-1}, a_{-i,}, \omega_{-i})$ cancels player i's instantaneous utility and leaves player i a negative (resp., positive) payoff when $x_{i-1} = G$ (resp., B)

If $\chi_i(h^{1^*}, h^{T^*+1}) = 1$, define

$$\hat{\pi}_i^{\theta}(x_{-i}^{\theta}, h^{T^*+1}) = \sum_{t=1}^{T^*} \pi_i^{\text{cancel}}\left(x_{i-1}^{\theta}, a_{-i,t}, \omega_{-i,t}\right). \tag{37}$$

If $\chi_i(h^{1^*}, h^{T^*+1}) = 0$, define

$$\hat{\pi}_i^{\theta}(x_{-i}^{\theta}, h^{T^*+1}) = \begin{cases} 1_{\{I_i^D(h^{1^*}, h^{T^*+1})=0\}} \bar{\varepsilon} T^* & \text{if } x_{i-1}^{\theta} = B, \\ -1_{\{I_i^D(h^{1^*}, h^{T^*+1})=1\}} 2\bar{u} T^* & \text{if } x_{i-1}^{\theta} = G. \end{cases}$$

That is, if $x_{i-1}^{\theta} = B$ then player i is rewarded if she follows the prescribed strategy; and if $x_{i-1}^{\theta} = G$ then she is punished if she detectably deviates.

Let

$$u_i\left(x^{\theta}, h^{1^*}\right) = \frac{1}{T^*} \mathbb{E}^{\sigma(\boldsymbol{x})}\left[\sum_{\tau=1}^{T^*} \hat{u}_i(\mathbf{a}_\tau) + \hat{\pi}_i^{\theta}(x_{-i}^{\theta}, h^{T^*+1}) + \operatorname{sign}\left(x_{i-1}^{\theta}\right) \bar{\varepsilon} T^* | h^{1^*}, \theta\right]. \tag{38}$$

Note that

$$\begin{aligned}
& \left| u_i\left(x^{\theta}, h^{1^*}\right) - \hat{u}_i\left(\mathbf{a}^{x^{\theta}}\right) \right| \\
= \;& \Pr\left(\chi_i(h^{1^*}, h^{T^*+1}) = 0 | \theta\right) \left| \frac{1}{T^*} \mathbb{E}^{\sigma(\boldsymbol{x})}\left[\left(\begin{array}{c} \sum_{\tau=1}^{T^*} \hat{u}_i(\mathbf{a}_\tau) \\ +\hat{\pi}_i^{\theta}(x_{-i}^{\theta}, h^{T^*+1}) \\ +\operatorname{sign}\left(x_{i-1}^{\theta}\right) \bar{\varepsilon} T^* \end{array}\right) | h^{1^*}, \theta, \chi_i(h^{1^*}, h^{T^*+1}) = 0\right] - \hat{u}_i\left(\mathbf{a}^{x^{\theta}}\right) \right| \\
& + \Pr\left(\chi_i(h^{1^*}, h^{T^*+1}) = 1 | \theta\right) \left| \frac{1}{T^*} \mathbb{E}^{\sigma(\boldsymbol{x})}\left[\left(\begin{array}{c} \sum_{\tau=1}^{T^*} \hat{u}_i(\mathbf{a}_\tau) \\ +\hat{\pi}_i^{\theta}(x_{-i}^{\theta}, h^{T^*+1}) \\ +\operatorname{sign}\left(x_{i-1}^{\theta}\right) \bar{\varepsilon} T^* \end{array}\right) | h^{1^*}, \theta, \chi_i(h^{1^*}, h^{T^*+1}) = 1\right] - \hat{u}_i\left(\mathbf{a}^{x^{\theta}}\right) \right| \\
\leq \;& \Pr\left(\chi_i(h^{1^*}, h^{T^*+1}) = 0 | \theta\right) \left(\frac{(3+K)T}{T^*} \bar{u} + 2\bar{\varepsilon}\right) + \Pr\left(\chi_i(h^{1^*}, h^{T^*+1}) = 1 | \theta\right) (2\bar{u} + \bar{\varepsilon}) \\
\leq \;& 2\bar{\varepsilon} + \left(\frac{3+K}{Z} + 2\Pr\left(\chi_i(h^{1^*}, h^{T^*+1}) = 1 | \theta\right)\right) \bar{u} \\
\leq \;& 2\bar{\varepsilon} + (3+K)\left(\frac{1}{Z} + 2\exp(-cZ)\right) \bar{u} \quad \text{(by Lemma 16)} \\
\leq \;& 3\bar{\varepsilon} \quad \text{(by (28))}.
\end{aligned} \tag{39}$$

