FUZZY CONVENTIONS

MARCIN PĘSKI

Abstract. We study binary coordination games with random utility played in networks. A typical equilibrium is fuzzy - it has positive fractions of agents playing each action. The set of average behaviors that may arise in an equilibrium typically depends on the network. The largest set (in the set inclusion sense) is achieved by a network that consists of a large number of copies of a large complete graph. The smallest set (in the set inclusion sense) is achieved on a lattice-type network. It consists of a single outcome that corresponds to a novel version of risk dominance that is appropriate for games with random utility.

1. Introduction

An individual's behavior in social or economic situations is often positively influenced by similar decisions made by their friends, acquaintances, or neighbors. Examples include the decision to maintain a neat front yard, to obey speed limits or tax laws, to engage in criminal activity, or to adopt a technology with network externalities. A substantial literature has shown that the details of the network of social interactions may affect which of the equilibria is more likely to arise (see, for example, references in [Jackson and Zenou(2015)]). A typical result in this literature establishes conditions under which a particular behavior is adopted by everybody and becomes a convention (see [Young(1993)], [Ellison(1993)], among many others). At the same time, a completely uniform behavior is very rare in the real world. Even in situations which clearly involve positive externalities, there will often be interactions in which neighbors make the opposite choices. For instance, there are families where some members use iPhone and others Android phones.

Date: December 17, 2021.

TBA.

An obvious reason for heterogeneous behavior is that individuals are different and their tastes and unique circumstances play just as important of a role in determining their decisions as the behavior of their neighbors. The goal of this paper is to analyze the impact of heterogeneity in a systematic way. A natural question is how adding heterogeneity in tastes affects our ability to predict the unique outcome. What can we say about the set of possible equilibrium conventions and how does it depend on the network, and other parameters of the model, like taste distribution?

To address these questions, we study a random utility coordination game played in a network. Each player chooses a binary action and the relative gain from the action is increasing in the fraction of neighbors who make the same choice. Additionally, payoffs are subject to individual i.i.d. shocks. The independence assumption is key for our results and it is appropriate for some, but not all applications. An individual's equilibrium action as well as the aggregate distribution of equilibrium actions depend on the realization of the entire profile of payoff shocks. We are interested in the asymptotic of the average (i.e., aggregate) behavior as the network becomes arbitrarily large and, importantly, the graph becomes sufficiently fine, i.e., the weight of the largest neighbor in an neighborhood of each player becomes sufficiently small. The latter ensures that no single individual has a disproportionate impact on another and it is the second key assumption in our model.

In contrast to simple model of coordination games, a typical equilibrium in our model is fuzzy - it has positive fractions of populations playing each action. Also, despite there being only two potential actions, a coordination game may have many more than two equilibria. To illustrate the latter point, consider a continuum toy version of the model, in which individual payoffs depend on the fraction x of agents choosing the high action in the entire population. Let $P(x)$ be the probability of a payoff shock for which the agent best response is to choose the high action as well. Function P has values between 0 and 1 and is increasing in x, but is otherwise arbitrary. An example is illustrated on Figure 1. Fixed points of P, i.e., intersections of the graph of P with 45° diagonal, correspond to equilibria of the toy model.

The goal of this paper is to study the set of all possible equilibrium conventions or, more precisely, the set of equilibrium average actions. Our results characterize the

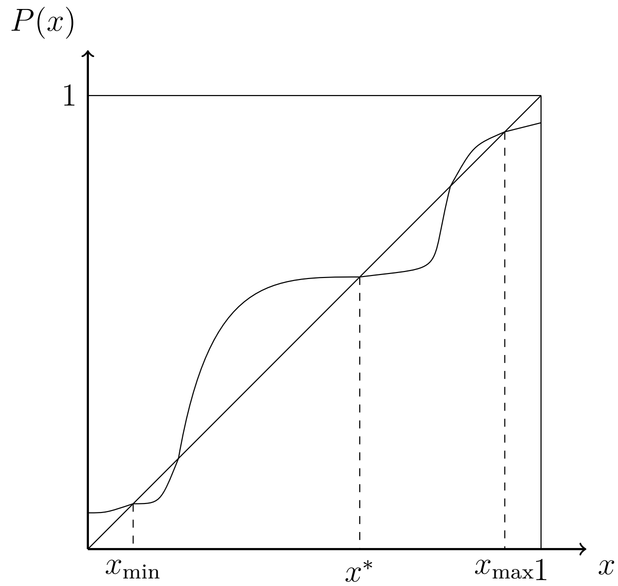

FIGURE 1. Continuum best response function P

asymptotic upper and lower bounds *in the sense of set inclusion* on the equilibrium sets, across all networks. Two results characterize the upper bound:

- Theorem 1 shows that if players live on a sufficiently large complete graph, all stable fixed points of P (essentially, fixed points where the graph of P crosses the diagonal from above) are arbitrarily close to average actions in some equilibrium. (This and all subsequent results are stated "with a probability arbitrarily close to 1.") That, generically, includes the largest $x_{\max}$ and the smallest $x_{\min}$ fixed point of P. The proof of Theorem 1 is straightforward.

  A corollary shows that when players live on sufficiently many disjoint copies of sufficiently large complete graphs, different equilibria on component networks can be mixed and matched so that the total average approximates arbitrary point on the interval $[x_{\min}, x_{\max}]$.
- Theorem 2 shows that for all sufficiently large and fine networks, there are no equilibria with average payoffs above $x_{\max}$ or below $x_{\min}$. Although the statement is very intuitive, our proof is surprisingly complicated. The difficulty is to show that none of the profiles with average payoffs outside of the range

is an equilibrium. There are many such candidate profiles and the claim must simultaneously address all of them. The difficulty is compounded by the lack of additional assumptions on the network.

Together, the two theorems show that the interval $[x_{\min}, x_{\max}]$ is a tight upper bound on the sets of equilibrium average actions across all networks. In this way, we obtain the strongest *partial identification* theory possible: without any further information about the network, an econometrician who uses observed average behavior x, can conclude that the parameters of the model must be such that the parameter-dependent set $[x_{\min}, x_{\max}]$ contains x.

In particular, $x_{\min} = x_{\max}$ is a sufficient condition for the uniqueness of an equilibrium convention, regardless of the network. As the subsequent results show, this condition is not necessary for some networks.

In order to characterize the lower bound on the equilibrium sets, define a random utility-dominant, or *RU*-dominant, outcome x^* as a solution to the maximization problem

$$x^* \in \arg\max_x \int_0^x \left(y - P^{-1}(y)\right) dy.$$

(See Figure 1.) Generically, an *RU*-dominant outcome is unique, in which case, it is always a stable fixed point of P. The notion of RU-dominance is one of the contributions of this paper. When the impact of payoff shocks on an individual utility converges to 0, the RU-dominant outcome converges to the risk-dominant outcome (as in [Harsanyi and Selten(1988)]) of the deterministic coordination 2×2 game.

We have two results:

- Theorem 3 shows that there exist networks where the average payoff in each equilibrium is arbitrarily close to x^*. One example of such a network is a 2-dimensional lattice. The idea of the proof is to show that for each profile with an average behavior that is not *RU*-dominant, contagion-like best response dynamics would bring the behavior close to x^*. The proof uses an idea from [Morris(2000)] to show how a contagion wave spreads across lattice networks.

This is supplemented with explicit calculations of (a) the likelihood that a favorable configuration of payoff shocks may initiate such a wave, and (b) the likelihood that such a wave would not be stopped by an unfavorable configuration of payoff shocks. The problem with the latter is the reason why the 1-dimensional network of [Ellison(1993)] is not a good example for the result and a 2-dimensional lattice is needed.
- Theorem 4 shows that any sufficiently large and fine network has an equilibrium with average payoffs close to x^*. The starting point of the proof is a beautiful idea from [Morris(2000)] where it is shown that it is not possible to spread risk-dominated actions by contagion. This idea is adapted to work for *RU*-dominance, random utilities, etc.

The two results together show that the single-element set $\{x^*\}$ is a tight lower bound on all sets of equilibrium average payoffs across all networks. This leads to an *equilibrium selection* theory: x^* is th eonly outcome that is robust to changes in the underlying network.

Coordination games form one of three main approaches in the literature that studies games in networks ([Jackson and Zenou(2015)]). The second set of results of this paper is very closely related, and it greatly benefits from the literature on contagion in networks, especially from two beautiful papers, [Ellison(1993)] and [Morris(2000)]. [Ellison(1993)] (see also [Ellison(2000)]) was the first to show that a risk-dominant action can spread from a small initial set of deviators to an entire 1-dimensional lattice network by a simple best response process. [Morris(2000)] describes properties of networks for which Ellison's contagion wave exists. Among others, any contagion wave from 1-dimensional lattices can also be used in higher dimensions. [Morris(2000)] also shows that risk-dominated actions cannot spread through a best response process no matter what is the geometry of the network.

Evolutionary game theory ([Kandori *et al.*(1993)Kandori, Mailath and Rob], [Young(1993)], [Blume(1993)], [Newton(2021)], and many others) studies the long-run behavior of perturbed best response processes, where players commit mistakes with a small probability, and instead of choosing a best response, take some other action. One of the key results of this literature is that the risk-dominant coordination is (uniquely) stochastically stable regardless of the underlying network ([Peski(2010)]). Our current results (specifically, Theorems 3 and 4) are closely related, but with some key differences. On the one hand, there is a relation between "noise" in the behavioral rules of the evolutionary literature and "noise" in the payoffs of the current paper. On the other hand, there are two important differences: We are interested here in static equilibria instead of a dynamic adjustment process and our payoff shocks are permanent instead of temporary mistakes. (The best response dynamic plays important role in the proofs as a tool to identify equilibria.) Finally, the evolutionary literature is subject to the criticism that one may need to wait for a really long time before reaching a stochastically stable outcome ([Ellison(1993)]). That criticism does not apply to our model.

