

Community Enforcement beyond the Prisoner's Dilemma*

Joyee Deb[†]

New York University

Julio González-Díaz[‡]

University of Santiago de Compostela

PRELIMINARY (April 2010)

Abstract

We study two-player games played by two communities in an infinitely repeated anonymous random matching setting. It is well-known that in this setting, for the prisoner's dilemma (PD), cooperation can be sustained in equilibrium through grim trigger strategies also called “contagion” or “community enforcement”. However, contagion does not work in general. Little is known about cooperation in games beyond the PD in this setting with minimal information transmission. In this paper we show that cooperation can indeed be sustained in the anonymous random matching setting, in a broad class of games beyond the PD.

***Acknowledgements.** The authors would like to thank Johannes Hörner for many insightful discussions and comments. We also thank seminar participants at GAMES 2008, Northwestern University, New York University and the Repeated Games workshop at SUNY StonyBrook for many suggestions. The second author gratefully acknowledges the support from the Sixth Framework Programme of the European Commission through a Marie Curie fellowship, the Ministerio de Ciencia e Innovación through a Ramón y Cajal fellowship and through project ECO2008-03484-C02-02. Support from Xunta de Galicia through project INCITE09-207-064-PR is also acknowledged.

[†]Email: joyee.deb@nyu.edu

[‡]Email: julio.gonzalez@usc.es

1 Introduction

We study infinitely repeated matching games where, in every period, players from two communities are randomly and anonymously matched to each other to play a two-player game. A common interpretation of such a setting is a large market where people are matched with strangers to engage in bilateral trades in which they may act in good faith or cheat. An interesting question is whether players can achieve cooperative outcomes in anonymous transactions. What payoffs of the stage-game can be achieved in equilibrium in the repeated random matching game?

The seminal papers by Kandori (1992) and Ellison (1994) showed that in this setting, for the Prisoner's Dilemma (PD), cooperation can be sustained by grim trigger strategies, also known as "community enforcement" or "contagion". In the PD, if a player ever faces a defection, she punishes all future rivals by switching to defection forever. By starting to defect, she spreads the information that someone has defected. The defection action spreads throughout the population, and cooperation eventually breaks down completely. The credible threat of such a breakdown of cooperation can deter players from defecting in the first place. However, these arguments rely critically on properties of the PD, in particular on the existence of a Nash equilibrium in strictly dominant strategies. The argument does not work in general. In an arbitrary game, on facing a deviation for the first time, players may not have the incentive to punish, because punishing can both lower future continuation payoffs and entail a short-term loss in that period. In the PD, the punishment action is dominant and so gives a current gain even if it lowers continuation payoffs.

An open question is whether cooperation can be sustained in this setting for games other than the PD, with minimal transmission of information, and this is the central question of this paper. We show that it is indeed possible to sustain a large range of payoffs in equilibrium in a wide range of games beyond the PD, provided the communities are large enough and all players are sufficiently patient. Cooperation can be sustained in any two-player game with a strict Nash equilibrium and where there is one action profile in which only one player has an incentive to deviate. In particular, we show that the idea of community enforcement can still be used to sustain cooperation in equilibrium.

To the best of our knowledge, this is the first paper to sustain cooperation in a random matching game beyond the PD, without adding any extra informational assumptions. Some papers that go beyond the PD introduce verifiable information about past play to sustain cooperation. For instance, Kandori (1992) assumes the existence of a mechanism that

assigns labels to players based on their history of play. Players who have deviated or have seen a deviation can be distinguished from those who have not, by their labels. This naturally enables transmission of information and cooperation can be sustained in a specific class of games.¹ More recently, Deb (2008) obtains a general folk theorem for any game by just adding unverifiable information (cheap talk).

Another important feature of our construction is that our strategies are fairly simple. Unlike recent work on repeated games with imperfect private monitoring (Ely and Välimäki, 2002; Piccione, 2002; Ely et al., 2005; Hörner and Olszewski, 2006) and, more specifically, in repeated random matching games (Takahashi, 2007; Deb, 2008), our equilibrium does not rely on belief-free ideas. In particular, players have strict incentives on and off the equilibrium path. Further, unlike most of the existing literature, our strategies are robust to changes in the discount factor.

Our methodological contribution lies in that we work explicitly with players' beliefs. We hope that the methods we use to study the evolution of beliefs will be of independent interest, and can be applied in other contexts.

The paper is organized as follows. Section 2 contains the model and the main result. In Section 3, we illustrate the result with a particular game, the product choice game. This section is useful in understanding the strategies and the intuition behind the result. Section 4 contains the formal equilibrium construction and proof. In section 5, we discuss the generality of our result and potential extensions.

2 Cooperation Beyond the PD

2.1 The Setting

There are $2M$ players, divided in two communities with M players each. In each period $t \in \{1, 2, \dots\}$, players are randomly matched into pairs with each player from Community 1 facing a player from Community 2. The matching is anonymous, independent and uniform over time.² After being matched, each pair of players plays a finite stage-game G . Players observe the actions and payoffs in their match. Then a new matching occurs in the next period.

¹For related approaches see Okuno-Fujiwara and Postlewaite (1995), Takahashi (2007), Dal Bó (2007), and Hasker (2007).

²Although the assumption of uniform matching greatly simplifies the calculations, we expect our results to hold for other matching technologies, in particular those close to the uniform one.

Players can observe only the transactions they are personally engaged in, *i.e.*, each player knows the history of action profiles played in each of his stage-games in the past. A player does not ever observe the identity of his opponent. Further, he gets no information about how other players have been matched or about the actions chosen by any other pair of players. All players have discount factor $\delta \in (0, 1)$ and their payoffs in the infinitely repeated random matching game are the normalized sum of the discounted payoffs from the stage-games. No public randomization device is assumed (Section 5 has a discussion of what can be gained with a randomization device). The solution concept used is sequential equilibrium (allowing for both pure and mixed strategies).

2.2 The Main Result

Our main objective is to establish that, in the anonymous random matching setting, it is feasible to sustain a wide range of equilibrium payoffs in a general class of games.

We define first, a particular class of games.

- **Class of Games:** We define \mathcal{G} , a class of finite two-player games with two properties:

P1. One-sided Incentives There exists a pure action profile (\hat{a}_1, \hat{a}_2) in which only one player has an incentive to deviate.

P2. Strict Nash equilibrium There exists a strict Nash equilibrium a^* .³

Given a game in $G \in \mathcal{G}$, the associated repeated random matching game, with communities each of size M and discount factor δ is denoted by G_δ^M .

We also define a set of payoffs for any stage-game.

- **Achievable Payoffs** Let G be a game and let $a^* \in A$. Let $A_{a^*} := \{a \in A : a_1 = a_1^* \iff a_2 = a_2^*\}$. Define $F_{a^*} := \text{conv}\{u(a) : a \in A_{a^*}\} \cap \{v \in \mathbb{R}^2 : v > u(a^*)\}$.

The main result of this paper says that given a game G in the class of games \mathcal{G} defined above, with strict Nash equilibrium a^* , it is indeed possible for players to achieve any payoff in F_{a^*} in equilibrium in the corresponding infinitely repeated random matching game G_δ^M , if the communities are large enough and players are sufficiently patient.

³A Nash equilibrium is strict if each player has a unique best response to his rival's strategies (Harsanyi (1973)). By definition, a strict equilibrium is a pure strategy equilibrium.

Theorem 1. *Let G be a game in \mathcal{G} with a strict Nash equilibrium a^* . There exists $\underline{M} \in \mathbb{N}$ such that, given any payoff profile $v \in F_{a^*}$, any $\varepsilon > 0$, and any $M \geq \underline{M}$, there exists $\underline{\delta} \in (0, 1)$ such that, there is a strategy profile in the repeated random matching game G_δ^M that constitutes a sequential equilibrium for each $\delta \in [\underline{\delta}, 1)$ with payoff within ε of v .*

Earlier literature (Kandori (1992), Ellison (1994)) establishes feasibility results for prisoner's dilemma games. The class of games \mathcal{G} is a bigger class of games, and in particular includes the PD. To get a better understanding, it will be useful to point out the games that are not included in it. \mathcal{G} excludes what we call aligned interests games.

Definition 1. A two-player game G is a **game of aligned interests** if at any action profile, a player has an incentive to deviate if and only if his opponent also has an incentive to deviate. e.g. Battle of the Sexes and Chicken are aligned interests games.

In fact, games in \mathcal{G} are generic in the class of non-aligned interests games with a pure strategy Nash equilibrium.

Note that the set of achievable payoffs includes payoffs arbitrarily close to efficiency for the prisoner's dilemma. In general, our strategies do not suffice to get a folk theorem for games in \mathcal{G} . However, we conjecture that it may be possible to obtain a Nash threats folk theorem for games in \mathcal{G} with some modifications to our strategies. The interested reader may refer to Section 5.1 for a discussion.

A noteworthy feature of our equilibrium is that the strategies are robust to changes in the discount factor. In other words, if our strategies constitute an equilibrium for a given discount factor, they do so for any higher discount factor as well. This is in contrast with existing literature. In games with private monitoring, strategies have to be fine-tuned based on the discount factor. In Ellison (1994), the severity of punishments depends on the discount factor. Moreover, unlike Ellison (1994), we do not need a common discount factor. We just need all players to be sufficiently patient.

Another feature of our equilibrium is that the continuation payoff is within ε of the target equilibrium payoff v , not just in the initial period, but throughout the game on the equilibrium path; in this sense, cooperation is sustained as a durable phenomenon, which contrasts with the results for reputation models where, for every δ , there exists a time after which cooperation collapses (see Cripps et al. (2004)).

It is worthwhile to explain the role of the community sizes in Theorem 1. Contrary to intuition, having a large population is helpful in our construction. However, the result

should not be viewed as a limiting result in M ; it turns out that in most games, fairly small community sizes suffice to sustain cooperation.

How is cooperation sustained? Below, we describe the equilibrium strategies.

2.3 Equilibrium Strategies

Consider any two-player stage-game $G \in \mathcal{G}$, with strict Nash equilibrium a^* . Let (\hat{a}_1, \hat{a}_2) denote a pure action profile in which only one player has an incentive to deviate. Without loss of generality, suppose only player 1 wants to deviate at (\hat{a}_1, \hat{a}_2) . In particular, let \bar{a}_1 be player 1's most profitable deviation from (\hat{a}_1, \hat{a}_2) . Let the target equilibrium payoff be $v \in F_{a^*}$. In what follows, we maintain the convention that players 1 and 2 of the stage-game belong to Community 1 and 2 respectively. Also, henceforth, when we refer to the Nash action, we mean the strict Nash equilibrium action profile a^* . We now construct the strategies that sustain v in equilibrium. As in the product choice game example, we divide the game into three phases (Figure 2). Phases I and II are *trust-building* phases and Phase III is the *target payoff* phase.

Equilibrium play: Phase I: During the first \bar{T} periods, action profile (\hat{a}_1, \hat{a}_2) is played. In every period in this phase, players from Community 1 have a short-run incentive to deviate, but those from Community 2 do not. **Phase II:** During the next \bar{T} periods, players play (a_1^*, a_2) , an action profile where players from Community 1 play their Nash action and players from Community 2 do not play their best response⁴. In every period in this phase, players from Community 2 have a short-run incentive to deviate. **Phase III:** For the rest of the game, the players play a sequence of pure action profiles that approximates target payoff v .

Off Equilibrium play: A player can be in one of two moods: *uninfected* and *infected*, with the latter mood being irreversible. At the start of the game all players are uninfected. We classify off equilibrium play into two types of actions. Any deviation by player 2 in Phase I is called a *non-triggering* action. Any other action that is off equilibrium is a *triggering* action. A player who has observed a triggering action is in the infected mood. Now, we specify off-path behavior. An uninfected player continues to play as if on-path (i.e. if a player from Community 2 observes a deviation

⁴Player 2's action a_2 can be any action other than a_2^* in the stage-game.

in Phase I, he ignores the deviation, and continues to play as if on-path.). An infected player acts as follows.

- A player who gets infected after facing a triggering action switches to his Nash action forever either from the beginning of Phase II or immediately from the next period, whichever is later. In other words, any player in Community 2 who faces a triggering action in Phase I switches to her Nash action forever from the start of Phase II, playing as if on-path in the meantime. A player i facing a triggering action at any other stage of the game immediately switches to the Nash action forever.
- A player who gets infected by playing a triggering action himself henceforth best responds to the strategies of the other players (this implies in particular, that for large enough \hat{T} , a player from Community 1 who deviates in the first period will continue to deviate until the end of Phase I, and then switch to playing the Nash action forever).

We show in this paper that the above strategies constitute a sequential equilibrium. The general proof of Theorem 1 follows in Section 4. First we present an example in some detail in the next section. We use the product choice game as an expository tool that will illustrate clearly how the equilibrium works. We hope this section will provide the reader with the main intuition behind the proof.

3 An Illustration: The Product Choice Game

Consider the following simultaneous move game between a buyer and a seller.

		Buyer	
		B_H	B_L
Seller	Q_H	1, 1	$-l, 1 - c$
	Q_L	$1 + g, -l$	0, 0

Figure 1: The product-choice game.

Suppose this game is played by a community of M buyers and a community of M sellers in the repeated anonymous random matching setting. In each period, every seller is

randomly matched with a buyer. After being matched, each pair plays the product-choice game above (Figure 1), where $g > 0$, $c > 0$, and $l > 0$.⁵ The seller can exert either high effort (Q_H) or low effort (Q_L) in the production of his output. The buyer, without observing the choice of the seller, can either buy a high-priced product (B_H) or a low-price product (B_L). The buyer prefers the high-priced product if the seller has exerted high effort and prefers the low-priced product if the seller has not. For the seller, exerting low effort is a dominant action. The efficient outcome of this game is the seller exerting high effort and the buyer buying the high-priced product, while the Nash equilibrium is (Q_L, B_L) . We denote a product-choice game by $\Gamma(g, l, c)$. The infinitely repeated random matching game associated with the product-choice game $\Gamma(g, l, c)$, with discount parameter δ and communities of size M is denoted by $\Gamma_\delta^M(g, l, c)$. We will show that it is indeed possible for players to achieve a payoff arbitrarily close to $(1, 1)$ in equilibrium.

Notice that the product choice game belongs to the class of games \mathcal{G} defined in Section 2. It represents a minimal departure from the PD. If we replace the payoff $1 - c$ with $1 + g$ we get the standard PD. However, even with this small departure from the PD, cooperation can not be sustained in equilibrium using the standard grim trigger strategies (also called “contagion” or “community enforcement”). The main difficulty in sustaining cooperation through community enforcement is that it is hard to provide buyers with the incentive to punish deviations. Proposition 1 shows that a straightforward adaptation of the contagion strategies as in Ellison (1994) to support cooperation in the PD does not work.

3.1 A negative result

Proposition 1. *Let $\Gamma(g, l, c)$ be a product-choice game with $c \leq 1$. Then, there is $\underline{M} \in \mathbb{N}$ such that, for each $M \geq \underline{M}$, regardless of the discount factor δ , the repeated random matching game $\Gamma_\delta^M(g, l, c)$ has no sequential equilibrium in which (Q_H, B_H) is played in every period on the equilibrium path, and in which players play the Nash action off the equilibrium path.*

Proof. Suppose there exists an equilibrium in which (Q_H, B_H) is played in every period on the equilibrium path, and in which players play the Nash action off the equilibrium path. Suppose a seller deviates in period 1. We argue below that for a buyer who observes this deviation, it will not be optimal to switch to the Nash action permanently from period 2. In

⁵A more detailed discussion of this game within the context of repeated games can be found in Mailath and Samuelson (2006).

particular, we show that playing B_H in period 2 followed by switching to B_L from period 3 onwards gives the buyer a higher payoff. The buyer who observes the deviation knows that, in period 2, with probability $\frac{M-1}{M}$ she will face a different seller who will play Q_H . Consider the short-run and long-run incentives of this buyer:

Short-run: The buyer's payoff in period 2 from playing B_H is $\frac{1}{M}(-l) + \frac{M-1}{M}$. Her payoff if she switches to B_L is $\frac{M-1}{M}(1-c)$. Hence, if M is large enough, she has no short-run incentive to switch to the Nash action.

Long-run: With probability $\frac{1}{M}$, the buyer will meet the deviant seller (who is already playing Q_L) in period 2. In this case, her action will not affect this seller's future behavior, and so her continuation payoff will be the same regardless of her action.

With probability $\frac{M-1}{M}$, the buyer will meet a different seller. Note that, since $1 - c \geq 0$, a buyer always prefers to face a seller playing Q_H . So, regardless of the buyer's strategy, the larger the number of sellers who have already switched to Q_L , the lower is her continuation payoff. Hence, playing B_L in period 2 will give her a lower continuation payoff than playing B_H , because action B_L will make a new seller switch permanently to Q_L . \square

Since there is no short-run or long-run incentive to switch to the Nash action in period 2, the buyer will not start punishing. Therefore, playing (Q_H, B_H) in every period on path, and playing the Nash action off path does not constitute a sequential equilibrium.

While standard community enforcement does not work, we show that it is still possible to sustain cooperation in the product choice game, using the strategies described in Section 2.3. We divide the game into three phases (see Figure 2). Phases I and II are *trust-building* phases and Phase III is the *target payoff* phase.

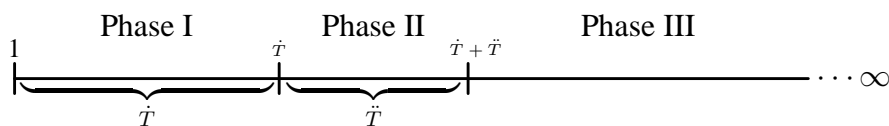


Figure 2: Different phases of the strategy profiles.

Equilibrium play: **Phase I:** During the first \bar{T} periods, the players play (Q_H, B_H) . **Phase II:** During the next \bar{T} periods, the players play (Q_L, B_H) . **Phase III:** For the rest of the game, the players play the efficient action profile (Q_H, B_H) .

Off Equilibrium play: A player can be in one of two moods: *uninfected* and *infected*, with the latter mood being irreversible. At the start of the game all players are uninfected. Action B_L in Phase I is a *non-triggering* action. Any other action that is off equilibrium is a *triggering* action. A player who has observed a triggering action is in the infected mood. An uninfected player continues to play as if on-path. An infected player acts as follows.

A player who gets infected after facing a triggering action switches to his Nash action forever from the beginning of Phase II or immediately from the next period, whichever is later. A player facing a triggering action at any other stage of the game immediately switches to the Nash action forever.