Here the first inequality follows because (i) when $\chi_i(h^{1^*}, h^{T^*+1}) = 0$, $\mathbf{a}_\tau = \mathbf{a}^{x^{\theta}}$ in main rounds (i.e., in all but $(3+K)T$ periods), (ii) the magnitude of $\hat{u}_i(\mathbf{a}_\tau)$ is bounded by $\frac{1}{2}\bar{u}$, and (iii) on-path (i.e., when $I_i^D(h^{1^*}, h^{T^*+1}) = 0$), the magnitude of $\hat{\pi}_i^{\theta}(x_{-i}^{\theta}, h^{T^*+1})$ is bounded by $\bar{\varepsilon} T^*$.

We now define the reward function

$$\begin{aligned}&\pi_i^\theta(x_{-i}^\theta, h^{1^*}, h^{T^*+1})\\ =\ & 1_{\{I_i^D(h^{1^*},h^{T^*+1})=0\}}\left(v_i^\theta(x_{i-1}^\theta) - u_i\left(\hat{x}_i^\theta(h^{T^*+1}), x_{-i}^\theta, h^{1^*}\right)\right)T^* + \hat{\pi}_i^\theta(x_{-i}^\theta, h^{T^*+1}) + \operatorname{sign}\left(x_{i-1}^\theta\right)\bar{\varepsilon}T^*.\end{aligned}$$

We verify that, with this reward function, Conditions (30)–(33) are satisfied. This will complete the proof. We first establish Conditions (31)–(33), deferring Condition (30) (sequential rationality) to the next subsection.

Since $I_i^D(h^{1^*}, h^{T^*+1}) = 0$ on path, (38) implies that expected per-period block payoffs given $|\theta| \geq \alpha N$ equal $v_i^\theta(x_{i-1}^\theta)$. Hence, (31) holds.

By (39) and (26), we have

$$\operatorname{sign}\left(x_{i-1}^\theta\right)\left(v_i^\theta(x_{i-1}^\theta) - u_i\left(\hat{x}_i^\theta(h^{T^*+1}), x_{-i}^\theta, h^{1^*}\right)\right) \geq 0.$$

Together with (36) and (37), this implies

$$\operatorname{sign}\left(x_{i-1}^\theta\right)\pi_i^\theta(x_{-i}^\theta, h^{1^*}, h^{T^*+1}) \geq \bar{\varepsilon}T, \tag{40}$$

and hence (32).

Moreover, if $I_i^D\left(h^{1^*}, h^{T^*+1}\right) = 0$ then

$$\left|\pi_i^\theta(x_{-i}^\theta, h^{1^*}, h^{T^*+1})\right| \leq \left|v_i^\theta(x_{i-1}^\theta) - u_i\left(\hat{x}_i^\theta(h^{T^*+1}), x_{-i}^\theta, h^{1^*}\right)\right|T^* + 2\bar{\varepsilon}T^* \leq 2\bar{u}T^*;$$

and if $I_i^D\left(h^{1^*}, h^{T^*+1}\right) = 1$ then

$$\left|\pi_i^\theta(x_{-i}^\theta, h^{1^*}, h^{T^*+1})\right| \leq 2\bar{u}T^*.$$

Hence, (33) holds.

