Communication and Community Enforcement*

Takuo Sugaya and Alexander Wolitzky

Stanford GSB and MIT

January 11, 2021

Abstract

We study the repeated prisoner's dilemma with random matching, a canonical model of community enforcement with decentralized information. We assume (1) with small probability, each player is a "bad type" who never cooperates, (2) players observe and remember their partners' identities, and (3) each player interacts with others frequently, but meets any particular partner infrequently. We show that these assumptions preclude cooperation in the absence of explicit communication, but that introducing within-match cheap talk communication restores cooperation. Thus, communication is essential for community enforcement.

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

JEL codes: C72, C73, D83

*For helpful comments, we thank Daron Acemoglu, Glenn Ellison, Jeff Ely, Drew Fudenberg, Yuval Heller, Matt Jackson, Emir Kamenica, George Mailath, Stephen Morris, Satoru Takahashi, Omer Tamuz, three anonymous referees, and seminar particiants at Princeton, Stanford, and Wharton. Wolitzky acknowledges financial support from the NSF and the Sloan Foundation.

"He that filches from me my good name
Robs me of that which not enriches him
And makes me poor indeed."
—*Othello*, III.3.

1 Introduction

Everyday experience and a wealth of evidence from across the social sciences indicate that communication about the reputation of third parties is a key mechanism of social cooperation.[1] No one doubts that if they misbehave in a relationship with one (trading, business, romantic) partner, word might spread and they may end up being excluded from valuable future relationships. The threat of communication tomorrow and ostracism the day after keeps us on good behavior today.

While the role of communication in the community enforcement of cooperation seems familiar, it is not well-captured by existing game theory models of cooperation in large societies. In the classic community enforcement models of Kandori (1992) and Ellison (1994), where players observe only their partners' actions, cooperation is supported without explicit communication by relying on *contagion strategies*, a form of *collective punishment*: whenever a player sees anyone defect, she starts defecting against everyone. Contagion strategies cause the fastest possible breakdown of cooperation following a defection, and therefore the harshest punishment for defection. Why then do real-world societies often rely on communication and individualized punishment rather than contagion-like strategies? Are communication and individualized punishment truly necessary for supporting cooperation in large communities, or are they merely quirks of the particular cooperative equilibrium in which we happen to find ourselves?

This paper establishes the necessity of communication and individualized punishment in a version of the standard community enforcement model with more realistic assumptions. First, with small probability, each player is a "bad type" who always defects. In a companion paper (Sugaya and Wolitzky, 2020), we show that in games with anonymous players—like the anonymous prisoner's dilemma studied by Kandori and Ellison—this assumption completely precludes cooperation

[1] Many references can be given. For instance, see Grief (1993), Dixit (2003), and Tadelis (2016) in economics; Raub and Weesie (1990) in sociology; Ostrom (1990) and Ellickson (1994) in political science and law; Gluckman (1963) in anthropology; Noon and Delbridge (1993) in organizational behavior; Baumeister, Zhang, and Vohs (2004), Dunbar (2004), and Feinberg, Willer, and Schultz (2014) in psychology; and Sommerfeld et al. (2007) in evolutionary biology. Many of these papers refer to communication about third parties' behavior as *gossip*. This terminology accords with the computer science literature on gossip protocols (Hedetniemi, Hedetniemi, and Liestman, 1988; Shah, 2009), which we draw on.

in large societies, intuitively because collective punishment is too likely to be triggered in the presence of bad types. We therefore consider here the more realistic case where players observe (and remember) their partners' identities, so that individualized punishments (i.e., strategies that condition on the partner's identity) are technologically feasible. With observable identities, the presence of a few bad types obviously poses no obstacle to cooperation when each pair of players interacts frequently, because players can treat the overall repeated game as a collection of two-player games, cooperating with each partner if and only if he has behaved well in their bilateral relationship. We instead assume that, while each player interacts with others frequently (i.e., the discount factor δ is close to 1), she meets any particular partner infrequently (i.e., the population size N is much larger than $1/(1-\delta)$). In sum, we consider community enforcement with (1) a small chance of bad types, (2) observable identities, and (3) patient players but infrequent bilateral interactions. We show that cooperation in this environment is impossible in the absence of explicit communication (Theorem 1), but becomes possible if within-match cheap talk—ordinary conversation between matched partners—is allowed (Theorems 2 and 3).

More precisely, our results hold fixed the payoff parameters of the prisoner's dilemma stage game as well as an ε probability that each player is a "commitment type" who always defects (independent across players) and consider sequences of repeated games where the discount factor δ and the population size N change together. By viewing the overall repeated game as a collection of two-player games, it is trivial to support cooperation among pairs of rational players along any sequence where $(1-\delta)N \to 0$—that is, whenever bilateral interactions become frequent.[2]

In stark contract, our first main result (Theorem 1) shows that average payoffs converge to the mutual defection payoff along any sequence where $(1-\delta)N \to \infty$—that is, whenever bilateral interactions become infrequent. The logic of this result combines ideas from repeated game theory and information theory. Roughly speaking, the presence of bad types renders collective punishment ineffective, so incentives can be provided only by individualized punishment. When $(1-\delta)N \to \infty$, the population is too large for individualized punishment to be executed bilaterally—instead, a player's misbehavior against a partner must affect third parties' behavior towards her. Therefore, to support cooperation, players' actions must convey information about specific individuals' past behavior. This step is where information theory enters the picture: when actions are binary and

[2] To see why $(1-\delta)N \to 0$ corresponds to frequent bilateral interactions, suppose players match once every Δ units of real time with fixed discount rate $r > 0$, so $\delta = e^{-r\Delta}$, and hence $(1-\delta) \approx r\Delta$. Since each pair of players interact $1/(\Delta \times (N-1)) \approx r/((1-\delta)N)$ times per unit of real time on average, $(1-\delta)N \to 0$ means that each pair of players interact frequently, while $(1-\delta)N \to \infty$ means that each of pair of players rarely interact.

explicit communication is not allowed, each player receives only one bit of information per period (i.e., the opponent's action). We show that $O(N)$ bits are required to provide significant information about N players' individual actions (Lemma 3). Hence, $O(N)$ periods of communication-via-actions are required to monitor N players' actions. But this speed of communication is too slow to provide meaningful incentives when $(1-\delta)N \to \infty$. Thus, neither collective nor individualized punishment is effective, and cooperation is impossible.[3]

We then show that allowing within-match cheap talk restores the possibility of cooperation. Within-match cheap talk makes it technologically feasible for information about everyone's behavior to spread through the population exponentially quickly, reaching all players within $O(\log N)$ periods with high probability. Indeed, we establish that cooperation is possible along any sequence where $(1-\delta)\log N \to 0$.[4] We first derive a relatively simple version of this result (Theorem 2), which shows that cooperation can be achieved as an approximate Nash equilibrium using realistic strategies where (1) each player keeps track of a "blacklist" of opponents whom she believes have ever defected against a rational player, (2) players share their blacklists with each other prior to taking actions, and (3) each player defects against the opponents on her blacklist. However, such strategies form only an approximate equilibrium, because they break down in the low-probability event that a large fraction of the population are bad types; moreover, the possibility that this event can occur may unravel the equilibrium even in situations where no one assigns a high probability to this event.

Our final result (Theorem 3) then shows that cooperation can be achieved as an exact sequential equilibrium by combining the simple blacklisting idea of Theorem 2 with more complicated, "block belief-free" strategies that prevent unraveling. In our construction, players cooperate only after learning through communication that a large enough fraction of the population is rational. Furthermore, a player who does not learn that there are enough rational types can defect without fear of being punished (in the event that there are many rational types), because in this event subsequent communication will reveal that her defection was "justified" by her failure to learn.

Finally, the version of our model with communication can be extended by de-coupling the rate at which players meet to engage in payoff-relevant interactions and the rate at which they meet to engage in cheap-talk communication. In this extended model, we give a fairly complete

[3] The same argument applies when each player receives K bits of information per period, for any number K fixed independently of N. Thus, Theorem 1 holds for many information structures besides the canonical one where players observe only their partners' actions.

[4] Conversely, it is straightforward to show that cooperation is impossible if there exists $\rho > 0$ such that $(1-\delta)^{1+\rho}\log N \to \infty$.

characterization of how the prospects for cooperation depend on the population size and both meeting rates.

1.1 Related Literature

This paper contributes to the literatures on community enforcement, the folk theorem in repeated games, and the role of communication in supporting cooperation. Its most novel features are analyzing how the rates at which $\delta \to 1$ and $N \to \infty$ affect the scope for cooperation in a repeated random matching game, and showing that introducing explicit communication dramatically affects the race between δ and N.

The literature on community enforcement in repeated games originates with Kandori (1992) and Ellison (1994).[5] These authors assume complete information (no "bad types") and show that the threat of collective punishment via contagion supports cooperation whenever $(1-\delta)\log N \to 0$. Our companion paper (Sugaya and Wolitzky, 2020) shows that collective punishment breaks down in the presence bad types, which motivates the current paper's focus on individualized punishment.[6]

Bad types are also considered in the community enforcement models of Ghosh and Ray (1996) and Heller and Mohlin (2018), but in these papers bad types *help* support cooperation, by making players less tempted to cheat their current partners and return to the matching pool (in Ghosh and Ray's voluntary separation model) or by stabilizing grim trigger-like strategies by making the observation that a partner defected in the past informative of his being a bad type (in Heller and Mohlin's model). These papers are therefore less closely related to ours.[7]

In many papers on the folk theorem in repeated games, implicit communication through actions is found to be just as effective as explicit communication through cheap talk when δ is close to 1. For example, this is the case in Hörner and Olszewski's (2006) folk theorem with almost perfect monitoring and in Deb, Sugaya, and Wolitzky's (2020) folk theorem for anonymous random matching games. In contrast, implicit and explicit communication are not equivalent in our model, because we take $\delta \to 1$ and $N \to \infty$ simultaneously (so communication speed matters) and explicit

[5] See also Harrington (1995) and Okuno-Fujiwara and Postlewaite (1995).

[6] Kandori and Ellison were well aware of the importance of bad types but did not include them in their models. For example, Ellison wrote, "If one player were 'crazy' and always played D [defect]... 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," (p. 578).

[7] 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. Dilmé (2016) considers a similar model where cooperation is robust to introducing a small measure of bad types.

communication allows more information to be transmitted in each meeting.[8]

Several papers on community enforcement and repeated games on networks fix $\delta < 1$ and show that cooperation is easier to support when the news that a defection occurred spreads more quickly (Raub and Weesie, 1990; Klein, 1992; Ahn and Suominen, 2001; Dixit, 2003; Lippert and Spagnolo, 2011; Ali and Miller, 2013; Wolitzky, 2013; Balmaceda and Escobar, 2017). This force differs from the role of communication in our model, where introducing within-match communication does not increase the speed at which the community learns that *someone* defected, but rather enriches the information that can be transmitted in each match, so that the community learns faster *which* players defected. The value of communication is thus tied to the need to use individualized rather than collective punishment, which in turn is necessitated by the presence of bad types (which are absent in the above papers).[9] Finally, a number of papers consider settings where the need to provide incentives for honest communication constrains community enforcement (Bowen, Kreps, and Skrzypacz, 2013; Wolitzky, 2015; Ali and Miller, 2016, 2020; Barron and Guo, 2019). While we of course also insist that communication is incentive compatible, Theorem 3 shows that this constraint is ultimately not binding in our model.

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 match in pairs, uniformly at random and independently across periods, to play the prisoner's dilemma:

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

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

Each player is *rational* with probability $1 - \varepsilon$ and *bad* with probability ε, for some $\varepsilon \in (0, 1)$, independently across players.[10] The number of bad players thus follows a binomial distribution.

[8] Deb (2020) establishes a folk theorem for anonymous random matching games with explicit communication. In these games, communication has the distinctive role of serving to relax anonymity, as players can identify each other via endogenous "names." This role does not arise in our model with observable player identities.

[9] A very different role for explicit communication with high δ (and small N) is analyzed by Awaya and Krishna (2016, 2019). They show that communication can in effect improve monitoring by exploiting correlation between players' signals.

[10] We discuss generalizations to multiple "commitment types" and to correlated types in Section 5.1.

Rational players maximize expected discounted payoffs with discount factor $\delta \in (0,1)$. Bad players always play D.

Matching is *non-anonymous*. That is, at the beginning of each period t, every player i observes the identity (but not the type) of her period-t partner, which we denote by $\mu_{i,t} \in I \backslash \{i\}$. A player then chooses her own action $a_{i,t} \in \{C, D\}$, and finally observes her partner's action $a_{\mu_{i,t},t}$ at the end of the period. Thus, player i's *history* at the beginning of period t is $h_i^t = \left(\left(\mu_{i,\tau}, a_{i,\tau}, a_{\mu_{i,\tau},\tau}\right)_{\tau=1}^{t-1}, \mu_{i,t}\right)$, with $h_i^1 = \mu_{i,1}$. In Section 4, we augment the game by allowing pre-play cheap talk communication within each match. The description of a history for player i will then also include the history of messages sent by player i to her partners and received by player i from her partners.

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 $\sigma_i\left(h_i^t\right)$ 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 she is rational, by U_i. Our measure of population average payoffs is $U = \sum_i U_i/N$, which is the average over players of their expected payoffs conditional on the event that they are rational.[11] Note that since the minmax payoff is 0, bad types always take D, and the maximum sum of stage game payoff is 2, we have $U \in [0,1]$ in any Nash equilibrium. Since the payoff from mutual defection is $(0,0)$, we say that (some) cooperation arises if and only if $U \neq 0$.

For all of our results, we fix the stage game payoff parameters G and L and the "commitment probability" ε, and simultaneously vary the population size N and the discount factor δ. It is fairly trivial to see that cooperation can occur in a Nash (or sequential) equilibrium if $(1-\delta)N \to 0$, even without cheap talk communication; and that cooperation cannot occur in any Nash equilibrium if there exists $\rho > 0$ such that $(1-\delta)^{1+\rho} \log N \to \infty$, even with within-match cheap talk.[12] In contrast, our main results show that without cheap talk cooperation cannot occur if $(1-\delta)N \to \infty$, and that with cheap talk a folk theorem holds if $(1-\delta)\log N \to 0$.

[11] By focusing on players' payoffs conditional on being rational, we avoid the need to specify utility functions for bad types. An earlier version of this paper instead assumed that bad types have the same utility function as rational types (while being constrained to always play D) and considered ex ante expected payoffs. Our results hold verbatim for this alternative notion of average payoffs, except that the payoffs supported in Theorem 3 must be adjusted by an $O(\varepsilon)$ term.

[12] We establish the former result in Section 3.3 and the latter (which is similar to Proposition 3 of Kandori (1992)) in Section 4.

3 No Cooperation without Communication

We first show that cooperation without communication is impossible when bilateral interactions are infrequent.

Theorem 1 *For any sequence* (N, δ) *where* $(1-\delta) N \to \infty$ *and any corresponding sequence of Nash equilibrium population payoffs* (U)*, we have* $U \to 0$*.*

The intuition is that (1) under incomplete information ($\varepsilon > 0$), cooperation requires separately monitoring the actions of $O(N)$ players, (2) implicit communication via actions can convey only one bit of information per period, (3) $O(N)$ bits—and hence $O(N)$ periods of communication—are required to monitor the actions of $O(N)$ players, and (4) the promise of reward or punishment $O(N)$ periods in the future is insufficient to motivate cooperation when $(1-\delta) N$ is large.

We prove Theorem 1 in Section 3.1, deferring some details to the appendix. In Section 3.2, we discuss how the theorem extends when players observe additional information beyond their current partner's action. We then present a partial converse—a folk theorem when $(1-\delta) N \to 0$—in Section 3.3.

3.1 Proof of Theorem 1

We establish the stronger conclusion that $U \to 0$ for any sequence of strategy profiles σ such that each player i obtains a higher expected payoff from playing σ_i than from always playing D. That is, we relax the requirement that σ is a Nash equilibrium to the weaker requirement that each player i prefers σ_i to the strategy *Always Defect*. Furthermore, we show that this conclusion holds even when we consider "extended" strategy profiles $\sigma = (\sigma_i)_{i \in I}$ where σ_i can condition player i's action not only on her own partners' identities but on the entire match realization. We thus allow more potential equilibrium strategies than are actually available to the players, while requiring that fewer potential deviations are unprofitable.

Slightly abusing notation, let $\mu^t = (\mu_{i,\tau})_{i \in I, \tau \leq t}$ denote the first t periods of the match realization, let $h_i^t = \left(a_{i,\tau}, a_{\mu_{i,\tau},\tau}\right)_{\tau=1}^{t-1}$ denote the history of player i's own actions and past opponents' actions at the beginning of period t, and let σ_i denote a mapping from $\left(h_i^t, \mu^t\right)$ to a mixed action. That is, $\sigma_i\left(h_i^t, \mu^t\right)$ 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.) Fix such an (extended) strategy profile $\sigma = (\sigma_i)_i$.

Let 0_i (resp., 1_i) denote the event that player i is rational (resp., bad). For any $x_i \in \{0_i, 1_i\}$ 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 first t periods of the match realization are given by μ^t.