A player who gets infected by playing a triggering action himself henceforth best responds to the strategies of the other players.

Note that a profitable deviation by a player is punished (ultimately) by the whole community of players, with the punishment action spreading like an epidemic. We refer to the spread of punishments as contagion.

The difference between our strategies and standard contagion (*e.g.*, Ellison (1994) and Kandori (1992)) is that here, the game starts with two trust-building phases. In Phase I, sellers build credibility by not deviating even though they have a short-run incentive to do so. The situation is reversed in Phase II, where buyers build credibility (and reward sellers for not deviating in Phase I), by not playing B_L even though they have a short-run incentive to do so. A deviation by a seller in Phase I is not punished in the seller's trust-building phase, but is punished as soon as the phase is over. Similarly, if a buyer deviates in her trust-building phase, she effectively faces punishment once the trust-building phase is over. Unlike the results for the PD, where the equilibria are based on trigger strategies, we have delayed trigger strategies. In Phase III, deviations immediately trigger Nash reversion.

Clearly, the payoff from the strategy profile above will be arbitrarily close to the efficient payoff $(1, 1)$ for δ large enough. We need to establish that the strategy profile constitutes a sequential equilibrium of the repeated random matching game $\Gamma_\delta^M(g, l, c)$ when M is large enough, \hat{T} and \bar{T} are appropriately chosen, and δ is close enough to 1.

3.2 Optimality of Strategies: Intuition

The incentives on-path are quite straightforward. Any short-run profitable deviation will eventually trigger Nash reversion that will spread and reduce continuation payoffs. Hence, given M , \dot{T} , and \ddot{T} , for sufficiently patient players, the future loss in continuation payoff will outweigh any current gain from deviation.

Establishing sequential rationality of the strategies off-path is the challenge. Below, we consider some key histories that may arise and argue why the strategies are optimal after these histories. We start with two observations.

First, a seller who deviates to make a short-term gain at the beginning of the game will find it optimal to revert to the Nash action immediately. A seller who deviates in period 1 knows that, regardless of his choice of actions, from period \dot{T} on, at least one buyer will start playing Nash and then, from period $\dot{T} + \ddot{T}$ on, contagion will spread exponentially fast. Thus, his continuation payoff after $\dot{T} + \ddot{T}$ will be quite low regardless of what he does in the remainder of Phase I. So, if \dot{T} is large enough, no matter how patient this seller is, the best thing he can do after deviating in period 1 is to play the Nash action forever.⁶

Second, the optimal action of a player after he observes a triggering action depends on the beliefs that he has about how the contagion has spread already. To see why, think of a buyer who observes a triggering action during, say, Phase III. Is Nash reversion optimal for her? If she believes that there are few people infected, then playing the Nash action may not be optimal. With high probability she will face a seller playing Q_H and playing the Nash action will entail a loss in that period. Moreover, she is likely to infect her opponent, hastening the contagion and lowering her own continuation payoff. The situation is different if she believes that almost everybody is infected (so, playing Nash). Then, there is a short-run gain by playing the Nash action in this period. Moreover, the effect on the contagion process and the continuation payoff will be negligible. Since the optimal action for a player after observing a triggering action depends on the beliefs he has about “how spread the contagion is”, we need to define a system of beliefs and check if Nash reversion is optimal after getting infected, given these beliefs.

We define beliefs as follows. If a player observes a triggering action, he thinks that some seller deviated in period 1 and contagion has been spreading since then (if an uninfected player observes a non-triggering action, then he just thinks that the opponent made a mistake and that no one is infected).

⁶Clearly, the best thing he can do in Phase II is to play the Nash action, as he was supposed to do on-path.

These beliefs, along with the fact that a deviant seller will play the Nash action forever, imply that any player who observes a triggering action thinks that, since contagion has been spreading from the start of the game, almost everybody must have got infected by the end of Phase I. This makes Nash reversion optimal for him after the end of Phase I. To gain some insight, consider the following histories.

- **Suppose I am a buyer who gets infected in Phase I.** I think that a seller deviated in the first period and that he will continue infecting buyers throughout Phase I. If M is large, in each of the remaining periods of Phase I, the probability of meeting the same seller again is low; so I prefer to play B_H during Phase I (since all other sellers are playing Q_H). Yet, if \dot{T} is large enough, once Phase I is over I will think that, with high probability, every buyer is now infected. Nash reversion thereafter is optimal.

It may happen that after I get infected I observe Q_L in most (possibly all) periods of Phase I. Then, I think that I met the deviant seller repeatedly, and so not all buyers are infected. However, it turns out that if \ddot{T} is large enough I will still revert to Nash play. Since I expect my continuation payoff to drop after $\dot{T} + \ddot{T}$ anyway, for \ddot{T} large enough, I prefer to play the myopic best reply in Phase II, to make some short-term gains (similar to the argument for a seller's best reply if he deviates in period 1).

- **Suppose I am a seller who faces B_L in Phase I.** (Non-triggering actions) Since such actions are never profitable (on-path or off-path), after observing such an action I will think it was a mistake and that no one is infected. Then, it is optimal to ignore it. The deviating player knows this, and so it is also optimal for him to ignore it.
- **Suppose I am a player who gets infected in Phase II, or shortly after period $\dot{T} + \ddot{T}$.** I know that contagion has been spreading since the first period. But, the fact that I was uninfected so far may indicate to me that possibly not so many people were infected. We show that if \dot{T} is large enough and $\dot{T} \gg \ddot{T}$, I will still think that, with high probability, I was just lucky not to have been infected so far, but that everybody is infected now. This makes Nash reversion optimal.
- **Suppose I get infected late in the game, at period $\bar{t} \gg \dot{T} + \ddot{T}$.** If $\bar{t} \gg \dot{T} + \ddot{T}$, it is not possible to rely any more on how large \dot{T} is to characterize my beliefs. However, for this and other related histories late in the game, it turns out that I will still believe that “enough” people are infected and already playing the Nash action, so that playing the Nash action is also optimal for me.

3.3 Choosing \underline{M} , \ddot{T} , \dot{T} , and $\underline{\delta}$

It is useful to clarify how the different parameters are chosen to construct the equilibrium. First, given a game, we find \underline{M} so that i) a buyer who is infected in Phase I does not revert to the Nash action before Phase II and ii) players who are infected very late in the game believe that enough people are infected for Nash reversion to be optimal. Then, we choose \ddot{T} so that, in Phase II, any infected buyer will find it optimal to revert the Nash action (even if she observed Q_L in all periods of Phase I). Then, we pick \dot{T} , with $\dot{T} \gg \ddot{T}$, so that the players infected in Phase II or early in Phase III believe that almost everybody is infected. Further \dot{T} must be large enough so that a seller who deviates in period 1 plays the Nash action ever after. Finally, we pick $\underline{\delta}$ large enough so that players do not deviate on the equilibrium path.⁷

The role of the discount factor δ requires further explanation. Clearly, a high δ deters players from deviating from cooperation. However, a high δ also makes players want to slow down the contagion. Then, why is it that very patient players are willing to spread the contagion after getting infected? A key observation is the following. Fix \underline{M} and consider a perfectly patient player ($\delta = 1$). Once this player gets infected he knows that, at some point, the contagion will start spreading exponentially and the expected payoffs in future periods will converge to 0 exponentially fast. There is an upper bound on the undiscounted sum of future gains any player can make by slowing down the contagion once it has started. Think of an infected player who is deciding whether to revert to the Nash action or not when the strategy asks to do so. In our construction, two things can happen. First, this player believes that so many people are already infected that, regardless of his action, his continuation payoff is already guaranteed to be very low. In this case, he is willing to play the Nash action and at least avoid a short-run loss. Second, this player does not believe that many people are infected. But he still knows that in Phase III his continuation payoff will drop exponentially fast and, in our construction, we ensure that there are enough periods in the immediate future when playing the Nash action will give him a short-run gain. In this case, he is again willing to play the Nash, as the immediate short-run gain outweighs the future loss in continuation payoff.

In the next section, we formalize the intuition above, and present the formal proof for a general game $G \in \mathcal{G}$. Some readers may prefer to skip the proof and go to Section 5, where we discuss the generality and robustness of our main result.

⁷Note that $\underline{\delta}$ must also be large enough so that the payoff achieved in equilibrium is close enough to $(1, 1)$.

4 Proof of the Main Result

We will show below that the strategy profile described in Section 2.3 constitutes a sequential equilibrium of the repeated random matching game G_δ^M when M is large enough, \dot{T} and \ddot{T} are appropriately chosen, and δ is close enough to 1.

Notice that if we compare the general game to our example of the product choice game, a player in Community 1 in a general game in \mathcal{G} is like the seller who is the only player with an incentive to deviate in Phase I, and a player in Community 2 is like the buyer who has no incentive to deviate in Phase I but has an incentive to deviate in Phase II. For the sake of clarity, for the rest of this section, we often refer to players from Community 1 as sellers and players from Community 2 as buyers.

4.1 Incentives on-path

The incentives on-path are straightforward, and so we omit the formal proof. The logic is that if players are sufficiently patient, on-path deviations can be deterred by the threat of eventual Nash reversion. It is easy to see that, if \dot{T} is large enough, the most “profitable” on-path deviation is that of a player in Community 1 (a seller) in period 1. Given M , \dot{T} and \ddot{T} , the discount factor δ can be chosen close enough to 1 to deter such deviations.

4.2 System of beliefs

We make the following assumptions on the system of beliefs of players. Beliefs are updated as usual using Bayes rule.

- i) **Assumption 1:** If a player observes a triggering action, then this player believes that some player in Community 1 (a seller) deviated in the first period of the game, and after that, play has proceeded as prescribed by the strategies.

This requirement on beliefs may seem extreme. However, the essential assumption is that players regard earlier deviations as more likely. Please refer to Section 5.3 for a detailed discussion of this assumption.

- ii) **Assumption 2:** If a player observes a non-triggering action, then this player believes that his opponent made a mistake.

iii) **Assumption 3:** If a player observes a history that is not consistent with either of the above beliefs (*erroneous history*), he will think that a seller deviated in the first period of the game, and further, one or more of his opponents in his earlier matches made a mistake⁸. Indeed, this player will think that there have been as many mistakes by his past rivals as needed to explain the history at hand.⁹ Erroneous histories include the following:

- A player observes an action in Phase II or III, that is neither part of the strict Nash profile a^* , nor the action that is supposed to be played in equilibrium.
- A player who, after being certain that all the players in the other community are infected, faces an opponent who does not play the Nash action (this can only happen in Phase III).

We refer the reader to the Appendix for the proof of the consistency of these beliefs.

4.2.1 Modeling Beliefs with Contagion Matrices

So far, we have not formally described the structure of a player's beliefs. The payoff relevant feature of a player's beliefs is the number of people he believes is currently infected. Accordingly, we let a vector $x^t \in \mathbb{R}^M$ denote the beliefs of player i about the number of infected people in the other community at the end of period t , where x_k^t denotes the probability he assigns to exactly k people being infected in the other community. To illustrate, when player i observes the first triggering action, Assumption 1 implies $x^1 = (1, 0, \dots, 0)$. In some abuse of notation, when it is known that a player assigns 0 probability to more than k opponents being infected, we work with $x^t \in \mathbb{R}^k$. We say a belief $x^t \in \mathbb{R}^k$ *first-order stochastically dominates* a belief y^t if x^t assigns higher probability to more people being infected; *i.e.*, for each $l \in \{1, \dots, k\}$, $\sum_{i=l}^k x_i^t \geq \sum_{i=l}^k y_i^t$. Let \mathcal{I}^t denote the random variable representing the number of infected people in the other community at the end of period t . Let k^t denote the event “ k people in the other community are infected by the end of period t ”, *i.e.*, k^t and $\mathcal{I}^t = k$ denote the same event.

As we will see below, the beliefs of players after different histories evolve according to simple Markov processes, and so can be studied using an appropriate transition matrix

⁸This assumption on beliefs is not crucial. We use it to ensure that beliefs are specified fully (after all possible histories).

⁹For the formation of a player's beliefs after erroneous histories, we assume that infected players are infinitely more likely to make mistakes than uninfected players. Hence, if a player observes an erroneous history, he will think that a seller deviated in period 1 and moreover other players have made mistakes.

and an initial belief. We define below a useful class of matrices: *contagion matrices*. Given a population size M , a contagion matrix C is an $M \times M$ matrix that represents the transitions between beliefs after a given history. The element c_{ij} of a contagion matrix C denotes the probability that the state “ i rivals infected” transitions to the state “ j rivals infected”. Formally, if we let \mathcal{M}_k denote the set of $k \times k$ matrices with real entries, we say that a matrix $C \in \mathcal{M}_k$ is a contagion matrix if it has the following properties:

- i) All the entries of C belong to $[0, 1]$ (represent probabilities).
- ii) C is upper triangular (being infected is irreversible).
- iii) All diagonal entries are strictly positive (with some probability, infected people meet other infected people and contagion does not spread in the current period).
- iv) For each $i > 1$, $c_{i-1,i}$ is strictly positive (With some probability, exactly one new person gets infected in the current period, unless everybody is already infected.).

A useful technical property is that, since contagion matrices are upper triangular, their eigenvalues correspond to the diagonal entries. Given $x \in \mathbb{R}^k$, let $\|x\| := \sum_{i \in \{1, \dots, k\}} x_i$. We will often be interested in the limit behavior of $\frac{x C^t}{\|x C^t\|}$, where C is a contagion matrix and x is a probability vector. Given a matrix C , let $C_{\downarrow l}$ denote the matrix derived by removing the last l rows and columns from C . Similarly, $C_{\uparrow k}$ is the matrix derived by removing the first k rows and columns and $C_{\uparrow k, \downarrow l}$ by doing both operations simultaneously.

4.3 Incentives off-path

To prove sequential rationality we need to examine incentives of players after all possible off-path histories, given the beliefs. This is the heart of the proof and the exposition proceeds as follows. We classify all possible off-path histories of a player i based on when player i observed off-path behavior for the first time.

- Given the beliefs described above, it will be important to first characterize the best response of a seller who deviates in the first period of the game.
- Next, we consider histories where player i observes a triggering action (gets infected) for the first time in the Target Play Phase (Phase III).

- We then consider histories where player i observes a triggering action for the first time during one of the two Trust-Building Phases.
- Finally, we discuss non-triggering actions.

We need some extra notation. Denote a t -period private history for a player i by h^t . At any time period, we denote by g (good behavior) the action an uninfected player would choose. Let b (bad behavior) denote any other action. In other words, b could be an action an infected player would choose, or any triggering action. For example, if player i observes a t -period history h^t followed by three periods of good behavior and then one period of bad behavior, we represent this by $h^t g g g b$. For any player i , let g_t denote the event “player i observed g in period t ”, and let U^t denote the event “player i is uninfected at the end of period t ”. In an abuse of notation, the history of a player is written omitting his own actions. In most of the paper, we discuss beliefs from the point of view of a fixed player i , and so often refer to player i in the first person.

4.3.1 Computing off-path beliefs

Since we work with off-path beliefs of players, it is useful to clarify at the outset, our approach to computing beliefs. As an example, consider the following history. I am on-path until period $\bar{t} \gg \bar{T} + \bar{T}$, when I observe a triggering action followed by on-path behavior at period $\bar{t} + 1$, *i.e.*, $h^{\bar{t}+1} = gg \dots gbg$. It is easy to see that, after period $\bar{T} + \bar{T} + 2$, the number of infected people will be the same in both communities. So it suffices to compute beliefs about the number of people infected in the rival community. These beliefs are represented by $x^{\bar{t}+1} \in \mathbb{R}^M$, where $x_k^{\bar{t}+1}$ is the probability of exactly k people being infected after period $\bar{t} + 1$, and must be computed using Baye’s rule and conditioning on my private history. What is the information I have after history $h^{\bar{t}+1}$? I know a seller deviated at period 1, so $x^1 = (1, 0, \dots, 0)$. I also know that, after any period $t < \bar{t}$, I was not infected (U^t). Moreover, since I got infected at period \bar{t} , at least one player in the rival community got infected in the same period. Finally, since I faced an uninfected player at $\bar{t} + 1$, at most $M - 2$ people were infected after any period $t < \bar{t}$ ($\mathcal{I}^t \leq M - 2$).

To compute $x^{\bar{t}+1}$ we compute a series of intermediate beliefs x^t , for $t < \bar{t} + 1$. We compute x^2 from x^1 by conditioning on U^2 and $\mathcal{I}^2 \leq M - 2$, then we compute x^3 from x^2 and so on. Note that, to compute x^2 , we do not use the information that “I did not get infected at any period $2 < t < \bar{t}$ ”. So, at each $t < \bar{t}$, x^t represents my beliefs when I condition on the fact that the contagion started at period 1 and that no matching that leads to

more than $M - 2$ people being infected could have been realized.¹⁰ Put differently, at each period, I compute my beliefs by eliminating (assigning zero probability to) the matchings I know could not have taken place. At a given period $\tau < \bar{t}$, the information that “I did not get infected at any period t , with $\tau < t < \bar{t}$ ” is not used. This extra information is added period by period, *i.e.*, only at period t we add the information coming from the fact that “I was not infected at period t ”. In Section A.2 in the Appendix we show that this method of computing $x^{\bar{t}+1}$ generates the required beliefs, *i.e.*, beliefs at period $\bar{t} + 1$ conditioning on the entire history. Now we are equipped to check the optimality of the equilibrium strategies.

4.3.2 A seller deviates at beginning of the game

The strategies specify that a player who gets infected by deviating to a triggering action henceforth plays his best response to the strategies of the other players. As we argued for the product choice game, it is easy to show that for a given M , if \dot{T} is large enough, a seller who deviates from equilibrium action \hat{a}_1 in period 1 will find it optimal to continue deviating and playing \bar{a}_1 until the end of Phase 1 (to maximize short-run gains), and then will switch to playing the Nash a_1^* forever.