B.2.9 Verifying Sequential Rationality (Condition (30))

Given (37) and $|\theta| \geq \alpha N$, if $\chi_i(h_{1^*}, h^{T^*+1}) = 1$, then any action is optimal for player i. Since $\Pr\left(\chi_i(h_{1^*}, h^{T^*+1}) = 1|\sigma_i, \theta\right)$ is independent of σ_i, it is without loss to verify sequential rationality conditional on the event $\{\chi_i(h^{1^*}, h^{T^*+1}) = 0 \vee |\theta| < \alpha N\}$. We thus restrict attention to pairs $\left(\boldsymbol{x}_{-i}, \tilde{h}_i^{b,t}\right)$ such that $\Pr\left(\{\chi_i(h^{1^*}, h^{T^*+1}) = 0 \vee |\theta| < \alpha N\}|\boldsymbol{x}_{-i}, \tilde{h}_i^{b,t}\right) > 0$. Note this implies that (34) and (35) hold conditional on the triple $\left(\{\chi_i(h^{1^*}, h^{T^*+1}) = 0 \vee |\theta| < \alpha N\}, \boldsymbol{x}_{-i}, \tilde{h}_i^{b,t}\right)$. We consider separately the cases $|\theta| < \alpha N$ and $\left\{\chi_i(h^{1^*}, h^{T^*+1}) = 0 \wedge |\theta| \geq \alpha N\right\}$.

Conditional on $|\theta| < \alpha N$, by (35), player i believes that $\left|\theta(h_n^{T+1})\right| < \alpha N$ for each $n \neq i$. Hence, she believes that players $-i$ take D throughout the block and $\pi_i^\theta(x_{-i}^\theta, h^{1^*}, h^{T^*+1}) = 0$ regardless of her own strategy. It is therefore optimal for player i to take D in each period and send any messages. And this behavior is indeed what is prescribed for player i, since (34) implies that $\left|\theta(h_i^{T+1})\right| < \alpha N$.

It remains to verify sequential rationality conditional on $\left\{\chi_i(h^{1^*}, h^{T^*+1}) = 0 \wedge |\theta| \geq \alpha N\right\}$. We proceed in three steps.

It is optimal to take D and send any message after player i detectably deviates after the 1^*-communication sub-block.

Let τ be the first period in which player i detectably deviated. First, suppose that τ is before supplemental round 0. Then, regardless of player i's behavior after period τ, the fact that $\chi_i(h^{1^*}, h^{T^*+1}) = 0$ (and hence matching is regular) implies that players $-i$ will become aware of player i's deviation at the end of supplemental round 0 and will then take D for the rest of the block. Moreover, the reward function is constant:

$$\pi_i^{\theta}(x_{-i}^{\theta}, h^{1^*}, h^{T^*+1}) = \hat{\pi}_i^{\theta}(x_{-i}^{\theta}, h^{T^*+1}) = \begin{cases} 0 & \text{if } x_{i-1}^{\theta} = B, \\ -2\bar{u}T^* & \text{if } x_{i-1}^{\theta} = G. \end{cases}$$

Hence, taking D and sending any messages is optimal for player i.

Second, suppose τ is in or after supplemental round 0. Then, regardless of player i's behavior after period τ, players $-i$ take $\mathbf{a}^{x^{\theta}}$ in the main sub-block and take D in other rounds until next supplemental round; and subsequently (since matching is regular) they will switch to D for the rest of the block. Again, the reward is constant. Hence, taking D and sending any messages is optimal.

It is optimal not to detectably deviate from the equilibrium strategy at on-path histories.