There is a literature that studies evolutionary equilibrium selection in games with heterogeneous populations. The interests in such games arises naturally from evolutionary biology like predator-prey models. For instance, [Friedman(1991)] describes a general framework with multiple continuum populations choosing actions and receiving payoffs and studies evolutionary study states of continuous time adjustment dynamics. Closer related to this paper is [Neary(2012)], who studies similar model to us but with two payoff shocks (more precisely, two subpopulations of deterministic size) and agents located on a complete graph. The paper presents conditions under which the evolutionary dynamics of [Kandori *et al.*(1993)Kandori, Mailath and Rob] selects a fuzzy convention, i.e., an equilibrium where members of different subpopulations play different actions. [Neary and Newton(2017)] study general payoff shock and present a sufficient condition under which the logit dynamics of [Blume(1993)] selects a fuzzy convention.

Section 2 contains the model. The next four sections state and discuss the four theorems mentioned above. The last section concludes.

2. Model

2.1. **Coordination game in a network.** There are N agents $i = 1, ..., N$ who live in the nodes of a network. The network is defined as an undirected weighted graph with weights $g_{ij} = g_{ji} \geq 0$ for $i, j \leq N$. We assume that $g_{ii} = 0$ and that $g_i = \sum_j g > 0$ for each player i. Let

$$d(g) = \max_{i,j} \frac{g_{ij}}{g_i} \text{ and } w(g) = \frac{\max_i g_i}{\min_i g_i},$$

where $d(g) \in [0,1]$ is a bound on the importance of a single player in another player's neighborhood and it describes how fine the network is, and $w(g) \geq 1$ is a rough measure of the degree inequality. A network is *balanced* if all players have the same degree $g_i = g_j$ for each i, j. In balanced networks, $w(g) = 1$.

The agents play a binary action coordination game. Each agent chooses an action $a_i \in \{0,1\}$ and receives a payoff

$$\frac{1}{g_i} \sum_j g_{ij} u(a_i, a_{-i}, \varepsilon_i), \tag{1}$$

which depends on the actions of her neighbors and a payoff shock $\varepsilon_i \in [0,1]$ drawn i.i.d. from a probability distribution $F(.)$. The payoffs are supermodular in actions: for each ε,

$$u(1,1,\varepsilon) + u(0,0,\varepsilon) > u(1,0,\varepsilon) + u(0,1,\varepsilon).$$

Mixed actions are represented by the probability $a \in [0,1]$ of pure action 1. Due to expected utility, payoffs are linear in mixed actions. We refer to the tuple (u, F) as the *random utility game.*

Example 1. In an additive payoff shock model, the payoffs of player i from interaction with j are equal to

$$u(a_i, a_j) + \Lambda \varepsilon_i \mathbf{1}(a_i = 1), \tag{2}$$

where u is a symmetric 2×2 coordination game. Although (1) seems more general than (2), the two models are equivalent in the sense that the payoff shocks can be matched so that the best responses to mixed strategies in both models are identical. Parameter Λ measures the importance of the payoff shocks. When $\Lambda \to 0$, the model converges to the deterministic game.

2.2. **Equilibria.** We assume that the payoff shocks are publicly observable, i.e., players know preferences of other players. Each network g, and each realization of payoff shocks ε leads to a many-player complete information static game $G(g, \varepsilon)$. Let (a_i) be a (possibly, mixed) profile of actions. Let

$$\mathrm{Av}\left(a\right)=\frac{1}{\sum_{i}g_{i}}\sum_{i}g_{i}a_{i}$$

be the average action weighted by each player's neighborhood size. This turns out to be the natural notion of average behavior. If $g_i \in \{0,1\}$, then g_i is a count of the interactions in which agent i participates, and $\mathrm{Av}\left(a\right)$ is the average number of interactions in which action 1 is played.

For each profile a, let $\beta_i^a = \frac{1}{g_i}\sum_j g_{ij}a_j$ be the weighted average number of agents in the neighborhood of i who play 1. We refer to profile $\beta^a = (\beta_i^a)$ as the profile of average neighborhood behaviors.

Profile a is a Nash equilibrium if for each player i, $u\left(a_i, \beta_i^a, \varepsilon_i\right) \geq u\left(1-a_i, \beta_i^a, \varepsilon_i\right)$. Denote the set of average behaviors attained in Nash equilibria as

$$\mathrm{Eq}\left(g,\varepsilon\right)=\left\{\mathrm{Av}\left(a\right):a\text{ is a Nash eq. of }\mathrm{G}(g,\varepsilon)\right\}\subseteq\left[0,1\right].$$

$\mathrm{Eq}\left(g\right)$ as a set-valued random variable, i.e., mapping from the space of payoff shock profiles to subsets of $[0,1]$. The goal of the paper is to analyze the behavior of $\mathrm{Eq}\left(g\right)$ as the network becomes larger and the importance of individual players decreases, $d\left(g\right)\rightarrow 0$.

For any $x\in[0,1]$ and any two compact subsets $A,B\subseteq[0,1]$, say A *is η-included in* B, write $A\subseteq_\eta B$, if $\max_{x\in A}\min_{y\in B}|x-y|\leq\eta$. If $A\subseteq_\eta B$ and $B\subseteq_\eta A$, then we write $A=_\eta B$.

2.3. **Continuum best response function.** For each $x\in[0,1]$, let

$$P\left(x\right)=F\left(\varepsilon:u\left(1,x,\varepsilon\right)\geq u\left(0,x,\varepsilon\right)\right).$$

$P\left(x\right)$ is the ex-ante probability that action 1 is a best response if a player faces x fraction of opponents who also play 1. A typical graph of P is illustrated on Figure 1. The assumptions imply that P is increasing, right-continuous, and that $P\left(x\right)\in[0,1]$. We do not assume that P is invertible (and it won't be, if, for instance, F has atoms). Instead, we define $P^{-1}\left(y\right)=\inf\left\{\left(x:P\left(x\right)\geq y\right)\right\}$.

It is helpful to think about $P(x)$ as a best response function in a continuum toy version of the game, where each agent's payoff depends on the fraction of the entire population who choose to play 1. Due to the continuum law of large numbers, $P(x)$ is the fraction of the population for whom 1 is a best response. Fixed points of P, i.e., intersections of the graph on Figure 1 with 45°-line, correspond to Nash equilibria in the continuum version of the game.

3. Equilibria on complete graphs

In this section, we consider a complete graph, i.e., network g such that $g_{ij} = 1$ for each $i \neq j$. For large N, such a graph should approximate well the continuum toy model.

We say that a fixed point $x = P(x)$ is *strongly stable* if there exist $\gamma < 1$ and a neighborhood $U \ni x$, such that for each $y \in U$, if $y \leq x$ (resp. $y \geq x$), then $P(y) \geq P(x) + \gamma(y - x)$ (resp., $P(y) \leq P(x) + \gamma(y - x)$).

Theorem 1. *Suppose that x is a strongly stable fixed point of P. Let g^N be a complete graph with N nodes. For each $\eta > 0$, there is $N > 0$, such that*

$$\mathbb{P}\left(\{x\} \subseteq_\eta Eq\left(g^N, \varepsilon\right)\right) \geq 1 - \eta.$$

Large complete graphs have equilibria that are close to strongly stable points of x. The result is a sanity check, as it confirms our interpretation of P as a best response function on the continuum toy model. The proof is straightforward (see Appendix B).

When there are (finitely many) multiple strongly stable points, Theorem 1 implies that, with a large probability, all of them are close to the average behavior in some equilibrium. In particular, if $x_{\min}$ and $x_{\max}$ are, respectively, the smallest and the largest of the fixed points of P, then $\{x_{\min}, x_{\max}\} \subseteq_\eta \mathrm{Eq}(g)$ with a large probability for a sufficiently large complete graph.

One can obtain other equilibrium averages by mixing and matching networks. By taking a large number of disjoint copies of large complete graphs (see Figure 2), and considering a variety of equilibria on component networks, we can an approximate an arbitrary point on the interval $[x_{\min}, x_{\max}]$.

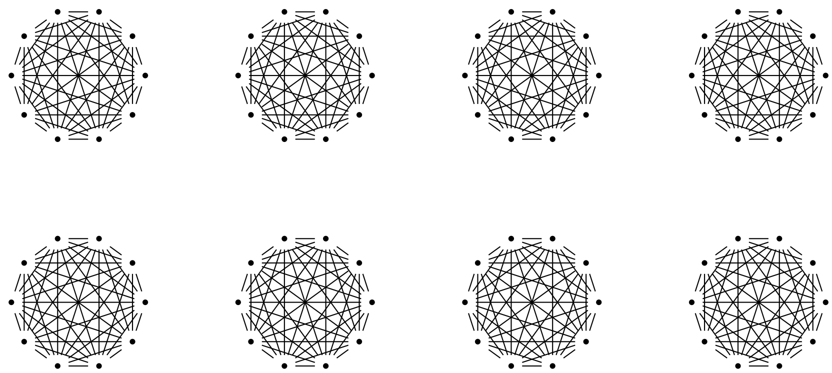

FIGURE 2. Multiple complete graphs

Corollary 1. *Suppose that $x_{\min}$ and $x_{\max}$ are strongly stable. For each $\eta > 0$, there exists a balanced network g such that*

$$\mathbb{P}\left([x_{\min}, x_{\max}] \subseteq_\eta Eq\left(g, \varepsilon\right)\right) \geq 1 - \eta.$$

4. Upper bound on equilibrium set

The next result shows that $[x_{\min}, x_{\max}]$ is an upper bound on the equilibrium set.

Theorem 2. *Suppose that $x_{\min}$ and $x_{\max}$ are strongly stable. For each $\eta > 0$ and $w < \infty$, there is $\delta > 0$ such that for each network g, if $d\left(g\right) \leq \delta$, $w\left(g\right) \leq w$, then*

$$\mathbb{P}\left(Eq\left(g\right) \subseteq_\eta [x_{\min}, x_{\max}]\right) \geq 1 - \eta.$$

The theorem yields a partial identification theory of the parameters of the model. Consider an econometrician who studies a coordination game on a network. The econometrician may not know the network g on which the game is played, nor the parameters of the random utility model, and she treats them as parameters. If she observes the average behavior x, she may reject all parameters for which $x \notin [x_{\min}, x_{\max}]$.