4.3.3 A player gets infected in Phase III

Case 1: Infection at the start of Phase III. Let $h^{\dot{T}+\ddot{T}+1}$ denote a history in which I am a player who gets infected in period $\dot{T} + \ddot{T} + 1$. The equilibrium strategies prescribe that I switch to the Nash action forever. For this to be optimal, I must believe that enough players in the other community are already infected. My beliefs depend on what I know about how contagion spreads in Phase I after a seller deviates in period 1. In Phase I, only one deviant seller is infecting buyers. (Recall that buyers cannot spread contagion in Phase 1.) The contagion then is a Markov process with state space $\{1, \dots, M\}$, where a state represents the number of infected buyers. The transition matrix is $S_M \in \mathcal{M}_M$, where a state k transits to $k + 1$ if the deviant seller meets an uninfected buyer, which has probability $\frac{M-k}{M}$. With the remaining probability, *i.e.*, $\frac{k}{M}$, state k remains at state k . When no confusion arises, we omit subscript M in matrix S_M . Let s_{kl} be the probability that state i transitions to state j . Then,

¹⁰The updating after period \bar{t} is different, since I know that I was infected at \bar{t} and that no more than $M - 1$ people could possibly be infected in the other community at the end of period \bar{t} .

$$S_M = \begin{pmatrix} \frac{1}{M} & \frac{M-1}{M} & 0 & 0 & \dots & 0 \\ 0 & \frac{2}{M} & \frac{M-2}{M} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & & \vdots \\ 0 & 0 & 0 & \frac{M-2}{M} & \frac{2}{M} & 0 \\ 0 & 0 & 0 & 0 & \frac{M-1}{M} & \frac{1}{M} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

The transition matrix represents how the contagion is expected to spread. To compute my current beliefs, I must also condition on the information from my own history. Consider any period $t < \hat{T}$. After observing history $h^{\hat{T}+\hat{T}+1} = g \dots gb$, I know that at the end of period $t + 1$ at most $M - 1$ buyers were infected and I was not infected. Therefore, to compute x^{t+1} , my intermediate beliefs about the number of buyers who were infected at the end of period $t + 1$ (*i.e.*, about \mathcal{I}^{t+1}), I need to condition on the following:

- i) My beliefs about \mathcal{I}^t : x^t .
- ii) I was uninfected at the end of $t + 1$: the event U^{t+1} (this is irrelevant if I am a seller, since sellers cannot get infected in Phase I).
- iii) At most $M - 1$ buyers were infected by the end of period $t + 1$: $\mathcal{I}^{t+1} \leq M - 1$ (otherwise I could not have been uninfected at the start of Phase III).

Therefore, given $l < M$, if I am a buyer, the probability that exactly l buyers are infected after period $t + 1$, conditional on the above information, is given by:

$$\begin{aligned} \mathrm{P}(l^{t+1} | x^t \cap U^{t+1} \cap \mathcal{I}^{t+1} \leq M - 1) &= \frac{\mathrm{P}(l^{t+1} \cap U^{t+1} \cap \mathcal{I}^{t+1} \leq M - 1 | x^t)}{\mathrm{P}(U^{t+1} \cap \mathcal{I}^{t+1} \leq M - 1 | x^t)} \\ &= \frac{x_{l-1}^t s_{l-1,l} \frac{M-l}{M-l+1} + x_l^t s_{l,l}}{\sum_{l=1}^{M-1} (x_{l-1}^t s_{l-1,l} \frac{M-l}{M-l+1} + x_l^t s_{l,l})}. \end{aligned}$$

The expression for seller would be analogous, but without the $\frac{M-l}{M-l+1}$ factors. Notice that we can express the transition from x^t to x^{t+1} using what we call the *conditional transition matrix*. Since we already know that $x_M^t = x_M^{t+1} = 0$, we can just work in \mathbb{R}^{M-1} . Let $C \in \mathcal{M}_M$ be defined, for each pair $k, l \in \{1, \dots, M - 1\}$, by $c_{kl} := s_{kl} \frac{M-l}{M-k}$; with the remaining entries being 0.

Recall that C_{\lfloor} and S_{\lfloor} denote the matrices obtained from C and S by removing the last row and the last column of each. So, the truncated matrix of conditional transitional

probabilities C_{1j} is as follows:

$$C_{1j} = \begin{pmatrix} \frac{1}{M} & \frac{M-1}{M} \frac{M-2}{M-1} & 0 & 0 & \dots & 0 \\ 0 & \frac{2}{M} & \frac{M-2}{M} \frac{M-3}{M-2} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & & \vdots \\ 0 & 0 & 0 & 0 & \frac{M-2}{M} & \frac{2}{M} \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & \frac{M-1}{M} \end{pmatrix}.$$

My beliefs are such that $x^1 = (1, \dots, 0)$, since the deviant seller infected one buyer at period 1. So, if I am a buyer, x^{t+1} can be computed as

$$x^{t+1} = \frac{x^t C_{1j}}{\|x^t C_{1j}\|} = \frac{x^1 C_{1j}^t}{\|x^1 C_{1j}^t\|}.$$

The expression for a seller would be analogous, with S_{1j} instead of C_{1j} . Hence, I can compute my beliefs about the situation at the end of Phase I by

$$x^{\dot{T}} = \begin{cases} \frac{x^1 C_{1j}^{\dot{T}-1}}{\|x^1 C_{1j}^{\dot{T}-1}\|} & \text{if I am a buyer,} \\ \frac{x^1 S_{1j}^{\dot{T}-1}}{\|x^1 S_{1j}^{\dot{T}-1}\|} & \text{if I am a seller.} \end{cases}$$

Lemma 1. Fix M . Then, $\lim_{\dot{T} \rightarrow \infty} x^{\dot{T}} = (0, \dots, 0, 1)$.

The intuition for the above lemma is as follows. Note that the largest diagonal entry in the matrix C_{1j} (or S_{1j}) is the last one. This means that the state $M - 1$ is more stable than any other state. Consequently, as more periods of contagion elapse in Phase I, state $M - 1$ becomes more and more likely. The formal proof is a straightforward consequence of some of the properties of contagion matrices (see Proposition A.1 in the Appendix).

Proposition 2. Fix \ddot{T} and M . If I observe history $h^{\dot{T}+\ddot{T}+1} = g \dots gb$ and \dot{T} is large enough, then it is sequentially rational for me to play the Nash action at period $\dot{T} + \ddot{T} + 2$.

Proof. Suppose I am a buyer. Since sellers always play the Nash action in Phase II, I cannot learn anything from play in Phase II. By Lemma 1, if \dot{T} is large enough, I assign very high probability to $M - 1$ players in my own community being infected at the end of Phase I. Then, at least as many sellers got infected during Phase II. If exactly $M - 1$ buyers were

infected by the end of Phase II, then the only uninfected seller must have got infected in period $\dot{T} + \ddot{T} + 1$, since in this period I was the only uninfected buyer and I met an infected opponent. So, if \dot{T} is large enough, I assign probability arbitrarily close to 1 to all sellers being infected by the end of $\dot{T} + \ddot{T} + 1$. Hence, it is optimal to play my Nash action a_2^* .

Next, suppose I am a seller. In this case, the fact that no player infected me in Phase II will make me update my beliefs about how contagion has spread. However, if \dot{T} is large enough relative to \ddot{T} , even if I factor in the information that I was not infected in Phase II, the probability I assign to $M - 1$ buyers being infected by the end of Phase I is arbitrarily higher than the probability I assign to k players being infected for any $k < M - 1$. By the same argument as above, playing Nash is optimal. \square

Next, we consider $(\dot{T} + \ddot{T} + 1 + \alpha)$ -period histories of the form $h^{\dot{T} + \ddot{T} + 1 + \alpha} = g \dots gbg \dots^\alpha$, with $1 \leq \alpha \leq M - 2$, *i.e.*, histories where I was infected at period $\dot{T} + \ddot{T} + 1$ and then I observed α periods of good behavior while I was playing the Nash action. For the sake of exposition, assume that I am a buyer (the arguments for a seller are analogous). Why are these histories significant? Notice that if I get infected in period $\dot{T} + \ddot{T} + 1$, I can believe that all other players in my community are infected. However, if after that, I observe an action that can be played on equilibrium path, I have to revise my beliefs, since it is not possible that all the players in my community were infected after period $\dot{T} + \ddot{T} + 1$. Can this alter my incentives to play the Nash action?

Suppose $\alpha = 1$. After history $h^{\dot{T} + \ddot{T} + 2} = g \dots gbg$, I know that at most $M - 2$ buyers were infected by the end of Phase I. So, for each $t \leq \dot{T}$, $x_M^t = x_{M-1}^t = 0$. My beliefs are no longer computed using C_{1j} , but rather with C_{2j} , which is derived by removing the last two rows and columns of C . By a similar argument as Lemma 1, we can show that $x^{\dot{T}} \in \mathbb{R}^{M-2}$ converges to $(0, 0, \dots, 1)$. In other words, the state $M - 2$ is the most stable and, for \dot{T} large enough, I assign very high probability to $M - 2$ buyers being infected at the end of Phase I. Consequently, at least as many sellers were infected during Phase II. This in turn implies (just as in Proposition 2) that I believe that, with high probability, all players are infected by now (at $t = \dot{T} + \ddot{T} + 2$). To see why, note that in the worst case, exactly $M - 2$ sellers were infected during Phase II. In that case, one of the uninfected sellers met an infected buyer at period $\dot{T} + \ddot{T} + 1$ and I infected the last one in the last period. So, I assign very high probability to everyone being infected by now, and it is optimal for me to play Nash. A similar argument holds for $(\dot{T} + \ddot{T} + 1 + \alpha)$ -period histories with $\alpha \in \{1, \dots, M - 1\}$.

We need not check incentives for Nash reversion after histories where I observe more than $M - 2$ periods of g after being infected. These are erroneous histories. To see why, note that when I get infected at period $\dot{T} + \ddot{T} + 1$, I believe that the minimum number of people infected in each community is two (at least one pair of players in period 1 and one pair of players in period $\dot{T} + \ddot{T} + 1$). Suppose after that, I face $M - 2$ instances of g . Since I switched to Nash after being infected, I infected my opponents in the last $M - 2$ periods. At the end of $M - 2$ observations of g , I am sure that even if there were only exactly two players infected in each community at the end of $\dot{T} + \ddot{T} + 1$, I myself have infected all the remaining uninfected people. So observing more g after that cannot be explained by a single deviation in period 1. Then, I will believe that in addition to the first triggering deviation by a seller in period 1, there have been as many mistakes by players as needed to be consistent with the observed history.

Finally, consider histories where, after getting infected, I observe a sequence of actions that includes both g and b , *i.e.*, histories starting with $h^{\dot{T}+\ddot{T}+1} = g \dots gb$ and where I faced b in one or more periods after getting infected. After such histories, I will assign higher probability to more people being infected compared to histories where I only observed g after getting infected. The intuition is that observing b reconfirms my belief that contagion is widely spread. Thus, it is optimal for me to play the Nash action after any such history.

We have thus shown that a player who observes a triggering action for the first time at the start of Phase III will find it optimal to revert permanently to the Nash action.

Case 2: Infection late in Phase III. Next, suppose I get infected after observing history $h^{\bar{t}+1} = g \dots gb$, with $\bar{t} \gg \dot{T} + \ddot{T}$. Now we need to study the contagion during Phase III. From period $\dot{T} + \ddot{T} + 2$ on, the same number of people is infected in both communities. The contagion can again be studied as a Markov process with state space $\{1, \dots, M\}$. In contrast to Phase I, all infected players spread the contagion in Phase III.

The new transition matrix is $\hat{S} \in \mathcal{M}_M$. For each pair $k, l \in \{1, \dots, M\}$, if $k > l$ or $l > 2k$, $\hat{s}_{kl} = 0$; otherwise, *i.e.*, if $k \leq l \leq 2k$, (see Figure 3)

$$\hat{s}_{kl} = \frac{\left(\binom{k}{l-k} \binom{M-k}{l-k} (l-k)! \right)^2 (2k-l)!(M-l)!}{M!} = \frac{(k!)^2 ((M-k)!)^2}{((l-k)!)^2 (2k-l)!(M-l)!M!}.$$

Consider any t such that $\dot{T} + \ddot{T} < t < \bar{t}$. If I observe history $h^{\bar{t}+1} = g \dots gb$, I know that “at most $M - 1$ people could have been infected in the rival community at the end of period

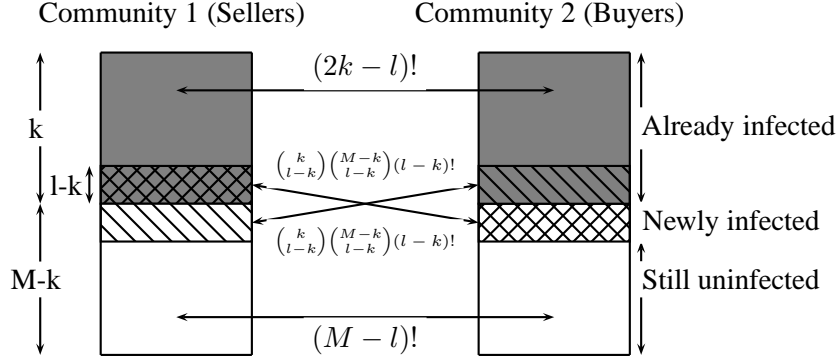


Figure 3: Spread of Contagion in Phase III. There are $M!$ possible matchings. For state k to transit to state l , exactly $(l-k)$ infected people from each community must meet $(l-k)$ uninfected people from the other community. The number of ways of choosing exactly $(l-k)$ buyers from k infected ones who will spread the contagion is $\binom{k}{l-k}$. The number of ways of choosing the corresponding $(l-k)$ uninfected sellers who will get infected is $\binom{M-k}{l-k}$, and the number of ways in which these sets of $(l-k)$ people can be matched is the total number of permutations of $l-k$ people, *i.e.*, $(l-k)!$. Analogously, we choose the $(l-k)$ infected sellers who will be matched to $(l-k)$ uninfected buyers. The number of ways in which the remaining infected buyers and sellers get matched to each other is $(2k-l)!$, and the uninfected ones is $(M-l)!$.

t ” ($\mathcal{I}^{t+1} \leq M-1$) and “I was not infected” (U^t). As before, let \hat{x}^t be my intermediate beliefs after period t . We want to compute the belief $\hat{x}^{\bar{t}+1}$, but first we study $\hat{x}^{\bar{t}}$ as $\bar{t} \rightarrow \infty$. Our limit results do not depend on the specific beliefs at the end of Phase II (as long as $\hat{x}_1^{\bar{t}+\bar{t}+1} > 0$, which is always true). Since, for each $t \leq \bar{t}$, $\hat{x}_M^t = 0$, we can just work with $\hat{x}^t \in \mathbb{R}^{M-1}$. As before, we want to compute $P(l^{t+1} | \hat{x}^t \cap U^{t+1} \cap \mathcal{I}^{t+1} \leq M-1)$ for $l \in \{1, \dots, M-1\}$.

$$\begin{aligned} P(l^{t+1} | \hat{x}^t \cap U^{t+1} \cap \mathcal{I}^{t+1} \leq M-1) &= \frac{P(l^{t+1} \cap U^{t+1} \cap \mathcal{I}^{t+1} \leq M-1) | \hat{x}^t}{P(U^{t+1} \cap \mathcal{I}^{t+1} \leq M-1) | \hat{x}^t} \\ &= \frac{\sum_{k \in \{1, \dots, M\}} x_k^t \hat{s}_{kl} \frac{M-l}{M-k}}{\sum_{l \in \{1, \dots, M-2\}} \left(\sum_{k \in \{1, \dots, M\}} x_k^t \hat{s}_{kl} \frac{M-l}{M-k} \right)}. \end{aligned}$$

These conditional transition probabilities can be expressed in matrix form. Let $\hat{C} \in \mathcal{M}_M$ be defined, for each pair $k, l \in \{1, \dots, M-1\}$, by $\hat{c}_{kl} := \hat{s}_{kl} \frac{M-l}{M-k}$; with the remaining entries being 0. Then, my intermediate beliefs at $t+1$ are given by

$$\hat{x}^{t+1} = \frac{\hat{x}^t \hat{C}_{1j}}{\|\hat{x}^t \hat{C}_{1j}\|}.$$

We show next that a result similar to Lemma 1 holds.

Lemma 2. *Fix $\dot{T} \in \mathbb{N}$, $\ddot{T} \in \mathbb{N}$, and $M \in \mathbb{N}$. Then, $\lim_{\bar{t} \rightarrow \infty} \hat{x}^{\bar{t}} = (0, 0, \dots, 0, 1) \in \mathbb{R}^{M-1}$.*

The lemma follows from properties of contagion matrices (see Proposition A.1 in the Appendix for the proof). The informal argument is as follows. The largest diagonal entries of the matrix \hat{C}_{1j} are the first and last ones (\hat{c}_{11} and $\hat{c}_{M-1, M-1}$), which are equal. Unlike in the Phase I transition matrix, state $M - 1$ is not the unique most stable state. Here, states 1 and $M - 1$ are equally stable, and more stable than any other state. Why do beliefs then converge to $(0, 0, \dots, 0, 1)$? In each period, many states transit to $M - 1$ with positive probability, while no state transits to state 1, and so the ratio $\frac{\hat{x}_{M-1}^t}{\hat{x}_1^t}$ goes to ∞ as t increases. So, late in the game, I assign arbitrarily high probability to state $M - 1$.

Proposition 3. *Fix $\dot{T} \in \mathbb{N}$, $\ddot{T} \in \mathbb{N}$, and $M \in \mathbb{N}$. If I observe history $h^{\bar{t}+1} = g \dots gb$ and \bar{t} is large enough, then it is sequentially rational for me to play the Nash action at period $\bar{t} + 2$.*

Proof. By Lemma 2, if \bar{t} is large enough, $\hat{x}^{\bar{t}}$ is such that I assign very high probability to $M - 1$ players in the other community being infected by the end of period \bar{t} . Now, to compute $\hat{x}^{\bar{t}+1}$ from $\hat{x}^{\bar{t}}$, I add the information that I got infected at $\bar{t} + 1$ and hence, the only uninfected person in the other community got infected too. So, now I assign very high probability to everyone being infected. Then, Nash reversion is optimal. \square

Suppose now that I observe $h^{\bar{t}+2} = g \dots gbg$ and that I played the Nash action at period $\bar{t} + 2$. Then, I will know that less than $(M - 1)$ people were infected at the end of period \bar{t} since, otherwise, I could not have faced g in period $\bar{t} + 2$. In other words, I have to recompute my beliefs using the information that, for each $t \leq \bar{t}$, $\mathcal{I}^t \leq M - 2$. I should now use the truncated transition matrix \hat{C}_{2j} . Since, for each any $t \leq \bar{t}$, $\hat{x}_M^t = \hat{x}_{M-1}^t = 0$, to obtain $\hat{x}^{\bar{t}}$ we just work with $\hat{x}^t \in \mathbb{R}^{M-2}$. Now we have

$$\hat{x}^{t+1} = \frac{\hat{x}^t \hat{C}_{2j}}{\|\hat{x}^t \hat{C}_{2j}\|}.$$

As before, we will study the limit behavior of $\hat{x}^{\bar{t}}$ as \bar{t} goes to ∞ , and use $\hat{x}^{\bar{t}}$ to compute beliefs at $\bar{t} + 2$. First, we show that $\hat{x}^{\bar{t}}$ indeed converges.