We compare the maximum gain in within-block payoffs from a detectable deviation to the minimum loss in the reward function. Since matching is regular, players $-i$ switch to D starting in the next main round. Hence, the maximum gain in within-block payoffs is at most $Z^2 \log_2 N \times \max\{G, L\}$. In contrast, if $x_{i-1}^{\theta} = B$, the loss in the reward function from switching $I_i^D\left(h^{1^*}, h^{T^*+1}\right)$ from 0 to 1 is at least $\bar{\varepsilon}T^*$; this comes from the $\hat{\pi}_i^{\theta}\left(x_{-i}^{\theta}, h^{T^*+1}\right)$ term in the reward function, noting that the term

$$1_{\left\{I_i^D(h^{1^*}, h^{T^*+1})=0\right\}}\left(v_i^{\theta}(x_{i-1}^{\theta}) - u_i\left(\hat{x}_i^{\theta}(h^{T^*+1}), x_{-i}^{\theta}, h^{1^*}\right)\right)$$

in the reward $\pi_i^{\theta}(x_{-i}^{\theta}, h^{1^*}, h^{T^*+1})$ is non-negative. By (27), $\bar{\varepsilon}T^* \geq Z^2 \log_2 N \times \max\{G, L\}$, so deviating is unprofitable when $x_{i-1}^{\theta} = B$. If instead $x_{i-1}^{\theta} = G$, the loss in the reward function from switching $I_i^D\left(h^{1^*}, h^{T^*+1}\right)$ from 0 to 1 is at least

$$2\bar{u}T^* - \left|v_i^{\theta}(x_{i-1}^{\theta}) - u_i\left(\hat{x}_i^{\theta}(h^{T^*+1}), x_{-i}^{\theta}, h^{1^*}\right)\right| T^* \geq \bar{u}T^* \geq Z^2 \log_2 N \times \max\{G, L\},$$

where the first inequality follows because $\left|v_i^{\theta}(x_{i-1}^{\theta}) - u_i\left(\hat{x}_i^{\theta}(h^{T^*+1}), x_{-i}^{\theta}, h^{1^*}\right)\right| \leq \bar{u}$. In total, for any x_{i-1}^{θ}, the net deviation gain is negative.

It is optimal to send message $\hat{\boldsymbol{x}}_i = \boldsymbol{x}_i$ in the $\boldsymbol{x}$-communication sub-block.

We show that, for any $\boldsymbol{x}_{-i}$, player i is indifferent among the block strategies $(\sigma_i(\boldsymbol{x}_i))_{\boldsymbol{x}_i}$. Player i's expected payoff conditional on $|\theta| \geq \alpha N$ equals

$$\mathbb{E}\left[\frac{1}{T^*}\mathbb{E}^{\sigma(\boldsymbol{x})}\left[\sum_{\tau=3T+1}^{T^*} \hat{u}_i(\mathbf{a}_\tau) + \hat{\pi}_i^{\theta}(x_{-i}^{\theta}, h^{T^*+1}) + \operatorname{sign}\left(x_{i-1}^{\theta}\right)\bar{\varepsilon}T^* | h^{1^*}, \theta\right] | \boldsymbol{x}_{-i}\right] = v_i^{\theta}(x_{i-1}^{\theta}).$$

Moreover, her payoff conditional on $\chi_i(h^{1^*}, h^{T^*+1}) = 1$ equals $\operatorname{sign}\left(x_{i-1}^{\theta}\right)\left(\frac{1}{2} + \bar{\varepsilon}T^*\right)$. Since these payoffs depend on $\boldsymbol{x}$ only through x_{i-1}^{θ}, and additionally $\Pr\left(\chi_i(h^{1^*}, h^{T^*+1}) = 1\right)$ is independent of $\boldsymbol{x}$, it follows that player i's expected payoff conditional on $\left\{\chi_i(h^{1^*}, h^{T^*+1}) = 0 \wedge |\theta| \geq \alpha N\right\}$ also depends on $\boldsymbol{x}$ only through x_{i-1}^{θ}. This completes the proof of Theorem 5.