Theorem 2 and Corollary 1 together show that the interval $[x_{\min}, x_{\max}]$ is a tight upper bound (in the sense of set inclusion) on the average behavior across all networks. In particular, the partial identification obtained from the result cannot be improved.

4.1. **Proof intuition.** Although the statement of Theorem 2 is intuitive, our proof is surprisingly complicated. To explain the issue, notice first that it is not difficult to show that the probability of any given profile being an equilibrium is small. In fact, one can easily find an exponential bound $\exp(-\rho N)$ on such a probability, where N is the number of agents and $\rho > 0$ is some constant. A brute force method to show that, with a large probability, *none* of the profiles with an average payoff $x > x_{\max} + \eta$ is an equilibrium, could be to multiply (a) the above exponential bound by (b) the number of such profiles. The problem is that the number (b) is exponentially large in N and there is no guarantee that the product of (a) and (b) converges to 0 as $N \to \infty$.

The brute force method relies on the worst-case scenario where events "profile a is an equilibrium" across different as are treated as disjoint. However, they are typically correlated, more so for profiles that are similar, in some way. The idea of the proof is to divide profiles a into groups of similar profiles such that (a) there exists an exponential bound on the probability that none of the profiles in a group is an equilibrium (Lemma 9 in the Appendix), and (b) the number of groups grows at a much slower rate than the exponent of the part from (a).

In order to explain the division into groups, define the notion of closeness of two profiles a and b as a weighted version of the Euclidean metric:

$$d(a,b) = \sqrt{\frac{1}{\sum g_i^2} \sum g_i^2 (a_i - b_i)^2}. \tag{3}$$

For some $\delta > 0$, and each profile a, let

$$B(a,\delta) = \left\{ b : d\left(\beta^a, \beta^b\right) \leq \delta \right\}.$$

$B(a,\delta)$ consists of profiles that have similar average neighborhood behavior for all agents. Lemma 10 in the Appendix shows that the number of sets $B(a,\delta)$ required to cover the whole space of profiles is exponentially large, but not bigger than $\exp\left(\frac{1}{\delta^2} c_{d(g)} N\right)$, where $c_{d(g)}$ is a constant that decreases to 0 as $d(g) \to 0$. Because the exponent converges to 0 for sufficiently fine networks, the division deals with step (b).

We give more details for step (a). We start with some preparatory remarks. First, observe that it is enough to work with a continuum best response function $P^*(x) = \max(x_{\max}, x_{\max} + \gamma(x - x_{\max}))$, where $\gamma < 1$ is chosen so that $P(x) \leq P^*(x)$ for each x. (Such a γ exists due to $x_{\max}$ being the largest fixed point that is also strongly stable.)

The reason is that one can construct game payoffs with continuum best response P^* and show that the equilibria of the latter dominate the equilibria of the original game.

Next, for each profile a, define the best response profile $b^*(a,\varepsilon)$=$(b_i^*(a,\varepsilon))$, where for each i, $b_i^*(a,\varepsilon)$ is the highest best response of agent i with payoff shock ε_i against profile a. The expected value of agent i's best response is equal to $\mathbb{E}\, b_i^*(a,\varepsilon_i) = P^*(\beta_i^a)$, or the continuum best response function applied to the average neighborhood behavior. Let $p^a = (P^*(\beta_i^a))_i$ be the profile of the expected best responses.

The form of the continuum best response function P^*, and specifically the fact that $x - P^*(x) \geq (1-\gamma)(x - x_{\max})$ for each $x > x_{\max}$, implies that the expected best response behavior p^a to a profile with average neighborhood behavior above $\beta_i^a \geq x_{\max}$ is strictly smaller than the average action in the average neighborhood behavior. More generally, we establish the following fact (Lemma 6): if profile a is close (in the sense of metric (3)) to some other profile b with the property that $b_i \geq x_{\max}$ for each i, the average action in the expected best response profile p^a is strictly smaller than the average neighborhood action β^a by a factor $(1-\gamma)(\mathrm{Av}(a) - x_{\max})$

The rest is divided into three steps.

- If agent i has sufficiently many neighbors and none of the neighbors is too important, the average best response action in the neighborhood of i is likely to be similar to the average expected best response, $\frac{1}{g_i}\sum_j g_{ij} b_i^*(a.\varepsilon_i) \approx \frac{1}{g_i}\sum_j g_{ij} P^*\left(\beta_j^a\right)$. More precisely, we show (Lemma 7 in the Appendix) that, if the network is sufficiently fine, then, with a large probability, the average neighborhood behavior $\beta^{b^*(a,\varepsilon)}$ is close (in the sense of metric (3)) to the average neighborhood behavior β^{p^a}obtained from the expected best response profile p^a. The probability bound is exponential, where the exponent depends on the measure of fineness of the network $d(g)$.
- Take any two profiles a and b and consider best responses $b^*(a,\varepsilon)$ and $b^*(b,\varepsilon)$ of agent i to such profiles. The probability that the payoff shock ε_i such that the best responses are different, $b_i^*(a,\varepsilon) \neq b_i^*(b,\varepsilon)$ is closely related to the difference between the neighborhood behavior witnessed by player i, i.e., β_i^a and β_i^b. The smaller the difference, the smaller the probability. For any two profiles

a and $b \in B(a,\delta)$, the great majority of agents observes a similar neighborhood behavior $\beta_i^a \sim \beta_i^b$, which means that their best responses are likely to be identical (Lemma 7 in the Appendix makes this argument precise). Further, it means that, with a large probability, the distance (3) between the average neighborhood behavior profiles $\beta^{b^*(a,\varepsilon)}$ and $\beta^{b^*(b,\varepsilon)}$ is small (and decreasing with δ), with a large probability. Together with the previous paragraph, this shows that, with a large probability, $\beta^{b^*(b,\varepsilon)}$ is close to β^{p^a} *uniformly* across all profiles $b \in B(a,\delta)$. A similar statement about the uniform closeness of β^{p^b} and β^{p^a} holds as well.

- Take a that $\mathrm{Av}(a) > x_{\max}$ and suppose that the above large probability events hold. Because of the form of the continuum best response function P^*, we clearly have $\beta_i^{p_a} \geq x_{\max}$ for each i. Suppose that $b \in B(a,\delta)$ is an equilibrium, $b^*(b,\varepsilon) = b$. On one hand, the above fact applied to profiles β^b and β^{p^a} implies that the average action in profile β^b must be strictly smaller than the average action in profile β^{p^b} and the gap between two average actions can be shown to be of the same order as $(1-\gamma)(\mathrm{Av}(a) - x_{\max})$. On the other hand, the probabilistic assumption implies that profile $\beta^b = \beta^{b^*(b,\varepsilon)}$ is close to β^{p^b} in the sense of metric (3). It is straightforward to show that the average action in profiles that are close in the sense of metric (3) is also close. The contradiction between two claims means that no profile in $B(a,\delta)$ is an equilibrium.

5. RU-dominant selection

In this section, we introduce an equilibrium selection tool appropriate for coordination games with random utility: the random utility-, or *RU*-dominant outcome. We show that there are networks on which the *RU*-dominant outcome is essentially the only equilibrium average.

5.1. *RU*-dominant outcome.

An equilibrium action $x^* \in [0,1]$ is *RU-dominant* if it is a maximizer of

$$x^* \in \arg\max_x \int_0^x \left(y - P^{-1}(y)\right) dy. \tag{4}$$

It is *strictly RU-dominant*, if it is a unique maximizer. Generically, any game with random utility has a RU-dominant action.

The following example shows that if the impact of the random utility impact disappears, the RU-dominant outcomes converge to standard risk dominance of [Harsanyi and Selten(1988)].

Example 2. (Cont. of Example 1) Suppose w.l.o.g. that 0 is the unique strictly risk-dominant action of the coordination game with payoffs u. Then, each player is indifferent between two actions if a fraction $\alpha > \frac{1}{2}$ of players plays action 1. When $\Lambda \to 0$, $P^{-1}(y) \to \alpha$ for each $y \in (0,1)$, and we have

$$\int_0^x \left(y - P^{-1}(y)\right) dy \to \int_0^x (y-\alpha)\, dy = \frac{1}{2}x^2 - \alpha x = x\left(\frac{1}{2}x - \alpha\right).$$

The last expression is maximized by $x = 0$. Hence the RU-dominant outcome(s) converge to 0, i.e., the risk-dominant action of deterministic game u.

The main result of this section shows that there are networks where, with a large probability, all equilibrium averages are close to the strictly RU-dominant outcome x^*.

Theorem 3. *Suppose that x^* is the strictly RU-dominant outcome and that either $x^* > 0$ and $P(0) > 0$, or $x^* < 1$ and $P(1) < 1$. For each $\eta > 0$, there is a network g such that*

$$\mathbb{P}\left(Eq(g) \subseteq_\eta \{x^*\}\right) \geq 1 - \eta.$$

5.2. **Proof intuition.** The network constructed in the proof is a 2-dimensional lattice, parameterized with M and m. There are M^2 agents living on square $\left[0, \frac{M}{m}\right]^2 \subseteq \mathbb{R}^2$ at fractional points $\left(\frac{k}{m}, \frac{l}{m}\right)$ for some $k, l = 1, ..., M$. Any two agents i and j are connected, $g_{ij} = 1$, if the (Euclidean) distance between them is no larger than 1. To make the network balanced and to simplify the arguments, we assume that all distance calculations are done mod$\frac{M}{m}$, which turns the square $\left[0, \frac{M}{m}\right]^2$ into a torus.

We show that, if m and $\frac{M}{m}$ are sufficiently large, then for a large probability set of realizations, there is no equilibrium in which the average action is significantly higher than x^*. Together with an analogous argument for the other side, this suffices to establish the theorem. Our argument should extend to K-dimensional lattices for $K > 2$, but not to $K = 1$.