Lemma 3. *For each $\dot{T} \in \mathbb{N}$, each $\ddot{T} \in \mathbb{N}$, and each $M \in \mathbb{N}$, $\lim_{\bar{t} \rightarrow \infty} \hat{x}^{\bar{t}} = \bar{x}$, where \bar{x} is the unique left eigenvector associated with the largest eigenvalue of \hat{C}_{2j} such that $\|\bar{x}\| = 1$. That is, $\bar{x} \hat{C}_{2j} = \frac{\bar{x}}{M}$.*

See Proposition A.1 in the Appendix for a proof. Note that in matrix \hat{C}_{2j} , the largest diagonal entry is the first one. So, a result like Lemma 1 no longer holds, *i.e.*, $\bar{x} \neq (0, \dots, 0, 1)$. However, we show that I will still believe that “enough” people are infected with “high enough probability”.

Lemma 4. *Let $\bar{x} = \lim_{\bar{t} \rightarrow \infty} \hat{x}^{\bar{t}}$, where $\hat{x}^{\bar{t}}$ denotes a player’s beliefs at the end of period \bar{t} after he observes history $h^{\bar{t}+2} = g \dots gbg$. Let $r \in (0, 1)$. Then, for each $\varepsilon > 0$, there is $\underline{M} \in \mathbb{N}$ such that, for each $M \geq \underline{M}$,*

$$\sum_{j=\lceil rM \rceil}^{M-2} \bar{x}_j > 1 - \varepsilon,$$

where $\lceil z \rceil$ is the ceiling function and is defined as the smallest integer not smaller than z . Indeed, for each $m \in \mathbb{N}$, there is $\underline{M} \in \mathbb{N}$ such that, for each $M \geq \underline{M}$,

$$\sum_{j=\lceil rM \rceil}^{M-2} \bar{x}_j > 1 - \frac{1}{M^m}.$$

The lemma can be interpreted as follows. Think of r as a proportion of people, say 0.9. If the population size is large enough, after observing history $h^{\bar{t}+2} = g \dots gbg$, my limiting belief \bar{x} will be such that I will assign probability at least $(1 - \varepsilon)$ to at least 90% of the population being infected. We can choose r close enough to 1 and ε small enough and then find $M \in \mathbb{N}$ large enough so that I believe that the contagion has spread enough that playing Nash action is optimal.

Figure 4 below represents the probabilities $\sum_{j=\lceil rM \rceil}^{M-2} \bar{x}_i$ for different values of r and M . In particular, it shows that they go to one very fast with M . From the results in this section it will follow that, after any history in which I have been infected late in Phase III, I will believe that the contagion is at least as spread as \bar{x} indicates. For instance, consider $M = 20$. Now \bar{x} is such that I believe that with probability at least 0.75, at least 90% of the people are infected. This may be enough to induce the right incentives for many games. In general, the incentives will hold even for fairly small population sizes.¹¹

In order to prove Lemma 4, we need to study more carefully the transitions between states. There are two opposing forces that affect how my beliefs evolve after I observe

¹¹The non-monotonicity in the graphs in Figure 4 may be surprising. To the best of our understanding, this can be essentially attributed to the fact that the states that are powers of 2 tend to be more likely and their distribution within the top $M - \lceil rM \rceil$ states varies in a non-monotone way.

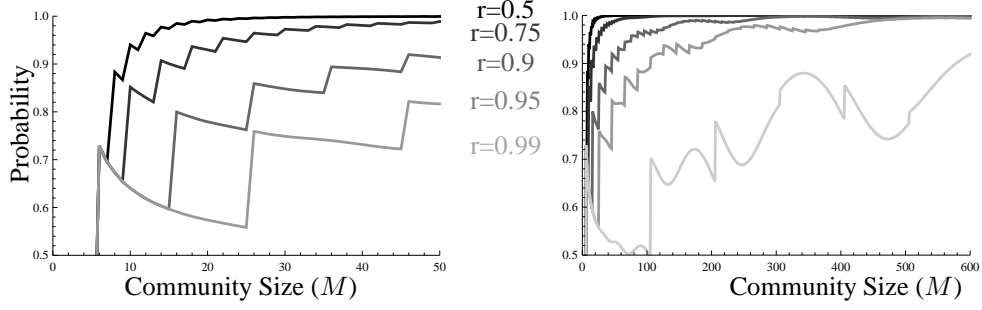


Figure 4: Illustration of Lemma 4.

$g \dots gbg$. First, each observation of g suggests to me that not too many people are infected, making me step back in my beliefs and assign higher weight to lower states (less people infected). On the other hand, since I believe that the contagion started at $t = 1$ and that it has been spreading exponentially during Phase III, every elapsed period makes me assign more weight to higher states (believe more people are infected). What we need to do is to compare the magnitudes of these two effects. Two main observations drive the proof. First, each time I observe g , my beliefs get updated with more weight assigned to lower states and, roughly speaking, this step back in beliefs turns out to be of the order of M . Second, we identify the the most likely transition from any given state k . It turns out the likeliest transition from k , say state k' , is about \sqrt{M} times as likely as the state k . Similarly, the most likely transition from k' , say state k'' , is \sqrt{M} times as likely as the state k' . Hence, given a proportion of people $r \in (0, 1)$, it is easy to see that, if M is large enough, for each state $k < rM$, we can find another state \bar{k} that is at least M^2 times more likely than state k . So the second effect on the beliefs is of an order of M^2 , which dominates the first one.

We need some preliminaries before we formally prove Lemma 4. Recall that

$$(\hat{c}_2)_{k,k+j} = \hat{s}_{k,k+j} \frac{M-k}{M-k-j} = \frac{(k!)^2((M-k)!)^2}{(j!)^2(k-j)!(M-k-j)!M!} \frac{M-k-j}{M-k}.$$

Given a state $k \in \{1, \dots, M-2\}$, consider the transition from k to state $\text{tr}(k) := k + \lfloor \frac{k(M-k)}{M} \rfloor$, where $\lfloor z \rfloor$ is the floor function defined as the largest integer not larger than z . It turns out that, for large M , this is a good approximation of the most likely transition from state k .

For analytical ease, we assume for now, that there is a continuum of states, *i.e.*, let the set of states be the interval $[0, M]$. We show later that the results also hold in the finite

setting. In the continuous setting, a state $z \in [0, M]$, can be represented as rM ; where $r = z/M$ can be interpreted as the proportion of infected people at state z . Let $\alpha \in \mathbb{R}$. We define a function $f_\alpha: [0, 1] \rightarrow \mathbb{R}$ as

$$f_\alpha(r) := \frac{rM(M - rM)}{M} + \alpha = (r - r^2)M + \alpha.$$

Note that f_α is continuous and further that $rM + f_0(r)$ would just be the extension of the function $\text{tr}(k)$ to the continuous case. We want to know how likely the transition from state r to $r + f_0(r)$ is. Define a function $g: [0, 1] \rightarrow [0, 1]$ as

$$g(r) := 2r - r^2.$$

The function g is continuous and strictly increasing. Given $r \in [0, 1]$, $g(r)$ represents the proportion of infected people if, at state rM , $f_0(r)$ people get infected. To see why, note that $rM + f_0(r) = rM + (r - r^2)M = (2r - r^2)M$. Let $g^2(r) := g(g(r))$ and define analogously any other power of g . Hence, for each $r \in [0, 1]$, $g^n(r)$ represents the fraction of people infected after n steps starting at rM when transitions are made according to $f_0(\cdot)$.

Claim 1. *Let $M \in \mathbb{N}$ and $a, b \in [0, 1]$, with $a > b > 0$. Then, $aM + f_0(a) > bM + f_0(b)$.*

Proof. Note that $aM + f_0(a) - bM - f_0(b) = aM + \frac{aM(M-aM)}{M} - bM - \frac{bM(M-bM)}{M} = 2aM - a^2M - 2bM + b^2M = (g(a) - g(b))M$. Since $g(\cdot)$ is strictly increasing in $(0, 1)$, the result follows. \square

Now we define a function $h_\alpha^M: (0, 1) \rightarrow (0, \infty)$ as

$$h_\alpha^M(r) := \frac{(rM!)^2((M - rM)!)^2}{(f_\alpha(r)!)^2(rM - f_\alpha(r))!(M - rM - f_\alpha(r))!M!} \frac{M - rM - f_\alpha(r)}{M - rM}.$$

This function is the continuous version of the transitions given by the matrix \hat{C}_2 . In particular, given $\alpha \in \mathbb{R}$ and $r \in [0, 1]$ the function $h_\alpha^M(r)$ represents the conditional probability of transition from state rM to state $rM + f_\alpha(r)$. In some abuse of notation, we apply the factorial function to non-integer real numbers. In such cases, the factorial can be interpreted as the corresponding Gamma function, *i.e.*, $a! = \Gamma(a + 1)$.

Claim 2. *Let $\alpha \in \mathbb{R}$ and $r \in (0, 1)$. Then, $\lim_{M \rightarrow \infty} Mh_\alpha^M(r) = \infty$. More precisely,*

$$\lim_{M \rightarrow \infty} \frac{Mh_\alpha^M(r)}{\sqrt{M}} = \frac{1}{r\sqrt{2\pi}}.$$

Proof. We prove the claim in two steps.

Step 1: $\alpha = 0$. Stirling's formula implies that $\lim_{n \rightarrow \infty} (e^{-n} n^{n+\frac{1}{2}} \sqrt{2\pi})/n! = 1$. Given $r \in (0, 1)$, to study $h_\alpha^M(r)$ in the limit, we use the approximation $n! = e^{-n} n^{n+\frac{1}{2}} \sqrt{2\pi}$. Substituting and simplifying, we get the following:

$$\begin{aligned} Mh_0^M(r) &= M \frac{((rM)!)^2(((1-r)M)!)^2}{M!(r^2M)!(((r-r^2)M)!)^2((1-r)^2M)!} (1-r) \\ &= \frac{M(rM)^{1+2rM}((1-r)M)^{1+2(1-r)M}(1-r)}{\sqrt{2\pi} M^{\frac{1}{2}+M}((1-r)^2M)^{1+2(1-r)^2M}((r-r^2)M)^{\frac{1}{2}+(r-r^2)M}(r^2M)^{\frac{1}{2}+r^2M}} \\ &= \frac{\sqrt{M}}{r\sqrt{2\pi}}. \end{aligned}$$

Step 2: Let $\alpha \in \mathbb{R}$ and $r \in (0, 1)$. Now,

$$\frac{h_0^M(r)}{h_\alpha^M(r)} = \frac{(r^2M - \alpha)!(((r-r^2)M + \alpha)!)^2((1-r)^2M - \alpha)!}{(r^2M)!(((r-r^2)M)!)^2((1-r)^2M)!} \frac{(1-r)^2M}{(1-r)^2M - \alpha}.$$

Applying Stirling's formula we get

$$\frac{(r^2M - \alpha)^{\frac{1}{2}+r^2M - \alpha}}{(r^2M)^{\frac{1}{2}+r^2M}} \frac{((r-r^2)M + \alpha)^{1+2(r-r^2)M + 2\alpha}}{((r-r^2)M)^{1+2(r-r^2)M}} \frac{((1-r)^2M - \alpha)^{\frac{1}{2}+(1-r)^2M - \alpha}}{((1-r)^2M)^{\frac{1}{2}+(1-r)^2M}} \frac{(1-r)^2M}{(1-r)^2M - \alpha}. \quad (1)$$

To compute the limit of the above expression as $M \rightarrow \infty$, we analyze the four fractions above separately. Clearly, $((1-r)^2M)/((1-r)^2M - \alpha) \rightarrow 1$ as $M \rightarrow \infty$. So, we restrict attention to the first three fractions. Take the first one:

$$\frac{(r^2M - \alpha)^{\frac{1}{2}+r^2M - \alpha}}{(r^2M)^{\frac{1}{2}+r^2M}} = \left(1 - \frac{\alpha}{r^2M}\right)^{\frac{1}{2}} \cdot \left(1 - \frac{\alpha}{r^2M}\right)^{r^2M} \cdot (r^2M - \alpha)^{-\alpha} = A_1 \cdot A_2 \cdot A_3,$$

where $\lim_{M \rightarrow \infty} A_1 = 1$ and $\lim_{M \rightarrow \infty} A_2 = e^{-\alpha}$. Similarly, the second fraction decomposes as $B_1 \cdot B_2 \cdot B_3$, where $\lim_{M \rightarrow \infty} B_1 = 1$, $\lim_{M \rightarrow \infty} B_2 = e^{2\alpha}$ and $B_3 = ((r-r^2)M + \alpha)^{2\alpha}$. The third fraction can be decomposed as $C_1 \cdot C_2 \cdot C_3$, where $\lim_{M \rightarrow \infty} C_1 = 1$, $\lim_{M \rightarrow \infty} C_2 = e^{-\alpha}$ and $C_3 = ((1-r)^2M - \alpha)^{-\alpha}$. The limit of expression (1) as $M \rightarrow \infty$ reduces to

$$\lim_{M \rightarrow \infty} \left(\frac{((r-r^2)M + \alpha)^2}{(r^2M - \alpha)((1-r)^2M - \alpha)} \right)^\alpha = 1. \quad \square$$

We are now ready to prove Lemma 4.

Proof of Lemma 4. Fix $r \in (0, 1)$ and $\varepsilon > 0$. We show that there is \underline{M} such that, for each

$M \geq \underline{M}$ and each state $k < \lceil rM \rceil$, there is a state $\bar{k} \in \{\lceil rM \rceil, \dots, M - 2\}$ such that $\bar{x}_{\bar{k}} > M^2 \bar{x}_k$. We show this first for state $k_0 = \lceil rM \rceil - 1$. Consider the state rM and let $\bar{r} := g^5(r)$. Recall that \bar{r} is the state reached (proportion of people infected) from initial state rM after 5 steps according to the function f_0 . Recall that functions f_0 and g are such that, $r < \bar{r} < 1$. Moreover, suppose M is large so that $\bar{r}M \leq M - 2$. Consider the state k_0 and let \bar{k} be the number of infected people after 5 steps according to function $\text{tr}(\cdot)$. Clearly, for each of these steps, there is $\alpha_j \in (-1, 0]$, with $j \in \{1, \dots, 5\}$, such that the step corresponds with that of function f_{α_j} . By Claim 1, since $k_0 < rM$, $\bar{k} < M - 2$. Moreover, it is trivial to see that $\bar{k} > \lceil rM \rceil$. Let k_1 be the state that is reached after the first step from k_0 according to function $\text{tr}(\cdot)$. By Lemma 3, $\bar{x} = M\bar{x}\hat{C}_2$. Then, $\bar{x}_{k_1} = M \sum_{k=1}^{M-2} \bar{x}_k (\hat{C}_2)_{kk_1} > M\bar{x}_{k_0} (\hat{C}_2)_{k_0 k_1} = \bar{x}_{k_0} M h_{\alpha_1}^M(r)$, which, if M is large enough, can be approximated by $\bar{x}_{k_0} \frac{\sqrt{M}}{r\sqrt{2\pi}}$. Repeating the same argument for the other intermediate states that are reached in each of the five steps, we get that, if M is large enough, $\bar{x}_{\bar{k}} > M^2 \bar{x}_{k_0}$. The proof for an arbitrary state $k < \lceil rM \rceil - 1$ is very similar, with the only difference that more steps might be needed to get to a state $\bar{k} \in \{\lceil rM \rceil, \dots, M - 2\}$; yet, the extra number of steps makes the difference between \bar{x}_k and $\bar{x}_{\bar{k}}$ even bigger.¹²

Let $\bar{k} \in \{\lceil rM \rceil, \dots, M - 2\}$ be a state such that $\bar{x}_{\bar{k}} > M^2 \max\{\bar{x}_k : k \in \{1, \dots, \lceil rM \rceil - 1\}\}$. Then, $\sum_{k=1}^{\lceil rM \rceil - 1} \bar{x}_k < rM \frac{\bar{x}_{\bar{k}}}{M^2} < \frac{1}{M}$. Therefore, if M is big enough, $\frac{1}{M} < \varepsilon$ and we get that $\sum_{i=\lceil rM \rceil}^{M-2} x_i > 1 - \varepsilon$.

The second part of the statement is now straightforward. Just take $\bar{r} := g^{2m+3}(r)$ and repeat the argument above. \square

Proposition 4. Fix a game $\Gamma(g, l, c)$. Fix $\dot{T} \in \mathbb{N}$ and $\ddot{T} \in \mathbb{N}$ and let $\bar{t} \gg \ddot{T} + \dot{T}$. Suppose I observe history $h^{\bar{t}+2} = g \dots gbg$. Then, if M is large enough, it is sequentially rational for me to play the Nash action at period $\bar{t} + 3$.

Proof. First, consider my beliefs $\hat{x}^{\bar{t}}$ computed conditioning only the information that at most $M - 2$ people were infected after period \bar{t} and that I was uninfected until period $\bar{t} + 1$. From Lemma 3, if \bar{t} is large enough, $\hat{x}^{\bar{t}}$ is very close to \bar{x} . In particular, I believe that, with probability at least $1 - \varepsilon$, at least rM people are infected. We can use $\hat{x}^{\bar{t}}$ to compute my beliefs at period $\bar{t} + 2$.

- After period $\bar{t} + 1$: I compute $\hat{x}^{\bar{t}+1}$ by updating $\hat{x}^{\bar{t}}$, conditioning on i) I observed b in period $\bar{t} + 1$ and ii) at most $M - 1$ people were infected after $\bar{t} + 1$ (I observed g at $\bar{t} + 2$).

¹²It is worth noting that we can do this argument uniformly for the different states and the corresponding α 's because we know that all of them lie inside $[-1, 0]$, a bounded interval; that is, we can take one M big enough so as to ensure that we can use the approximation given by Lemma 2 for any α in $[-1, 0]$.

Let $\tilde{x}^{\bar{t}+1}$ be the belief computed from $\hat{x}^{\bar{t}}$ by conditioning instead on i) I observed g and ii) at most $M - 2$ people are infected. Clearly, $\hat{x}^{\bar{t}+1}$ first order stochastically dominates $\tilde{x}^{\bar{t}+1}$, in the sense of placing higher probability on more people being infected. Now, recall that the belief $\hat{x}^{\bar{t}}$ is very close to \bar{x} and, by definition, the belief $\tilde{x}^{\bar{t}+2}$ is even closer to \bar{x} .