When player i's opponents play $(\sigma_j)_{j\neq i}$, the distribution over paths of play of the repeated game when $x_i = 0_i$ but i deviates to *Always Defect* is the same as that when $x_i = 1_i$ (in which case i is forced to play *Always Defect*). We call the condition that the rational type of player i prefers her equilibrium strategy σ_i to the strategy *Always Defect* "incentive compatibility," since it requires precisely that this type prefers to follows her own equilibrium strategy rather than the bad type's strategy. This condition appears in the appendix as equation (5). The first step of the proof of Theorem 1 puts this condition in a more convenient form and averages it over players $i \in I$. (Proofs of lemmas are deferred to the appendix. In the following, for $x \in \mathbb{R}$, $(x)_+ := \max\{x, 0\}$.)

Lemma 1 *If each player i prefers strategy σ_i to* Always Defect *(that is, if equation (5) in the appendix holds for all $i \in I$), then*

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

$$\leq \quad (1-\varepsilon)(1-\delta)\sum_t \delta^{t-1}\sum_{\mu^t}\Pr\left(\mu^t\right)\sum_{i,j\neq i}\frac{1}{N(N-1)}\sum_{h_j^t}\begin{pmatrix}\Pr\left(h_j^t|0_i,0_j,\mu^t\right)\\ -\Pr\left(h_j^t|1_i,0_j,\mu^t\right)\end{pmatrix}_+(1+G). \tag{1}$$

Intuitively, the left-hand side of (1) is a lower bound on the average over players i of the "cooperation cost" that player i incurs by following strategy σ_i rather than *Always Defect*; and the right-hand side is an upper bound on the average over players i of the "benefit from averted punishment" that player i gains by following σ_i rather than *Always Defect*. The heart of the proof of Theorem 1 consists of showing that the average benefit from averted punishment goes to 0 if $(1-\delta)N \to \infty$. Roughly speaking, this amounts to showing that the (expected, discounted, per-player average) influence of player i's type on the histories of players $j \neq i$ is small.

Lemma 2 *If $(1-\delta)N \to \infty$ then the "average benefit from averted punishment" (the right-hand side of (1)) goes to* 0*.*

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

$$U_i \leq (1-\delta)\sum_t \delta^{t-1}\sum_{\mu^t}\Pr\left(\mu^t\right)\sum_{j\neq i}\frac{1}{N-1}\sum_{h_j^t}(1-\varepsilon)\Pr\left(h_j^t|0_i,0_j,\mu^t\right)\Pr\left(\sigma_j(h_j^t,\mu^t)=C\right)(1+G),$$

because player i's payoff would equal the right-hand side of this inequality if her stage-game payoff were $1+G$ whenever her partner takes C and 0 whenever her partner takes D. By Bayes' rule, $\Pr\left(h_j^t|0_i,0_j,\mu^t\right) \leq \Pr\left(h_j^t|0_j,\mu^t\right)/(1-\varepsilon)$. Hence,

$$U_i \leq (1+G)(1-\delta)\sum_t \delta^{t-1}\sum_{\mu^t}\Pr\left(\mu^t\right)\sum_{j\neq i}\frac{1}{N-1}\sum_{h_j^t}\Pr\left(h_j^t|0_j,\mu^t\right)\Pr\left(\sigma_j(h_j^t,\mu^t)=C\right).$$

Averaging over $i \in I$, we have

$$U \leq (1+G)(1-\delta)\sum_t \delta^{t-1}\sum_{\mu^t}\Pr\left(\mu^t\right)\sum_{j}\frac{1}{N}\sum_{h_j^t}\Pr\left(h_j^t|0_j,\mu^t\right)\Pr\left(\sigma_j(h_j^t,\mu^t)=C\right).$$

Since Lemmas 1 and 2 imply that the right-hand side of this inequality goes to 0 if $(1-\delta)N \to \infty$, we have $U \to 0$ as well.

The proof of Lemma 2 relies on the following information theory result, which implies that the average influence of N players' types on a k-dimensional binary random variable is non-negligible only if $k/N \nrightarrow 0$.[13]

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

$$\sum_{i=1}^{N}\sum_{s\in\{0,1\}^k}\left(\Pr(s|0_i)-\Pr(s|1_i)\right)_+ \leq \sqrt{\frac{\log(2)\,kN}{\underline{\varepsilon}}}. \tag{2}$$

Let us provide some intuition and a proof sketch for Lemma 3. The $k=1$ case is the relatively familiar result that the average influence of N independent "votes" on a binary outcome is maximized by majority rule: if we let $s=1$ if and only if $\#\{i : X_i = 1\} \geq \varepsilon N$, each voter i is pivotal with probability approximately $1/\sqrt{N}$, and summing over voters gives a "total influence" of $\sum_{i=1}^{N}\sum_{s\in\{0,1\}}\left(\Pr(s|0_i)-\Pr(s|1_i)\right)_+ \approx \sum_{i=1}^{N} 1/\sqrt{N} = \sqrt{N}$.[14] The lemma asserts that in general the total influence of N independent votes on k binary outcomes is bounded by approximately $\sqrt{kN}$. This bound can be attained by splitting the population into k equal-sized groups and running majority rule within each group: each voter is then pivotal with probability approximately $1/\sqrt{N/k}$,

[13] We thank Omer Tamuz for suggesting a proof of this lemma.

[14] See for example Lemma A of Fudenberg, Levine, and Pesendorfer (1998) and Theorem 2 of Al-Najjar and Smorodinsky (2000).

and summing over voters gives a total influence of approximately $\sum_{i=1}^{N} 1/\sqrt{N/k} = \sqrt{kN}$. Intuitively, the bound is tight because more complex signals introduce correlation between the different signal dimensions, which reduces the average influence of a vote.

The proof of Lemma 3 proceeds as follows. First, we use Pinsker's inequality and some manipulations to show that the influence of i's vote, $\sum_{s\in\{0,1\}^k} \left(\Pr(s|0_i) - \Pr(s|1_i)\right)_+$, is at most $\sqrt{\log(2)\, I(S;X_i)/\underline{\varepsilon}}$, where $I(S;X_i)$ denotes the mutual information between S and X_i, measured in bits. Elementary properties of mutual information, together with independence of the X_i's, imply that $\sum_i I(S;X_i)$ is at most the entropy of S, which in turn is at most k since S is a k-dimensional binary random variable. Therefore, the sum of the squared influences is at most $\log(2) \sum_i I(S;X_i)/\underline{\varepsilon} \le \log(2)\, k/\underline{\varepsilon}$, and hence the sum of the influences is at most $\sqrt{\log(2)\, kN/\underline{\varepsilon}}$ by the $\ell_1 - \ell_2$ norm inequality.[15]

The proof of the theorem is completed by showing that Lemma 3 implies Lemma 2. This final step uses one more mathematical fact, which is that $\sum_{t=1}^{\infty} \delta^t \sqrt{t} \le (1-\delta)^{-3/2}$ for all $\delta \in (0,1)$.[16] Now, since $h_j^t = \left(a_{j,\tau}, a_{\mu_{j,\tau},\tau}\right)_{\tau=1}^{t-1}$ 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 $(X_i)_{i\neq j}$ (which are themselves independent conditional on $x_j = 0_j$ and μ^t), we have

$$
\begin{aligned}
& (1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}\sum_{\mu^t}\Pr\left(\mu^t\right)\sum_{i,j\neq i}\frac{1}{N(N-1)}\sum_{h_j^t}\left(\Pr\left(h_j^t|0_i,0_j,\mu^t\right)-\Pr\left(h_j^t|1_i,0_j,\mu^t\right)\right)_+ \\
= \; & (1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}\sum_{\mu^t}\Pr\left(\mu^t\right)\sum_{j}\frac{1}{N(N-1)}\sum_{i\neq j}\sum_{h_j^t}\left(\Pr\left(h_j^t|0_i,0_j,\mu^t\right)-\Pr\left(h_j^t|1_i,0_j,\mu^t\right)\right)_+ \\
\le \; & (1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}\frac{N}{N(N-1)}\sqrt{\frac{2\log(2)(t-1)(N-1)}{\underline{\varepsilon}}} \qquad \text{(by (2))} \\
= \; & \sqrt{\frac{2\log(2)}{\underline{\varepsilon}}}\frac{1}{\sqrt{N-1}}(1-\delta)\sum_{t=1}^{\infty}\delta^t\sqrt{t} \\
\le \; & \sqrt{\frac{2\log(2)}{\underline{\varepsilon}}}\frac{1}{\sqrt{N-1}}\frac{1}{\sqrt{1-\delta}}. \qquad (3)
\end{aligned}
$$

Hence, if $(1-\delta)N \to \infty$ then the right-hand side of (1) goes to 0.

[15] One could also try to prove Lemma 3 by induction on k. This approach easily gives a bound of order $k\sqrt{N}$. This approach is used by Awaya and Krishna (2016, Lemma 4; 2019, Lemma A.1), and is also reminiscent of Proposition 1 of Fudenberg, Levine, and Pesendorfer (1998). However, Lemma 3 requires a bound of order $\sqrt{kN}$, which seems difficult to establish by induction.

[16] This follows because, since $\sqrt{\cdot}$ is a concave function, Jensen's inequality gives $\frac{1-\delta}{\delta}\sum_{t=1}^{\infty}\delta^t\sqrt{t} \le \sqrt{\frac{1-\delta}{\delta}\sum_{t=1}^{\infty}\delta^t t} = \sqrt{\frac{1-\delta}{\delta}\frac{\delta}{(1-\delta)^2}} = \frac{1}{\sqrt{1-\delta}}$, and hence $\sum_{t=1}^{\infty}\delta^t\sqrt{t} \le \delta(1-\delta)^{-3/2} \le (1-\delta)^{-3/2}$.

3.2 Additional Signals

While Theorem 1 concerns the canonical repeated prisoner's dilemma with non-anonymous random matching, where players observe only their current partner's identity and action, it also extends to richer settings where players observe additional signals. Suppose that at the beginning of every period t, each player i privately observes an additional K-dimensional binary signal $s_{i,t} \in \{0,1\}^K$, where the distribution of $s_{i,t}$ depends on the entire history of matches, actions, and signals up to the end of period $t-1$, $\left(\mu_{i,\tau}, a_{i,\tau}, s_{i,\tau}\right)_{i \in I, \tau \leq t-1}$, and these signals can be arbitrarily correlated across players. For example, each player might observe the most recent action of each of K randomly chosen players in the population, or she might observe in which one out of 2^K possible "bins" lies the total number of players who cooperated last period, $|\{j \in I : a_{j,t-1} = C\}\ |$.[17] Theorem 1 extends to this more general model as follows.

Theorem 1' *Suppose each player observes an additional K-dimensional binary signal in every period, for some $K \in \mathbb{N}$ that may vary together with N and δ. For any sequence (N, δ, K) where $(1-\delta) N/K \to \infty$ and any corresponding sequence of Nash equilibrium population payoffs (U), we have $U \to 0$.*

Proof. If we re-define h_i^t as $\left(a_{i,\tau}, a_{\mu_{i,\tau},\tau}, s_{i,\tau}\right)_{\tau=1}^{t-1}$ rather than $\left(a_{i,\tau}, a_{\mu_{i,\tau},\tau}\right)_{\tau=1}^{t-1}$, the proof of Theorem 1 goes through as written, except that h_i^t now has dimension $(2+K)(t-1)$ rather than $2(t-1)$. This means that (3) must be replaced by

$$\sqrt{\frac{(2+K)\log(2)}{\underline{\varepsilon}}} \frac{1}{\sqrt{N-1}} \frac{1}{\sqrt{1-\delta}}.$$

If $(1-\delta) N/K \to \infty$ then this expression goes to 0, which implies the conclusion of Theorem 1. ∎

Since bilateral interactions between any two players become infrequent if $(1-\delta) N \to \infty$, a corollary of Theorem 1' is that, if players observe K-dimensional signals for any $K \in \mathbb{N}$ fixed independently of N and δ, cooperation is impossible when bilateral interactions are infrequent.

Theorem 1' also implies that the cheap talk communication we introduce in Section 4 can make a difference only if communication can convey an unbounded amount of information in a single period. This can be achieved by, for example, having each player convey a binary summary of

[17] Note that the signal is not allowed to depend on the identity of player i's period-t partner, $\mu_{i,t}$. This seemingly small point is actually very important. For example, if each player is exogenously informed of the "social standing" of her current partner, then cooperation can be supported even if $(1-\delta)N$ is large (and, indeed, even if the population is infinite). Models of this type, which are quite different from ours, are studied by Kandori (1992, Section 5), Okuno-Fujiwara and Postlewaite (1995), and Clark, Fudenberg, and Wolitzky (2020), among others.

the "reputation" of every player in the population to her current partner. The same conclusion also applies in situations (which are outside our model) where players have rich action sets and implicitly communicate through the fine details of their actions, as when bidders in an auction communicate via the trailing digits of their bids (Cramton and Schwartz, 2000).

3.3 A Converse: Cooperation with Frequent Bilateral Interactions

We provide a partial converse to Theorem 1, which shows that the presence of bad types does not hinder cooperation when bilateral interactions are frequent. The idea is simply to view the overall repeated game as a collection of $N(N-1)/2$ bilateral relationships, one for each pair of players, and use 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 prisoner's dilemma. 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 parameters (N, δ), let $E \subseteq \mathbb{R}^N$ denote the set of rational-player sequential equilibrium payoff vectors: that is, $(v_i)_i \in E$ if there exists a sequential equilibrium where $U_i = v_i$ for all $i \in I$.

Proposition 1 *Fix a constant $\eta > 0$ and a sequence $(N, \delta)_l$ indexed by $l \in \mathbb{N}$ satisfying $\lim_{l\to\infty}(1-\delta_l) N_l = 0$. 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 $l > \bar{l}$, in the game with parameters (N_l, δ_l) there is a payoff vector $v \in E_l$ satisfying*

$$\left| \left(\frac{1}{N_l - 1} \sum_{j \in I_l \setminus \{i\}} (1-\varepsilon) v_{i,j} \right) - v_i \right| < \varepsilon\eta \quad \textit{for all } i \in I_l.$$

Proof. See Appendix A.3. ■

The proof of Proposition 1 uses bilateral grim trigger strategies to show that any profile of bilaterally feasible and strictly individually rational payoff pairs is sustainable in sequential equilibrium when bilateral interactions are frequent. We conjecture that an even larger set of payoffs can be supported using more complex strategies. For example, player 1 may be willing to accept a negative present value payoff in her relationship with player 2 if she is compensated with a positive payoff in her relationship with player 3. In principle, such payoff vectors can be supported by having players occasionally communicate implicitly via actions.[18] We do not pursue such a result here.

One can also consider the case where $\delta \to 1$ and $N \to \infty$ at the same rate, so the bilateral in-

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

teraction frequency stays constant. Here partial cooperation is possible: the maximum equilibrium value of U can exceed 0, but a folk theorem typically does not hold (as follows from applying the proof of Theorem 1 when $(1-\delta)N$ is constant but large).

4 Cooperation with Communication

We now show that if players can exchange cheap talk messages with their partners before taking actions, cooperation is possible whenever $(1-\delta)\log N \to 0$. We assume that the set of possible messages is finite but can be taken arbitrarily large relative to the population size, N. Rational types communicate strategically to maximize their expected utility. As for bad types, we require that their communication strategy does not affect the distribution of opposing actions they face in any period. This implies that if we endowed bad types with utility functions that depend on the sequence of opposing actions they face, the strategies we construct would form an equilibrium for any such functions (although our analysis does not involve specifying utilities for bad types).[19]

We prove versions of our result for two solution concepts: *robust η-Nash equilibrium* and *robust (exact) sequential equilibrium*, where in both cases the word "robust" refers to the property that a bad type's communication strategy does not affect the distribution of actions she faces in any period. In this section, a strategy σ_i for player i specifies, as a function of her type and history, both a distribution over messages to send to her partner and a distribution of actions, where the specified action is always D when the player's type is bad.

Definition 1 *For any $\eta > 0$, a strategy profile $\sigma = (\sigma_i)_{i \in I}$ is a* robust η-Nash equilibrium *if the following conditions hold for each $i \in I$:*

1. *If player i is rational ($x_i = 0$), any deviation by player i improves her expected per-period payoff by at most η: for any strategy $\tilde{\sigma}_i$,*

$$\mathbb{E}^{\sigma}\left[(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}u\left(a_{i,t},a_{\mu_{i,t},t}\right)|x_i=0\right] \geq \mathbb{E}^{\tilde{\sigma}_i,\sigma_{-i}}\left[(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}u\left(a_{i,t},a_{\mu_{i,t},t}\right)|x_i=0\right]-\eta.$$

2. *If player i is bad ($x_i = 1$), her communication strategy does not affect the distribution of*

[19] An even more demanding equilibrium concept would require that the rational types' strategies remain optimal against any communication strategies for the bad types. Our constructions do not satisfy this stronger requirement.

actions she faces in any period: for any strategy $\tilde{\sigma}_i$ (which always takes D when $x_i = 1$),

$$\Pr^{\sigma}\left(a_{\mu_{i,t},t} = C | x_i = 1\right) = \Pr^{\tilde{\sigma}_i,\sigma_{-i}}\left(a_{\mu_{i,t},t} = C | x_i = 1\right).$$

This is a permissive version of approximate equilibrium, because (rational) players' gains from deviating are required to be small only in ex ante terms. This permissiveness allows a simple equilibrium construction, which we subsequently complement with a more complicated construction for exact sequential equilibrium.