The proof relies on the idea of a contagion through best response dynamics introduced in [Ellison(1993)] and further expanded in [Morris(2000)]. If the lattice is sufficiently large, with a large probability, there exists a contiguous group of agents for whom 0 is strictly dominant. We refer to these agents as *initial infectors.* Assume that, initially, all the other agents play 1. Consider a best response process, in which agents, in some order, are offered an opportunity to revise their actions to a myopic best response. Because of the payoff complementarities, the revisions will always go in the same direction, i.e., towards action 0. The process must eventually stop and the profile of action at which it stops is the highest equilibrium for a given realization of payoff shock.

It is helpful to imagine downwards revision of actions as a wave of 0s moving away from the set of initial infectors. We are going to show that, for almost all realizations, the contagion wave spreads throughout the entire network, in a way so that, eventually, in almost each neighborhood, the average fraction of agents who play 1 is not much higher than x^*.

The rest of the argument is divided into two parts.

5.2.1. *Contagion wave on line.* First, we explain how the existence of the contagion wave is related to the maximization problem (4). It is convenient to explain this part of the argument using a version of the line network from [Ellison(1993)]. Suppose that agents are located along a line at equally spaced locations. We assume that each location in the network contains a continuum population of mass 1. We further assume that the weight of connection between agents in locations i and j depends only on their distance $g_{ij} = g_{i-j} =: g_{j-i}$, where we take no connections between agents in the same location, i.e., $g_0 = 0$. We normalize the weights so that $\sum g_d = 1$. The continuum assumption allows us to use the law of large numbers to compute the average best response action of agents in node i as equal to $P\left(\sum_d g_d a_{i+d}\right)$, where a_j is the average current action played by agents in node j. Initially, locations $i \leq 0$ consist of initial infectors - those agents play 0. All locations $i > 0$ initially play action 1.

Let a_i^* be the average action in location $i \geq 0$ when the best response process stops. Due to payoff complementarities, a_i^* must be increasing in i. Let $a_i = \max(x^*, a_i^*)$ and $a = \lim_{i\to\infty} a_i = \sup_i a_i^*$. Suppose that that $a > x^*$. Because $x^* \leq P(y)$ for each $y \geq x^*$ and because a_i^* is the limit action of the best response process, the average action cannot be larger than the best response

$$a_i = \max(x^*, a_i^*) = \max\left(x^*, P\left(\sum_d g_d a_{i+d}^*\right)\right) \leq \max\left(x^*, P\left(\sum_d g_d a_{i+d}\right)\right) = P\left(\sum_d g_d a_{i+d}\right).$$

Taking inverse, we obtain

$$P^{-1}(a_i) \leq \sum_d g_d a_{i+d} = x^* + \sum_j \left(\sum_{d \geq j-i} g_d\right)(a_{j+1} - a_j),$$

where the equality is due to a discrete version of the "integration by parts" formula and the fact that $a_i \geq x^*$ for each i. After multiplying by $a_{i+1} - a_i \geq 0$, and summing up across all locations i, we obtain

$$\sum_i \left(P^{-1}(a_i) - x^*\right)(a_{i+1} - a_i) \leq \sum_{i,j}\left(\sum_{d\geq j-i} g_d\right)(a_{i+1} - a_i)(a_{j+1} - a_j). \tag{5}$$

The left-hand side of the inequality is approximately equal to $\int_{x^*}^a (P^{-1}(y) - x^*)\,dy$. To compute the right hand side, notice that we can switch the roles of i and j in the summation, and using the fact that $\sum_{d\geq j-i} g_d + \sum_{d \geq i-j} g_d = \sum g_d = 1$, we have

$$\begin{aligned}
\sum_{i,j}\left(\sum_{d\geq j-i} g_d\right)(a_{i+1} - a_i)(a_{j+1} - a_j) &= \frac{1}{2}\left(\sum_{i,j}\left(\sum_{d\geq j-i} g_d + \sum_{d\geq i-j} g_d\right)(a_{i+1} - a_i)(a_{j+1} - a_j)\right) \\
&= \frac{1}{2}\left(\sum_{i,j}(a_{i+1} - a_i)(a_{j+1} - a_j)\right) = \frac{1}{2}(a - x^*)^2 \\
&= \frac{1}{2}(a - x^*)^2 = \int_{x^*}^{a} (y - x^*)\,dy.
\end{aligned}$$

Putting the two sides together, inequality (5) implies that

$$\int_{x^*}^{a} \left(y - P^{-1}(y)\right) dy \geq 0,$$

which contradicts the fact that x^* is the unique maximizer of the integral on the right-hand side. Hence the contagion must spread across the entire network.

More formally, Lemma 12 in the Appendix uses similar calculations to establish, for each $\eta > 0$, the existence of $\delta > 0$, $L < \infty$, and a δ-contagion wave: a strategy profile σ_i with the property that (a) $\sigma_i \leq x^* + \eta$ for $i \leq 0$ and $\sigma_i \geq 1$ for $i \geq L$ and (b), for each location i, the best response strategy of the agents to $\sigma + \delta$ in that location is smaller than the current strategy $\sigma_{i-\delta}$ in the location that is δ away on the left side. The idea is that, starting from σ, the best response dynamics moves the population strategy σ rightwards by at least δ at each stage of the dynamics - spreading eventually the behavior of at most $x^* + \eta$ to the entire network. We refer to L as the length of the wave.

The spread of a contagion wave from a small set of initial infectors extends from the line to higher-dimensional lattices due to an elegant argument from [Morris(2000)]. The idea is that if the front of the wave is sufficiently smooth, i.e., with a sufficiently low curvature, then it can be locally approximated by a hyperplane. The spread of the wave in the direction that is orthogonal to its front can be analyzed using the same techniques as the spread of the wave on one-dimensional line.

5.2.2. *Obstacles.* The continuum assumption used in the above argument ensures that the average best response action of agents in a location is given by the continuum best response function $P(.)$. The assumption makes it easy to compute average best responses. At the same time, it ignores a positive probability of a contiguous group of "bad" agents for whom 1 is strictly dominant action. If sufficiently large, such a group of "bad" agents will stop the best response revisions towards action 0 and block the contagion wave (see the left panel on Figure 3).

One could try to compare the relative frequency of initial infectors necessary to start the wave versus the sets of "bad" agents who may block it. Unfortunately, for some Ps, the latter are more frequent. As a result, the line network is not a good candidate example for Theorem 3.

At the same time, the "bad" sets are intuitively less likely to block the contagion wave on higher-dimensional lattices (see the right panel of Figure 3). The reason is

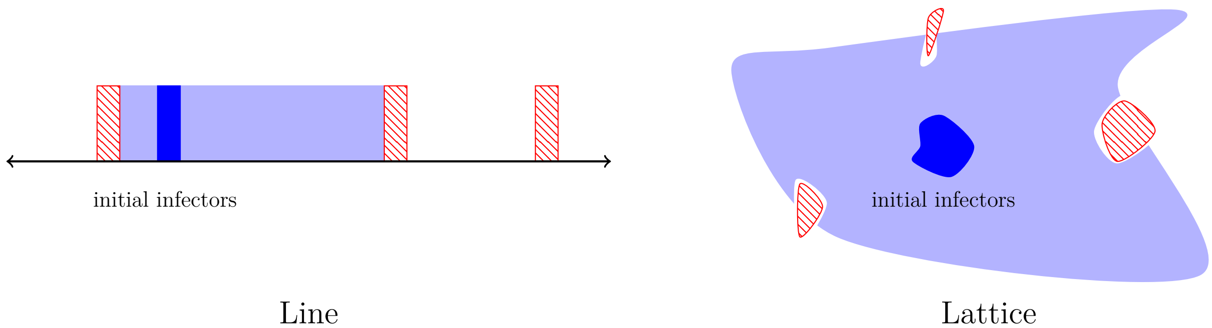

FIGURE 3. Obstacles to contagion wave

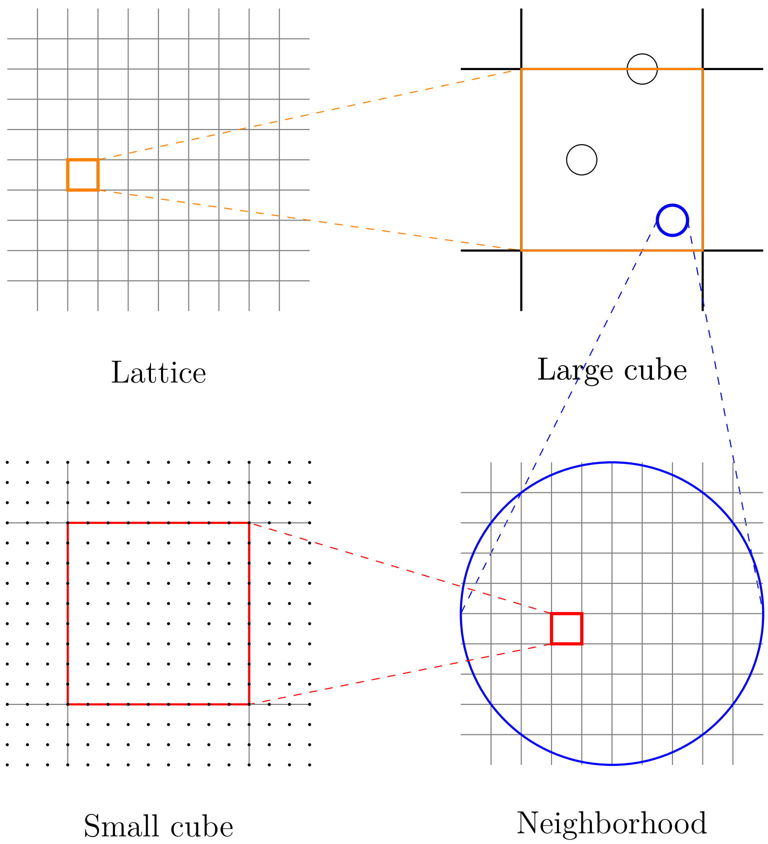

FIGURE 4.

that to block the wave, the "bad" sets would have to be arranged so to surround it. Even if the number of "bad" set is much larger than the number of initial infectors, the probability of a bad arrangement can be quite small.