- After period $\bar{t} + 2$: I compute $x^{\bar{t}+2}$ based on $x^{\bar{t}+1}$ and conditioning on i) I observed g , ii) I infected my opponent by playing the Nash action at $\bar{t} + 2$, and iii) at most M have been infected after $\bar{t} + 2$. Again, this updating leads to beliefs that first order stochastically dominate beliefs we would obtain if we instead conditioned on i) I observed g and ii) at most $M - 2$ people were infected after $\bar{t} + 2$. Again, the belief $\tilde{x}^{\bar{t}+2}$ would be very close to \bar{x} .

Hence, if it is optimal for me to play the Nash action when my beliefs are given by \bar{x} , it is also optimal to do so after observing the history $h^{\bar{t}+2} = g \dots gbg$ (provided \bar{t} is large enough). Lemma 4 ensures that, if M is large enough, the beliefs \bar{x} can be made as extreme as desired (in the sense of many people being infected), ensuring that players have the incentive to switch to Nash. \square

As we did in Case 1, we still need to show that Nash reversion is optimal after histories of the form $h^{\bar{t}+1+\alpha} = g \dots gbg \dots g$. The intuition is as follows. If α is small, then I will think that a lot of players were already infected when I got infected; the argument being similar to that in Proposition 4. If α is large, I may learn that there were not so many players infected when I got infected. However, the number of players I myself have infected since then, together with the exponential spread of the contagion, will be enough to convince me that, at the present period, contagion is widely spread anyway.

To formalize the above intuition, we need the following strengthening of Lemma 4. We omit the proof, as it involves a minor elaboration of the arguments in Lemma 4.

Lemma 5. *Let $r \in (0, 1)$. Then, for each $\varepsilon > 0$, there are $\hat{r} \in (r, 1)$ and $\underline{M} \in \mathbb{N}$ such that, for each $M \geq \underline{M}$,*

$$\frac{\sum_{j=\lceil rM \rceil}^{\lfloor M-\hat{r}M \rfloor} x_j}{1 - \sum_{j=\lfloor M-\hat{r}M \rfloor+1}^M x_j} > 1 - \varepsilon$$

Indeed, for each $m \in \mathbb{N}$, there is $\underline{M} \in \mathbb{N}$ such that, for each $M \geq \underline{M}$,

$$\frac{\sum_{j=\lceil rM \rceil}^{\lfloor M - \hat{r}M \rfloor} x_j}{1 - \sum_{j=\lfloor M - \hat{r}M \rfloor + 1}^M x_j} > 1 - \frac{1}{M^m}.$$

Proposition 5. Fix a game $\Gamma(g, l, c)$. Fix $\dot{T} \in \mathbb{N}$ and $\ddot{T} \in \mathbb{N}$ and let $\bar{t} \gg \dot{T} + \ddot{T}$. Let $\alpha \geq 1$. Suppose I observe history $h^{\bar{t}+1+\alpha} = g \dots gbg \dots g$. Then, if M is large enough, it is sequentially rational for me to play the Nash action at period $\bar{t} + 2 + \alpha$.

Proof. The proof is similar to that of Proposition 4. First, I know that at most $M - \alpha - 1$ people were infected after period \bar{t} . The new limit vector $\tilde{x} \in \mathbb{R}^{M-\alpha-1}$ must be computed using matrix $\hat{C}_{(\alpha+1)}$. If we define $y := (\bar{x}_1, \dots, \bar{x}_{M-\alpha-1})$, it is easy to see that $\tilde{x} = \frac{y}{\|y\|}$.

Second, consider the following scenario. Suppose that, late in Phase III, an infected player believes that exactly two people were infected in each community, and then he played the Nash action for multiple periods while observing only g . In each period he infected a new person. At the same time, contagion continued to spread exponentially. If the number of periods in which the player infected people is large enough, Nash reversion would be the best reply, irrespective of what he observes in the meantime (he would have infected enough people himself). Let $\phi(M)$ denote this number of periods. Since the contagion spreads exponentially, the threshold $\phi(M)$ is some logarithmic function of M . Hence, for each $\hat{r} \in (0, 1)$, there is \hat{M} such that for $M > \hat{M}$, $\hat{r}M > \phi(M)$. Now, given $\varepsilon > 0$, we can find \hat{r} and \underline{M} such that Lemma 5 holds. For the rest of the proof we work with $M > \max\{\hat{M}, \underline{M}\}$. There are two cases.

$\alpha < \phi(M)$: In this case, we can repeat the arguments in the proof of Proposition 4 to show that my beliefs $\hat{x}^{\bar{t}+1+\alpha}$ first order stochastically dominate \tilde{x} . Since $\hat{r}M > \phi(M)$, $\lfloor M - \hat{r}M \rfloor < M - \phi(M)$. Now we rely on Lemma 5 to get the desired result.

$\alpha \geq \phi(M)$: In this case I played the Nash action for α periods. By definition of $\phi(M)$, playing Nash is the unique best reply after observing $h^{\bar{t}+1+\alpha}$. \square

Finally, we consider histories in which after getting infected, I observe actions that include both g and b , i.e., histories starting with $h^{\bar{t}+1} = g \dots gb$ and in which I have observed b in one or more periods after getting infected. The reasoning is the same as Case 1. After such histories, I will assign higher probability to more people being infected compared to histories where I only observed g after getting infected.

Case 3: Infection in other periods of Phase III (“Monotonicity” of Beliefs). So far we

have shown that if a player is infected early in Phase III, he thinks that he was the last player to be infected and that everybody is infected, making Nash reversion optimal. Also, if a player is infected late in Phase III, he will believe that enough players are already infected, for it to be optimal to play the Nash action (with his limit belief being given by \bar{x}). Next, we show that the belief of a player who is infected not very late in Phase III will be somewhere in between. The earlier a player gets infected in Phase III, the closer his belief will be to $(0, \dots, 0, 1)$ and, the later he gets infected, the closer his belief will be to \bar{x} .

Proposition 6. *Fix a game $\Gamma(g, l, c)$. Fix $\dot{T} \in \mathbb{N}$. There is $\underline{M} \in \mathbb{N}$ such that, for each $M > \underline{M}$, if \dot{T} is large enough, then it is sequentially rational for me to play the Nash action after any history in which I get infected in Phase III.*

Proof. In Cases 1 and 2 we showed that if I get infected at the start of Phase III (at $\dot{T} + \ddot{T} + 1$) or late in Phase III (at $\tilde{t} \gg \dot{T} + \ddot{T}$), I will switch to the Nash action. What remains to be shown is that the same is true if I get infected at some intermediate period in Phase III. We prove this for histories in Phase III of the form $h^{\tilde{t}+2} = g \dots gbg$. The proof can be extended to include other histories just as in Cases 1 and 2. Recall that \bar{x} denotes the limit belief when \tilde{t} goes to infinity.

We want to compute my belief $\hat{x}^{\tilde{t}+2}$ after history $h^{\tilde{t}+2} = g \dots gbg$. We first compute intermediate belief \hat{x}^t , for $t \leq \tilde{t}$. Take \underline{M} such that Proposition 4 holds for all $M > \underline{M}$.

Beliefs are computed using matrix C_{2j} in Phase I, and \hat{C}_{2j} in Phase III. We know (from Case 1) that for \dot{T} large enough, $\hat{x}^{\dot{T} + \ddot{T}} \in \mathbb{R}^{M-1}$ is close to $(0, \dots, 0, 1)$. Moreover, by taking \dot{T} large enough, we also get that $\hat{x}_k^{\dot{T} + \ddot{T} + 1} \gg \hat{x}_{k-1}^{\dot{T} + \ddot{T} + 1} \gg \dots \gg \hat{x}_1^{\dot{T} + \ddot{T} + 1} > 0$ and, for each $i > j$, $\hat{x}_i^{\dot{T} + \ddot{T} + 1} / \hat{x}_j^{\dot{T} + \ddot{T} + 1} > \bar{x}_i / \bar{x}_j$. Using the properties of the contagion matrix \hat{C}_{2j} , we can show that if we start Phase III with such a belief $\hat{x}^{\dot{T} + \ddot{T} + 1}$ we also get that, for each $i > j$, $\hat{x}_i^{\tilde{t}} / \hat{x}_j^{\tilde{t}} > \bar{x}_i / \bar{x}_j$ (see Proposition A.2 in the Appendix). This means that $\hat{x}^{\tilde{t}}$ first order stochastically dominates \bar{x} , in the sense of placing higher probability on more people being infected. Now, my beliefs need to be updated from $\hat{x}^{\tilde{t}}$ to $\hat{x}^{\tilde{t}+1}$ and then from $\hat{x}^{\tilde{t}+1}$ to $\hat{x}^{\tilde{t}+2}$. We can use similar arguments as in Proposition 4 to show that $\hat{x}^{\tilde{t}+2}$ first order stochastically dominates \bar{x} . Hence, if it is sequentially rational for me to play the Nash action when my beliefs are \bar{x} , it is also sequentially rational to do so when my belief is $\hat{x}^{\tilde{t}+2}$. \square

Hence, we have established that if a player observes a triggering action any time during Phase III, it is sequentially rational for him to revert to the Nash action.

4.3.4 A player observes a triggering action in Phases I or II

It remains to check the incentives for a player who is infected during the initial Trust-Building Phases. We argued informally in Section 3.2 why players would find it optimal to switch to the Nash action. We omit the formal proofs, as the arguments are very similar to those used for Case 1 above.

4.3.5 A player observes a non-triggering action

An uninfected player who observes a non-triggering action knows that his opponent will not get infected, and will continue to play as if on-path. Since he knows that contagion will not start, clearly, the best thing to do is also to ignore this off-path behavior.

4.4 Choosing \underline{M} , \ddot{T} , \dot{T} , and $\underline{\delta}$

We have shown that if \underline{M} , \ddot{T} , \dot{T} , and $\underline{\delta}$ are chosen appropriately, the prescribed strategies are sequentially rational. We show now that it is possible to choose these parameters to satisfy all incentive constraints simultaneously. Fix any game $G \in \mathcal{G}$ with strict Nash equilibrium a^* and a target payoff $v \in F_{a^*}$.

The first step is to choose \underline{M} large enough so that incentive constraints in Phase III are satisfied, *i.e.*, a player who observes a triggering action late in Phase III believes that enough people are already infected so that Nash reversion is optimal.

There is one subtle issue. Once \underline{M} , \ddot{T} , and \dot{T} are chosen, we need players to be patient enough (δ large) to prevent deviations on-path. But, we also need to check that, a very patient player does not want to slow down the contagion once it has started. We do this below. The essence of the argument is in observing that, for a fixed population size, once contagion has started, the expected future stage-game payoffs go down to the Nash payoff $u(a^*)$ exponentially fast. That is, regardless of how an infected player plays in the future, the undiscounted sum of possible future gains he can make relative to the payoff from Nash reversion is bounded above. Thus, in some sense, even a perfectly patient player becomes effectively impatient.

Let \bar{m} be the maximum possible gain any player can make by a unilateral deviation from any action profile of the stage-game. Let \underline{l} be the minimum loss that a player suffers in the stage-game when he does not play his best response to his rival's strict Nash action a^* (Recall that since a^* is a strict Nash equilibrium, $\underline{l} > 0$). Suppose we are in Phase III and

take a player i who knows that the contagion has started. For the analysis in this section, we use the continuation payoffs of the Nash reversion strategy as the benchmark.

For any given continuation strategy of a player i , let $v(M)$ denote player i 's (expected) undiscounted sum of future gains relative to his payoff from the Nash reversion strategy. It is easy to see that $v(M)$ is finite. Player i knows that contagion is spreading exponentially and, hence, his payoff will drop to the Nash payoff $u(a^*)$ in the long run. In fact, although $v(M)$ increases with M , since contagion spreads exponentially fast, $v(M)$ grows at a slower rate than M .

Similarly, for any continuation strategy, and for each $r \in (0, 1]$, let $v(r, M)$ denote the (expected) undiscounted sum of future gains player i can make from playing this continuation strategy relative to the payoff from Nash reversion, when he is in Phase III and knows that at least rM people are infected in each community. In the result below we show that $v(r, M)$ is uniformly bounded on M .

Lemma 6. *Fix a game $G \in \mathcal{G}$. Let $r > 1/2$. Then, there is \bar{U} such that, for each $\bar{r} > r$ and each $M \in \mathbb{N}$, $v(\bar{r}, M) \leq \bar{U}$.*

Proof. Let $P(r, M)$ denote the probability of the following event: “If $\lceil rM \rceil$ people are infected at period t , the number of infected people at period $t+1$ is, at least, $\lceil rM + \frac{(1-r)}{2}M \rceil$ ”, i.e., $P(r, M)$ is the probability that at least half the uninfected people get infected in the present period, given that rM people are infected already. We claim that $P(\frac{1}{2}, M) \geq \frac{1}{2}$.

Suppose M is even and $\frac{M}{2}$ people are infected at the start of period t (i.e. $r = \frac{1}{2}$). What is the probability of at least $\frac{M}{4}$ more people getting infected in this period? It is easy to see from the contagion matrix, that there is a symmetry in the transition probabilities of different states. The probability of no one getting infected in this period is the same as that of no one remaining uninfected. The probability of one person being infected is the same as that of only one person remaining uninfected. In general, probability of $k < M/4$ players getting infected is the same as that of k players remaining uninfected. This symmetry implies immediately that $P(\frac{1}{2}, M) \geq \frac{1}{2}$.

It is easy to see that for $r > \frac{1}{2}$, $\lceil rM \rceil > \frac{M}{2}$, and further $P(r, M) > P(\frac{1}{2}, M) \geq \frac{1}{2}$. Thus, for each $r > \frac{1}{2}$ and each $M \in \mathbb{N}$, $P(r, M) > \frac{1}{2}$. Also, for any fixed $r > \frac{1}{2}$, $\lim_{M \rightarrow \infty} P(r, M) = 1$; intuitively, if more than half the population is already infected, then the larger M is, the more unlikely it is that more than half of the uninfected people remain uninfected. Hence, given $r > \frac{1}{2}$, there is $\hat{p} < \frac{1}{2}$ such that, for each $M \in \mathbb{N}$, $P(r, M) > 1 - \hat{p}$.

Given any continuation strategy, we want to compute $v(r, M)$. Note that if a player meets an infected opponent, then not playing the Nash action involves a minimal loss of l ; and if he meets an uninfected opponent, his maximal gain relative to playing the Nash action is \bar{m} . So, we first compute the probability of meeting an uninfected player in any future period. Suppose $\lceil rM \rceil$ people are infected at the start of period t . Let $s_0 = 1 - r$.

Current period t : There are $\lceil rM \rceil$ infected players. So the probability of meeting an uninfected opponent is $(1 - r) = s_0$.

Period $t + 1$: With probability at least $1 - \hat{p}$, at least half the uninfected players got infected in period t and with probability at most \hat{p} , less than half the uninfected players got infected. So, the probability of meeting an uninfected opponent in this period is, at most, $(1 - \hat{p})\frac{s_0}{2} + \hat{p}s_0 < \frac{s_0}{2} + \hat{p}s_0$.

Period $t + 2$: Similarly, the probability of meeting an uninfected opponent in period $t + 2$ is, at most,

$$(1 - \hat{p})\left(\frac{s_0}{4} + \hat{p}\frac{s_0}{2}\right) + \hat{p}\left((1 - \hat{p})\frac{s_0}{2} + \hat{p}s_0\right) \leq \frac{s_0}{4} + 2\hat{p}\frac{s_0}{2} + \hat{p}^2s_0 = s_0\left(\frac{1}{2} + \hat{p}\right)^2$$

Period $t + \tau$: Probability of meeting an uninfected opponent in any future period $t + \tau$ is, at most, $s_0\left(\frac{1}{2} + \hat{p}\right)^\tau$.

So, we have

$$\begin{aligned} & v(r, M) \\ & \leq s_0\bar{m} - (1 - s_0)l + s_0\left(\frac{1}{2} + \hat{p}\right)\bar{m} - (1 - s_0\left(\frac{1}{2} + \hat{p}\right))l + s_0\left(\frac{1}{2} + \hat{p}\right)^2\bar{m} - (1 - s_0\left(\frac{1}{2} + \hat{p}\right)^2)l + \dots \\ & = \sum_{\tau=0}^{\infty} s_0\left(\frac{1}{2} + \hat{p}\right)^\tau\bar{m} - (1 - s_0\left(\frac{1}{2} + \hat{p}\right)^\tau)l \leq \sum_{\tau=0}^{\infty} s_0\left(\frac{1}{2} + \hat{p}\right)^\tau\bar{m} \\ & = s_0\bar{m} \sum_{\tau=0}^{\infty} \left(\frac{1}{2} + \hat{p}\right)^\tau = \bar{U}, \end{aligned}$$

Convergence of the series follows from the fact that $\frac{1}{2} + \hat{p} < 1$. Clearly, for $\bar{r} > r$, $v(\bar{r}, M) \leq v(r, M)$, and hence $v(\bar{r}, M)$ is uniformly bounded¹³, i.e., for each $\bar{r} \geq r$ and each $M \in \mathbb{N}$, $v(\bar{r}, M) \leq \bar{U}$. \square

¹³The upper bound obtained here is a very loose one. Notice that contagion proceeds at a very high rate. In the proof of Lemma 4, we showed that if the state of contagion is such that number of infected people is rM , the most likely state in the next period is $(2r - r^2)M$. This implies that the fraction of uninfected people goes down from $(1 - r)$ to $(1 - 2r + r^2) = (1 - r)^2$. More generally, if we consider contagion evolving along this path of ‘‘most likely’’ transitions for t consecutive periods, the number of uninfected people would go down to $(1 - r)^{2^t}$, i.e., the contagion spreads at a highly exponential rate. In particular, this rate of convergence is independent of M .

Proposition 7. Fix a game $G \in \mathcal{G}$. Then, there exist $\underline{r} \in (0, 1)$ and $\underline{M} \in \mathbb{N}$ such that, for each $r \geq \underline{r}$ and each $M \geq \underline{M}$, a player who gets infected very late in the game will not slow down the contagion, regardless of how patient he is.

Proof. Take $r > \frac{1}{2}$. By Lemma 4, if M is big enough, a player who gets infected late in Phase III believes that “with probability at least $1 - \frac{1}{M^2}$ at least rM people in each community are infected”. Consider a continuation strategy where he deviates and does not play the Nash action.

- i) With probability $1 - \frac{1}{M^2}$ at least rM people are infected. So, with probability at least r he meets an infected player, makes a loss of at least \underline{l} by not playing Nash, and does not slow down the contagion. With probability $1 - r$ he gains at most \bar{m} in the current period and $v(r, M)$ in the future.
- ii) With probability $\frac{1}{M^2}$, less than rM people are infected, and the player’s gain is at most \bar{m} in the current period and at most, $v(M)$ in the future.