Theorem 2 *Fix a sequence $(N, \delta)_l$ satisfying $(1 - \delta)\log N \to 0$. With cheap talk, for any $\eta > 0$ there exists $\bar{l} > 0$ such that, for every $l \geq \bar{l}$, in the game with parameters (N_l, δ_l) there is a robust η-Nash equilibrium in which rational players always cooperate with each other along the equilibrium path of play.*

Our proof of this result (in Appendix A.4) involves strategies that seem fairly realistic. Each player keeps track of a "blacklist" of players (other than herself) who she believes have ever previously played D against a rational opponent. Every period, players communicate their blacklists to their partners and then update their own blacklists 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 equilibrium whenever $(1 - \delta)\log N \to 0$, note that since the only consequence of communication is that players stop cooperating with opponents on their blacklists, a player is always indifferent between reporting any two blacklists to her partner (recalling that a player never appears on her own blacklist). Moreover, in equilibrium all blacklisted players are indeed bad, so it is optimal to take D against players on one's blacklist. It remains to show that it is approximately optimal to take C against players who are not on one's blacklist (on the equilibrium path).

To show this, observe 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. We formalize this observation with a lemma on diffusion processes. Consider random matching among N agents. For each $i \in I$, agent i's *state* at time t is $s_{i,t} = (s_{i,j,t})_{j \in I} \in \{0, 1\}^N$. When agents i and i' match at time t, the j^{th} component of each of their states updates to

$$
\begin{aligned}
s_{i,j,t+1} = s_{i',j,t+1} = \max\left\{s_{i,j,t}, s_{i',j,t}\right\} &\quad \text{if } j \notin \{i, i'\}, \\
s_{i,j,t+1} = s_{i,j,t} \text{ and } s_{i',j,t+1} = s_{i',j,t} &\quad \text{if } j \in \{i, i'\}.
\end{aligned}
$$

An interpretation is that $s_{i,j,t} = 1$ means that agent i knows a "rumor" about agent j at time t, and agents share all the rumors they know with their partners, except for rumors concerning themselves. The following lemma says that if initially agents i and $i+1$ know a rumor about i (for each i), then the probability that all N agents know all N rumors by time T converges to 1 exponentially in T.

Lemma 4 *Consider the above diffusion process with $s_{i,i,1} = s_{i,i+1,1} = 1$ and $s_{i,j,1} = 0$ for all i and $j \notin \{i, i+1\}$. There exist constants $c > 0$ and $Z > 0$ (independent of N) such that, for all $T > Z \log N$,*

$$
\Pr\left(s_{i,j,T} = 1 \textit{ for all } i, j \in I\right) \geq 1 - \exp\left(-\frac{cT}{\log N}\right).
$$

Frieze and Grimmett (1985) prove a similar result when each agent shares a rumor with a randomly selected receiver, rather than having players meet in pairs as in the current model.[20] 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 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. The spread of each rumor thus approximates a logistic curve, as in standard diffusion models.

By Lemma 4, 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 $(1-\delta)\log N \approx 0$, taking D against a rational opponent is unprofitable (at an on-path history where one has not yet been blacklisted). Moreover, Lemma 4 also implies that all bad types are very likely to be blacklisted within $O(\log N)$ periods. Hence, if $(1-\delta)\log N \approx 0$, a rational player's payoff loss from cooperating with not–yet-blacklisted bad types is small. The prescribed strategies thus form an η-Nash equilibrium.

[20] 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 4.

Several challenges arise in attempting to transform this η-Nash equilibrium into an (exact) sequential equilibrium. The most serious concerns the low-probability event where a player learns that the fraction of rational types in the population is actually much smaller than $1-\varepsilon$. In the extreme, suppose player 1 observes (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. When player 1 subsequently meets player 2, if $(1-\delta)N \approx \infty$ she should play D against him even if he is not on her blacklist, because players 1 and 2 now effectively find themselves in a two-player repeated game with discount factor $\delta^{N-1} \approx 0$ (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: under such strategies a player must assess whether her opponents believe that there are few rational players, whether they believe that their opponents believe this, and so on, and the equilibrium can easily unravel.

We therefore need a more sophisticated approach to construct an exact Nash or sequential equilibrium. The basic idea is to prescribe cooperation only after a player learns through communication that a large enough fraction of the population is rational, while preventing unraveling by excusing players who defect at "erroneous" histories where they failed to learn that there are enough rational types. To identify when a player's history was erroneous in this sense, her opponents must aggregate their information about her history, which by Lemma 4 can be achieved in $O(\log N)$ periods with high probability. Note that a given player's opponents can collectively identify her history through their own past observations; moreover, by making their continuation payoffs independent of whether she is rewarded or punished, they can be induced to communicate this information honestly.[21] This proof approach lets us support a wide range of payoffs as sequential equilibria. However, not surprisingly, the strategies used in the proof (in Online Appendix B.2) are much more complicated than those used to prove Theorem 2.

To state this more general theorem, first fix N and δ, and denote the (random) set of rational players by $\theta^* \subseteq I$. For each θ^*, let $F(\theta^*) \subseteq \mathbb{R}^{|\theta^*|}$ denote the set of payoff profiles for the rational players that are feasible when rational types always take D against bad types. 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 then define $F(\theta^*) = \mathrm{co}(\{\hat{u}_i(\mathbf{a})\}_{i\in\theta^*, \mathbf{a}\in\mathbf{A}(\theta^*)}) \subseteq \mathbb{R}^{|\theta^*|}$, where $\mathbf{A}(\theta^*) = \{\mathbf{a} : \mathbf{a}_i(j) = D \text{ if } i \notin \theta^* \text{ or } j \notin \theta^*\}$.

[21] This approach to incentivizing communiation was introduced by Compte (1998) and Kandori and Matsushima (1998) in the context of general repeated games with private monitoring.

Let $F^*(\theta^*) = F(\theta^*) \cap \mathbb{R}_+^{|\theta^*|}$ denote the set of individually rational payoff profiles in $F(\theta^*)$. Note that $F^*(\theta^*)$ implicitly depends on N, but not on δ. Finally, for any $\theta^* \in \{0,1\}^N$, $\alpha \in (0, 1-\varepsilon)$, and $\eta \in (0,1)$, we define the set $F^{\alpha,\eta}(\theta^*) \subseteq F^*(\theta^*)$ as follows:

1. If $|\theta^*| \geq \alpha N$ then

$$F^{\alpha,\eta}(\theta^*) = \left\{ v^{\theta^*} \in \mathbb{R}_+^{|\theta^*|} : \prod_{i\in\theta^*} \left[v_i^{\theta^*} - \eta, v_i^{\theta^*} + \eta \right] \subseteq F^*(\theta^*) \text{ for all } i \in \theta^* \right\}.$$

2. If $|\theta^*| < \alpha N$ then $F^{\alpha,\eta}(\theta^*)$ is the 1-element set consisting of the 0 vector in $\mathbb{R}^{|\theta^*|}$.

Intuitively, if $|\theta^*| \geq \alpha N$ then $F^{\alpha,\eta}(\theta^*)$ is the set of payoff profiles for the rational players which are feasible and individually rational with η slack, while if $|\theta^*| < \alpha N$ then $F^{\alpha,\eta}(\theta^*)$ consists of the payoff vector that results from mutual defection. It is not hard to show that $F^{\alpha,\eta}(\theta^*)$ is non-empty whenever $\eta < \alpha/2$.[22]

Theorem 3 *Fix a sequence $(N,\delta)_l$ satisfying $(1-\delta)\log N \to 0$, and fix any $\alpha \in (0, 1-\varepsilon)$, $\eta \in (0, \alpha/2)$, and $\gamma > 0$. With cheap talk, there exists $\bar{l} > 0$ such that, for any $l > \bar{l}$ and any $\left(v^{\theta^*}\right)_{\theta^*}$ satisfying $v^{\theta^*} \in F^{\alpha,\eta}(\theta^*)$ for all $\theta^* \subseteq I_l$, in the game with parameters (N_l, δ_l) there is a robust sequential equilibrium σ satisfying*

$$\left| \mathbb{E}^{\sigma}\left[(1-\delta) \sum_{t=1}^{\infty} \delta^{t-1} u\left(a_{i,t}, a_{\mu_{i,t},t}\right) |\theta^* \right] - v_i^{\theta^*} \right| < \gamma \text{ for all } \theta^* \subseteq I_l \text{ and } i \in \theta^*.$$

Proof. See Appendix B.2. ■

The following proposition (which is similar to Proposition 3 of Kandori, 1992) shows that the $(1-\delta)\log N \to 0$ sufficient condition in Theorem 3 is nearly the best possible.

Proposition 2 *Fix a sequence $(N,\delta)_l$ satisfying $(1-\delta)^{1+\rho}\log N \to \infty$ for some $\rho > 0$. There exists $\bar{l} > 0$ such that, for any $l > \bar{l}$, in the game with parameters (N_l, δ_l) the unique Nash equilibrium is* Always Defect.

Proof. The maximum number of players who can possibly learn about a single player's deviation within t periods is $\min\{2^t, N\}$. Hence, the difference in a player's total continuation payoff when

[22] This is true by defintion when $|\theta^*| < \alpha N$, so suppose that $|\theta^*| \geq \alpha N$. Since a matched pair of rational players can obtain payoffs $(0,0)$ or $(1,1)$ (among others), we have $(w,\ldots,w) \in F^*(\theta^*)$ for any $w \in \left[0, p^{\theta^*}\right]$, where $p^{\theta^*} := (|\theta^*|-1)/(N-1)$ is the probability that a rational player meets another rational player. Since $p^{\theta^*} > |\theta^*|/N \geq \alpha > 2\eta$, we have $\prod_{i\in\theta^*}[\alpha/2-\eta, \alpha/2+\eta] \subset F^*(\theta^*)$, and hence $(\alpha/2,\ldots,\alpha/2) \in F^{\alpha,\eta}(\theta^*)$.

she conforms rather than deviates is at most

$$\sum_{t=1}^{\infty} \delta^t \min\left\{\frac{2^t}{N}, 1\right\}(1+G+L),$$

where here $1+G+L$ represents the maximum difference between any two stage game payoffs. Note that, for any $\eta \in (0,1)$,

$$\begin{aligned}
\sum_{t=1}^{\infty} \delta^t \min\left\{\frac{2^t}{N}, 1\right\} &= \sum_{t=1}^{\lfloor \eta \log_2 N \rfloor} \delta^t \frac{2^t}{N} + \sum_{t=\lfloor \eta \log_2 N \rfloor + 1}^{\infty} \delta^t \min\left\{\frac{2^t}{N}, 1\right\} \\
&\leq \sum_{t=1}^{\lfloor \eta \log_2 N \rfloor} \frac{2^t}{N} + \frac{\delta^{\lfloor \eta \log_2 N \rfloor}}{1-\delta} \\
&\leq \frac{\lfloor \eta \log_2 N \rfloor \times N^{\eta}}{N} + \frac{\exp\left(-(1-\delta)\lfloor \eta \log_2 N \rfloor\right)}{1-\delta}, \qquad (4)
\end{aligned}$$

where the last inequality follows because the first term in the second line is the sum of $\lfloor \eta \log_2 N \rfloor$ terms that are each less than $2^{\lfloor \eta \log_2 N \rfloor} \leq N^{\eta}$, and $\delta^{\lfloor \eta \log_2 N \rfloor} = \exp\left(\lfloor \eta \log_2 N \rfloor \log \delta\right) \leq \exp\left(-(1-\delta)\lfloor \eta \log_2 N \rfloor\right)$. Note that (4) goes to 0 whenever there exists $\rho > 0$ such that $(1-\delta)^{1+\rho} \log N \to \infty$. Thus, if $(1-\delta)^{1+\rho} \log N \to \infty$ for some $\rho > 0$, the unique Nash equilibrium is *Always Defect*. ■

The remainder of this section previews the equilibrium construction for Theorem 3 (Section 4.1) and extends the model to de-couple the frequencies at which players interact and communicate (Section 4.2).

4.1 Sketch of the Equilibrium Construction for Theorem 3

The proof of Theorem 3 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 anonymous random matching games by Deb, Sugaya, and Wolitzky (2020), and to ex post equilibria in games with incomplete information by Sugaya and Yamamoto (2020). 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 few rational types, we must keep track of players' beliefs about the number of rational types. In particular, the equilibrium cannot be ex post with respect to the set of rational types. On the other hand, the availability of cheap talk makes providing incentives for truthful communication easier, as compared to the case where communication can be executed only through payoff-relevant actions.

The proof shows that strategies of the following form give a sequential equilibrium:

In the very first period of the repeated game, all rational players are supposed to cooperate. Given the realized period 1 action profile, we let $\theta \subseteq I$ denote the set of players who cooperated in period 1. Thus, θ is always a subset of θ^*, and in equilibrium θ equals θ^*. As we will see, all players will eventually abandon cooperation in the event that $|\theta| < \alpha N$. Period 1 thus plays a distinguished role in the equilibrium construction.

Following period 1, the repeated game is viewed as an infinite sequence of finite blocks of T^{**} consecutive periods, where T^{**} is a large number specified in the proof. At the beginning of each block, each player i selects her *state profile* $\left(x_i^\theta\right)_{\theta\subseteq I} \in (\{G,B\})_{\theta\subseteq I}$ for the block, which specifies a state $x_i^\theta \in \{G,B\}$ for each possible realization of θ, the set of players who cooperated in period 1.[23] Intuitively, even if at some point in the game player i comes to believe with probability 1 that the set of players who cooperated in period 1 was θ, she continues to entertain the possibility that the set of period 1 cooperators was actually some $\theta' \neq \theta$, and she keeps track of a state $x_i^{\theta'} \in \{G,B\}$ for each possible set θ'.

The interpretation of player i's state is as follows: As in Hörner and Olszewski (2006), Deb, Sugaya, and Wolitzky (2020), and Sugaya and Yamamoto (2020), player i can be viewed as the *arbiter* of player $i+1$'s payoff, meaning that player $i+1$'s equilibrium continuation payoff is high when player i is in the good state G, and player $i+1$'s equilibrium continuation payoff is low when player i is in the bad state B. Specifically, $x_i^\theta = G$ means that, if in the coming block the players reach agreement that the set of period 1 cooperators was θ, then player i prescribes a high continuation payoff for player $i+1$ (which is delivered both by player i cooperating with player $i+1$ herself, and also by player i "instructing" other players to cooperate with player $i+1$); similarly, $x_i^\theta = B$ means that, if agreement is reached that the set of period 1 cooperators was θ, then player i prescribes a low continuation payoff for player $i+1$ (and thus defects against player $i+1$ herself while also instructing others to defect against player $i+1$).[24] While it might seem unnecessary for the players to form a belief about θ anew in every block (since, of course, the true value of θ is determined once and for all by the period 1 action profile), this approach conveniently preserves the equilibrium's recursive structure.[25]

[23] These states have nothing to do with the states of the diffusion process analyzed in Lemma 4.

[24] "Agreement" here means that all players learn the same the set of period 1 cooperations, according to a communication protocol described below. The players can reach agreement without it being common knowledge that they have done so.

[25] This aspect of the construction is facilitated by specifying that "trembles" are much more likely in earlier blocks. Thus, if the players' communications in a prior block indicated that the set of period 1 cooperators is θ, while their communications in the current block indicate that it is $\theta' \neq \theta$, all players believe that the communications in the

The defining feature of a block belief-free equilibrium is that, for each i and θ, all players other than player $i+1$ (including player i herself) are indifferent as to whether player i selects state $x_i^\theta = G$ and $x_i^\theta = B$, while player $i+1$ is better-off when player i selects $x_i^\theta = G$. Player i can thus be prescribed to randomize between $x_i^\theta = G$ and $x_i^\theta = B$ with a probability depending on her history in the previous block, so as to provide incentives for player $i+1$ in the previous block. We now describe how a block is constructed so as to ensure that players are incentivized to both take the prescribed actions and communicate honestly.

Each block is divided into several sub-blocks. In the first portion of the block, players defect while communicating about who cooperated in period 1 so as to reach agreement about θ, as well as communicating their state profiles $\left(x_i^\theta\right)_{\theta \subseteq I}$.[26] Then, in each of K "main sub-blocks" (which together comprise the vast majority of the block, and hence determine the equilibrium payoffs), for multiple periods players take their prescribed actions (which depend on their states, the period 1 history, and the history within the block so far), and players then communicate about their observations within the block so far. Importantly, if the agreed-upon state θ satisfies $|\theta| < \alpha N$ then all players are prescribed defection in the main sub-blocks.

Communication is always executed through a protocol that facilitates truthtelling. In essence, when player i meets player j, she reports her past direct observations to him (i.e., her past actions and her past opponents' identities and actions); and also, for each third party $k \notin \{i, j\}$, she tells him all information that she has previously learned *via a chain of players that excludes player* k. Players thus "tag" each piece of information with the identities of the players who have previously conveyed it: for example, one piece of information might be, "I heard from Alice that she heard from Bob that Carol defected against David in period 5." Since tagging occurs at each step, so long as players other than player k have not deviated, player j can trust any information he receives that is tagged as coming from a chain that excludes k. As a consequence, a player cannot unilaterally affect others' inferences about any variable other than her direct observations. Moreover, since a player's direct observations in any period t are also observed by her period t partner, a player cannot unilaterally affect others' inferences about these variables, either.

The communication protocol thus ensures that each player cannot prevent her opponents from aggregating information about her own behavior. Together with the threat of punishment (which

earlier block were erroneous and proceed in the current block as if the true set were θ'.