We show that, indeed, the likelihood of "bad" sets surrounding the initial infectors is very small if the lattice is sufficiently large. Below, we sketch the main ideas. The details of the proof can be found in the Appendix D.

First, the lattice is divided into large and small cubes (see Figure 4) - so that the number of the large cubes in the lattice is very large, each large cube contains a very large number of disjoint neighborhoods, each neighborhood contains a very large number of small cubes, and each small cube contains a very large number of agents. These numbers are chosen so that a series of claims holds:

(1) The numbers of small cubes in a neighborhood and the number of agents in a small cube are sufficiently large, so that the fraction of shared agents and the fraction of shared small cubes in neighborhoods of any two agents i and j is well approximated by the area of the intersection of two 1-radius circles with centers at i and j (Lemma 13).

(2) The size of the small cube is sufficiently large so that, for each small cube, with a probability close to 1, the empirical distribution of payoff shocks within the cube is close to the true distribution. We say that a small cube is *(γ-)bad* if, for some fraction x, the average best response action of the agents within the cube is (γ-)larger than $P(x)$. Agents in bad cubes may tilt toward higher best responses than a statistical agent. Agents in a small cube that is not bad are well-approximated by the continuum assumption in the following sense: the average best response in the small cube is not higher than to $P(\beta)$, where β is the average "belief" (i.e., the average neighborhood action) for members of the cube (Lemma 15).

(3) A large cube is *good* if it contains no bad small cubes. The ratio of the size of the small cube (i.e., the number of agents within) to the number of small cubes in a large cube is sufficiently large, so that the probability p that a large cube is not good is arbitrarily close to 0.

    A large cube is *extraordinary*, if it contains only agents for whom 0 is the strictly dominant action. Extraordinary cubes play the role of the set of the initial infectors. The number of large cubes is sufficiently large, so that the probability that an extraordinary large cube exists is arbitrarily close to 1.

(4) Two large cubes are *connected* if they share a wall. The number of large cubes is sufficiently large, and the probability p that a large cube is not good is sufficiently small, so that there exists a giant component of good large cubes - a set of good large cubes that contains almost all large cubes on the lattice and such that all of its elements are connected with each other by paths of large cubes that share a wall. This argument is the content of Lemma 17) and it relies on definitions and results from the percolation theory ([Bollobás *et al.*(2006)Bollobás, Riordan and Riordan]).
   (a) First, we show that each connected set S can be surrounded by a connected "boundary" ∂S that isolates set S (and, possibly, some other large cubes) from the remaining large cubes. The total number of large cubes isolated away from set S is not larger than $|S|^2$. (On a two-dimensional lattice, the worst-case scenario bound comes from elements of set S arranged in a way that surrounds an interior of size that is proportional to the square of its perimeter.)
   (b) For a collection of connected sets $S_1, ..., S_J$ that are *not* connected with each other, the giant connected component that omits all sets S_j contains all but at most $\sum |S_j|^2$ large cubes.
   (c) Let $S_1, ..., S_J$ be the collection of all maximally connected collections of large *bad* cubes. We estimate the expected value of $\sum |S_j|^2$ as proportional to the number of all large cubes multiplied by the probability p that a single large cube is bad (Lemma 24). An application of the Markov inequality shows that, if p is sufficiently small, the giant connected component that contains only good cubes contains a fraction of all large cubes that is arbitrarily close to 1.

(5) If the curvature of the two-dimensional contagion wave is sufficiently small relative to the curvature of an individual neighborhood, the contagion wave will spread, as long as its path contains only good small cubes (Lemma 16).

Putting it together, the contagion wave is going to spread through a great majority of the giant connected component of good large cubes, which means a great majority of the lattice. Hence, with a large probability, the average action in the largest equilibrium on a sufficiently large two-dimensional lattice is close to x^*.

6. RU-dominant equilibrium in each network

The previous section identified an *RU*-dominant outcome as a candidate solution for equilibrium selection theory. Next, we ask whether there are other potential candidates, i.e., whether there are other outcomes that can be unique equilibria on some networks.

The next result shows that the answer is negative.

Theorem 4. *Suppose that x^* is the strictly RU-dominant outcome. For each $\eta > 0$, there is $d > 0$ such that, for each network g, if $d(g) \leq d$, then*

$$\mathbb{P}\left(\{x^*\} \subseteq_\eta Eq(g)\right) \geq 1 - \eta.$$

If the network is sufficiently fine, then, for almost all realizations of payoff shocks, there is an equilibrium with action distribution close to the RU-dominant action. In particular, no other outcome than the *RU*-dominant outcome can be a unique equilibrium in some network.

Theorems 3 and 4 lead to an *equilibrium selection* theory: only the *RU*-dominant outcome x^* is robust to changes in the underlying network. This claim is made precise by the proof of Theorem 4. In the proof, we consider a profile in which almost all players choose best responses as if x^* neighbors play action 1. We show that any best response dynamics starting from such a profile will stop in an equilibrium profile in which a great majority of players never revise their actions. It follows that, if players play such an equilibrium under one network, and then the network is changed (in a manner independent of actions and payoff shocks), then the best response process will end up with a very similar profile as an equilibrium.

6.1. Proof intuition.

We start with an *initial profile* a^0 in which all players choose best responses as if fraction x^* of their opponents plays 1,

$$a_i^0 \in \arg\max u\left(a_i, x^*, \varepsilon_i\right).$$

Although each agent chooses depending on their payoff shock, the law of large numbers and the fact that x^* is an equilibrium of the continuum game imply that the average action in the population is unlikely to be far from x^*.

Starting from the initial profile, we consider an *upper* best response dynamics, where at each stage, a single player is allowed to revise their action towards the best response, but only upwards, i.e., if the best response is the action 1. Such dynamics must stop eventually, and the resulting profile a^U does not depend on the order in which players revise their actions, as long as all players for whom 1 is the best response has the opportunity to revise. We argue below that the average action under a^U is not too far from the average action under a^0, and hence from x^*. Similarly, an analogous results holds when we analyze a downward counterpart of the best response dynamics. Because of payoff complementarities, there must be an equilibrium action profile sandwiched between the limit profiles obtained by the upward and downward best response dynamics. The two results imply that such an equilibrium is not far away from x^*.

In order to motivate explain the key step, it is helpful to begin with a special case when the game is (almost) deterministic and $x^* \approx 0$ (i.e., a small perturbation of Example 1). In this case, Theorem 4 follows from an argument that is based on the proof of Proposition 3 in [Morris(2000)]. First, because $x^* \approx 0$ and the definition of a^0, the initial profile a^0 has a small number ($\sim x^* N$) of agents who play action 1. Second, let a^t be the tth stage of the upper best response dynamics. At each stage, we define the *infection capacity* of profile a^t as the mass of links that connect agents who play action 1 with agents who play action 0,

$$\mathcal{F}_0\left(a\right)=\sum_{i,j:a_i^t=1,a_j^t=0} g_{ij}. \tag{6}$$

If, at stage $t+1$, player i revises her action upwards, then (a) the capacity will increase by the weight $\sum_{j:a_j^t=0} g_{ij}$ links that i has agents who play 0, and (b) decrease by the weight $\sum_{j:a_j^t=1} g_{ij}$ of links that i has with agents who play 1 in profile a^t. Recall that $\alpha > \frac{1}{2}$ is a fraction of neighbors that makes players indifferent between two actions. Since action 1 is i's best response, we have $\sum_{j:a_j^t=1} g_{ij} \approx \alpha \sum_j g_{ij} = \alpha g_i$ and

$\sum_{j:a_j^t=0} g_{ij} \approx (1-\alpha)g_i$, which implies that the capacity in each stage must decrease by $\alpha g_i - (1-\alpha)\, g_i = (2\alpha - 1)\, g_i > 0$. Because the capacity cannot fall below 0, we obtain a bound on the total mass of players who switch actions

$$(2\alpha - 1) \sum_{i:0=a_i^0<a_i^U=1} g_i \leq \mathcal{F}_0\left(a^0\right).$$

As the initial profile had a small number of agents playing 1, the number of agents who revise their actions must be small as well.

There are two important features of the above argument: the initial capacity is small and it must appropriately decrease with each action revised upward. The proof of Theorem 4 preserves the two features, but with a modified notion of capacity. We cannot use (6), because, for general payoff shocks and $x^* \in (0,1)$, a substantial fraction of the population plays each action and (6) is too large. Instead, we replace actions a_i by their expected best response versions $p_i = P\left(\frac{1}{g_i}\sum_j g_{ij} a_j\right) = P\left(\beta_i^s\right)$ and define

$$\mathcal{F}\left(p\right) = \frac{1}{2}\sum_{i,j} g_{ij}\left(p_i - p_j\right)^2. \tag{7}$$

(To motivate the definition, notice that if we replace p_i by a_i, then (6) and (7) are equal.)

The law of large numbers implies that, under the initial profile a^0, the average action among the neighbors, $\beta_i^0 = \frac{1}{g_i}\sum_j g_{ij} a_j^0$, and hence the expected best response p_i must also be close to x^*. Thus, the capacity of the initial profile is appropriately small and the first required feature of capacity is preserved.

The second feature is preserved as well. We sketch the idea here and leave the details to the Appendix. Due to symmetry in the weights $g_{ij} = g_{ji}$ for each i, j, we have for each t,

$$\begin{aligned}
&\mathcal{F}\left(p^{t+1}\right)-\mathcal{F}\left(p^{t}\right)\\
=&\sum_i g_i\left(p_i^{t+1}\right)^2-\sum_i g_i\left(p_i^t\right)^2-\sum_i\left(p_i^{t+1}-p_i^t\right)\sum_j g_{ij}\sum_{s=t,t+1}p_j^s\\
=&\sum_i g_i\left(p_i^{t+1}\right)^2-\sum_i g_i\left(p_i^t\right)^2-\sum_i g_i\left(p_i^{t+1}-p_i^t\right)\sum_{s=t,t+1}\beta_i^s\\
&+\sum_i\left(p_i^{t+1}-p_i^t\right)\sum_j g_{ij}\sum_{s=t,t+1}\left(a_j^s-p_j^s\right),
\end{aligned}$$

where in the third line we used $\beta_i^s=\frac{1}{g_i}\sum_j g_{ij}a_j^s$. Because $p_i^s=P\left(\beta_i^s\right)$, we have

$$\left(p_i^{t+1}-p_i^t\right)\sum_{s=t,t+1}\beta_i^s\approx 2\int_{p_i^t}^{p_i^{t+1}}P^{-1}(y)\,dy+\text{small terms},$$

where, here and below, the "small terms" depend on the stage increase in $\beta_i^{t+1}-\beta_i^t$, which is small due to our assumption that at most one agent revises her action per period and because the impact of a single agent in the neighborhood of another is smaller than $d(g)$. They may also depend on the difference $\beta_i^0-x^*$, which is small because the initial profile is close to x^*.