Hence, by Lemma 6, the gain from not playing the Nash action instead of doing so is bounded above by:

$$\frac{\bar{m} + v(M)}{M^2} + \left(1 - \frac{1}{M^2}\right) (-rl + (1-r)(\bar{m} + v(M, r))) < \frac{M + \bar{m}}{M^2} + \left(1 - \frac{1}{M^2}\right) (-rl + (1-r)(\bar{m} + \bar{U})).$$

The inequality follows from the facts that $v(M)$ is finite and increases slower than the rate of M and that \bar{U} is a uniform bound for $v(r, M)$ for any $r > \frac{1}{2}$. If M is large enough and r is close to 1, the expression becomes negative. So there is no incentive to slow down the contagion. \square

Once \underline{M} is chosen, we pick \dot{T} . \ddot{T} is chosen large enough so that a buyer who is infected in Phase I and knows that not all buyers were infected by the end of Phase I still has an incentive to play B_L in Phase II. This buyer knows that contagion will spread from Phase III anyway, and playing B_L gives him a short-term gain in Phase II. If \ddot{T} is long enough, she will want to play B_L in Phase II. Because of the finiteness of $v(M)$, we can pick \ddot{T} such that the incentive constraint holds even for a perfectly patient buyer.

Next, we choose \dot{T} . \dot{T} must be chosen large enough so that i) a buyer infected in Phase I who has observed Q_H in most periods of Phase I believes that with high probability all buyers were infected during Phase I, ii) a seller infected in Phase II believes that with high probability at least $M - 1$ buyers were infected during Phase I, iii) a seller who deviates in

Phase I believes that, with high probability, he met all the buyers in Phase I, and iv) players infected in Phase III believe with high probability that “enough” people were infected by the end of Phases I and II.

Finally, once \underline{M} , \ddot{T} , and \dot{T} have been chosen, we find the threshold $\underline{\delta}$ such that for discount factors $\delta > \underline{\delta}$, players will not deviate on-path.

5 Discussion and Extensions

The main contribution of this paper lies in showing that, in the repeated anonymous random matching setting, it is possible to sustain a wide range of payoffs in equilibrium in a class of games much beyond the PD. Moreover, we show that the ideas of community enforcement can be applied to sustain such cooperation. The main goal of this section is to discuss the versatility of the trust-building ideas we use here. We provide intuition for the way in which our approach might be used to get more general results such as a Nash-threats folk theorem, or similar results when the role of each player is randomly assigned in each period. We present these as conjectures, as the formal proofs would require analysis similar to what we have already developed without adding new insights.

Note that throughout the paper we use symmetric strategies, and so get symmetric payoffs for players within a community. In the discussion below, when we talk about the set of equilibrium payoffs, we restrict attention to such symmetric payoffs, i.e., we do not consider feasible payoff vectors where players of the same community get different payoffs.

5.1 Can we get a Folk Theorem?

Our strategies do not suffice to get a folk theorem for all games in \mathcal{G} . For a game $G \in \mathcal{G}$ with strict Nash equilibrium a^* , the set of equilibrium payoffs, F_{a^*} does not include action profiles where only one player is playing the Nash action a_i^* . For instance, in the product choice game, this means we cannot achieve payoffs close to $(1 + g, -l)$ or $(-l, 1 - c)$.

However, we believe that the concept of trust-building that we develop is powerful enough to take us farther. We conjecture that it is possible to obtain a Nash threats folk theorem for two-player games by modifying our strategies with the addition of further trust-building phases. We do not prove a folk theorem here, but hope that the informal argument will illustrate how this may be done.

To fix ideas, let us restrict attention to the product choice game. Consider a feasible

and individually rational target payoff that can be achieved by playing short sequences of (Q_H, B_H) (10% of the time) alternating with longer sequences of (Q_H, B_L) (90% of the time). It is not possible to sustain this payoff in Phase III with our strategies. To see why not, consider a long time window in Phase III where the prescribed action profile is (Q_H, B_L) . Suppose a buyer faces Q_L for the first time in a period of this phase followed by many periods of Q_H . Notice that since the action for a buyer is B_L in this time window, she cannot infect any sellers herself. So, with more and more observations of Q_H , she will ultimately get convinced that few people are infected. So, it may not be optimal to keep playing Nash any more. Contrast this with the original situation where the target action is (Q_H, B_H) . In that case, a player who gets infected starts infecting players himself and so at most, after $M - 1$ periods of infecting opponents, he is convinced that everyone is infected.

What modification to our strategies might enable us to attain these payoffs? We can use additional trust-building phases. Consider a target payoff phase that involves alternating sequences of (Q_H, B_L) for T_1 periods and (Q_H, B_H) for $T_2 = \frac{1}{9}T_1$ periods. In the modified equilibrium strategies, in Phase III, the windows of (Q_H, B_L) and (Q_H, B_H) will be separated by trust-building phases. To illustrate, we start the game as before, with two phases: \dot{T} periods of (Q_H, B_H) and \ddot{T} periods of (Q_L, B_H) . In Phase III, players play the action profile (Q_H, B_L) for T_1 periods, followed by a new trust-building phase of T' periods during which (Q_L, B_H) is played. Then players switch to playing the sequence of (Q_H, B_H) for T_2 periods. The new phase is chosen to be short enough (*i.e.*, $T' \ll T_1$) to have no significant payoff consequences. Yet, it is chosen long enough so that a player who is infected during the T_1 period window but thinks that very few people are infected, will still want to revert to Nash punishments to make short-term gains during the new phase.¹⁴ We conjecture that adding appropriate trust-building phases in the target payoff phase in this way can guarantee that players have the incentive to revert to Nash punishments off-path for any beliefs they may have about the number of people infected.

¹⁴For example, think of a buyer who observes a triggering action for the first time in Phase III (while playing (Q_H, B_L)) and then observes only good behavior for a long time while continuing to play (Q_H, B_L) . Even if this buyer is convinced that very few people are infected, she knows that the contagion has begun, and ultimately her continuation payoff will become very low. So, if there is a long enough phase of playing (Q_L, B_H) ahead, she will choose to revert to Nash because this is the myopic best response, and would give her at least some short-term gains.

5.2 Interchangeable Populations

So far we have assumed that there are two independent communities; *i.e.*, each player belongs either to Community 1 or to Community 2. Alternatively, we could have assumed that there is one population whose members are matched in pairs in every period and, in each match, the roles of players are randomly assigned. At the start of every period, each player has an equal probability of playing as Player 1 or 2 in the stage-game.

A first implication of this alternative modeling is that a negative result like that of Proposition 1 may not be true any more.¹⁵ We believe that the trust-building ideas that underlie this paper can be adapted to this new setting.

Suppose we want to get cooperation in the repeated product-choice game when roles are randomly assigned in each period. We conjecture that the following version of our strategies can be used to get as close as desired to the efficient payoff $(1, 1)$. There are two phases. Phase I is the trust-building phase: sellers play Q_L and buyers play B_H ; the important features of this profile being that i) only buyers have an incentive to deviate and ii) sellers are playing a Nash action. Phase II is the target payoff phase and (Q_H, B_H) is played. Deviations are punished through Nash reversion; there is no delay in the punishment now. The main difference now is that contagion also takes place in Phase I; whenever an “infected” player is in the role of a buyer he will play B_L and spread the contagion, so we do not have a single player infecting people in this phase. This implies that we do not need a second trust-building phase, since its primary goal was to give the infected buyers the right incentives to “tell” the sellers that there had been a deviation.

The arguments for the incentives would be very similar to those in the setting with independent populations. After getting infected, a player would form his beliefs based on the fact that a buyer deviated in period one and that punishments have been going on ever since. Proving formally that players have the right incentives after all histories is a hard exercise for which we cannot rely on the analysis of the independent populations case. The fact that players’ roles are not fixed has two main consequences for the analysis. First, the contagion is not the same and a slightly different mathematical object would be needed to model it. Second, the set of histories a player may have observed would depend on the roles he played in the past periods, so it is harder to characterize all possible histories. We think that this exercise would not add any new economic insights, and so leave it as a conjecture.

¹⁵A buyer infected in period 1 might become a seller in period 2 and he might indeed have the right incentives to punish.

5.3 Alternative Systems of Beliefs

We assume that a player who observes a triggering action believes that some player from Community 1 deviated in the first period of the game. This ensures that an infected player thinks that the contagion has been spreading long enough that, after Phase I, almost everybody is infected. It is easy to see that alternate (less extreme) assumptions on beliefs would still have delivered this property. We work with this case mainly for tractability. Also, since our equilibrium is based on communities building trust in the initial phases of the game, it is plausible that players regard deviations to be more likely earlier rather than later.

Further, the assumption we make is a limiting one in the sense that it yields the weakest bound on M . With other assumptions, for a given game $G \in \mathcal{G}$ and given \dot{T} and \ddot{T} , the threshold population size \underline{M} required to sustain cooperation would be weakly greater than the threshold we obtain. Why is this so? On observing a triggering action, my belief about the number of infected people is determined by two factors: my belief about when the first deviation took place and the subsequent contagion process (described by the matrices of transition probabilities). Formally, on getting infected at period t , my belief x^t can be expressed $x^t = \sum_{\tau=1}^t \mu(\tau) y^t(\tau)$, where $\mu(\tau)$ is the probability I assign to the first deviation having occurred at period τ and $y^t(\tau)$ is my belief about the number of people infected if I know that the first deviation took place at period τ . Since contagion is not reversible, every elapsed period of contagion results in a weakly greater number of infected people. Thus, my belief if I think the first infection occurred at $t = 1$, first order stochastically dominates my belief if I think the first infection happened later, at any $t > 1$, *i.e.*, For each τ , for each $k \in \{1, \dots, M\}$, $\sum_{i=k}^M y_i^t(\tau) \geq \sum_{i=k}^M y_i^t(1)$. Now consider any belief \tilde{x}^t that I might have had, with alternate assumptions on when I think the first deviation occurred. This belief will be some convex combination of $y^t(\tau)$, for $\tau = 1, \dots, t$. Since we know that $y^t(1)$ first order stochastically dominates $y^t(\tau)$ for any $\tau > 1$, it follows that $y^t(1)$ will also first order stochastically dominate \tilde{x}^t . This in turn implies that with most alternate belief assumptions we would have needed, at least, the population size to be larger in order to ensure that my limit beliefs in Phase III assigned enough weight to a large number of people being infected.

5.4 Stability and Robustness to Introduction of Noise

A desirable feature of an equilibrium could be global stability. A globally stable equilibrium is one where after any finite history, play finally reverts to cooperative play (Kandori

(1992)). The notion is appealing because it implies that a single mistake does not entail permanent reversion to punishments. The equilibrium here fails to satisfy this property. However, global stability can be obtained if a public randomization device is introduced. This is similar to Ellison (1994). The role of the randomization device would be to allow for the possibility of restarting the game in any period, with a low but positive probability.

A related question is to see if the equilibrium can be sustained in a model with some noise. First note that since players have strict incentives in equilibrium, our strategies are robust to the introduction of some noise in the payoffs. However, if we consider a setting where players make mistakes, or there is noisy observation of one's opponents' actions, our equilibrium is no longer robust. Consider a setting where players are constrained to make mistakes (play an off-equilibrium action) with probability at least $\varepsilon > 0$ at every possible history. We can ask if the equilibrium survives for small ε . Our construction is not robust to this modification. The incentive compatibility of our strategies crucially relies on the fact that players believe that early deviations are more likely. If players make mistakes with positive and equal probability in all periods, this property is lost. To see a particularly problematic case, consider the following situation in the product choice game in a setting with noise. If a buyer makes a mistake late in Phase II, no matter what she does after that, she will start phase III knowing that not many people are already infected. Hence, if she is very patient, it may be optimal for her to continue play as if on equilibrium path and slow down the contagion. Suppose a seller observes a triggering action in the last period of Phase II. This seller will think that, it is very likely that his opponent was uninfected and has just made a mistake, and so will not punish. In this case, neither player reverts to Nash punishments. This implies that a buyer may profitably deviate in the last period of Phase II, since her deviation would go unpunished.

5.5 Uncertainty about Calendar Time

In the equilibrium strategies here, players condition behavior on calendar time. On-path, players in Community 1 switch their action in a coordinated way at the end of Phases I and II. Off-path, players coordinate the start of the punishment phase. The calendar time and timing of phases (\dot{T} and \ddot{T}) are commonly known and used to coordinate behavior. Arguably, in modeling large communities, the need to switch behavior with precise coordination is an unappealing feature. It may be interesting to investigate if cooperation can be sustained if players were not sure about calendar time or the precise time to switch actions.

A complete analysis of this issue is beyond the scope of this paper, but we conjecture that a modification of our strategies would be robust to the introduction of *small* uncertainty about timing. The reader may refer to the Appendix Section A.4, where we consider an altered environment in which players are slightly uncertain about the timing of the different phases. We conjecture equilibrium strategies in this setting, and provide the main intuition behind why the efficient payoff might still be achieved.

References

- CRIPPS, M. W., G. J. MAILATH, AND L. SAMUELSON (2004): “Imperfect Monitoring and Impermanent Reputations,” *Econometrica*, 72, 407–432.
- DAL BÓ, P. (2007): “Social norms, cooperation and inequality,” *Economic Theory*, 30, 89–105.
- DEB, J. (2008): “Cooperation and Community Responsibility: A Folk Theorem for Random Matching Games with Names,” Mimeo.
- ELLISON, G. (1994): “Cooperation in the Prisoner’s Dilemma with Anonymous Random Matching,” *Review of Economic Studies*, 61, 567–88.
- ELY, J. C., J. HÖRNER, AND W. OLSZEWSKI (2005): “Belief-Free Equilibria in Repeated Games,” *Econometrica*, 73, 377–415.
- ELY, J. C. AND J. VÄLIMÄKI (2002): “A Robust Folk Theorem for the Prisoner’s Dilemma,” *Journal of Economic Theory*, 102, 84–105.
- HARSANYI, J. (1973): “Oddness of the number of equilibrium points: A new proof,” *International Journal of Game Theory*, 2, 235–250.
- HASKER, K. (2007): “Social norms and choice: a weak folk theorem for repeated matching games,” *International Journal of Game Theory*, 36, 137–146.
- HÖRNER, J. AND W. OLSZEWSKI (2006): “The folk theorem for games with private almost-perfect monitoring,” *Econometrica*, 74, 1499–1544.
- KANDORI, M. (1992): “Social Norms and Community Enforcement,” *Review of Economic Studies*, 59, 63–80.

MAILATH, G. J. AND L. SAMUELSON (2006): *Repeated Games and Reputations: Long-Run Relationships*, Oxford University Press.

OKUNO-FUJIWARA, M. AND A. POSTLEWAITE (1995): “Social Norms and Random Matching Games,” *Games and Economic Behavior*, 9, 79–109.

PICCIONE, M. (2002): “The Repeated Prisoner’s Dilemma with Imperfect Private Monitoring,” *Journal of Economic Theory*, 102, 70–83.

TAKAHASHI, S. (2007): “Community enforcement when players observe partners’ past play,” Mimeo.

A Appendix

A.1 Properties of the Conditional Transition Matrices

In Section 4.2 we introduced a class of matrices, contagion matrices, which turns out to be very useful in analyzing the beliefs of players. First note that, since contagion matrices are upper triangular, their eigenvalues correspond with the diagonal entries. Given $x \in \mathbb{R}^k$, let $\|x\| := \sum_{i \in \{1, \dots, k\}} x_i$. We are often interested in the limit behavior of $x^t := \frac{x C^t}{\|x C^t\|}$, where C is a contagion matrix and x is a probability vector. We present below a few results about this limit behavior. We distinguish three special types of contagion matrices that will deliver different limiting results.

Property C1: $\{c_{11}\} = \operatorname{argmax}_{i \in \{1, \dots, k\}} c_{ii}$.

Property C2: $c_{kk} \in \operatorname{argmax}_{i \in \{1, \dots, k\}} c_{ii}$.

Property C3: For each $l < k$, $C_{[l]}$ satisfies C1 or C2.

Lemma A.1. *Let C be a contagion matrix and let λ be its largest eigenvalue. Then, the left eigenspace associated with λ has dimension 1. That is, the geometric multiplicity of λ is one, irrespective of its algebraic multiplicity.*

Proof of Lemma A.1. Let l be the largest index such that $c_{ll} = \lambda > 0$ and let x be a left eigenvector associated with λ . We claim that, for each $i < l$, $x_i = 0$. Suppose not and let i be the largest index smaller than l such that $x_i \neq 0$. If $i < l - 1$, we have that $x_{i+1} = 0$ and, since $c_{i,i+1} > 0$, we get $(xC)_{i+1} > 0$, which contradicts that x is an eigenvector associated with λ . If $i = l - 1$, then $(xC)_l \geq c_{ll}x_l + c_{l-1,l}x_{l-1} > c_{ll}x_l = \lambda x_l$, which, again, contradicts that x is an eigenvector associated with λ . Then, we can restrict attention to matrix $C_{[(l-1)]}$. Now, also λ is the largest eigenvalue of $C_{[(l-1)]}$ but, by definition of l , only one diagonal entry of $C_{[(l-1)]}$ equals λ and, hence, its multiplicity is one. Then, $y \in \mathbb{R}^{k-(l-1)}$ is a left eigenvector associated with λ for matrix $C_{[(l-1)]}$ if and only if $(0, \dots, 0, y) \in \mathbb{R}^k$ is a left eigenvector associated with λ for matrix C . \square

Given a contagion matrix C with largest eigenvalue λ , we denote by \check{x} the unique left eigenvector associated with λ such that $\|\check{x}\| = 1$.

Proposition A.1. *Let $C \in \mathcal{M}_k$ be a contagion matrix satisfying C1 or C2. Then, for each nonnegative vector $x \in \mathbb{R}^k$ with $x_1 > 0$, we have $\lim_{t \rightarrow \infty} \frac{x C^t}{\|x C^t\|} = \check{x}$. In particular, under C2, $\check{x} = (0, \dots, 0, 1)$.*

Proof of Proposition A.1. Clearly, since C is a contagion matrix, if t is large enough all the components of x^t are positive. Then, for the sake of exposition, we assume that all the components of x are positive. We distinguish two cases.