[26] Technically, the first portion of the block is divided into three sub-blocks: one to reach agreement about θ, one to reach agreement about the state profiles, and one to ensure that, if communication in the first two sub-blocks was unsuccessful (e.g., if a player deviated, or if the matching process took on an unlikely realization that prevented all players from meeting and thus reaching agreement), this fact becomes known to all players with high probability.

takes the form of both "blacklisting" within the block and reduced continuation payoffs at the beginning of the next block), this strategy ensures incentive compatibility at on-path histories. To ensure incentive compatibility at off-path histories, other players reward player j for punishing player i in period t if and only if they confirm through a chain that excludes player j that player j's period t history was one where he was prescribed to punish player i. Since player j's continuation payoff is determined solely by the state of player $j-1$, which does not affect the payoffs of players other than j, it is optimal for player j's opponents to communicate her history honestly.

Finally, at the end of the block, each player i learns the player $i+1$'s history in the block through a chain that excludes $i+1$. She then uses this information to adjust her state mixing probability at the beginning of the next block, so as to deliver the promised continuation payoff to player $i+1$.

4.2 De-Coupling Interaction Frequency and Communication Frequency

Suppose that rather than matching every period to communicate and interact (i.e., play the PD) with the same partner, players instead form "communication matches" (where they exchange cheap talk messages but do not play the PD) every Δ_M units of time and "interaction matches" (where they play the PD but do not exchange messages) every Δ_A units of time, where players re-match uniformly at random in each meeting. The proofs of Theorems 2 and 3 do not rely on the assumption that players communicate and interact with the same partner each period, so if $\Delta_M = \Delta_A = \Delta$ and we fix a real-time discount rate r, we recover these results for any sequence (N, Δ) satisfying $\Delta \log N \to 0$ (because $1-\delta = 1-e^{-r\Delta} \approx r\Delta$). More interestingly, if $\Delta_M \neq \Delta_A$ it is straightforward to show that the conclusions of Theorems 2 and 3 hold for any sequence (N, Δ_M, Δ_A) satisfying $\Delta_A \to 0$ and $\Delta_M \log N \to 0$. That is, what these results really require is that players interact frequently (so that the discounting between interaction matches, $1 - e^{-r\Delta_A}$, is small), and that information spreads through the population quickly when players share all their information in every communication match. Since information spreads exponentially in the number of communication matches (by Lemma 4), the latter condition requires only that $\Delta_M \log N \to 0$. We explain this point further in Appendix B.3.

Putting our results together, we can characterize the prospects for cooperation for a wide range of combinations of the parameters N, Δ_M, and Δ_A. In particular, a folk theorem holds if either

1. $\Delta_A N \to 0$ (regardless of Δ_M; this is a version of Proposition 1), or

2. $\Delta_A \to 0$ and $\Delta_M \log N \to 0$ (by modifying Theorem 2 or 3 as just discussed).

Conversely, population average payoffs converge to 0 in any Nash equilibrium if either

1. $\Delta_A \to \infty$ (regardless of Δ_M; this is obvious), or
2. $(\min\{\Delta_A, \Delta_M\})^{1+\rho} \log N \to \infty$ for some $\rho > 0$ (this is a version of Proposition 2), or
3. $\Delta_A N \to \infty$ and $\Delta_A N/\Delta_M \to 0$ (this is an extension of Theorem 1—which directly applies when $\Delta_A N \to \infty$ and $\Delta_M = \infty$—to the case where Δ_M is large but finite; we omit the proof).

Thus, if we focus on the case where $\Delta_A/\Delta_M \nrightarrow 0$, so the communication frequency is not much less than the interaction frequency (as seems realistic), we obtain either a folk theorem or an impossibility result whenever Δ_A converges to either 0 or ∞, $\Delta_A N$ converges to either 0 or ∞, and either $\Delta_M \log N \to 0$ or $\Delta_M^{1+\rho} \log N \to \infty$ for some $\rho > 0$.

5 Discussion

5.1 Possible Extensions

We discuss the prospects for extending our model in some technical directions, deferring a broader discussion of future research to the next subsection.

Multiple commitment types: While the simple "bad types" we consider seem natural and realistic, there is little reason to rule out additional behavioral types that are committed to strategies other than *Always Defect*. Let us continue to assume that each player is bad with probability ε, while introducing a probability ε' that each player may be committed to some other (arbitrary) repeated game strategy. Theorem 1 extends immediately to this more general setting, with the modified conclusion that $\liminf U \leq \varepsilon' (1 + G)$ (to account for the fact that the ε' commitment types may cooperate). Theorem 2 also extends, provided that ε' is sufficiently small: if instead ε' is large, then the other commitment type strategies could provide incentives that overturn those in our construction, for example by rewarding players for defecting against rational opponents.[27] Finally, we conjecture that Theorem 3 extends if ε' is sufficiently small and in addition the set of commitment types has the property that there exists a pure strategy for the rational types and a

[27] However, if each commitment type takes a deterministic sequence of actions and messages (rather than responding to its opponents' behavior), then Theorem 2 holds for any ε'.

finite time T such that, when the rational types follow this strategy, each commitment type takes a different action than the rational types at some time $t < T$. (Under this property, the first T periods may collectively replace period 1 in the proof of Theorem 3.) But we have not verified this conjecture.

Correlated types: Our analysis of anonymous games in Sugaya and Wolitzky (2020) allows players' types to be correlated, so long as the distribution of the number of bad types satisfies a smoothness condition. In the present paper, independence is used critically in Lemma 3. As correlation is introduced, the exponent on N in equation (2) increases and the required condition on δ and N in Theorem 1 becomes more stringent, eventually becoming impossible to satisfy when types are perfectly correlated.[28] In contrast, the proofs of Theorems 2 and 3 can easily accommodate correlated types.

Independent noise: An interesting open question is how Theorem 1 might extend with i.i.d. noise, where each player is forced to play D with independent probability ε in every period, rather than with probability ε being forced to play D in all periods. Ellison (1994; Proposition 2) shows that contagion strategies (which require only $(1-\delta)\log N \to 0$) are robust to i.i.d. noise if $\varepsilon \to 0$ for a fixed N. If instead $N \to \infty$ for a fixed $\varepsilon > 0$, contagion strategies break down, but also our proof of Theorem 1 does not apply.[29] In this case, determining the critical discount factor for supporting cooperation as a function of N seems to require a more intricate analysis of sequential rationality constraints.

5.2 Conclusion

This paper has analyzed community enforcement in the presence of "bad types" who never cooperate. We established two main results. First, without explicit communication, community enforcement is ineffective, in that cooperation is sustainable only if bilateral interactions are frequent. Second, introducing ordinary conversation (cheap talk) between matched partners enables cooperation with infrequent bilateral interaction, so long as the population size is not exponentially greater than $1/(1-\delta)$. Together, these results show that communication is essential for supporting cooperation in large populations. We believe that our model and results provide a more realistic perspective on large-group cooperation than earlier analyses which focused on anonymous agents

[28] Specifically, independence implies that $\sum_i I(S;X_i) \leq k$, where $I(S;X_i)$ is the mutual information between S and X_i (see equation (7) in the appendix). As the X_i's become correlated, the upper bound of k increases towards N, which increases the upper bound in (2).

[29] With i.i.d. noise, the probability that a player is forced to play D for k consecutive periods is ε^k. We could thus apply Lemma 3 to this model with $\underline{\varepsilon}^k$ in place of $\underline{\varepsilon}$. But this bound is too loose to yield the conclusion of Theorem 1.

and collective punishment.

There are a few promising ways in which the theory could be brought even closer to reality. First, while we have shown that communication enables cooperation under realistic-seeming strategies that are approximately optimal (Theorem 2), our proof for exact sequential equilibrium (Theorem 3) relies on much more complicated strategies that should not be taken literally as a description of real-world behavior. A natural next question is whether and how cooperation can be supported in sequential equilibrium using simpler strategies, perhaps while allowing communication devices that are more powerful than plain cheap talk.

Richer communication devices must also be introduced to support cooperation in populations where N is exponentially greater than $1/(1-\delta)$ (or where $N=\infty$), as well as to model real-world informational institutions such as fiat money, credit bureaus, and online ratings systems. Recent papers on this topic include Heller and Mohlin (2018), Bhaskar and Thomas (2019), and Clark, Fudenberg, and Wolitzky (2020). Individuals' incentives to provide information to such institutions remain relatively poorly understood, as does these institutions' robustness to dishonest or malicious reporting (e.g., what happens if some agents are "bad communication types," in addition to the "bad action types" we considered?).[30]

Finally, in reality large cooperative groups may not be well-approximated by the canonical uniform random matching model studied here. Introducing incomplete information (e.g., "bad types") into more structured population models—such as models with voluntary separation, assortative matching, or network structure—is another interesting direction for future research.

[30] There is however an interesting empirical literature on these issues in the context on online ratings systems, which is surveyed in Section 5 of Tadelis (2016).

A Appendix: Omitted Proofs

The condition in the proof of Theorem 1 that player i (when rational) prefers strategy σ_i to *Always Defect* is given by

$$\begin{aligned}
&\underbrace{(1-\delta)\sum_t \delta^{t-1}\sum_{\mu^t}\Pr\left(\mu^t\right)\sum_{j\neq i}\frac{1}{N-1}\sum_{h_i^t,h_j^t}\left(\begin{array}{c}(1-\varepsilon)\Pr\left(h_i^t,h_j^t|0_i,0_j,\mu^t\right)u\left(\sigma_i(h_i^t,\mu^t),\sigma_j(h_j^t,\mu^t)\right)\\ +\varepsilon\Pr\left(h_i^t,h_j^t|0_i,1_j,\mu^t\right)u\left(\sigma_i(h_i^t,\mu^t),D\right)\end{array}\right)}_{i\text{'s expected payoff when rational and playing }\sigma_i}\\
\geq\ &\underbrace{(1-\delta)\sum_t \delta^{t-1}\sum_{\mu^t}\Pr\left(\mu^t\right)\sum_{j\neq i}\frac{1}{N-1}\sum_{h_i^t,h_j^t}\left(\begin{array}{c}(1-\varepsilon)\Pr\left(h_i^t,h_j^t|1_i,0_j,\mu^t\right)u\left(D,\sigma_j(h_j^t,\mu^t)\right)\\ +\varepsilon\Pr\left(h_i^t,h_j^t|1_i,1_j,\mu^t\right)u\left(D,D\right)\end{array}\right)}_{i\text{'s expected payoff when rational and playing \textit{Always Defect}}}, \quad (5)
\end{aligned}$$

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

A.1 Proof of Lemma 1

Since $u(D,D)=0$ and $u(C,D)=-L$, (5) 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 inequality may be rewritten as

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

Since $u(C,a) - u(D,a) \leq -\min\{G,L\}$ and $u(D,a) \in \{0, 1+G\}$ for each $a \in \{C,D\}$, this inequality implies that

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

or equivalently

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

Summing this inequality over i and dividing by N yields (1).

A.2 Proof of Lemma 3

For random variables A and B taking values in sets $\mathcal{A}$ and $\mathcal{B}$, we denote entropy by $H(A)$, conditional entropy by $H(A|B)$, and mutual information by $I(A;B)$. We have

$$\begin{aligned} H(A) &= -\sum_{a\in\mathcal{A}}\Pr(A=a)\log_2(\Pr(A=a)), \\ H(A|B) &= -\sum_{a\in\mathcal{A}}\sum_{b\in\mathcal{B}}\Pr(A=a,B=b)\log_2(\Pr(A=a|B=b)), \\ I(A;B) &= \sum_{a\in\mathcal{A}}\sum_{b\in\mathcal{B}}\Pr(A=a,B=b)\log_2\left(\frac{\Pr(A=a,B=b)}{\Pr(A=a)\Pr(B=b)}\right) \\ &= H(A)-H(A|B)=H(B)-H(B|A)=I(B;A). \end{aligned}$$

Recall that, for any random variables $A_1,...,A_n,B$, we have $H(A_1,...,A_n|B) \leq \sum_{i=1}^n H(A_i|B)$.[31]

Let $X=(X_i)_{i=1}^N$ be a collection of 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. Recall that

[31] This and other basic facts about entropy and mutual information used in the proof can be found in, for example, Cover and Thomas (2006).

$H(X) = \sum_i H(X_i)$ (by independence) and $H(S) \le k$. Denote the "influence" of X_i on S by

$$M_i(S) = \sum_{s \in \{0,1\}^k} \left(\Pr(S = s | X_i = 0) - \Pr(S = s | X_i = 1)\right)_+ .$$

Letting $\underline{\varepsilon} = \min\{\varepsilon, 1-\varepsilon\}$, we wish to show that

$$\sum_{i=1}^{N} M_i(S) \le \sqrt{\frac{\log(2) kN}{\underline{\varepsilon}}}.$$

Note that

$$\begin{aligned} M_i(S) &= \sum_{s \in \{0,1\}^k} \left(\Pr(S = s | X_i = 0) - \Pr(S = s) - \left(\Pr(S = s | X_i = 1) - \Pr(S = s)\right)\right)_+ \\ &= \sum_{x \in \{0,1\}} \sum_{s \in \{0,1\}^k} \left(\Pr(S = s | X_i = x) - \Pr(S = s)\right)_+ . \end{aligned}$$

Let $P(s) = \Pr(S = s)$, $P^0(s) = \Pr(S = s | X_i = 0)$, and $P^1(s) = \Pr(S = s | X_i = 1)$. By Pinsker's inequality,

$$\begin{aligned} \sum_{s \in \{0,1\}^k} \left(\Pr(S = s | X_i = 0) - \Pr(S = s)\right)_+ &\le \sqrt{\frac{\log(2)}{2} D_{KL}(P^0 || P)} \qquad \text{and} \\ \sum_{s \in \{0,1\}^k} \left(\Pr(S = s | X_i = 1) - \Pr(S = s)\right)_+ &\le \sqrt{\frac{\log(2)}{2} D_{KL}(P^1 || P)}, \end{aligned}$$

where $D_{KL}(\cdot || \cdot)$ denotes Kullback-Leibler divergence (measured in bits). Note that

$$\begin{aligned} D_{KL}\left(P^0 || P\right) &= \sum_{s \in \{0,1\}^k} \Pr(S = s | X_i = 0) \log_2 \frac{\Pr(S = s | X_i = 0)}{\Pr(S = s)} \\ &= \frac{1}{1-\varepsilon} \sum_{s \in \{0,1\}^k} \Pr(S = s, X_i = 0) \log_2 \left(\frac{\Pr(S = s, X_i = 0)}{\Pr(S = s) \Pr(X_i = 0)}\right), \end{aligned}$$

and

$$D_{KL}\left(P^1 || P\right) = \frac{1}{\varepsilon} \sum_{s \in \{0,1\}^k} \Pr(S = s, X_i = 1) \log_2 \left(\frac{\Pr(S = s, X_i = 0)}{\Pr(S = s) \Pr(X_i = 1)}\right).$$

Hence, we have

$$\begin{aligned} M_i(S) &\le \sqrt{\frac{\log(2)}{2} D_{KL}(P^0||P)} + \sqrt{\frac{\log(2)}{2} D_{KL}(P^1||P)} \\ &= \sqrt{\frac{\log(2)}{2(1-\varepsilon)} \sum_{s\in\{0,1\}^k} \Pr(S=s, X_i=0) \log_2\left(\frac{\Pr(S=s, X_i=0)}{\Pr(S=s)\Pr(X_i=0)}\right)} \\ &\quad + \sqrt{\frac{\log(2)}{2\varepsilon} \sum_{s\in\{0,1\}^k} \Pr(S=s, X_i=1) \log_2\left(\frac{\Pr(S=s, X_i=0)}{\Pr(S=s)\Pr(X_i=1)}\right)} \\ &\le \sqrt{\frac{\log(2)}{2\underline{\varepsilon}}} \left(\begin{array}{c} \sqrt{\sum_{s\in\{0,1\}^k} \Pr(S=s, X_i=0) \log_2\left(\frac{\Pr(S=s,X_i=0)}{\Pr(S=s)\Pr(X_i=0)}\right)} \\ +\sqrt{\sum_{s\in\{0,1\}^k} \Pr(S=s, X_i=1) \log_2\left(\frac{\Pr(S=s,X_i=0)}{\Pr(S=s)\Pr(X_i=1)}\right)} \end{array} \right) \\ &\le \sqrt{\frac{\log(2)}{\underline{\varepsilon}}} \sqrt{\sum_{x\in\{0,1\}} \sum_{s\in\{0,1\}^k} \Pr(S=s, X_i=x) \log_2\left(\frac{\Pr(S=s, X_i=x)}{\Pr(S=s)\Pr(X_i=x)}\right)} \\ &= \sqrt{\frac{\log(2) I(S;X_i)}{\underline{\varepsilon}}}, \end{aligned} \tag{6}$$

where the last inequality follows because $D_{KL}(\cdot||\cdot)$ is non-negative and $\sqrt{a}+\sqrt{b} \le \sqrt{2}\sqrt{a+b}$ for non-negative a, b by the $\ell_1 - \ell_2$ norm inequality, and the last equality is the definition of mutual information.