B.2.10 Proof of Theorem 4

Since $F^{\alpha,\eta}$ is compact, the following lemma is sufficient:

Lemma 17 *For any $\tilde{v} \in F^{\alpha,\eta}$, there exist $\gamma > 0$ such that $B^{\gamma}(\tilde{v}) \subseteq E^*$ for sufficiently large l.*

Proof. Fix $\tilde{v} \in F^{\alpha,\eta}$ and the associated state-contingent payoffs $\left(\tilde{v}^{\theta}\right)_{\theta}$. Define $\left(v^{\theta}\right)_{\theta}$ as in (17), and then fix $\left(v_i^{\theta}\left(x_{i-1}^{\theta}\right)\right)_{i\in I,\theta\in\Theta,x_{i-1}\in\{G,B\}}$ and $\bar{\varepsilon} > 0$ to satisfy (26). The proof of Theorem 5 shows that a convex set $\mathbf{V} \subset \mathbb{R}^{N|\Theta|}$ is self-generating for sufficiently large l if it satisfies

$$\begin{aligned} V_i^{\theta} &\subseteq \left(v_i^{\theta}(B) + 4\bar{\varepsilon}, v_i^{\theta}(G) - 4\bar{\varepsilon}\right) \text{ for } |\theta| \geq \alpha N, \\ V_i^{\theta} &= \{0\} \text{ for } |\theta| < \alpha N. \end{aligned} \tag{41}$$

Define the set $\mathbf{V}^{\hat{\gamma}} \subset \mathbb{R}^{N|\Theta|}$ by

$$\begin{aligned} V_i^{\theta,\hat{\gamma}} &= \left[v_i^{\theta} - \hat{\gamma}, v_i^{\theta} + \hat{\gamma}\right] \text{ for } |\theta| \geq \alpha N, \\ V_i^{\theta} &= \{0\} \text{ for } |\theta| < \alpha N. \end{aligned}$$

Fix $\hat{\gamma} > 0$ sufficiently small so that $\mathbf{V}^{\hat{\gamma}}$ satisfies (41) and (24); such $\hat{\gamma} > 0$ exists since $V^{\theta} = \{v^{\theta}\}$ satisfies (41) and $\left(v^{\theta}\right)_{\theta}$ satisfies (24) with strict inequality. Now fix $\bar{l}_1$ sufficiently large so that, for each $l \geq \bar{l}_1$, the set $\mathbf{V}^{\hat{\gamma}}$ is self-generating. Since (24) holds, rational players take C in period 1^*. Hence, for each $\left(\hat{v}^{\theta}\right)_{\theta} \in \mathbf{V}^{\hat{\gamma}}$, there exists a sequential equilibrium such that, for each $\theta \in \Theta$, the resulting payoff when $\theta^* = \theta$ equals $\hat{v}^{\theta}$.

Let $\hat{V}^{\hat{\gamma}} \subset \mathbb{R}^N$ be the set of $\hat{v} \in \mathbb{R}^N$ such that there exists $\left(\hat{v}^{\theta}\right)_{\theta} \in \mathbf{V}^{\hat{\gamma}}$ satisfying

$$\left((1-\delta)\left(p^0 + p^1\left(\frac{1+G-L}{2}\right)\right) + \delta\sum_{\theta^*} p\left(I \setminus \theta^*\right)\hat{v}_i^{\theta^*}\right)_i = \hat{v}_i.$$

(Note that this set $\hat{V}^{\hat{\gamma}}$ depends on l.) Since any payoff vector in $\mathbf{V}^{\hat{\gamma}}$ is implementable, so is any expected payoff in $\hat{V}^{\hat{\gamma}}$. Since $V_i^{\theta,\hat{\gamma}} = \left[v_i^{\theta} - \hat{\gamma}, v_i^{\theta} + \hat{\gamma}\right]$ for $|\theta| \geq \alpha N$, by taking $\gamma > 0$ sufficiently small and $\bar{l}_2$ sufficiently large, we have $B^{\gamma}(\tilde{v}) \subseteq \hat{V}^{\hat{\gamma}}$ for all $l \geq \bar{l}_2$. For such $\gamma > 0$, we have $B^{\gamma}(\tilde{v}) \subseteq E^*$ for all $l \geq \max\left\{\bar{l}_1, \bar{l}_2\right\}$, as desired. ■