C satisfies C1. In this case λ has multiplicity 1. We show that, for each pair $i, j \in \{1, \dots, k\}$, $\lim_{t \rightarrow \infty} \frac{x_i^t}{x_j^t} = \frac{\check{x}_i}{\check{x}_j}$. Once this is established, the result immediately follows from the fact that, for each $t \in \mathbb{N}$, $\|x^t\| = \|\check{x}\| = 1$. We already know that, $\check{x}C = \lambda\check{x}$, where λ is the largest eigenvalue of C . Then, the vector x can be written as $x = \alpha\check{x} + v$, where v is a vector orthogonal to \check{x} . Since \check{x} is a nonnegative vector different from 0 and all the components of x are positive, \check{x} and x are not orthogonal. Hence, $\alpha > 0$. Then,

$$\frac{x_i^t}{x_j^t} = \frac{(xC^t)_i / \|xC^t\|}{(xC^t)_j / \|xC^t\|} = \frac{(xC^t)_i}{(xC^t)_j} = \frac{\lambda^t \alpha \check{x}_i + (vC^t)_i}{\lambda^t \alpha \check{x}_j + (vC^t)_j} = \frac{\alpha \check{x}_i + (v(\frac{1}{\lambda}C)^t)_i}{\alpha \check{x}_j + (v(\frac{1}{\lambda}C)^t)_j}$$

Since λ is the largest eigenvalue of C and has multiplicity one, as $t \rightarrow \infty$, the second terms in both the numerator and denominator vanish. Then the limit as $t \rightarrow \infty$ is $\frac{\check{x}_i}{\check{x}_j}$.

C satisfies C2. We show that, for each $i < k$, $\lim_{t \rightarrow \infty} x_i^t = 0$. We prove this by induction on i . Let $i = 1$. Then, for each $t \in \mathbb{N}$,

$$\frac{x_1^{t+1}}{x_k^{t+1}} = \frac{c_{11}x_1^t}{\sum_{l \leq k} c_{lk}x_l^t} < \frac{c_{11}x_1^t}{c_{kk}x_k^t} \leq \frac{x_1^t}{x_k^t},$$

where the first inequality follows from the facts that $x_{k-1} > 0$ and $c_{k-1,k} > 0$ (C is a contagion matrix); the second inequality follows from C2. Hence, the ratio $\frac{x_1^t}{x_k^t}$ is strictly decreasing in t . Moreover, since all the components of x^t lie in $[0, 1]$, it is not hard to see that, as far as x_1^t is bounded away from 0, the speed at which the above ratio decreases is also bounded away from 0.¹⁶ Therefore, we must have $\lim_{t \rightarrow \infty} x_1^t = 0$. Suppose the claim holds for each $i < j < k - 1$. Now,

$$\frac{x_j^{t+1}}{x_k^{t+1}} = \frac{\sum_{l \leq j} c_{lj}x_l^t}{\sum_{l \leq k} c_{lk}x_l^t} < \frac{\sum_{l \leq j} c_{lj}x_l^t}{c_{kk}x_k^t} = \sum_{l < j} \frac{c_{lj}}{c_{kk}} \frac{x_l^t}{x_k^t} + \frac{c_{jj}}{c_{kk}} \frac{x_j^t}{x_k^t} \leq \sum_{l < j} \frac{c_{lj}}{c_{kk}} \frac{x_l^t}{x_k^t} + \frac{x_j^t}{x_k^t}.$$

By the induction hypothesis, for each $l < j$, the term $\frac{x_l^t}{x_k^t}$ can be made arbitrarily small for large enough t . Then, the first term in the above expression can be made arbitrarily small.

¹⁶Roughly speaking, this is because the state k will always get some probability from state 1 via the intermediate states, and this probability will be bounded away from 0 as far as the probability of state 1 is bounded away from 0.

Hence, it is easy to see that, for large enough t , the ratio $\frac{x_j^t}{x_k^t}$ is strictly decreasing in t . As above, this can only happen if $\lim_{t \rightarrow \infty} x_j^t = 0$. \square

Recall the matrices used to represent a player's beliefs after he observes history $h^t = g \dots gb$. At the beginning of Phase III, the beliefs evolved according to matrices C_{1j} and S_{1j} , and late in Phase III, according to \hat{C}_{1j} . Note that these three matrices all satisfy the conditions of the above proposition. This is what drives Lemmas 1 and 2 in the text. Consider the truncated matrix \hat{C}_{2j} that gave the transition of beliefs of a player who observes history $h^t = g \dots bg$. This matrix also satisfies the conditions of the above proposition and this suffices for Lemma 3.

Proposition A.2. *Let $C \in \mathcal{M}_k$ be a contagion matrix satisfying C1 and C3. Let $x \in \mathbb{R}^k$ be a nonnegative vector. Then, if x is close enough to $(0, \dots, 0, 1)$, we have that, for each $t \in \mathbb{N}$ and each $l \in \{1, \dots, k\}$, $\sum_{i=l}^k x_i^t \geq \sum_{i=l}^k \check{x}_i^t$.*

Whenever two vectors are as x^t and \check{x} above, we say that x^t *first order stochastically dominates* \check{x} (in the sense of more people being infected).

Proof of Proposition A.2. For each $i \in \{1 \dots, k\}$, let e_i denote the i -th element of the canonical basis in \mathbb{R}^k . By C1, c_{11} is larger than any other diagonal entry of C . Let \check{x} be the unique left eigenvector associated with c_{11} such that $\|\check{x}\| = 1$. Clearly, $\check{x}_1 > 0$ and, hence, $\{\check{x}, e_2, \dots, e_k\}$ is a basis in \mathbb{R}^k . With respect to this basis, the matrix C looks like

$$\left(\begin{array}{c|c} c_{11} & 0 \\ \hline 0 & C_{\bar{1}} \end{array} \right).$$

Now we distinguish two cases.

$C_{\bar{1}}$ satisfies C2. In this case we can apply Proposition A.1 to $C_{\bar{1}}$ to get that, for each nonnegative vector $y \in \mathbb{R}^{k-1}$ with $y_1 > 0$, $\lim_{t \rightarrow \infty} \frac{y C_{\bar{1}}^t}{\|y C_{\bar{1}}^t\|} = (0, \dots, 0, 1)$. Now, let $x \in \mathbb{R}^k$ be the vector in the statement of this result. Since x is very close to $(0, \dots, 0, 1)$. Then, using the above basis, it is clear that $x = \alpha \check{x} + v$, with $\alpha \geq 0$ and $v \approx (0, \dots, 0, 1)$. Let $t \in \mathbb{N}$. Then, for each $t \in \mathbb{N}$,

$$x^t = \frac{x C^t}{\|x C^t\|} = \frac{\lambda^t \alpha \check{x} + v C^t}{\|x C^t\|} = \frac{\lambda^t \alpha \check{x} + \|v C^t\| \frac{v C^t}{\|v C^t\|}}{\|x C^t\|}.$$

Clearly, $\|x^{C^t}\| = \|\lambda^t \alpha \check{x} + \|vC^t\| \frac{vC^t}{\|vC^t\|}\|$ and, since all the terms are positive,

$$\|x^{C^t}\| = \|\lambda^t \alpha\| \|\check{x}\| + \|vC^t\| \left\| \frac{vC^t}{\|vC^t\|} \right\| = \|\lambda^t \alpha\| + \|vC^t\|$$

and, hence, we have that x^t is a convex combination of \check{x} and $\frac{vC^t}{\|vC^t\|}$. Since $v \approx (0, \dots, 0, 1)$ and $\frac{vC^t}{\|vC^t\|} \rightarrow (0, \dots, 0, 1)$, it is clear that, for each $t \in \mathbb{N}$, $\frac{vC^t}{\|vC^t\|}$ first order stochastically dominates \check{x} in the sense of more people being infected. Therefore, also x^t will first order stochastically dominate \check{x} .

$C_{\bar{1}}$ satisfies C1. By C1, the first diagonal entry of $C_{\bar{1}}$ is larger than any other diagonal entry. Let \check{x}^2 be the unique associated left eigenvector such that $\|\check{x}^2\| = 1$. It is easy to see that \check{x}^2 first order stochastically dominates \check{x} ; the reason is that \check{x}^2 and \check{x} are the limit of the same contagion process with the only difference that the state in which only one person has been infected is known to have probability 0 when using obtaining \check{x}^2 from $C_{\bar{1}}$. Clearly, $\check{x}_2^2 > 0$ and, hence, $\{\check{x}, \check{x}^2, e_3, \dots, e_k\}$ is a basis in \mathbb{R}^k . With respect to this basis, the matrix C looks like

$$\left(\begin{array}{c|c|c} c_{11} & 0 & 0 \\ \hline 0 & c_{22} & 0 \\ \hline 0 & 0 & C_{\bar{2}} \end{array} \right).$$

Again, we can distinguish two cases.

- $C_{\bar{2}}$ satisfies C2. In this case we can repeat the arguments above to show that x^t is a convex combination of \check{x} , \check{x}^2 and $\frac{vC^t}{\|vC^t\|}$. Since both \check{x}^2 and $\frac{vC^t}{\|vC^t\|}$ first order stochastically dominate \check{x} , also x^t does.
- $C_{\bar{2}}$ satisfies C1. Now we would get a vector \check{x}^3 and the procedure would continue until a truncated matrix satisfies C2 or until we get a basis of eigenvectors, one of them being \check{x} and all the others first order stochastically dominate \check{x} . In both situations the result immediately follows from the above arguments. \square

Note that the matrix $\hat{C}_{\bar{2}}$ which gave the transition of beliefs of a player conditional on history $h^t = g \dots gbg$ late in the game, satisfies the conditions of the above proposition. This property is useful in proving Proposition 6.

A.2 Updating of Beliefs Conditional of Observed Histories

Suppose player i observes history $h^{\bar{t}+1} = g \dots gbg$ in Phase III, and we want to compute her beliefs at period $\bar{t} + 1$ conditional on $h^{\bar{t}+1}$, namely $x^{\bar{t}+1}$. Recall our method for computing $x^{\bar{t}+1}$. We first compute a set of intermediate beliefs x^t for $t < \bar{t} + 1$. For any period τ , we compute $x^{\tau+1}$ from x^τ by conditioning on the event that “I was uninfected in period $\tau + 1$ ” and that “ $\mathcal{I}^{\tau+1} \leq M - 2$ ” (\mathcal{I}^t is the random variable representing the number of infected people after period t). We do not use the information that “I remained uninfected after any period t with $\tau + 1 < t < \bar{t}$ ”. This information is added later period by period, *i.e.*, only at period t we add the information coming from the fact that “I was not infected at period t ”. Below, we show that this method of computing beliefs is equivalent to the standard updating of beliefs conditioning on the entire history at once.

Let $\alpha \in \{0, \dots, M-2\}$ and let $h^{t+1+\alpha}$ denote the $(t+1+\alpha)$ -period history $g \dots gbg \dots^\alpha g$. Recall that U^t denotes the event that i is uninfected at the end of period t . Let $b_t(g_t)$ denote the event that player i faced $b(g)$ in period t . We introduce some additional notation.

- $I_{i,k}^{(t)}$ denotes the event $i < \mathcal{I}^t < k$, *i.e.*, the number of infected people at the end of t periods is at least i and at most k .
- $E_\alpha^t := I_{0, M-\alpha}^t \cap U^t$
- $E_\alpha^{t+1} := E_\alpha^t \cap I_{1, M-\alpha+1}^{t+1} \cap b^{t+1}$
- For each $\beta \in \{1, \dots, \alpha - 1\}$,

$$E_\alpha^{t+1+\beta} := E_\alpha^{t+\beta} \cap I_{\beta+1, M-\alpha+\beta+1}^{t+1+\beta} \cap g^{t+1+\beta}$$
- $E_\alpha^{t+1+\alpha} := E_\alpha^{t+\alpha} \cap g^{t+1+\alpha} = h^{t+1+\alpha}$.

Let H^t be a history of the contagion process up to period t . Let \mathcal{H}^t be the set of all H^t histories. \mathcal{H}_k^t denotes the set of t -period histories of the stochastic process where $\mathcal{I}^t = k$. We say $H^{t+1} \Rightarrow h^{t+1}$ if history H^{t+1} implies that I observed history h^{t+1} . Let $i \xrightarrow{t+1+\beta} k$ denote the event that state i transits to state k in period $t + 1 + \beta$, consistently with $h^{t+1+\alpha}$, or equivalently, consistently with $E_\alpha^{t+1+\beta}$.

The probabilities of interest are $P(\mathcal{I}^{t+1+\alpha} = k |_{h^{t+1+\alpha}}) = P(\mathcal{I}^{t+1+\alpha} = k |_{E_\alpha^{t+1+\alpha}})$. We want to show that we can obtain the probabilities after $t + 1 + \alpha$ conditional on $h^{t+1+\alpha}$ by starting with the probabilities after t conditional on E_α^t and then let the contagion elapse one more period at a time conditioning on the new information, *i.e.*, adding the “local”

information that player i observed g in the next period and that infected one more person. Precisely, we want to show that, for each $\beta \in \{0, \dots, \alpha\}$,

$$P(\mathcal{I}^{t+1+\beta} = k |_{E_\alpha^{t+1+\beta}}) \stackrel{?}{=} \frac{\sum_{i=1}^M P(i \xrightarrow{t+1+\beta} k) P(\mathcal{I}^{t+\beta} = i |_{E_\alpha^{t+\beta}})}{\sum_{j=1}^M \sum_{i=1}^M P(i \xrightarrow{t+1+\beta} j) P(\mathcal{I}^{t+\beta} = i |_{E_\alpha^{t+\beta}})}.$$

Fix $\beta \in \{0, \dots, \alpha\}$. For each $H^{t+1+\beta} \in \mathcal{H}^{t+1+\beta}$, let $H^{t+1+\beta, \beta}$ denote the unique $H^{t+\beta} \in \mathcal{H}^{t+\beta}$ that is compatible with $H^{t+1+\beta}$, *i.e.*, the restriction of $H^{t+1+\beta}$ to the first $t + \beta$ periods. Let $F^{1+\beta} := \{\tilde{H}^{t+1+\beta} \in \mathcal{H}^{t+1+\beta} : \tilde{H}^{t+1+\beta} \Rightarrow E_\alpha^{t+1+\beta}\}$. Let $F_k^{1+\beta} := \{\tilde{H}^{t+1+\beta} \in F^{1+\beta} : \tilde{H}^{t+1+\beta} \in \mathcal{H}_k^{t+1+\beta}\}$. Clearly, the $F_k^{1+\beta}$ sets define a ‘‘partition’’ of $F^{1+\beta}$ (one or more sets in the partition might be empty). Let $F_k^\beta := \{\tilde{H}^{t+1+\beta} \in F^{1+\beta} : \tilde{H}^{t+1+\beta, \beta} \in \mathcal{H}_k^{t+\beta}\}$. Clearly, also the F_k^β sets define a ‘‘partition’’ of $F^{1+\beta}$. Note that, for each pair $H^{t+1+\beta}, \tilde{H}^{t+1+\beta} \in F_k^{1+\beta} \cap F_i^\beta$, $P(H^{t+1+\beta} |_{H^{t+1+\beta, \beta}}) = P(\tilde{H}^{t+1+\beta} |_{\tilde{H}^{t+1+\beta, \beta}})$. Denote this probability by $P(F_i^\beta \xrightarrow{t+1+\beta} F_k^{1+\beta})$. Let $|i \xrightarrow{t+1+\beta} k|$ denote the number of ways in which i can transition to k at period $t + 1 + \beta$ consistently with $h^{t+1+\alpha}$ or, equivalently, consistently with $E_\alpha^{t+1+\beta}$. Clearly, this number is independent of the history that led to i people being infected. Now, $P(i \xrightarrow{t+1+\beta} k) = P(F_i^\beta \xrightarrow{t+1+\beta} F_k^{1+\beta}) |i \xrightarrow{t+1+\beta} k|$. Then,

$$\begin{aligned} & P(\mathcal{I}^{t+1+\beta} = k |_{E_\alpha^{t+1+\beta}}) = \\ &= \sum_{H^{t+1+\beta} \in \mathcal{H}_k^{t+1+\beta}} P(H^{t+1+\beta} |_{E_\alpha^{t+1+\beta}}) = \sum_{H^{t+1+\beta} \in F_k^{1+\beta}} P(H^{t+1+\beta} |_{E_\alpha^{t+1+\beta}}) \\ &= \sum_{H^{t+1+\beta} \in F_k^{1+\beta}} \frac{P(H^{t+1+\beta} \cap E_\alpha^{t+1+\beta})}{P(E_\alpha^{t+1+\beta})} = \frac{1}{P(E_\alpha^{t+1+\beta})} \sum_{H^{t+1+\beta} \in F_k^{1+\beta}} P(H^{t+1+\beta}) \\ &= \frac{1}{P(E_\alpha^{t+1+\beta})} \sum_{i=1}^M \sum_{H^{t+1+\beta} \in F_k^{1+\beta} \cap F_i^\beta} P(H^{t+1+\beta}) \\ &= \frac{1}{P(E_\alpha^{t+1+\beta})} \sum_{i=1}^M \sum_{H^{t+1+\beta} \in F_k^{1+\beta} \cap F_i^\beta} P(H^{t+1+\beta} |_{H^{t+1+\beta, \beta}}) P(H^{t+1+\beta, \beta} |_{E_\alpha^{t+\beta}}) P(E_\alpha^{t+\beta}) \\ &= \frac{P(E_\alpha^{t+\beta})}{P(E_\alpha^{t+1+\beta})} \sum_{i=1}^M P(F_i^\beta \xrightarrow{t+1+\beta} F_k^{1+\beta}) \sum_{H^{t+1+\beta} \in F_k^{1+\beta} \cap F_i^\beta} P(H^{t+1+\beta, \beta} |_{E_\alpha^{t+\beta}}), \end{aligned}$$

which equals

$$\begin{aligned}
&= \frac{P(E_\alpha^{t+\beta})}{P(E_\alpha^{t+1+\beta})} \sum_{i=1}^M P(F_i^{\beta \rightarrow t+1+\beta} F_k^{1+\beta} | i \xrightarrow{t+1+\beta} k) \sum_{H^{t+\beta} \in \mathcal{H}_i^{t+\beta}} P(H^{t+\beta} | E_\alpha^{t+\beta}) \\
&= \frac{P(E_\alpha^{t+\beta})}{P(E_\alpha^{t+1+\beta})} \sum_{i=1}^M P(F_i^{\beta \rightarrow t+1+\beta} F_k^{1+\beta} | i \xrightarrow{t+1+\beta} k) P(\mathcal{I}^{t+\beta} = i | E_\alpha^{t+\beta}) \\
&= \frac{P(E_\alpha^{t+\beta})}{P(E_\alpha^{t+1+\beta})} \sum_{i=1}^M P(i \xrightarrow{t+1+\beta} k) P(\mathcal{I}^{t+\beta} = i | E_\alpha^{t+\beta})
\end{aligned}$$

It is easy to see that $P(E_\alpha^{t+1+\beta}) = \sum_{j=1}^M P(E_\alpha^{t+\beta}) \sum_{i=1}^M P(i \xrightarrow{t+1+\beta} j) P(\mathcal{I}^{t+\beta} = i | E_\alpha^{t+\beta})$ and the result follows.