We next show that

$$\sum_i I(S;X_i) \le k. \tag{7}$$

To see this, first note that

$$k \ge H(S) \ge I(S;X) = H(X) - H(X|S) = \sum_i H(X_i) - H(X|S),$$

where the last equality follows from independence of $(X_i)_i$. Hence, $H(X|S) \ge \sum_i H(X_i) - k$, and therefore

$$\sum_i H(X_i|S) \ge H(X|S) \ge \sum_i H(X_i) - k.$$

Since $H(X_i|S) = H(X_i) - I(S;X_i)$, we have

$$\sum_i (H(X_i) - I(S;X_i)) \ge \sum_i H(X_i) - k \Leftrightarrow \sum_i I(S;X_i) \le k.$$

Combining (6) and (7), we have

$$\sum_i M_i(S)^2 \le \log(2) \sum_i \frac{I(S;X_i)}{\varepsilon} \le \frac{\log(2)\,k}{\varepsilon},$$

or

$$\sqrt{\sum_i M_i(S)^2} \le \sqrt{\frac{\log(2)\,k}{\varepsilon}}.$$

Finally, by the $\ell_1 - \ell_2$ norm inequality, in an N-dimensional space $|x|_1 \le \sqrt{N}\,|x|_2$. Hence,

$$\sum_i M_i(S) \le \sqrt{N}\sqrt{\sum_i M_i(S)^2} \le \sqrt{\frac{\log(2)\,kN}{\varepsilon}}.$$

A.3 Proof of Proposition 1

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 $\left(a_t^{i,j}\right)_{t\in\mathbb{N}}$.

Suppose each player i conditions her behavior against each player $j \neq i$ only 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

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

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

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

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

It remains to show that, for l sufficiently high, each player i's expected payoff (when rational) in the resulting sequential equilibrium satisfies

$$v_i \in \left[\frac{1}{N-1}\sum_{j\neq i}\left((1-\varepsilon)\,v_{i,j} - \varepsilon\eta\right), \frac{1}{N-1}\sum_{j\neq i}(1-\varepsilon)\,v_{i,j}\right].$$

When player j is rational, player i obtains payoff $v_{i,j}$ against player j. 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 $\left(1-\delta^t\right)u_{i,j}^{<t}+\delta^t\left(1-\delta\right)\left(-L\right)$, 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 $\left(1-\delta\right)L\leq\eta/2$, i's payoff against j is at least

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

Moreover, i's payoff against j is non-positive, since j always defects. Hence, i's expected payoff against j is at least $\left(1-\varepsilon\right)v_{i,j}-\varepsilon\eta$ and at most $\left(1-\varepsilon\right)v_{i,j}$. Averaging over $j\neq i$ yields the desired bounds for v_i.

A.4 Proof of Theorem 2

Consider the following strategies, which do not depend on l.

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

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

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

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

Fix $\eta>0$. We prove that, for sufficiently large l, these strategies form an η-Nash equilibrium. To do so, we (1) compute lower bounds on the equilibrium payoffs of rational and bad types, (2) compute upper bounds on the payoffs of rational and bad types from any unilateral deviation, and (3) show that the latter cannot exceed the former by more than η.

Rational type equilibrium payoff. Suppose i is rational, let S denote the set of bad players, and suppose that $|S| = n$. Fix any $T, Z \in \mathbb{N}$ with $Z > \bar{Z}$ (with $\bar{Z}$ defined as in the statement of Lemma 4). 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 4, $I^D_{i,T+Z\log_2 N} = S$ 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

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

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

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

where $p_n = \binom{N}{n}\varepsilon^n (1-\varepsilon)^{N-n}$ denotes the probability that there are n bad types.

We will show that, for sufficiently large l,

$$\begin{aligned} & \sum_n p_n \frac{N-1-n}{N-1} - \sum_n p_n \min_{T\in\mathbb{N}, Z>\bar{Z}} \left\{ \left(1 - \delta^{T+Z\log_2 N}\right) + n\left(\frac{n-1}{N-1}\right)^T + \exp(-cZ) \right\} (1+L) \\ \geq \; & \sum_n p_n \frac{N-1-n}{N-1} - \frac{2}{3}\eta. \end{aligned} \tag{8}$$

First, fix some $\hat{\alpha} \in (0, 1-\varepsilon)$, and fix some $Z > \bar{Z}$ such $\exp(-cZ) \leq \frac{\eta}{4(1+L)}$. Here $\hat{\alpha}$ and Z are independent of l. By the central limit theorem, for sufficiently large N (or l), the probability that there are more than $(1-\hat{\alpha})N$ bad types, $\sum_{n\geq(1-\hat{\alpha})N} p_n$, is less than $\frac{1}{12}\frac{1}{1+N+\exp(-cZ)}\frac{\eta}{1+L}$. Since $\left(1 - \delta^{T+Z\log_2 N}\right) + n\left(\frac{n-1}{N-1}\right)^T + \exp(-cZ) \leq 1 + N + \exp(-cZ)$ for each N, $n \leq N$, and $T \in \mathbb{N}$, to establish (8) it suffices to show that, for sufficiently large l, there exists T such that, for each

$n \leq (1-\hat{\alpha})\, N$, we have

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

By definition of Z, $\exp(-cZ)(1+L) \leq \frac{1}{4}\eta$. Hence, it remains to show that, for sufficiently large l, there exists T such that, for each $n \leq (1-\hat{\alpha})\, N$, we have

$$\left(\left(1-\delta^{T+Z\log_2 N}\right)+n\left(\frac{n-1}{N-1}\right)^T\right)(1+L) \leq \frac{1}{3}\eta. \tag{9}$$

To establish (9), for a given value of l, let $T \in \mathbb{N}$ be the smallest integer such that $N(1-\hat{\alpha})^T \leq \frac{\eta}{4(1+L)}$. Note that $T \leq \hat{c}\log_2 N$ for some constant $\hat{c}$. Now, for all $n \leq (1-\hat{\alpha})\, N$, we have

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

Hence, to establish (9), it suffices to show that, for sufficiently large l, we have

$$\left(1-\delta^{T+Z\log_2 N}\right)(1+L) \leq \frac{1}{12}\eta. \tag{10}$$

Finally, we have

$$1-\delta^{T+Z\log_2 N} \leq (1-\delta)\left(T+Z\log_2 N\right) \leq (1-\delta)\left(\hat{c}+Z\right)\log_2 N,$$

and $(1-\delta)\log_2 N \to 0$ as $l \to \infty$ by hypothesis. This establishes (10) (and hence (8)), as desired.

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 chain 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 $j_\tau = \mu_{i,\tau}$, $j \in \{j_\tau, \ldots, j_{t-1}\}$, $i \notin \{j_\tau, \ldots, j_{t-1}\}$, and $j_{t'+1} = \mu_{j_{t'},t'+1}$ for each $t' \in \{\tau, \ldots, t-2\}$. By Lemma 4, if player i takes D against a rational player in period τ, then, for every $Z > \bar{Z}$, 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

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

First fixing Z such that $\exp(-cZ) \leq \frac{\eta}{3(1+G)}$ and then taking $l \to \infty$, we see that $1 - \delta^{Z \log_2 N} \leq (1 - \delta) Z \log_2 N \to 0$, so 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 (8), we see that 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, for every $Z > \bar{Z}$, 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 payoff against bad opponents is non-positive, her payoff under any unilateral deviation is at most

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

As we have seen, this converges to 0 as $l \to \infty$. 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. "Enforcing Cooperation in Networked Societies." *Working Paper* (2013).

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

[4] Ali, S. Nageeb, and David A. Miller. "Communication and Cooperation in Markets." *Working Paper* (2020).

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

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

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

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

[9] Barron, Daniel, and Yingni Guo. "The Use and Misuse of Coordinated Punishments." *Working Paper* (2019).

[10] Baumeister, Roy F., Liqing Zhang, and Kathleen D. Vohs. "Gossip as Cultural Learning." *Review of General Psychology* 8.2 (2004): 111-121.

[11] Bhaskar, V., and Caroline Thomas. "Community Enforcement of Trust with Bounded Memory." *Review of Economic Studies* 86.3 (2019): 1010-1032.

[12] Bowen, T. Renee, David M. Kreps, and Andrzej Skrzypacz. "Rules with Discretion and Local Information." *Quarterly Journal of Economics* 128.3 (2013): 1273-1320.

[13] Clark, Daniel, Drew Fudenberg, and Alexander Wolitzky. "Record-Keeping and Cooperation in Large Societies." *Working Paper* (2020).

[14] Compte, Olivier. "Communication in Repeated Games with Imperfect Private Monitoring." *Econometrica* (1998): 597-626.

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

[16] Cramton, Peter, and Jesse A. Schwartz. "Collusive Bidding: Lessons from the FCC Spectrum Auctions." *Journal of Regulatory Economics* 17.3 (2000): 229-252.

[17] Deb, Joyee. "Cooperation and Community Responsibility," *Journal of Political Economy* 128.5 (2020): 1976-2009.

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

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

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

[21] Dunbar, Robin. "Gossip in Evolutionary Perspective," *Review of General Psychology* 8.2 (2004): 100-110.

[22] Ellickson, Robert C. *Order Without Law: How Neighbors Settle Disputes.* Harvard University Press, 1994.

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

[24] Feinberg, Matthew, Robb Willer, and Michael Schultz. "Gossip and Ostracism Promote Cooperation in Groups." *Psychological Science* 25.3 (2014): 656-664.

[25] 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.

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

[27] Fudenberg, Drew, and Eric Maskin. "On the Dispensability of Public Randomization in Discounted Repeated Games." *Journal of Economic Theory* 53.2 (1991): 428-438.

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

[29] Gluckman, Max. "Papers in Honor of Melville J. Herskovits: Gossip and Scandal." *Current Anthropology* 4.3 (1963): 307-316.

[30] Greif, Avner. "Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition." *American Economic Review* (1993): 525-548.

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

[32] Hedetniemi, Sandra M., Stephen T. Hedetniemi, and Arthur L. Liestman. "A Survey of Gossiping and Broadcasting in Communication Networks." *Networks* 18.4 (1988): 319-349.

[33] Heller, Yuval, and Erik Mohlin. "Observations on Cooperation." *Review of Economic Studies* 85.4 (2018): 2253-2282.

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

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

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

[37] Klein, Daniel B. "Promise Keeping in the Great Society: A Model of Credit Information Sharing." *Economics & Politics* 4.2 (1992): 117-136.

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

[39] Noon, Mike, and Rick Delbridge. "News from Behind my Hand: Gossip in Organizations," *Organization Studies* 14.1 (1993): 23-36.

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

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

[42] Raub, Werner, and Jeroen Weesie. "Reputation and Efficiency in Social Interactions: An Example of Network Effects." *American Journal of Sociology* 96.3 (1990): 626-654.

[43] Shah, Devavrat. *Gossip Algorithms.* Now Publishers Inc. (2009).

[44] 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.

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

[46] Sugaya, Takuo, and Yuichi Yamamoto. "Common Learning and Cooperation in Repeated Games." *Theoretical Economics* 15.3 (2020): 1175-1219.

[47] Tadelis, Steven. "Reputation and Feedback Systems in Online Platform Markets." *Annual Review of Economics* 8 (2016): 321-340.

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

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

[50] Wolitzky, Alexander. "Communication with Tokens in Repeated Games on Networks." *Theoretical Economics* 10.1 (2015): 67-101.

B Online Appendix

B.1 Proof of Lemma 4

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. Letting $T = Z \log_2 N$, it suffices 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 5 *There exists $\bar{\gamma} \in \left(0, \frac{1}{2}\right]$ such that, for every N and $n \leq \frac{2}{3}N$,*

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

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

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

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

$$\underbrace{\binom{n}{n'}}_{\substack{\text{who in } I_t \text{ meets} \\ \text{players in } I \backslash (I_t \cup \{i\})}} \times \underbrace{\frac{n-1}{N-1}}_{\substack{\text{"first" player in } I_t \\ \text{meets someone in } I_t}} \times \underbrace{\frac{n-3}{N-3}}_{\substack{\text{"second" (remaining) player in } I_t \\ \text{meets some (remaining) player in } I_t}} \times \cdots \times \frac{n'+1}{N-n+n'+1}$$

$$= \binom{n}{n'} \prod_{k=1}^{\frac{n-n'}{2}} \frac{n-2k+1}{N-2k+1}.$$

This expression 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

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

$$= \binom{n-1}{n'} \prod_{k=1}^{\frac{n-n'+1}{2}} \frac{n-2k+2}{N-2k+1} \leq \binom{n}{n'} \prod_{k=1}^{\frac{n-n'+1}{2}} \frac{n-2k+2}{N-2k+1}.$$

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

$$\begin{pmatrix} n \\ n' \end{pmatrix} \leq \frac{en^{n+\frac{1}{2}}e^{-n}}{2\pi\left(\gamma n\right)^{\gamma n+\frac{1}{2}}e^{-\gamma n}\left(\left(1-\gamma\right)n\right)^{(1-\gamma)n+\frac{1}{2}}e^{-(1-\gamma)n}} \leq \frac{e}{2\pi\left(\gamma\right)^{\gamma n+\frac{1}{2}}\left(1-\gamma\right)^{(1-\gamma)n+\frac{1}{2}}}.$$

We also have

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

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

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

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

and hence

$$\begin{aligned}\Pr\left(N_{t+1} \leq n+\gamma n | N_t = n\right) &\leq \frac{e\left(\gamma n+1\right)}{2\pi\left(\gamma\right)^{\gamma n+\frac{1}{2}}\left(1-\gamma\right)^{(1-\gamma)n+\frac{1}{2}}}\left(\frac{n}{N-1}\right)^{\frac{(1-\gamma)n}{2}} \\ &= \frac{e}{2\pi\gamma^{\frac{1}{2}}\left(1-\gamma\right)^{\frac{1}{2}}}\left(\frac{\left(\gamma n+1\right)^{\frac{2}{\gamma n}}}{\gamma^2\left(1-\gamma\right)^{2\frac{1-\gamma}{\gamma}}}\left(\frac{n}{N-1}\right)^{\frac{1-\gamma}{\gamma}}\right)^{\frac{\gamma n}{2}} \\ &\leq \frac{e}{2\pi\left(\gamma\right)^{\frac{1}{2}}\left(1-\gamma\right)^{\frac{1}{2}}}\left(\frac{e^2}{\gamma^2\left(1-\gamma\right)^{2\frac{1-\gamma}{\gamma}}}\left(\frac{2}{3}\frac{N}{N-1}\right)^{\frac{1-\gamma}{\gamma}}\right)^{\frac{\gamma n}{2}}. \quad (11)\end{aligned}$$

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

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

Such a $\bar{\gamma}$ exists as the left-hand side of this inequality goes to 0 as $\bar{\gamma} \to 0$. Since $N \geq 4$, we have $\frac{2}{3}\frac{N}{N-1} \leq \frac{8}{9}$. Hence, substituting $\gamma = \bar{\gamma}$ in (11), we have, for every N and $n \leq \frac{2}{3}N$,

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

as desired. ■

Fix $\bar{\gamma}$ satisfying the conditions of Lemma 5. Let $n^*(N)$ satisfy

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

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 5.

Lemma 6 *For every n satisfying $n^*(N) \leq n \leq \frac{2}{3}N$,*

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

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

Lemma 7 *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 $\gamma = \frac{1}{2}$ in (11), we have

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

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 8 *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 $\binom{6}{3} \times 12 \times \left(\frac{12}{N-1}\right)^3$. 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 $\bar{\gamma} \leq \frac{1}{2}$, we have the following lemma:

Lemma 9 *For every $N \geq \max\{\bar{N}_1, \bar{N}_2\}$,*

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

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

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

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

We now provide a symmetric bound for the case where $N_t^{-i}(j)$ is "large" for each $i \neq j$. 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 10 *There exists $\bar{N}_3$ such that, for each $N \geq \bar{N}_3$,*

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

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

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

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

Proof. Lemmas 5–9 provide an upper bound for the probability that fraction $\bar{\gamma}$ 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 $\bar{\gamma}$ 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 9 and 10 to prove Lemma 4. We first assume $N \geq \max\{\bar{N}_1, \bar{N}_2, \bar{N}_3\}$.

We have the following properties. First, if $\min_{i,j} N_t^{-i}(j) < 12$, then $\min_{i,j} N_{t+6}^{-i}(j) \geq 12$ with probability at least $\frac{3}{4}$. Second, if $12 \leq \min_{i,j} N_t^{-i}(j) \leq \frac{2}{3}N$, then $\min_{i,j} N_{t+1}^{-i}(j) \geq (1+\bar{\gamma}) \min_{i,j} N_t^{-i}(j)$ with probability at least $\frac{3}{4}$. (And note that $\log_{(1+\bar{\gamma})} \frac{2}{3}N$ "increases" by a factor of $(1+\bar{\gamma})$ suffice to raise $\min_{i,j} N_t^{-i}(j)$ to $\frac{2}{3}N$.) Third, if $\frac{2}{3}N \leq \min_{i,j} N_t^{-i}(j) \leq N-13$—or equivalently $12 \leq \max_{i,j} M_t^{-i}(j) \leq \frac{1}{3}N$—then $\max_{i,j} M_{t+1}^{-i}(j) \leq (1-\bar{\gamma}) \max_{i,j} M_t^{-i}(j)$ with probability at least $\frac{3}{4}$. (Note that $\log_{(1-\bar{\gamma})} 3\frac{1}{N}$ "decreases" suffice to reduce $\max_{i,j} M_t^{-i}(j)$ to 12.) Finally, if $\max_{i,j} M_t^{-i}(j) \leq 12$, 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 $2+\log_{(1+\bar{\gamma})}\frac{2}{3}N+\log_{(1-\bar{\gamma})}3\frac{1}{N}$ of them equal 1. By Hoeffding's inequality, this probability is at least

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

If $T=Z\log_2 N$, then

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

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

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

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 3

Fix $\alpha\in(0,1-\varepsilon)$ and $\eta\in(0,\alpha/2)$ throughout the proof. We first describe a protocol for the community to circulate messages. This protocol has the feature that, with high probability, the number of periods required for everyone to learn the message is of order $\log N$; moreover, no single player can prevent 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 the players obey the following protocol for 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"

$$\left(\zeta_{j,t}^{I,-i}, \zeta_{j,t}^{M,-i}\right)_{i \neq j} \subseteq \times_{i \neq j} \left(I \times \left(\times_{n \neq i} M_n\right)\right).$$

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 $\zeta_{j,t}^{M,-i}|_k \subseteq M_k$ is the set of messages reported to j as having been sent by k via a path that excludes i.[32] Note that $\zeta_{j,t}^{M,-i}$ does not include player i's message since it is infeasible to share player i's message via a path that excludes i.