Summing across $t\leq T$, and noting that $\left(p_i^{t+1}\right)^2-\left(p_i^t\right)^2=2\int_{p_i^t}^{p_i^{t+1}}ydy$, we obtain

$$\mathcal{F}\left(p^{T+1}\right)-\mathcal{F}\left(p^0\right)=\sum_{t\leq T}\left(\mathcal{F}\left(p^{t+1}\right)-\mathcal{F}\left(p^t\right)\right)$$

$$=-2\sum_i g_i\left[\int\limits_{p_i^0}^{p_i^{T+1}}\left(P^{-1}(y)-y\right)dy\right]+\sum_{t\leq T}\sum_i\left(p_i^{t+1}-p_i^t\right)\sum_j g_{ij}\sum_{s=t,t+1}\left(a_j^s-p_j^s\right)\tag{8}$$

$$+\ \text{small terms.}$$

The details of the calculations can be found in Appendix E.

The second term of the right-hand side is small for probabilistic reasons. Notice that the probability that action 1 is a player's j best response in period s is not higher than the expected action p_j^s. In fact, the probability is not higher even it is conditioned on the actions of other agents. The reason is is that agent j's behavior positively affects the actions of other players only after she revises her action. This observation, together with the fact that each agent j is small in the neighborhood of i, allows us to show that the second last term is small, with a large probability, due to a version of the finite law of large numbers.

Ignoring all the small terms (including terms discussed in the previous paragraph), summing across t, and remembering that $p_i^t = P\left(\beta_i^t\right)$ and that $\beta_i^0 \approx x^* = P\left(x^*\right) \approx P\left(\beta_i^0\right)$, we obtain

$$\mathcal{F}\left(p^0\right) \geq 2\sum_i g_i \left[\int_{x^*}^{P\left(\beta_i^T\right)} \left(P^{-1}\left(y\right) - y\right) dy\right].$$

The definition of the RU-dominant outcome implies that, at least locally, the integral is increasing in β_i^T. Hence, if the original capacity is small, then, for each T, the average behavior in the neighborhood of a great majority of players cannot be too far away from x^*. Hence, the limit of the upper best response dynamics cannot be to far away from x^*, which concludes the argument.

7. Discussion

7.1. **Unweighted average.** Our definition of the average action stated in Section 2.2 weights individuals by their neighborhood size g_i. An alternative is to use the unweighted average

$$\mathrm{Av}_{\text{unweighted}}\left(a\right) = \frac{1}{N}\sum_i a_i.$$

When the network is balanced, i.e., when $g_i = g_j$ for each i and j, the two notions of average are identical.

Because Theorem 1, Corollary 1, and Theorem 3 are proven using balanced networks, they continue to hold *verbatim* if we change the notion of average to unweighted one. A version of Theorem 4 holds with the following modification : *for each* $w > 0$, *and each* $\eta > 0$, *there is* $d > 0$ *such that, for each network* g, *if* $d\left(g\right) \leq d$ *and* $w\left(d\right) \leq w$, *then*

$$\mathbb{P}\left(\left\{x^*\right\} \subseteq_\eta Eq\left(g,\right)\right) \geq 1 - \eta.$$

The required modification of the proof is very minor and it can be found in Appendix E.8.

We were not able to find an immediate way of extending Theorem 2.

7.2. **Small number of links.** The results of this paper focus on the limit case $d(g) \to 0$ and they apply to networks with a large number of connections (i.e., large degrees), like networks of acquaintances. If $d(g) > 0$, none of the results hold. The small-degree case requires different techniques and separate analysis and we leave it for future research.

7.3. **Independence.** Another key assumption of the model is that the payoff shocks are independent across agents. An alternative and natural assumption is that the payoff shocks of directly connected agents can be correlated. If imperfect, such a correlation dies out exponentially with the distance between agents, making distant agents roughly independent. For this reason, we suspect that the results of this paper continue to hold. However, the proper analysis of this case is left to future research.

Appendix A. Monotonicity

This part of the Appendix shows that if P is a continuum best response function of random utility game (u, F), then, for any increasing and right-continuous function $P' \geq P$, there is a random utility game that has P' as a continuum best response function, and such that the distribution of equilibria first-order stochastically dominates the distribution of equilibria in the original game.

Formally, the space of (mixed) action profiles $\mathcal{A} = [0,1]^N$ is a lattice with coordinate-wise comparison: for any $a, b \in \mathcal{A}$, we have $a \leq b$ iff $a_i \leq b_i$ for each i. Let $\leq_S$ denote the strong set order on subsets or $\mathbb{R}$ and, as a lattice extension, of $\mathcal{A}$. We say that a probability distribution $\mu \in \Delta\mathcal{A}$ is dominated by $\mu' \in \Delta\mathcal{A}$ in the sense of first-order stochastic dominance, and write $\mu \leq_{FOS} \mu'$, if for each $a \in \mathcal{A}$, $\mu(\{a' : a' \geq a\}) \leq \mu'(\{a' : a' \geq a\})$.

Let (u, F) be a random utility game. Let $E(u, \varepsilon)$ denote the set of equilibrium profiles in random utility game (u, F). We compare sets using the strong set order. Let $\mu(u, F)$ denote the probability distribution over the sets of equilibrium profiles induced by distribution over profiles of payoffs shocks. We say that random utility game (u, F) is dominated by game (u', F') if $\mu(u, F) \leq_{FOS} \mu(u', F')$.

Lemma 1. *Suppose that P is a continuum best response function of random utility game (u, F). Then, for each increasing, right-continuous $P' \geq P$, there exists random utility game (u', F') such that (a) P' is a continuum best response function of (u', F'), and (b) random utility game (u, F) is dominated by game (u', F').*

Proof. First, observe that any two random utility models with the same continuum best response function P have the same distributions over sets of equilibria. Second, we show that we can construct different models over the same probability space. Let $\Omega = [0,1]^N$ and let λ^N be the product uniform measure on Ω. For each increasing, right-continuous P, define utility function u_P so that

$$u_P(1, x, \varepsilon) = x - P^{-1}(\varepsilon) \text{ and } u_P(0, x, \varepsilon) = 0.$$

Then, the continuum best response function of (u_P, λ) is equal to P. Finally, notice that if $P' \geq P$, and we consider two games u_P and u'_P on the same probability space $\left(\Omega, \lambda^N\right)$, then the best resposne of each player in the second model is always higher (in the sense of strong set order) than the best response of the player in the first model. A consequence is that, for each ε, $E(u_P, \varepsilon) \leq E(u_{P'}, \varepsilon)$, which concludes the proof of the result. $\square$

Appendix B. Proof of Theorem 1 and Corollary 1

B.1. **Proof of Theorem 1.** Let U be an open set from the definition of a strongly stable x. Fix $\delta > 0$ and $N < \infty$ such that $[x - 2\delta, x + 2\delta] \subseteq U$ and $\gamma \frac{1}{N-1} \leq \frac{1}{2}(1-\gamma)\delta$. Let $\eta = \frac{1}{2}(1-\gamma)\delta$. Then,

$$x - \delta \leq x - \delta + \left((1-\gamma)\delta - \gamma\frac{1}{N-1}\right) - \eta \leq P(x) - \gamma\left(\delta + \frac{1}{N-1}\right) - \eta \leq P\left(x - \delta - \frac{1}{N-1}\right) - \eta,$$

and similarly, $P\left(x + \delta + \frac{1}{N-1}\right) + \eta \leq x + \delta$. Additionally, choose a sufficiently large N so that $2\exp\left(-2N\eta^2\right) \leq \eta$.

Let

$$P_{\varepsilon}(x)=\frac{1}{N}\sum_{i}\mathbf{1}\left(\beta\left(\varepsilon_{i}\right)\leq x\right),$$

be the empirical distribution of best response thresholds. Define event $\mathcal{P}=\left\{\sup_{x}\left|P_{\varepsilon}(x)-P(x)\right|\leq\eta\right\}$. By the Dvoretsky-Kiefer–Wolfowitz–Massart inequality, for each $\eta>0$,

$$\operatorname{Prob}(\text{not }\mathcal{P})\leq 2\exp\left(-2N\eta^{2}\right)\leq\eta.$$

For each profile a, define $\beta_{i}^{a}=\frac{1}{N-1}\sum_{j\neq i}a_{j}$ as the average action in player i's neighborhood. The average action is not far from the average action in the population, $\left|\beta_{i}^{a}-\operatorname{Av}(a)\right|\leq\frac{1}{N-1}$.

Suppose that event $\mathcal{P}$ holds. Let $b(a,\varepsilon)$ be the best response profile to profile a, where, in a case of a tie, we assume that an agent chooses 1. Then,

$$\operatorname{Av}(b(a,\varepsilon))=\frac{1}{N}\sum\mathbf{1}\left\{\beta\left(\varepsilon_{i}\right)\leq\beta_{i}^{a}\right\}.$$

If $\operatorname{Av}(a)\in[x-\delta,x+\delta]$, the above inequalities imply that

$$\begin{aligned}
&x-\delta\\
\leq&P\left(\operatorname{Av}(a)-\frac{1}{N-1}\right)-\eta\leq\frac{1}{N}\sum\mathbf{1}\left\{\beta\left(\varepsilon_{i}\right)\leq\operatorname{Av}(a)-\frac{1}{N-1}\right\}\\
\leq&\operatorname{Av}(b(a,\varepsilon))\\
\leq&\frac{1}{N}\sum\mathbf{1}\left\{\beta\left(\varepsilon_{i}\right)\leq\operatorname{Av}(a)+\frac{1}{N-1}\right\}\leq\eta+P\left(\operatorname{Av}(a)+\frac{1}{N-1}\right)\\
\leq&x-\delta.
\end{aligned}$$

Hence, mapping $b(,.\varepsilon)$ maps the set of profiles a s.t. $\operatorname{Av}(a)\in[x-\delta,x+\delta]$ into itself. The result follows from the fixed-point theorem.