Similar arguments apply to histories $h^{t+1+\alpha} = g \dots g b g \dots^\alpha$. where player i observes both g and b in the α periods following the first triggering action.

A.3 Sequential Equilibrium - Consistency of Beliefs

In the construction of the equilibrium, we focused on sequential rationality of strategies. Below we prove the consistency of beliefs. Recall our two assumptions on beliefs.

- i) **Assumption 1:** If a player observes a triggering action, then this player believes that some player in Community 1 (seller) deviated in the first period of the game, and after that, play has proceeded as prescribed by the strategies.
- ii) **Assumption 2:** If a player observes a non-triggering action (in Phase 1), then this player believes that his opponent made a mistake.
- iii) **Assumption 3:** If a player observes a history that is not consistent with either of the above beliefs (*erroneous history*), he will think that the first triggering action was by a seller in the first period of the game, and further, one or more of his opponents in his earlier matches made a mistake. Indeed, this player will think that there have been as many mistakes by his past rivals as needed to explain the history at hand. Erroneous histories include the following:

- A player observes an action in Phase II or III, that is neither part of the strict Nash profile a^* , nor an action that can be played in equilibrium.

- A player who, after being certain that all the players in the other community are infected, faces an opponent who does not play the Nash action (this can only happen in Phase III).

Proof. Fix any player i . Perturb the equilibrium strategies as follows. Fix $\varepsilon > 0$ small. In any period t of the game, each player plays the prescribed equilibrium action with probability $(1 - \varepsilon^t)$, and plays the wrong action with probability ε^t . We need to show that, given any t -period private history off-path for player i , as perturbations vanish ($\varepsilon \rightarrow 0$), the strategies converge to the prescribed equilibrium, and player i believes that, with probability 1, the first deviation occurred at $t = 1$. Moreover, we require that this convergence in beliefs be uniform in t .

Consider a history h^{t^*} , late in Phase III ($t^* \gg \dot{T} + \ddot{T}$) such that player i observes the first triggering action at time t^* , *i.e.*, $h^{t^*} = g \dots gb$. Denote any sequence of matches up to period t^* by H^{t^*} . We say $H^{t^*} \implies h^{t^*}$, to mean that the sequence of matches H^{t^*} is consistent with history h^{t^*} being observed. Further, let $H^{t^*}(\tau)$ denote a realization of the matching technology, that is consistent with the observed history $h^{t^*} = g \dots gb$, and where the first triggering action occurred at period τ . Clearly, there exists a corresponding event (sequence of matches), denoted by $\tilde{H}^{t^*}(\tau)$, that satisfies the following:

- The first triggering action occurred at $t = 1$,
- The two players who got infected at period $t = 1$ were matched to each other in each period until τ , and
- The realized matches in $\tilde{H}^{t^*}(\tau)$ and $H^{t^*}(\tau)$ are the same from period τ until t^* .

We first show that, conditional on observed history h^{t^*} , player i assigns arbitrarily higher probability to the event $\tilde{H}^{t^*}(\tau)$ compared to the event $H^{t^*}(\tau)$.

$$\frac{P(\tilde{H}^{t^*}(\tau) | h^{t^*})}{P(H^{t^*}(\tau) | h^{t^*})} = \frac{P(\tilde{H}^{t^*}(\tau) \cap h^{t^*})}{P(H^{t^*}(\tau) \cap h^{t^*})} = \frac{\varepsilon(1 - \varepsilon)^{M-1} \frac{M-1}{M} \left[\prod_{k=2}^{\tau} \frac{1}{M} (1 - \varepsilon^k)^M \right] X}{(1 - \varepsilon)^M \left[\prod_{k=2}^{\tau-1} (1 - \varepsilon^k)^M \right] \varepsilon^{\tau} (1 - \varepsilon^{\tau})^{M-1} \frac{1}{M} X},$$

where X is the probability of the event (matches) that was realized from period τ until t^* in the events $\tilde{H}^{t^*}(\tau)$ and $H^{t^*}(\tau)$. The above expression simplifies to

$$\frac{1 - \varepsilon^{\tau}}{1 - \varepsilon} \frac{M - 1}{M} \frac{1}{M} \frac{1}{(\varepsilon M)^{\tau-1}}.$$

Clearly, for a fixed M , the above expression goes to infinity as ε goes to zero, uniformly in τ . To summarize, we have shown above that, for any possible sequence of matches $H^{t^*}(\tau)$ that is consistent with the observed history h^{t^*} and where the first triggering action occurred at some period $\tau \neq 1$, there exists a corresponding sequence of matches $\tilde{H}^{t^*}(\tau)$ which is also consistent with h^{t^*} , where the first triggering action occurred at $t = 1$, and that is arbitrarily more likely than $H^{t^*}(\tau)$. This implies in particular, that on observing a triggering action, a player will assign arbitrarily high probability to the event that the first deviation was by a seller in the first period of the game. To see why, note that

$$\frac{\text{P(First dev. at } t=1 | h^{t^*})}{\text{P(First dev. at } t \neq 1 | h^{t^*})} = \frac{\text{P(First dev. at } t=1 | h^{t^*})}{\sum_{\tau=2}^{t^*} \sum_{H^{t^*}(\tau): H^{t^*}(\tau) \Rightarrow h^{t^*}} \text{P}(H^{t^*}(\tau))} \geq \frac{\sum_{\tau=2}^{t^*} \sum_{H^{t^*}(\tau): H^{t^*}(\tau) \Rightarrow h^{t^*}} \text{P}(\tilde{H}^{t^*}(\tau))}{\sum_{\tau=2}^{t^*} \sum_{H^{t^*}(\tau): H^{t^*}(\tau) \Rightarrow h^{t^*}} \text{P}(H^{t^*}(\tau))}.$$

We know now that the above expression goes to infinity as ε goes to zero, uniformly in τ . Consequently, player i on observing h^{t^*} assigns arbitrarily high probability to the first deviation having occurred in the first period of the game. Exactly similar arguments can be used for other histories h^{t^*} with t^* not late in Phase III. \square

We omit here the proof for the cases covered by Assumptions 2 and 3. As mentioned earlier, to prove consistency in these cases, it suffices to assume that infected players are infinitely more likely to make mistakes than uninfected players. Hence, if a player observes an erroneous history, he will think that a player in Community 1 deviated in period 1 and moreover other players have made mistakes.

A.4 Uncertainty about Calendar Time

In this section, we investigate what happens in a setting in which players are not sure about the calendar time or about the precise timing of the different phases. We conjecture that a modification of our strategies would be robust to the introduction of *small* uncertainty about timing. To provide some intuition for this conjecture, we consider an altered environment where players are slightly uncertain about the timing of the different phases. For the purpose of this example, we restrict attention to the product-choice game and try to sustain a payoff arbitrarily close to the efficient outcome $(1, 1)$.

Given the product-choice game and community size M , we choose \dot{T} and \ddot{T} appropriately. At the start of the game, each player receives an independent, noisy but informative signal about the timing of the trust-building phases (values of \dot{T} and \ddot{T}). Each player receives a signal $\omega_i = (\dot{d}_i, \dot{\Delta}_i, \ddot{d}_i, \ddot{\Delta}_i)$, which is interpreted as follows. Player i on receiv-

ing signal ω_i can bound the values of \dot{T} and \ddot{T} with two intervals; *i.e.*, she knows that $\dot{T} \in [\dot{d}_i, \dot{d}_i + \dot{\Delta}_i]$ and $\ddot{T} \in [\ddot{d}_i, \ddot{d}_i + \dot{\Delta}_i]$. The signal generation process is described below. The idea is that players are aware that there are two trust-building phases followed by the target payoff phase. Moreover, signals are informative in that the two intervals are non-overlapping and larger intervals (imprecise estimates) are less likely than smaller ones.

- i) $\dot{\Delta}_i$ is drawn from a Poisson distribution with parameter $\dot{\gamma}$, and then \dot{d}_i is drawn from the discrete uniform distribution over $[\dot{T} - \dot{\Delta}_i, \dot{T} - \dot{\Delta}_i + 1, \dots, \dot{T}]$. If either 1 or \dot{T} lie in the resulting interval $[\dot{d}_i, \dot{d}_i + \dot{\Delta}_i]$, then $\dot{\Delta}_i$ and \dot{d}_i are drawn again.
- ii) After $\dot{\Delta}_i$ and \dot{d}_i are drawn as above, $\ddot{\Delta}_i$ is drawn from a Poisson distribution with parameter $\ddot{\gamma}$. Finally, \ddot{d}_i is drawn from the discrete uniform distribution over $[\ddot{T} - \dot{\Delta}_i, \ddot{T} - \dot{\Delta}_i + 1, \dots, \ddot{T}]$. If the resulting interval $[\ddot{d}_i, \ddot{d}_i + \ddot{\Delta}_i]$ overlaps with the first interval $[\dot{d}_i, \dot{d}_i + \dot{\Delta}_i]$ (*i.e.*, $\dot{d}_i + \dot{\Delta}_i \in [\ddot{d}_i, \ddot{d}_i + \ddot{\Delta}_i]$), then \ddot{d}_i is redrawn.

In this setting, players are always uncertain about the start of the trust-building phases and precise coordination is impossible. However, we conjecture that with a modification to our strategies, sufficiently patient players will be able to attain payoffs arbitrarily close to (1, 1), if the uncertainty about timing is very small. We describe below the modified strategies.

Equilibrium play: Phase I: Consider any player i with signal $\omega_i = (\dot{d}_i, \dot{\Delta}_i, \ddot{d}_i, \ddot{\Delta}_i)$. During the first $\dot{d}_i + \dot{\Delta}_i$ periods, he plays the cooperative action (Q_H or B_H). **Phase II:** During the next $\ddot{d}_i - (\dot{d}_i + \dot{\Delta}_i)$ periods, he plays as if he were in Phase II, *i.e.*, a seller plays Q_L and a buyer B_H . **Phase III:** For the rest of the game (*i.e.*, from period \ddot{d}_i on), he plays the efficient action (Q_H or B_H).

Off Equilibrium play: As before, a player can be in one of two moods: *uninfected* and *infected*, with the latter mood being irreversible. We define the moods a little differently. At the beginning of the game all players are uninfected. Any action (observed or played) that is not consistent with play that can arise on-path, given the signal structure, is called a deviation. We classify deviations into two types. Deviations that definitely entail a short-run loss for the deviating player are called *non-triggering* deviations (e.g. a buyer deviating in the first period of the game). Any other deviation is called a *triggering* deviation (*i.e.*, these are deviations that with positive probability give the deviating player a short-run gain). A player who is aware of a triggering deviation is said to be infected. Below, we specify off-path behavior. We do not

completely specify play after all possible histories, but we think the description below will suffice to provide the intuition behind the conjecture.

An uninfected player continues to play as if on-path. An infected player acts as follows.

- Deviations observed before $\dot{d}_i + \dot{\Delta}_i$: A buyer i who gets infected before period \dot{d}_i switches to her Nash action forever at some period between \dot{d}_i and $\dot{d}_i + \dot{\Delta}_i$ when she believes that enough buyers are infected and have also switched. Note that buyers cannot get infected between \dot{d}_i and $\dot{d}_i + \dot{\Delta}_i$, since any action observed during this period is consistent with equilibrium play (*i.e.*, a seller j playing Q_L at time $t \in [\dot{d}_i, \dot{d}_i + \dot{\Delta}_i]$ may have received a signal such that $\dot{d}_j + \dot{\Delta}_j = t$).
A seller i who faces B_L before period \dot{d}_i , ignores it (this is a non-triggering deviation, as the buyer must still be in Phase I, which means that the deviation entails a short-term loss for her). If a seller observes B_L between periods \dot{d}_i and $\dot{d}_i + \dot{\Delta}_i$, he will switch to Nash immediately.
- Deviations observed between $\dot{d}_i + \dot{\Delta}_i + 1$ and \ddot{d}_i : A player who gets infected in the interval $[\dot{d}_i + \dot{\Delta}_i + 1, \ddot{d}_i]$ will switch to Nash forever from period \ddot{d}_i . Note that buyers who observe Q_H ignore such deviations as they are non-triggering.
- Deviations observed after \ddot{d}_i : A player who gets infected after \ddot{d}_i switches to the Nash action immediately and forever.

We argue below why these strategies can constitute an equilibrium by analyzing some important histories. **Incentives of players on-path:** If triggering deviations are definitely detected and punished by Nash reversion, then, for sufficiently patient players, the short-run gain from a deviation will be less than the long-term loss in payoff from starting the contagion. So, we need to check that all deviations are detected (though, possibly with probability < 1 in this setting), and that the resultant punishment that is triggered is enough to deter the deviation.

- Seller i deviates (plays Q_L) at $t = 1$: With probability 1, his opponent will detect the deviation, and ultimately his payoffs will drop to a very low level. A sufficiently patient player will therefore not deviate.
- Seller i deviates at $2 \leq t < \dot{d}_i + \dot{\Delta}_i$: With positive probability, his opponent j has $\dot{d}_j > t$, and will detect the deviation and punish him. But, because of the uncertainty

about the values of \dot{T} and \ddot{T} , with positive probability, the deviation goes undetected and unpunished. The probability of detection depends on the time of the deviation (detection is more likely earlier than later, because early on, most players are outside their first interval). So, the deviation gives the seller a small current gain with probability 1, but a large future loss (from punishment) with probability less than 1. If the uncertainty about \dot{T} and \ddot{T} is small enough (*i.e.*, signals are very precise), then the probability of detection (and future loss) will be high. For a sufficiently patient player, the current gain will then be outweighed by the expected future loss.

- Seller i deviates (plays Q_L) at $t \geq \ddot{d}_i$: With positive probability, his opponent j has signal $\ddot{d}_j = \ddot{d}_i$, and will detect the deviation.
- All deviations by buyers (playing B_L) are detected, since B_L is never consistent with equilibrium play. If a buyer plays a triggering deviation B_L , she knows that with probability 1, her opponent will start punishing immediately. The buyer's incentives in this case are exactly as in the setting without uncertainty. For appropriately chosen \dot{T} and \ddot{T} , buyers will not deviate on-path.

Optimality of Nash reversion off-path: Now, because players are uncertain about the true values of \dot{T} and \ddot{T} , there are periods when they cannot distinguish between equilibrium play and deviations. We need to consider histories where a player can observe a triggering deviation, and check that it is optimal for him to start punishments.

We assume that players on observing a deviation believe that some seller deviated in the first period of the game. This assumption on beliefs serves the same purpose as before, *i.e.*, conditional on observing a deviation, when it is time to start playing the Nash action, players will think that enough people are already infected for the Nash action to be optimal.

First, consider incentives of a seller i . We argue that a seller who deviates at $t = 1$ will find it optimal to continue deviating. Further, a seller who gets infected by a triggering deviation at any other period will find it optimal to revert immediately to the Nash action.

- Suppose seller i deviates at $t = 1$, and plays Q_L . He knows that his opponent will switch to the Nash action at most at the end of her first interval (close to the true \dot{T} with high probability), and the contagion will spread exponentially from some period close to the true $\dot{T} + \ddot{T}$. Thus, if seller i is sufficiently patient, his continuation payoff will drop to a very low level after $\dot{T} + \ddot{T}$, regardless of his play in his Phase I (until

period $\dot{d}_i + \dot{\Delta}_i$). Therefore, for a given M , if \dot{T} is large enough (and so $\dot{d}_i + \dot{\Delta}_i$ is large), the optimal continuation strategy for seller i will be to continue playing Q_L .

- Seller i observes a triggering deviation of B_L : If a seller observes a triggering deviation of B_L by a buyer (in Phase II), he thinks that the first deviation occurred at period 1, and by now all buyers are infected. Since, his play will have a negligible effect on the contagion process, it is optimal to play Q_L .

Now, consider the incentives of a buyer.

- Buyer i observes Q_L at $1 \leq t < \dot{d}_i$: This must be a triggering deviation. A seller j should switch to Q_L only at the end of his first interval ($\dot{d}_j + \dot{\Delta}_j$), and this cannot be the case, because then, the true \dot{T} does not lie in player i 's first interval. On observing this triggering deviation, the buyer believes that the first deviation occurred at $t = 1$ and the contagion has been spreading since then. Consequently, she will switch to her Nash action forever at some period between \dot{d}_i and $\dot{d}_i + \dot{\Delta}_i$ when she begins believing that enough other buyers are infected and have switched as well (It is easily seen that at worst, buyer i will switch at period $\dot{d}_i + \dot{\Delta}_i$).
- Buyer i observes Q_L at $t \geq \ddot{d}_i + \ddot{\Delta}_i$. Since i is at the end of her second interval, she knows that every rival must have started his second interval, and should be playing Q_H . So, this is a triggering deviation. She believes that the first deviation occurred at $t = 1$, and so most players must be infected by now. This will make Nash reversion optimal for her.

Note that in any other period, buyers cannot distinguish a deviation from equilibrium play.

- Any action observed by buyer i in her first interval (*i.e.*, for t such that $\dot{d}_i \leq t < \dot{d}_i + \dot{\Delta}_i$) is consistent with equilibrium play. A seller j playing Q_H could have got signal $\dot{d}_j > t$, and a seller playing Q_L could have got signal $\dot{d}_j + \dot{\Delta}_j \leq t$.
- Any action observed by buyer i between her two intervals (*i.e.*, at t such that $\dot{d}_i + \dot{\Delta}_i \leq t < \ddot{d}_i$) is consistent with equilibrium play. Q_L is consistent with a seller j who got $\dot{d}_j + \dot{\Delta}_j \leq t$, and Q_H is consistent with a seller with signal such that $t < \dot{d}_j + \dot{\Delta}_j$.
- Any action observed by buyer i within her second interval (*i.e.*, at t such that $\ddot{d}_i \leq t < \ddot{d}_i + \ddot{\Delta}_i$) is consistent with equilibrium play. Q_L is consistent with a seller j who got $\ddot{d}_j > t$ (say $\ddot{d}_j = \ddot{d}_i + \ddot{\Delta}_i$), and Q_H is consistent with a seller with signal such that $\ddot{d}_j < t$ (say j got the same signal as buyer i).