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 $\left(\zeta_{j,t}^{I,-i}, \zeta_{j,t}^{M,-i}\right)_{i \neq j}$, if player j meets player k, player j sends message $\left(\zeta_{j,t}^{I,-i}, \zeta_{j,t}^{M,-i}\right)_{i \notin \{j,k\}}$. That is, player j passes all of his information to player k, except for the "$-k$" information being circulated by players $-k$. Given his 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 $n \neq i$ (recall that $n \neq i$ since $\zeta_{j,t}^{M,-i} \subseteq \times_{n \neq i} M_n$). For each $i \in \{j,k\}$, 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 him as having been sent by n (note that $k \neq i$ by definition: only player $k \neq i$ hears a message via a path that excludes i). (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 $m_{-i}(j) = \texttt{error}$.[33] 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 4 the probability that the match is erroneous decreases exponentially in Z. We thus have

Lemma 11 *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*

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

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

B.2.2 Parameters Used in the Construction

We fix some parameters that will be used in the construction. Let

$$K := \left\lceil \frac{\max\{G, L\}}{\eta/16} \right\rceil. \tag{12}$$

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

[33] Note that it is possible that $m_{-i}(j)|_n \neq m_{-i'}(j)|_n$ for some $i \neq i' \neq n \neq i$. Intuitively, $m_{-i}(j) = \texttt{error}$ means that j fails to infer a message through a chain of players excluding i, but not necessarily that the messages she infers through all chains are mutually consistent.

Fix Z sufficiently large such that $Z \geq \bar{Z}$ (with c and $\bar{Z}$ given in Lemma 11),

$$\exp(-cZ) \geq \frac{1}{2}, \text{ and} \tag{13}$$

$$(K+3)\left(\frac{1}{Z} + \exp(-cZ)\right)\bar{u} \leq \frac{\eta}{16}, \tag{14}$$

where $\bar{u} = 2\max\{L, 1+G\}$. Next, let

$$\begin{aligned} T &= Z\log_2 N, \\ T^* &= (3 + (K-1)(Z+1) + Z)T, \text{ and} \\ T^{**} &= (3 + K(1+Z))T. \end{aligned} \tag{15}$$

Note that, by (12) and (15), we have

$$\frac{\eta}{16}T^* \geq ZT \times \max\{G, L\}. \tag{16}$$

Finally, take l sufficiently large such that

$$(1-\delta)(-L) + \delta\eta \sum_{\theta^*:|\theta^*|\geq\alpha N} \Pr(\theta^*|i \in \theta^*) \geq (1-\delta)(1+G), \tag{17}$$

$$\frac{1}{T^*}\left(\left|\frac{1-\delta^{T^{**}}}{1-\delta} - T^*\right| + 2\sum_{t=1}^{T^*}\left(1-\delta^{t-1}\right)\right)\bar{u} \leq \frac{1}{8}\eta, \text{ and} \tag{18}$$

$$\frac{1-\delta}{\delta^{T^{**}}}\left(\left|\frac{1-\delta^{T^{**}}}{1-\delta} - T^*\right| + 2\sum_{t=1}^{T^*}\left(1-\delta^{t-1}\right) + 4T^*\right)\bar{u} \leq \frac{1}{2}\eta. \tag{19}$$

B.2.3 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 (see Section B.2.4 for the definition of the block). In period 1^*, every normal player is supposed to play C. Given the outcome of period 1^*, let θ denote the set of players who took $a_{i,1^*} = C$ as prescribed. (Note that $\theta \subseteq \theta^*$, 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, $h_{i,1^*} = \left(\mu_{i,1^*}, a_{i,1^*}, a_{\mu_{i,1^*},1^*}\right)$, is directly informative of θ; for this reason, players' period-1^* histories will play a distinguished role in our construction.[34]

Fix $\left(v^{\theta^*}\right)_{\theta^*}$ such that $v^{\theta^*} \in F^{\alpha,\eta}(\theta^*)$ for each θ^*; and let $\left(\tilde{v}^{\theta,\theta^*}\right)_{\theta,\theta^*}$ be such that, for each θ and θ^* with $\theta \subseteq \theta^*$, we have

$$\tilde{v}_i^{\theta,\theta^*} = \begin{cases} v_i^\theta & \text{if } i \in \theta, \\ 0 & \text{if } i \in \theta^* \setminus \{\theta\}. \end{cases}$$

Note that $\tilde{v}^{\theta,\theta^*} \in \mathbb{R}^{|\theta^*|}$. Since $v^\theta \in F^{\alpha,\eta}(\theta)$, the vector $\left(\tilde{v}_i^{\theta,\theta^*}\right)_{i\in\theta} \in \mathbb{R}^{|\theta|}$ satisfies the following

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

conditions:

1. For each θ satisfying $|\theta| \geq \alpha N$ and each $i \in \theta$, we have $\prod_{i\in\theta}\left[\tilde{v}_i^{\theta,\theta^*} - \eta, \tilde{v}_i^{\theta,\theta^*} + \eta\right] \subseteq F^*(\theta)$. Since $\mathbf{A}(\theta)$ specifies that players take D against opponents who are not in θ, we have

$$\prod_{i\in\theta}\left[\tilde{v}_i^{\theta,\theta^*} - \eta, \tilde{v}_i^{\theta,\theta^*} + \eta\right] \times \prod_{i\in\theta^*\setminus\{\theta\}}\{0\} \subseteq F^*(\theta^*).$$

2. For each θ satisfying $|\theta| < \alpha N$ and each $i \in \theta$, $v_i^\theta = 0$.

We will construct a sequential equilibrium where, for each θ^*, $(h_{i,1^*})_{i\in I}$, θ, and $i \in \theta^*$, player i's continuation payoff starting from the second period of the repeated game equals $\tilde{v}_i^{\theta,\theta^*}$. To see why this suffices to prove the theorem, note that in state θ^* a rational type's equilibrium payoff in period 1^* is $p^{\theta^*} - \left(1 - p^{\theta^*}\right)L$, and hence her repeated game payoff is $(1-\delta)\left(p^{\theta^*} - \left(1-p^{\theta^*}\right)L\right) + \delta v_i^{\theta^*}$. Since $v_i^{\theta^*} \geq \eta$ whenever $i \in \theta^*$ and $|\theta^*| \geq \alpha N$, player i's expected payoff from taking C in period 1^* is

$$\begin{aligned}
& (1-\delta)\sum_{\theta^*}\Pr(\theta^*|i\in\theta^*)\left(p^{\theta^*} - \left(1-p^{\theta^*}\right)L\right) + \delta\sum_{\theta^*:|\theta^*|\geq\alpha N}\Pr(\theta^*|i\in\theta^*)\,v_i^{\theta^*} \\
\geq\ & (1-\delta)(-L) + \delta\eta\sum_{\theta^*:|\theta^*|\geq\alpha N}\Pr(\theta^*|i\in\theta^*),
\end{aligned}$$

while her expected payoff from taking D in period 1^* is at most $(1-\delta)(1+G)$. By (17), taking C in period 1^* is optimal.

B.2.4 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. 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 $\boldsymbol{x}_i = \left(x_i^\theta\right)_{\theta\in\Theta}$ and player i's period-1^* history $h_{i,1^*}$, player i plays a behavioral 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* $h_i^t = \left(\left(\mu_{i,\tau}, m_{i,\tau}, m_{\mu_{i,\tau},\tau}, a_{i,\tau}, a_{\mu_{i,\tau},\tau}\right)_{\tau=1}^{t-1}, \mu_{i,t}\right)$. 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.

Target Payoffs. For each θ, $\boldsymbol{x}$, and $i \in \theta^*$, let

$$v_i^\theta(x_{i-1}^\theta) = \begin{cases} v_i^\theta - \operatorname{sign}\left(x_{i-1}^\theta\right)\frac{\eta}{4} & \text{if } |\theta| \geq \alpha N \text{ and } i \in \theta, \\ 0 & \text{otherwise,} \end{cases} \tag{20}$$

where $\operatorname{sign}\left(x_{i-1}^\theta\right) := 1_{\{x_{i-1}^\theta = B\}} - 1_{\{x_{i-1}^\theta = G\}}$. Note that

$$v_i^\theta \begin{cases} \in (v_i^\theta(B), v_i^\theta(G)) & \text{if } |\theta| \geq \alpha N \text{ and } i \in \theta, \\ = v_i^\theta(B) = v_i^\theta(G) & \text{otherwise.} \end{cases} \tag{21}$$

As will be seen, continuation payoffs at the beginning of a block conditional on state θ and strategy state profile $\boldsymbol{x}^\theta$ will equal $\left(v_i^\theta(x_{i-1}^\theta)\right)_{i\in I}$.

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 $\left(h_i^t\right)_{i\in I, t\le T^{**}}$.[35] Denote this number by $\tilde{c}$. Beliefs are derived from Bayes' rule along a sequence of completely mixed strategy profiles $\left(\sigma^l\right)_{l\in\mathbb{N}}$, in which each player i "trembles" 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 a block strategy profile $\left(\sigma_i^*\left(\boldsymbol{x}_i, h_i^{1^*}\right)\right)_{i\in I, \boldsymbol{x}_i\in\{G,B\}^{|\Theta|}, h_i^{1^*}\in H_i^{1^*}}$, and continuation payoffs $\left(w_i^\theta(x_{i-1}^\theta, h_{i-1}^{T^{**}+1})\right)_{i\in\theta,\theta\in\Theta, x_{i-1}^\theta\in\{G,B\}, h_{i-1}^{T^{**}+1}\in H_{i-1}^{T^{**}+1}}$. Since $\theta\subseteq\theta^*$, payoffs are well-defined.

We present a set of conditions which will guarantee the existence of a robust sequential equilibrium where payoffs from period 2 on are given by $\left(\tilde{v}^{\theta,\theta^*}\right)_{\theta,\theta^*}$. In what follows, $\Pr^\sigma(\cdot|\cdot)$ and $\mathbb{E}^\sigma[\cdot|\cdot]$ denote conditional probability and expectation, respectively, 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 $\theta\in\Theta$, the following two conditions hold:

   (a) For each $i\notin\theta$, $\boldsymbol{x}\in\{G,B\}^{N|\Theta|}$, $h_i^{1^*}\in H_i^{1^*}$, $t\in\{1,\ldots,T^{**}\}$, $b\in\mathbb{N}$, and $\tilde{h}_i^{b,t}\in\tilde{H}_i^{b,t}$,

$$\sum_{h_{-i}^{1*}\in H_{-i}^{1*}}\beta_i\left(h_{-i}^{1^*}|\boldsymbol{x}_{-i},\tilde{h}_i^{b,t}\right)\Pr{}^{\sigma^*\left(\boldsymbol{x},h^{1^*}\right)}\left(\left(a_{i,t},a_{\mu_{i,t},t}\right)=(D,D)\,|\boldsymbol{x},h^{1^*},\tilde{h}_i^{b,t}\right)=1.$$

[35] 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.

(That is, each player $i \notin \theta$ takes D and expects her opponents to take D towards her.)

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

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

(Here, the sum $\sum_{\tau=1}^{T^{**}}$ is taken over all periods in the current block b, where the current period $t \in \{(b-1) T^{**} + 2, \ldots bT^{**} + 1\}$ is some period in block b.

Note that both sequential rationality conditions are 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 $\theta \in \Theta$, $i \in \theta$, $x_{i-1}^\theta \in \{G, B\}^N$, $b \in \mathbb{N}$, and $\tilde{h}^{b,0} \in \tilde{H}^{b,0}$,

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

(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 $\theta \in \Theta$, $i \in \theta$,

(a) If θ satisfies $|\theta| \geq \alpha N$, we have $w_i^\theta (x_{i-1}^\theta, h_{i-1}^{T^{**}+1}) \in (v_i^\theta (B), v_i^\theta (G))$ for each $x_{i-1}^\theta \in \{G, B\}$ and $h_{i-1}^{T^{**}+1} \in H_{i-1}^{T^{**}+1}$.

(b) If θ satisfies $|\theta| < \alpha N$, we have $w_i^\theta (x_{i-1}^\theta, h_{i-1}^{T^{**}+1}) = 0$ for each $x_{i-1}^\theta \in \{G, B\}$ and $h_{i-1}^{T^{**}+1} \in H_{i-1}^{T^{**}+1}$.

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 $\theta \in \Theta$, the following two conditions hold:

(a) For each $i \notin \theta$, $\boldsymbol{x} \in \{G, B\}^{N|\Theta|}$, $h_i^{1^*} \in H_i^{1^*}$, $t \in \{1, \ldots, T^{**}\}$, $b \in \mathbb{N}$, and $\tilde{h}_i^{b,t} \in \tilde{H}_i^{b,t}$,

$$\sum_{h_{-i}^{1^*} \in H_{-i}^{1^*}} \beta_i \left(h_{-i}^{1^*} | \boldsymbol{x}_{-i}, \tilde{h}_i^{b,t}\right) \Pr{}^{\sigma^*\left(\boldsymbol{x}, h^{1^*}\right)} \left(\left(a_{i,t}, a_{\mu_{i,t},t}\right) = (D, D) | \boldsymbol{x}, h^{1^*}, \tilde{h}_i^{b,t} \right) = 1. \quad (22)$$

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

$$\sum_{h_{-i}^{1^*} \in H_{-i}^{1^*}} \beta_i \left(h_{-i}^{1^*} | \boldsymbol{x}_{-i}, \tilde{h}_i^{b,t}\right) \mathbb{E}^{\left(\sigma_i, \sigma_{-i}^*\left(\boldsymbol{x}_{-i}, h_{-i}^{1^*}\right)\right)} \left[\sum_{\tau=1}^{T^{**}} \delta^{\tau-1} \hat{u}_i \left(\mathbf{a}_\tau\right) + \pi_i^\theta (x_{i-1}^\theta, h_{i-1}^{T^{**}+1}) | \boldsymbol{x}_{-i}, \tilde{h}_i^{b,t} \right]. \quad (23)$$

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

$$v_i^{\theta}(x_{i-1}^{\theta}) = \mathbb{E}^{\sigma^*(\boldsymbol{x})}\left[\frac{1-\delta}{1-\delta^{T^{**}}}\sum_{t=1}^{T^{**}}\delta^{t-1}\hat{u}_i(\mathbf{a}_t) + \pi_i^{\theta}(x_{i-1}^{\theta}, h_{i-1}^{T^{**}+1})|\tilde{h}^{b,0}, \theta\right]. \tag{24}$$

3. *[Self-Generation]* For each $\theta \in \Theta$, $i \in \theta$, $x_{i-1}^{\theta} \in \{G,B\}$ and $h_{i-1}^{T^{**}+1} \in H_{i-1}^{T^{**}+1}$,

   (a) If θ satisfies $|\theta| \geq \alpha N$, we have

$$\text{sign}\left(x_{i-1}^{\theta}\right)\pi_i^{\theta}(x_{i-1}^{\theta}, h_{i-1}^{T^{**}+1}) > 0 \text{ and } \left|\frac{1-\delta}{\delta^{T^{**}}}\pi_i^{\theta}(x_{i-1}^{\theta}, h_{i-1}^{T^{**}+1})\right| < \frac{\eta}{2}. \tag{25}$$

   (b) If θ satisfies $|\theta| < \alpha N$, we have

$$\pi_i^{\theta}(x_{i-1}^{\theta}, h_{i-1}^{T^{**}+1}) = 0. \tag{26}$$

Lemma 12 *For each l, suppose there exist $\left(\sigma_i^*\left(\boldsymbol{x}_i, h_i^{1^*}\right)\right)_{i\in I, \boldsymbol{x}_i\in\{G,B\}^{|\Theta|}, h_i^{1^*}\in H_i^{1^*}}$ and $\left(\pi_i^{\theta}(x_{i-1}^{\theta}, h_{i-1}^{T^{**}+1})\right)_{i\in\theta, \theta\in\Theta, x_{i-1}^{\theta}\in\{G,B\}, h_{i-1}^{T^{**}+1}\in H_{i-1}^{T^{**}+1}}$ such that Conditions (22)–(26) are satisfied. Then there exists a robust sequential equilibrium such that, for each θ^*, $(h_{i,1^*})_{i\in I}$, and θ, payoffs from the second period of the repeated game equal $\left(\tilde{v}_i^{\theta,\theta^*}\right)_{i\in\theta^*}$.*

Proof. For $i \in \theta$, Conditions (24) and (25) imply that payoffs $\left(v_i^{\theta}(x_{i-1}^{\theta})\right)_{i\in I, \theta\in\Theta, x_{i-1}^{\theta}\in\{G,B\}}$ can be delivered at the beginning of each block with full-support state transition probabilities. For $i \in \theta^* \setminus \theta$, Conditions (22) and (26) imply that payoff 0 can be delivered at the beginning of each block with full-support state transition probabilities. Condition (21) 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 $\left(v_i^{\theta}\right)_{i\in\theta}$ can be delivered. This is as in, for example, Hörner and Olszewski (2006). In total, payoffs from the second period of the repeated game equal $\left(\tilde{v}_i^{\theta,\theta^*}\right)_{i\in\theta^*}$.