B.2. **Proof of Corollary 1.** By Theorem 1, for each $\eta>0$ and for sufficiently large N,

$$\mathbb{P}\left(\left\{x_{\min},x_{\max}\right\}\subseteq_{\frac{1}{8}\eta}\operatorname{Eq}\left(g^{N},\varepsilon\right)\right)\leq\sum_{x\in\left\{x_{\min},x_{\max}\right\}}\mathbb{P}\left(\{x\}\subseteq_{\frac{1}{8}\eta}\operatorname{Eq}\left(g^{N},\varepsilon\right)\right)\leq\frac{1}{4}\eta.$$

Let $g=g^{K,N}$ be a balanced network that consists of K copies of complete N-person graphs. Let g_{k} denote the kth copy. Let $A(g,\varepsilon)=\left\{k:\left\{x_{\min},x_{\max}\right\}\subseteq_{\frac{1}{8}\eta}\operatorname{Eq}\left(g_{k},\varepsilon\right)\right\}$

be the set of copies that contain equilibria with averages close to the largest and the smallest of the fixed points. By the choice of N and the Central Limit Theorem, for sufficiently large K,

$$\text{Prob}\left(\frac{1}{K}\left|A\left(g,\varepsilon\right)\right| \leq 1-\frac{3}{8}\eta\right) \leq \eta.$$

Let $\psi_{\max}, \psi_{\min} : \{1, ..., K\} \to [0,1]$ be functions such that $\psi_s(k) \in \text{Eq}(g_k, \varepsilon)$ for each $s = \max, \min$ and each k, and, if $k \in A(g, \varepsilon)$, then $|\psi_s(k) - x_s| \leq \frac{1}{8}\eta$. Then, for each subset $B \subseteq \{1, ..., K\}$ of copies, there is an equilibrium a with average payoffs equal to

$$\text{Av}\left(a\right) = \frac{1}{K}\left(\sum_{k\in B}\psi_{\max}\left(k\right) + \sum_{k\notin B}\psi_{\min}\left(k\right)\right).$$

Because of the choice of $\psi_{.}(.)$,

$$\frac{|B|}{K}x_{\max} + \frac{K-|B|}{K}x_{\min} - \frac{1}{2}\eta \leq \text{Av}\left(x\right) \leq \frac{|B|}{K}x_{\max} + \frac{K-|B|}{K}x_{\min} + \frac{1}{2}\eta.$$

If $K \geq \frac{2}{\eta}$, for any $x \in [x_{\min}, x_{\max}]$, we can choose B, and hence arrive at equilibrium a, so that the average payoffs in a are at most η-far from x, $|\text{Av}(a) - x| \leq \eta$.

Appendix C. Proof of Theorem 2

The first subsection introduces notation and metric d. Section C.2 derives various deterministic inequalities connecting metric d and average behavior. Section C.3 derives probabilistic bounds. The next two sections contain steps (a) and (b) described in the introduction. The last section concludes the proof of the theorem.

C.1. **Preliminary remarks.** We begin with few preliminary remarks to simplify the problem. First, it is enough to establish one side of the probability bound: for each $\eta > 0$ and $w < \infty$, there is $\delta > 0$, such that for each network g, if $d(g) \leq \delta$, $w(g) \leq w$, then

$$\mathbb{P}\left(\max \text{Eq}\left(g^N, \varepsilon\right) \geq x_{\max} + \eta\right) \leq \eta.$$

The proof of the other probability bound is analogous and the two bounds together combine to the statement of the theorem.

Second, say that a is an upper equilibrium if, whenever indifferent, each agent plays action 1. Because of supermodularity, if a is an equilibrium, there exists an upper equilibrium a' st. $a' \geq a$. Thus, it is enough to show the above probability bound when set $\mathrm{Eq}(g, \varepsilon)$ contains only the average payoffs in all upper equilibria.

Third, because $x_{\max}$ is strongly stable, there exists a constant $\gamma < 1$ such that for each x,

$$P(x) \leq \max\left(x_{\max}, x_{\max} + (1-\gamma)(x - x_{\max})\right) = P^*(x).$$

(Such constant exists locally due to the definition of strong stability. The existence for all x follows from compactness and the fact that $x_{\max}$ is the largest fixed point of P.) Because P^* is increasing and right-continuous, Lemma 1 implies that there exists a random utility game (u^*, F^*) with continuum best response function P^* that dominates (u, F). Thus, it is enough to show the second claim in Theorem 2 for game (u^*, F^*). Henceforth, we assume that P^* is the continuum best response function. Notice that P^* is Lipschitz with a Lipschitz constant equal to γ.

Finally, we will use the following notation. Let $b_i(a, \varepsilon) = \max\left(\arg\max_{a_i} u_i\left(a_i, \beta_i^a, \varepsilon\right)\right)$ be the largest best response action of agent i against a_{-i} given payoff shock ε_i. Let $b(a, \varepsilon)$ be the profile of best responses. If a is an upper equilibrium given ε, then $b(a, \varepsilon) = a$. Also, we denote $p^a = \left(P^*\left(\beta_i^a\right)\right)_i$ to be the profile of expected best responses.

Let $\mathcal{A} = [0,1]^N$ be the space of (mixed) action profiles. Let $\mathcal{B} = \{\beta^a : a \in \mathcal{A}\}$ be the set of profiles $\beta^a = (\beta_i^a)$ of neighborhood behaviors that can be generated from the profiles. We assume that $\mathcal{A}$ is a subset of a normed space $\mathbb{R}^N$ with a metric (3). Under this metric, $\mathrm{diam}\mathcal{A} = 1$.

Let $g_{\min} = \min_i g_i$ and $g_{\max} = \max_i g_i$.

C.2. **Deterministic relationships.** This section discusses deterministic relationships and bounds on profiles.

Lemma 2. *For each profile $a \in \mathcal{A}$,*

$$Av(a) = Av(\beta^a).$$

Proof. Notice that

$$\mathrm{Av}(\beta^a) = \frac{1}{\sum_i g_i} \sum_i g_i \frac{1}{g_i} \sum_j g_{ij} a_j = \frac{1}{\sum_i g_i} \sum_i \sum_j g_{ij} a_j = \frac{1}{\sum_i g_i} \sum_j a_j g_j = \mathrm{Av}(a).$$

□

Lemma 3. *For any profiles $a, b \in \mathcal{A}$,*

$$\left|Av\left(P^*\left(a\right)\right) - Av\left(P^*\left(b\right)\right)\right| \leq \left|Av\left(a\right) - Av\left(b\right)\right|.$$

Proof. The inequality follows from P^* being Lipschitz with a constant $\gamma < 1$. □

Lemma 4. *For any profiles $a, b \in \mathcal{A}$,*

$$\left|Av\left(a\right) - Av\left(b\right)\right| \leq \sqrt{w\left(g\right)}d\left(a, b\right).$$

Proof. Notice that

$$\begin{aligned}\left|\mathrm{Av}\left(a\right) - \mathrm{Av}\left(b\right)\right| &\leq \frac{1}{\sum g_i}\sum g_i\left|a_i - b_i\right| \leq \sqrt{\frac{1}{\sum g_i}\sum g_i\left(a_i - b_i\right)^2}\\ &\leq \sqrt{w\left(g\right)\frac{1}{\sum g_i^2}\sum g_i^2\left(a_i - b_i\right)^2} = \sqrt{w\left(g\right)}d\left(a, b\right),\end{aligned}$$

where the second inequality follows from the Jensen's inequality, and the third one from $\sum g_j^2 \leq g_{\max}\sum g_j \leq w\left(g\right)g_i\sum g_j$ for each i. □

Lemma 5. *Suppose that profile b is such that $b_i \geq x_{\max}$ for each i. Then, for each profile a,*

$$\frac{1}{\sum_i g_i}\sum_i g_i \max\left(x_{\max} - a_i, 0\right) \leq \sqrt{w\left(g\right)}d\left(a, b\right)$$

Proof. For each profile a, define profile $\min\left(x_{\max}, a\right)$ so that $\left(\min\left(x_{\max}, a\right)\right)_i = \min\left(x_{\max}, a_i\right)$. Then, because function $f\left(y\right) = \min\left(y, x_{\max}\right)$ is Lipschitz with constant 1, we have

$$d\left(\min\left(a, x_{\max}\right), x_{\max}\right) = d\left(\min\left(a, x_{\max}\right), \min\left(b, x_{\max}\right)\right) \leq d\left(a, b\right),$$

where, abusing notation, we write $x_{\max}$ to denote the constant profile, and we use the fact that $\min\left(b, x_{\max}\right) = x_{\max}$. By Lemma 4,

$$\begin{aligned}\frac{1}{\sum_i g_i}\sum_i g_i \max\left(x_{\max} - a_i, 0\right) &= \mathrm{Av}\left(x_{\max}\right) - \mathrm{Av}\left(\min\left(a, x_{\max}\right)\right)\\ &= \mathrm{Av}\left(\min\left(b, x_{\max}\right)\right) - \mathrm{Av}\left(\min\left(a, x_{\max}\right)\right)\\ &\leq \sqrt{w\left(g\right)}d\left(\min\left(a, x_{\max}\right), \min\left(b, x_{\max}\right)\right) \leq \sqrt{w\left(g\right)}d\left(a, b\right).\end{aligned}$$

□

Lemma 6. *Suppose that profile b is such that $b_i \geq x_{\max}$ for each i. Then, for each profile a,*

$$Av(a) - Av(P^*(a)) \geq (1-\gamma)(Av(a) - x_{\max}) - 2(w(g))^{\frac{1}{4}}\sqrt{d(a,b)},$$

where $P^*(a)$ is a profile of actions $P^*(a_i)$ for each agent i.