For $i \in \theta$, Condition (23) 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. For $i \notin \theta$, Condition (22) implies that the distribution of the action sequence that player i faces is independent of her strategy for each realization of $\boldsymbol{x}_{-i}$. It is therefore optimal for rational player i to take D in each period and send any messages, and this behavior is indeed what is prescribed for player i by (22). For bad player i, Condition (22) implies that her strategy does not affect the action sequence she faces, as required. ■

To prove Theorem 3, it thus suffices to show that there exist $\left(\sigma_i^*\left(\boldsymbol{x}_i, h_i^{1^*}\right)\right)_{i\in I, \boldsymbol{x}_i\in\{G,B\}^{|\Theta|}, h_i^{1^*}\in H_i^{1^*}}$ 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 (22)–(26) are satisfied.

B.2.5 Target Actions

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

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 $v_i^{\theta}(G) = v_i^{\theta}(B) = 0$ for each $i \in \theta$. Note that, to satisfy (25), 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 $x^\theta \in \{G, B\}^N$, there exists a distribution $\hat{\boldsymbol{\alpha}}^{x^\theta} \in \Delta(\mathbf{A}(\theta))$ such that, for each $i \in \theta$, $\hat{u}_i\left(\hat{\boldsymbol{\alpha}}^{x^\theta}\right) > v_i^\theta + \frac{1}{2}\eta$ if $x_{i-1}^\theta = G$, and $\hat{u}_i\left(\hat{\boldsymbol{\alpha}}^{x^\theta}\right) < v_i^\theta - \frac{1}{2}\eta$ if $x_{i-1}^\theta = B$; and for each $i \notin \theta$, $a_{\mu_i} = D$ with probability one for all x^θ. Here existence follows from the fact that $\left(v_i^\theta\right)_{i\in\theta}$ satisfies $\prod_{i\in\theta}\left[v_i^\theta - \eta, v_i^\theta + \eta\right] \subseteq F^*(\theta)$ and (20).

Since $\hat{\boldsymbol{\alpha}}^{x^\theta} \in \Delta(\mathbf{A}(\theta))$, we have

$$\hat{u}\left(\hat{\boldsymbol{\alpha}}^{x^\theta}\right) = \sum_{\mathbf{a}\in\mathbf{A}(\theta)} \boldsymbol{\alpha}^{x^\theta}(\mathbf{a}) \sum_{\mu} \frac{2}{N(N-1)} \left(u_i\left(\mathbf{a}_i(\mu_i), \mathbf{a}_{\mu_i}(i)\right)\right)_{i\in I}.$$

(Here, $\frac{2}{N(N-1)}$ is the probability of each match realization $\mu = (\mu_i)_{i\in I}$.)

Let $A_{i,j}(\theta) = \{(a_i, a_j) \in \{C, D\}^2 : \left(\mathbf{a}_i(\mu_i), \mathbf{a}_{\mu_i}(i)\right) = (D, D) \text{ if } i \notin \theta \text{ or } j \notin \theta\}$. Since $\mathbf{A}(\theta)$ specifies that players take D towards opponents outside θ, given $\hat{\boldsymbol{\alpha}}^{x^\theta}$, when players i and j meet they take $(a_i, a_j) \in A_{i,j}(\theta)$ with probability one. Hence, we have

$$u\left(\hat{\boldsymbol{\alpha}}^{x^\theta}\right) = \left(\sum_{\mu_i \neq i} \frac{1}{N-1} \sum_{\mathbf{a}_i(\mu_i), \mathbf{a}_{\mu_i}(i) \in A_{i,\mu_i}(\theta)} \boldsymbol{\alpha}_{i,\mu_i}^{x^\theta}\left(\mathbf{a}_i(\mu_i), \mathbf{a}_{\mu_i}(i)\right) u_i\left(\mathbf{a}_i(\mu_i), \mathbf{a}_{\mu_i}(i)\right)\right)_{i\in I},$$

where $\boldsymbol{\alpha}_{i,\mu_i}^{x^\theta}$ is the distribution of action pairs of player i's and player μ_i's given $\hat{\boldsymbol{\alpha}}^{x^\theta}$, conditional on the event that players i and μ_i match.

Therefore, for θ satisfying $|\theta| \geq \alpha N$, for each $x^\theta \in \{G, B\}^N$, there exists $\boldsymbol{\alpha}_{i,j}^{x^\theta} \in \Delta(A_{i,j}(\theta))$ for each $i \neq j$ such that the following conditions hold: letting $\boldsymbol{\alpha}^{x^\theta} \in \Delta A$ be the distribution of action profiles when μ is drawn uniformly and then (a_i, a_j) is drawn from $\boldsymbol{\alpha}_{i,j}^{x^\theta}$ given $j = \mu_i$, for each $i \in \theta$, $\hat{u}_i\left(\boldsymbol{\alpha}^{x^\theta}\right) > v_i^\theta + \frac{3}{4}\eta$ if $x_{i-1}^\theta = G$, and $\hat{u}_i\left(\boldsymbol{\alpha}^{x^\theta}\right) < v_i^\theta - \frac{3}{4}\eta$ if $x_{i-1}^\theta = B$; and for each $i \notin \theta$, $\left(a_i, a_{\mu_i}\right) = D$ with probability one for all x^θ.

Since $\hat{u}_i\left(\boldsymbol{\alpha}^{x^\theta}\right) > v_i^\theta + \frac{3}{4}\eta$ and $\hat{u}_i\left(\boldsymbol{\alpha}^{x^\theta}\right) < v_i^\theta - \frac{3}{4}\eta$ are strict inequalities, we can assume that $\boldsymbol{\alpha}_{i,j}^{x^\theta}(a_i, a_j)$ is a rational number: there exists $M^A \in \mathbb{N}$ such that, for all $(a_i, a_j) \in \{C, D\}^2$,

$$\boldsymbol{\alpha}_{i,j}^{x^\theta}(a_i, a_j) = \frac{m_{a_i,a_j}}{M^A} \text{ for some } m_{a_i,a_j} \in \mathbb{N}. \tag{27}$$

Arbitrarily label the four elements of $\{C, D\}^2$ by $\left(a_i^1, a_j^1\right)$, $\left(a_i^2, a_j^2\right)$, $\left(a_i^3, a_j^3\right)$, and $\left(a_i^4, a_j^4\right)$. For each x^θ and (i, j), we can see $\boldsymbol{\alpha}_{i,j}^{x^\theta}$ as a uniform distribution over $\left\{1, ..., M^A\right\}$. Given the realization $m \in \left\{1, ..., M^A\right\}$, there exists a unique $k \in \{1, 2, 3, 4\}$ such that

$$\sum_{0\leq\tilde{k}<k} m_{a_i^{\tilde{k}}, a_j^{\tilde{k}}} < m \leq \sum_{0\leq\tilde{k}\leq k} m_{a_i^{\tilde{k}}, a_j^{\tilde{k}}}. \tag{28}$$

For such k, we say that the realization of $\boldsymbol{\alpha}_{i,j}^{x^\theta}$ is $\left(a_i^k, a_j^k\right)$.

Note that, for each $i \in \theta$, we have

$$\left(\max_{x^\theta : x_{i-1}^\theta = B} \hat{u}_i\left(\boldsymbol{\alpha}^{x^\theta}\right)\right)_+ < v_i^\theta(B) - \frac{\eta}{4} < v_i^\theta < v_i^\theta(G) + \frac{\eta}{4} < \min_{x^\theta : x_{i-1}^\theta = G} \hat{u}_i\left(\boldsymbol{\alpha}^{x^\theta}\right). \tag{29}$$

B.2.6 Structure of the Block

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 (i,j) who are paired draw a joint controlled lottery according to the target action distribution $\boldsymbol{\alpha}_{i,j}^{x^\theta}$. Given the realization, they take the pure action.
   (Formally, players use their inference of θ and x^θ based on 1^*-communication and $\boldsymbol{x}$-communication sub-blocks. Moreover, players draw the joint controlled lottery by cheap talk. These are the only cheap talk messages sent in the main round).
   (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.

Given (15), T^{**} is in fact the length of the block, or equivalently the last period of supplemental round K; and T^* is the last period of main round K.

B.2.7 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 $\left(x_{-i}^\theta, h^{1^*}, h^{T^*+1}\right)$ rather than $\left(x_{i-1}^\theta, h_{i-1}^{T^{**}+1}\right)$ (when $|\theta| \geq \alpha N$).[36]

Let $\Sigma_i^{T^*}$ denote the set of i's block strategies up to period T^*. We show that the following conditions guarantee the existence of a robust sequential equilibrium where payoffs from period 2 on equal $\left(\tilde{v}_i^{\theta,\theta^*}\right)_{i\in\theta^*}$.

1. *[Sequential Rationality]* For each $\theta \in \Theta$, the following two conditions hold:
   (a) For each $i \notin \theta$, $\boldsymbol{x} \in \{G,B\}^{N|\Theta|}$, $h_i^{1^*} \in H_i^{1^*}$, $t \in \{1,\ldots,T^*\}$, $b \in \mathbb{N}$, and $\tilde{h}_i^{b,t} \in \tilde{H}_i^{b,t}$,

$$\sum_{h_{-i}^{1^*}\in H_{-i}^{1^*}} \beta_i\left(h_{-i}^{1^*}|\boldsymbol{x}_{-i},\tilde{h}_i^{b,t}\right)\Pr^{\sigma^*\left(\boldsymbol{x},h^{1^*}\right)}\left(\left(a_i,a_{\mu_{i,t}}\right)=(D,D)\,|\boldsymbol{x}_{-i},h^{1^*},\tilde{h}_i^{b,t}\right)=1. \quad (30)$$

[36] A similar but more complicated argument appears in Deb, Sugaya, and Wolitzky (2020).

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

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

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

$$v_i^\theta(x_{i-1}^\theta) = \mathbb{E}^{\sigma^*\left(\boldsymbol{x}, h^{1^*}\right)}\left[\begin{array}{cc} 1_{\{\theta:|\theta|\geq\alpha N\}} \frac{1}{T^*}\left(\sum_{\tau=1}^{T^*} \hat{u}_i\left(\mathbf{a}_\tau\right) + \pi_i^\theta(x_{-i}^\theta, h^{1^*}, h^{T^*+1})\right) & |\theta, \tilde{h}^{b,0} \\ +1_{\{\theta:|\theta|<\alpha N\}} \frac{1-\delta}{1-\delta^{T^*}} \sum_{\tau=1}^{T^*} \delta^{\tau-1} \hat{u}_i\left(\mathbf{a}_\tau\right) & \end{array}\right]. \tag{32}$$

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

$$\text{sign}\left(x_{i-1}^\theta\right) \pi_i^\theta(x_{-i}^\theta, h^{1^*}, h^{T^*+1}) \left\{\begin{array}{ll} \geq \frac{1}{8}\eta T^* & \text{for } \theta \text{ satisfying } |\theta| \geq \alpha N, \\ = 0 & \text{for } \theta \text{ satisfying } |\theta| < \alpha N. \end{array}\right. \tag{33}$$

and

$$\left|\pi_i^\theta(x_{-i}^\theta, h^{1^*}, h^{T^*+1})\right| \left\{\begin{array}{ll} \leq 2\bar{u} T^* & \text{for } \theta \text{ satisfying } |\theta| \geq \alpha N, \\ = 0 & \text{for } \theta \text{ satisfying } |\theta| < \alpha N. \end{array}\right. \tag{34}$$

Lemma 13 *Suppose 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* $\left(\pi_i^\theta(x_{-i}^\theta, h^{1^*}, h^{T^*+1})\right)_{i \in I, x_{-i}^\theta \in \{G,B\}^{N-1}, h^{1^*}}$, *such that Conditions (31)–(34) are satisfied. Then there exists a robust sequential equilibrium such that, for each* θ^*, $(h_{i,1^*})_{i \in I}$, *and* θ, *payoffs from the second period of the repeated game equal* $\left(\tilde{v}_i^{\theta,\theta^*}\right)_{i \in \theta^*}$.

Proof. We have fixed $v_i^\theta(x_{i-1}^\theta)$. Fix $\sigma_i^{T^*}$ and π_i^θ satisfying (31)–(34). We will construct $\tilde{\sigma}_i^*$, and $\tilde{\pi}_i^\theta$ that satisfy (22)–(26).

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 $\tilde{\pi}_i^\theta(x_{i-1}^\theta, h_{i-1}^{T^{**}+1})$ as follows. (i) If $m_{-i}(i-1) = \texttt{error}$, then $\tilde{\pi}_i^\theta(x_{i-1}^\theta, h_{i-1}^{T^{**}+1}) = 0$ for each $x_{i-1}^\theta, h_{i-1}^{T^{**}+1}$. (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

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

It remains to show that $\tilde{\sigma}_i^*\left(\boldsymbol{x}_i, h_i^{1^*}\right)$, $\tilde{\beta}^*$, $\tilde{\pi}_i^\theta$, and $v_i^\theta(x_{i-1}^\theta)$ satisfy (22)–(26).

For $i \notin \theta$, since players take D in supplemental round K, (30) implies (22). For $i \in \theta$, note that player i's payoff depends on the outcome of play in supplemental round K only through her stage game payoffs (rational types maximize them 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 $t \le T^*$, the expected value of

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

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

$$\begin{aligned} & 1_{\{\theta:|\theta|\ge\alpha N\}}\left(\left(\frac{1-\delta^{T^{**}}}{1-\delta}-T^*\right)v_i^\theta(x_{i-1}^\theta)+\sum_{\tau=t}^{T^*}\hat{u}_i\left(\mathbf{a}_\tau\right)+\pi_i^\theta(x_{-i}^\theta, h^{1^*}, h^{T^*+1})\right) \\ & +1_{\{\theta:|\theta|<\alpha N\}}\sum_{\tau=t}^{T^*}\delta^{\tau-1}\hat{u}_i\left(\mathbf{a}_\tau\right). \end{aligned}$$

Ignoring the constant $\left(\frac{1-\delta^{T^{**}}}{1-\delta}-T^*\right)v_i^\theta(x_{i-1}^\theta)$, (31) implies (23).

For $|\theta| < \alpha N$, (24)–(26) hold since all the payoffs and rewards are zero regardless of x_{i-1}^θ and $h_{i-1}^{T^{**}+1}$. For $|\theta| \ge \alpha N$, (24) follows from (32). In addition, (25) holds since

$$\Pr\left(m_{-i}(i-1) \neq \texttt{error}\right) \ge \frac{1}{2} \text{ by Lemma 11 and (13),}$$

$$\begin{aligned} & \text{sign}\left(x_{i-1}^\theta\right)\left(\left(\frac{1-\delta^{T^{**}}}{1-\delta}-T^*\right)v_i^\theta(x_{i-1}^\theta)+\frac{\sum_{t=1}^{T^*}\left(1-\delta^{t-1}\right)\hat{u}_i(\mathbf{a}_t(i-1))+\pi_i^\theta(\boldsymbol{x}_{-i}(i-1),h^{1^*}(i-1),h^{T^*+1}(i-1))}{\Pr(m_{-i}(i-1)\neq\texttt{error})}\right) \\ \ge & -\left(\left|\frac{1-\delta^{T^{**}}}{1-\delta}-T^*\right|\bar{u}+\frac{\sum_{t=1}^{T^*}\left(1-\delta^{t-1}\right)\bar{u}}{\frac{1}{2}}\right)+\frac{1}{8}\eta T^* \text{ by (33)} \\ > & \ 0 \text{ by (18),} \end{aligned}$$

and

$$\begin{aligned} & \frac{1-\delta}{\delta^{T^{**}}}\left|\left(\frac{1-\delta^{T^{**}}}{1-\delta}-T^*\right)v_i^\theta(x_{i-1}^\theta)+\frac{\sum_{t=1}^{T^*}\left(1-\delta^{t-1}\right)\hat{u}_i(\mathbf{a}_t(i-1))+\pi_i^\theta(\boldsymbol{x}_{-i}(i-1),h^{1^*}(i-1),h^{T^*+1}(i-1))}{\Pr(m_{-i}(i-1)\neq\texttt{error})}\right| \\ \le & \frac{1-\delta}{\delta^{T^{**}}}\left(\left|\frac{1-\delta^{T^{**}}}{1-\delta}-T^*\right|\bar{u}+\frac{\sum_{t=1}^{T^*}\left(1-\delta^{t-1}\right)\bar{u}}{\frac{1}{2}}\right)+\frac{1-\delta}{\delta^{T^{**}}}\frac{2\bar{u}T^*}{\frac{1}{2}} \text{ by (34)} \\ < & \frac{\eta}{2} \text{ by (19).} \end{aligned}$$

■

B.2.8 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 except for the joint controlled lottery drawn at the beginning of each period of the main round. 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 triple $\left(h_i^t, m_{i,t}, m_{\mu_{i,t},t}\right)$, an action $a_{i,t}$ is a *detectable deviation* if there does not exist a strategy state $\hat{\boldsymbol{x}}_i$ such that $\left(h_i^t, m_{i,t}, m_{\mu_{i,t},t}, a_{i,t}\right)$ 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, $m_i = \left\{i, \mu_{i,1^*}\right\} \cap \theta$.

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 $m_{-j}(i)|_k \neq m_{-j'}(i)|_k$ (i.e., i receives inconsistent messages through a path excluding j and through a path excluding j'). 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 11 immediately implies the following result.

Lemma 14 *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*

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

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 $\boldsymbol{x}_{-i} \in \{G, B\}^{|\Theta|(N-1)}$, player i believes that $\theta \supseteq \theta(h_i^{T+1})$:

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

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 $\boldsymbol{x}_{-i} \in \{G, B\}^{|\theta|(N-1)}$, player i believes that $\theta \supseteq \theta(h_j^{T+1})$ for each $j \neq i$:

$$\sum_{\theta \supseteq \theta(h_j^{T+1})} \beta_{i,t}\left(\theta | \boldsymbol{x}_{-i}, \tilde{h}_i^{b,t}\right) = 1. \tag{36}$$

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 $\emptyset$ (this is because, to have $\theta(h_j^{T+1}) \neq \emptyset$, player j needs to receive messages from every player except for i through the path excluding i; hence, by telling a lie, player i can only create an inconsistency in the messages that player j receives), 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 $\boldsymbol{x}(i) = \boldsymbol{B}$.[37]

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

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 $I^D\left(h_i^{<k}\right)$: At the beginning of each period, suppose $j = \mu_i$. Players (i, j) draw a joint controlled lottery, using the cheap talk. Specifically, each player draws $m_{i,t} \in \{1, ..., M^A\}$ uniformly at random. Define $m_{(i,j),t} = m_{i,t} + m_{j,t} \pmod{M^A}$. Note that, regardless of player i's strategy, $m_{(i,j),t}$ is distributed uniformly at random over $\left\{1, ..., M^A\right\}$. As seen in (28), given $x(i)$, we can see that $m_{(i,j),t}$ uniquely determines the realization of $\boldsymbol{\alpha}_{i,j}^{x(i)}$. Let $a_i^{x(i)}(m_{i,t}, m_{j,t})$ be the realization.

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 if player i has not detectably deviated during the current main round, player i takes $a_i^{x(i)}(m_{i,t}, m_{j,t})$. Otherwise, she takes D.

[37] Note that this can occur even if $m_{-j}(i) \neq \texttt{error}$, as in the situation noted in footnote 33.

Let $h_i^{\leq k}$ denote player i's history at the end of main round k. Note that we can recursively define that player i detectably deviates

1. if $I^D\left(h_i^{<k}\right) = 0$, she has not detectably deviated during the current main round, but she takes $a_i \neq a_i^{x(i)}(m_{i,t}, m_{j,t})$, or
2. if either $I^D\left(h_i^{<k}\right) = 1$ or she has detectably deviated during the current main round, but she takes $a_i \neq D$.

Supplemental round communication as a function of $h_i^{\leq k}$: Players circulate message $m = (m_i)_i = \left(h_i^{\leq k}\right)_i$. For each $j \neq i$, we define $m_{-j}(i)$ as in supplemental round 0.

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. $m_{-j}(i) = \texttt{error}$ for some $j \neq i$ during supplemental round k.
4. For each $j \neq i$, $m_{-j}(i) = \times_{n \neq i} h_n^{\leq k}$ 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 $\left(h_i^{\leq k}, m_{-j}(i)\right)$, player j detectably deviated during main round k.

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

B.2.9 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 $i \notin \theta$, π_i^θ is not defined (see in Lemma 13). For θ satisfying $|\theta| < \alpha N \wedge i \in \theta$, 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 (32)–(34); we verify Condition (31) (sequential rationality) in the next subsection. For the remainder of this section, assume $|\theta| \geq \alpha N \wedge i \in \theta$.

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 11 immediately implies the following result.

Lemma 15 *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*

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

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.

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

$$\begin{cases} \hat{u}_i(\boldsymbol{a}) + \pi_i^{\text{cancel}}\left(x_{i-1}^{\theta}, \boldsymbol{a}\right) = \text{sign}\left(x_{i-1}^{\theta}\right) \frac{1}{2}\bar{u} \\ \text{sign}\left(x_{i-1}^{\theta}\right) \pi_i^{\text{cancel}}\left(x_{i-1}^{\theta}, \boldsymbol{a}\right) \geq 0 \end{cases} \tag{37}$$

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

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

$$\hat{\pi}_i^{\theta}(x_{-i}^{\theta}, h^{T^*+1}) = \sum_{t=1}^{T^*} \pi_i^{\text{cancel}}\left(x_{i-1}^{\theta}, \boldsymbol{a}_t\right). \tag{38}$$

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

$$\hat{\pi}_i^{\theta}(x_{-i}^{\theta}, h^{T^*+1}) = \begin{cases} 1_{\{I_i^D(h^{1^*}, h^{T^*+1})=0\}} \frac{\eta}{16} T^* & \text{if } x_{i-1}^{\theta} = B, \\ -1_{\{I_i^D(h^{1^*}, h^{T^*+1})=1\}} \bar{u} T^* & \text{if } x_{i-1}^{\theta} = G. \end{cases}$$

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. Note that, for each (h^{1^*}, h^{T^*+1}), we have

$$\text{sign}\left(x_{i-1}^{\theta}\right) \hat{\pi}_i^{\theta}(x_{-i}^{\theta}, h^{T^*+1}) \geq 0. \tag{39}$$

Let

$$u_i\left(x^{\theta}, h^{1^*}\right) = \frac{1}{T^*} \mathbb{E}^{\sigma(\boldsymbol{x})}\left[\sum_{\tau=1}^{T^*} \hat{u}_i(\mathbf{a}_\tau) + \hat{\pi}_i^{\theta}(x_{-i}^{\theta}, h^{T^*+1}) | h^{1^*}, \theta\right]. \tag{40}$$

Note that the right hand side of (40) depends only on x^{θ} and h^{1^*} since (i) when $\chi_i(h^{1^*}, h^{T^*+1}) = 0$, the action pair $\left(a_i, a_{\mu_i}\right)$ is drawn from $\boldsymbol{\alpha}_{i,\mu_i}^{x^{\theta}}$ given μ in main rounds, (ii) when $\chi_i(h^{1^*}, h^{T^*+1}) = 1$, (37) implies that player i's payoff from $\hat{u}_i$ and π_i^{cancel} depends only on x_{i-1}^{θ} in main rounds, (iii) the distribution of $\chi_i(h^{1^*}, h^{T^*+1})$ is determined by the match realization, and (iv) in non-main rounds, players take defection for sure.

Note that

$$\begin{aligned}
& \left|u_i\left(x^{\theta}, h^{1^*}\right) - \hat{u}_i\left(\boldsymbol{\alpha}^{x^{\theta}}\right)\right| \\
= \ & \Pr\left(\chi_i(h^{1^*}, h^{T^*+1}) = 0 | \theta\right) \left|\frac{1}{T^*} \mathbb{E}^{\sigma(\boldsymbol{x})}\left[\left(\begin{array}{c} \sum_{\tau=1}^{T^*} \hat{u}_i(\mathbf{a}_\tau) \\ + \hat{\pi}_i^{\theta}(x_{-i}^{\theta}, h^{T^*+1}) \end{array}\right) | h^{1^*}, \theta, \chi_i(h^{1^*}, h^{T^*+1}) = 0\right] - \hat{u}_i\left(\boldsymbol{\alpha}^{x^{\theta}}\right)\right| \\
& + \Pr\left(\chi_i(h^{1^*}, h^{T^*+1}) = 1 | \theta\right) \left|\frac{1}{T^*} \mathbb{E}^{\sigma(\boldsymbol{x})}\left[\left(\begin{array}{c} \sum_{\tau=1}^{T^*} \hat{u}_i(\mathbf{a}_\tau) \\ + \hat{\pi}_i^{\theta}(x_{-i}^{\theta}, h^{T^*+1}) \end{array}\right) | h^{1^*}, \theta, \chi_i(h^{1^*}, h^{T^*+1}) = 1\right] - \hat{u}_i\left(\boldsymbol{\alpha}^{x^{\theta}}\right)\right| \\
\leq \ & \Pr\left(\chi_i(h^{1^*}, h^{T^*+1}) = 0 | \theta\right) \left(\frac{(3+K) T}{T^*} \bar{u} + \frac{\eta}{16}\right) + \Pr\left(\chi_i(h^{1^*}, h^{T^*+1}) = 1 | \theta\right) \left(\bar{u} + \frac{\eta}{16}\right) \\
\leq \ & \frac{\eta}{16} + \left(\frac{3+K}{Z} + \Pr\left(\chi_i(h^{1^*}, h^{T^*+1}) = 1 | \theta\right)\right) \bar{u} \\
\leq \ & \frac{\eta}{16} + (3+K)\left(\frac{1}{Z} + \exp(-cZ)\right) \bar{u} \quad \text{(by Lemma 15)} \\
\leq \ & \frac{\eta}{8} \quad \text{(by (14))}.
\end{aligned} \tag{41}$$

Here the first inequality follows because (i) when $\chi_i(h^{1^*}, h^{T^*+1}) = 0$, the action pair $\left(a_i, a_{\mu_i}\right)$ is drawn from $\boldsymbol{\alpha}^{x^\theta}_{i,\mu_i}$ given μ in main rounds, (i.e., in all but $(3+K)T$ periods), (ii) the magnitude of $\hat{u}_i\left(\mathbf{a}_\tau\right)$ is bounded by $\bar{u}/2$, 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 $\frac{\eta}{16}T^*$.

We now define the reward function

$$\pi_i^\theta(x_{-i}^\theta, h^{1^*}, h^{T^*+1}) = \left(v_i^\theta(x_{i-1}^\theta) - u_i\left(\hat{x}_i^\theta(h^{T^*+1}), x_{-i}^\theta, h^{1^*}\right)\right)T^* + \hat{\pi}_i^\theta(x_{-i}^\theta, h^{T^*+1}). \tag{42}$$

We verify that, with this reward function, Conditions (31)–(34) are satisfied. This will complete the proof. We first establish Conditions (32)–(34), deferring Condition (31) (sequential rationality) to the next subsection.

Since $I_i^D(h^{1^*}, h^{T^*+1}) = 0$ on path, (40) implies that expected per-period block payoffs given $|\theta| \geq \alpha N \wedge i \in \theta$ equal $v_i^\theta(x_{i-1}^\theta)$. Hence, (32) holds.

By (41) and (29), we have

$$\text{sign}\left(x_{i-1}^\theta\right)\left(v_i^\theta(x_{i-1}^\theta) - u_i\left(\hat{x}_i^\theta(h^{T^*+1}), x_{-i}^\theta, h^{1^*}\right)\right) \geq \frac{1}{8}\eta.$$

Together with (39), this implies

$$\text{sign}\left(x_{i-1}^\theta\right)\pi_i^\theta(x_{-i}^\theta, h^{1^*}, h^{T^*+1}) \geq \frac{1}{8}\eta T^*, \tag{43}$$

and hence (33).

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

$$\left|\pi_i^\theta(x_{-i}^\theta, h^{1^*}, h^{T^*+1})\right| \leq \left|v_i^\theta(x_{i-1}^\theta) - u_i\left(\hat{x}_i^\theta(h^{T^*+1}), x_{-i}^\theta, h^{1^*}\right)\right|T^* + \frac{\eta}{16}T^* \leq 2\bar{u}T^*;$$

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

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

Hence, (34) holds.

B.2.10 Verifying Sequential Rationality (Conditions (30) and (31))

Given (38) and $|\theta| \geq \alpha N \wedge i \in \theta$, 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 $\{\chi_i(h^{1^*}, h^{T^*+1}) = 0 \vee |\theta| < \alpha N \vee i \notin \theta\}$. We thus restrict attention to pairs $\left(\boldsymbol{x}_{-i}, \tilde{h}_i^{b,t}\right)$ such that $\Pr\left(\{\chi_i(h^{1^*}, h^{T^*+1}) = 0 \vee |\theta| < \alpha N \vee i \notin \theta\}|\boldsymbol{x}_{-i}, \tilde{h}_i^{b,t}\right) > 0$. Note this implies that (35) and (36) hold conditional on the triple $\left(\{\chi_i(h^{1^*}, h^{T^*+1}) = 0 \vee |\theta| < \alpha N \vee i \notin \theta\}, \boldsymbol{x}_{-i}, \tilde{h}_i^{b,t}\right)$. We consider separately the cases $|\theta| < \alpha N \vee i \notin \theta$ and $\left\{\chi_i(h^{1^*}, h^{T^*+1}) = 0 \wedge |\theta| \geq \alpha N \wedge i \in \theta\right\}$.

Conditional on $|\theta| < \alpha N \vee i \notin \theta$, by (36), player i believes that $\left|\theta(h_n^{T+1})\right| < \alpha N \vee i \notin \theta(h_n^{T+1})$ for each $n \neq i$. Hence, player $i \in \theta$ believes that players $-i$ take D throughout the block regardless of her own strategy. In addition, player $i \in \theta$ believes that $\pi_i^\theta(x_{-i}^\theta, h^{1^*}, h^{T^*+1}) = 0$ regardless of her own strategy (recall that π_i^θ is defined only for $i \in \theta$). It is therefore optimal for rational player i to take D in each period and send any messages, and this behavior is indeed what is prescribed for player i, since (35) implies that $\left|\theta(h_i^{T+1})\right| < \alpha N \vee i \notin \theta(h_i^{T+1})$. For bad player i, her strategy does not affect the action sequence that she faces.

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

$$\begin{aligned}\pi_i^\theta(x_{-i}^\theta, h^{1^*}, h^{T^*+1}) &= \left(v_i^\theta(x_{i-1}^\theta) - u_i\left(\hat{x}_i^\theta(h^{T^*+1}), x_{-i}^\theta, h^{1^*}\right)\right)T^* + \hat{\pi}_i^\theta(x_{-i}^\theta, h^{T^*+1}) \\ &= \left(v_i^\theta(x_{i-1}^\theta) - u_i\left(\hat{x}_i^\theta(h^{T^*+1}), x_{-i}^\theta, h^{1^*}\right)\right)T^* + \left\{\begin{array}{cl} 0 & \text{if } x_{i-1}^\theta = B, \\ -\bar{u}T^* & \text{if } x_{i-1}^\theta = G.\end{array}\right.\end{aligned}$$

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 $\hat{\boldsymbol{\alpha}}^{x^\theta}$ 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 $ZT \times \max\{G, L\}$. 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 $\frac{\eta}{16}T^*$; this comes from the $\hat{\pi}_i^\theta\left(x_{-i}^\theta, h^{T^*+1}\right)$ term in the reward function. By (16), $\frac{\eta}{16}T^* \geq ZT \times \max\{G, L\}$, 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 $\bar{u}T^* \geq ZT \times \max\{G, L\}$. 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 $(\sigma_i(\boldsymbol{x}_i))_{\boldsymbol{x}_i}$. By (40) and (42), player i's expected payoff conditional on $|\theta| \geq \alpha N \wedge i \in \theta$ equals

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

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 $\left\{\chi_i(h^{1^*}, h^{T^*+1}) = 0 \wedge |\theta| \geq \alpha N \wedge i \in \theta\right\}$ also depends on $\boldsymbol{x}$ only through x_{i-1}^θ. This completes the proof of Theorem 3.

B.3 De-Coupling Interaction and Community Frequency in Theorem 3

We explain how the proof of Theorem 3 must be modified when interaction and communication meetings are de-coupled.

In the proof of Theorem 3 (for our main model with coupled interaction and communication), the length of each communication round (i.e., the 1^*-communication sub-block, the $\boldsymbol{x}$-communication sub-block, and the supplemental rounds) must be much greater than $\log N$, so that messages spread throughout the population with high probability, while the number of communication rounds is independent on N. The proof requires that discounting in each communication round is negligible, which holds if $(1-\delta)\log N \to 0$. In addition, since players take D during communication rounds, the length of the main rounds must also increase with N, so that the vast majority of periods occur in main rounds rather than communication rounds (and thus players' overall repeated game payoffs are determined by their payoffs in the main rounds). The proof also requires that discounting in each main round is negligible; since the length of a main round can be taken to be of the same order as the length of a communication round, this also holds if $(1-\delta)\log N \to 0$.

If interaction and communication are de-coupled, so that players communicate every Δ_M periods and interact every Δ_A periods (with a constant real-time discount rate), the length (i.e., the number of meetings) in each communication round must again be much greater than $\log N$, and discounting in each communication round must be negligible. This now holds if $\Delta_M \log N \to 0$. Moreover, since now players do not accrue payoffs during communication rounds, the length of the main rounds may be fixed independently of N. This proof again requires that discounting in each main round is negligible, but since the length of the main rounds is now independent of N, this requires only that $\Delta_A \to 0$.

Given these observations, the proof of the de-coupled version of Theorem 3 under the assumptions that $\Delta_M \log N \to 0$ and $\Delta_A \to 0$ is very similar to the proof for the coupled version, but is somewhat simpler because we no longer need to account for the payoffs that players accrue during the communication rounds.

(One subtlety is that the proof of Theorem 3 relies on cheap talk within the main rounds to coordinate matched partners on a correlated action distribution $\boldsymbol{\alpha}_{i,j}^{x(\theta)}$. This main round communication can be dispensed with at the cost of complicating the messages players exchange during the communication rounds. It can also be dispensed with without changing the proof if the target payoff v_i^θ lies in $[0,1]$ for each $i \in \theta$ and $|\theta| \geq \alpha N$, because there then exists a degenerate distribution $\boldsymbol{\alpha}^{x^\theta}$ that satisfies the relevant condition, inequality (29).)