Proof. Lemma 5 implies that

$$\frac{1}{\sum_i g_i}\sum_i g_i \max(x_{\max} - a_i, 0) \leq \sqrt{w(g)}d(a,b) =: \delta,$$

where the last equality defines δ. Let $A = \sum_{i:a_i \leq x_{\max}-\sqrt{\delta}} g_i$ and notice that

$$\begin{aligned} A &= \sum_{i:a_i\leq x_{\max}-\sqrt{\delta}} g_i \leq \frac{1}{\sqrt{\delta}}\sum_{i:a_i\leq x_{\max}-\sqrt{\delta}} g_i(x_{\max}-a_i) \leq \frac{1}{\sqrt{\delta}}\sum_i g_i \max(x_{\max}-a_i,0) \\ &\leq \frac{\delta}{\sqrt{\delta}}\sum g_i = \sqrt{\delta}\sum g_i. \end{aligned}$$

Hence

$$\begin{aligned} &\mathrm{Av}(a) - \mathrm{Av}(P^*(a)) = \frac{1}{\sum g_i}\sum_i g_i(a_i - P^*(a_i)) \\ &\geq -\frac{1}{\sum g_i}\sum_{i:a_i\leq x_{\max}-\sqrt{\delta}} g_i - \sqrt{\delta}\frac{1}{\sum g_i}\sum_{i:x_{\max}\geq a_i\geq x_{\max}-\sqrt{\delta}} g_i + \frac{1}{\sum g_i}\sum_{i:a_i\geq x_{\max}} g_i(a_i - P^*(a_i)) \\ &\geq -\frac{1}{\sum g_i}A - \sqrt{\delta} + (1-\gamma)\frac{1}{\sum g_i}\sum_{i:a_i\geq x_{\max}} g_i(a_i - x_{\max}) \\ &\geq (1-\gamma)(\mathrm{Av}(a) - x_{\max}) - 2\sqrt{\delta}. \end{aligned}$$

The first inequality is due to $P^*(a_i) = x_{\max}$ for $a_i \leq x_{\max}$ and the fact that $x_{\max} - a_i \leq 1$. □

C.3. **Probability bounds on distances between profiles.** This subsection contains probabilistic bounds on the distances between profiles of neighborhood behaviors. First, we show that the distance between neighborhood behaviors obtained from the best response profile $b(a,\varepsilon)$ and the expected best response profile p^a is likely to be small.

Lemma 7. *There exists a universal constant $c<\infty$ such that, for each profile a,*

$$\mathbb{P}\left(d\left(\beta^{b(a,\varepsilon)},\beta^{p^a}\right)\geq\eta\right)\leq\exp\left(-\frac{c}{\left(w\left(g\right)\right)^4}N\left(\eta^2-d\left(g\right)\right)^2\right).$$

Proof. Notice that

$$\begin{aligned}\sum g_i^2\left(d\left(\beta^{b(a,\varepsilon)},\beta^{p^a}\right)\right)^2&=\sum g_i^2\left(\beta_i^{b(a,\varepsilon)}-\beta_i^{p^a}\right)^2\\ =&\sum_i\left(\sum_j g_{ij}\left(b_j\left(a,\varepsilon\right)-p_j^a\right)\right)^2\\ =&\sum_{j\neq k}\left(\sum_i g_{ji}g_{ik}\right)\left(b_j\left(a,\varepsilon\right)-p_j^a\right)\left(b_k\left(a,\varepsilon\right)-p_k^a\right)+\sum_j\left(\sum_i g_{ij}^2\right)\left(b_j\left(a,\varepsilon\right)-p_j^a\right)^2.\end{aligned}$$

Because $g_{ij}\leq d\left(g\right)g_i$, the second term is not larger than $d\left(g\right)\sum g_i^2$. Let $x_j=b_j\left(a,\varepsilon\right)-p_j^a$ for each j. Then,

$$\mathbb{P}\left(d\left(\beta^{b(a,\varepsilon)},\beta^{p^a}\right)\geq\eta\right)\leq\mathbb{P}\left(\sum_{j\neq k}\left(\sum_i g_{ji}g_{ik}\right)x_jx_k\geq\left(\eta^2-d\left(g\right)\right)\sum_i g_i^2\right).$$

Let $g_{jk}^{(2)}=\sum_i g_{ji}g_{ik}$ and let $G^{(2)}$ be the symmetric matrix of elements $g_{jk}^{(2)}$. Observe that

$$g_{jk}^{(2)}=\sum_i g_{ji}g_{ik}=\sum_i\frac{g_{ji}g_{ik}}{g_jg_i}g_jg_i\leq\left(w\left(g\right)\right)^2g_{\min}^2\pi_{jk},$$

where we denote $\pi_{jk}=\sum_i\frac{g_{ji}}{g_j}\frac{g_{ik}}{g_i}$. Note that, for each j, $\sum_k\pi_{jk}=\sum_{k,i}\frac{g_{ji}}{g_j}\frac{g_{ik}}{g_i}=\sum_i\frac{g_{ji}}{g_j}=1$. Hence $\pi_{jk}\leq 1$.

Because the best response of each player i depends only on independent shock ε_i (and not on other payoff shocks), x_j and x_k are independent for $j\neq k$. Hence the expected value of $\sum_{j\neq k}\left(\sum_i g_{ji}g_{ik}\right)x_jx_k$ is equal to 0, and we can use the Hansen-Wright inequality (Theorem 6.2.1 in [Vershynin(2018)]):

$$\mathbb{P}\left(\sum_{j\neq k}\left(\sum_i g_{ji}g_{ik}\right)x_jx_k\geq t\right)\leq 2\exp\left(-ct^2\left\|G^{(2)}\right\|_F^{-2}\right),$$

where c is some universal constant (note that the random variables x_j are bound by 2), and where $\left\|G^{(2)}\right\|_F$ is the Frobenius norm of matrix $G^{(2)}$:

$$\begin{aligned}\left\|G^{(2)}\right\|_F^2 &= \sum_i \sum_j \left(g_{ij}^{(2)}\right)^2 \leq (w(g))^4 g_{\min}^4 \sum_i \sum_j \pi_{ij}^2 \\ &\leq (w(g))^4 g_{\min}^4 \sum_i \sum_j \pi_{jk} \leq (w(g))^4 g_{\min}^4 N.\end{aligned}$$

Take $t = (\eta^2 - d(g)) \sum_i g_i^2$, and notice that $\sum_i g_i^2 \geq N g_{\min}^2$ to obtain the inequality in the statement of the lemma. □

The second result shows that, for any fixed profile a_0, the maximum distance between neighborhood behaviors obtained as the best response to a_0 and the best response to some other profile a, across all profiles a that have similar neighborhood behaviors to a_0, is small.

Lemma 8. *For each profile a_0,*

$$\mathbb{P}\left(\sup_{a:d(\beta^a,\beta^{a_0})\leq\delta} d\left(\beta^{b(a_0,\varepsilon)}, \beta^{b(a,\varepsilon)}\right) \geq \eta\right) \leq \exp\left(-\frac{1}{2(w(g))^4} N \left(\eta - 3\delta^{2/3}\right)^2\right).$$

Proof. For payoff shock ε, let $\beta(\varepsilon)$ such that $\beta(\varepsilon) = \max\{\beta : u(1,\beta,\varepsilon) \leq u(0,\beta,\varepsilon)\}$ with $\beta = -\infty$ if the set is empty. For each profile a and player i, $b_i(a,\varepsilon) \neq b(a_0,\varepsilon)$ if and only if either $\beta_i^a \leq \beta(\varepsilon) < \beta_i^{a_0}$ or $\beta_i^{a_0} \leq \beta(\varepsilon) < \beta_i^a$. Denote a random variable $\beta \geq \beta(\varepsilon) X_i = \mathbf{1}\left\{\beta(\varepsilon_i) \in \left[\beta_i^{a_0} - \delta^{2/3}, \beta_i^{a_0} + \delta^{2/3}\right]\right\}$. Then, for any profile a,

$$|b_i(a,\varepsilon) - b_i(a_0,\varepsilon)| \leq X_i \mathbf{1}\left\{|\beta_i^a - \beta_i^{a_o}| \leq \delta^{2/3}\right\} + \mathbf{1}\left\{|\beta_i^a - \beta_i^{a_o}| > \delta^{2/3}\right\}.$$

Hence

$$
\begin{aligned}
&\sup_{a:d(\beta^a,\beta^{a_0})\le\delta}\sum_i g_i^2\left(b_i\left(a,\varepsilon\right)-b_i\left(a_0,\varepsilon\right)\right)^2\\
&\le\sum g_i^2X_i^2+\sup_{a:d(\beta^a,\beta^{a_0})\le\delta}\sum_{i:\left|\beta_i^a-\beta_i^{a_o}\right|>\delta^{2/3}}g_i^2\\
&\le\sum g_i^2X_i^2+\sup_{a:d(\beta^a,\beta^{a_0})\le\delta}\delta^{-4/3}\sum_{i:\left|\beta_i^a-\beta_i^{a_o}\right|>\delta^{2/3}}g_i^2\left(\beta_i^a-\beta_i^{a_o}\right)^2\\
&\le\sum g_i^2X_i^2+\delta^{-4/3}\sup_{a:d(\beta^a,\beta^{a_0})\le\delta}\sum g_i^2\left(d\left(\beta^a,\beta^{a_0}\right)\right)^2,\\
&\le\sum g_i^2X_i^2+\delta^2\delta^{-4/3}\sum g_i^2=\sum g_i^2X_i^2+\delta^{2/3}\sum g_i^2.
\end{aligned}
$$

Variables $X_i^2=X_i$ are independent Bernoulli variables with parameter $\mathbb{E}\,X_i=P^*\left(\beta_i^a+\delta^{2/3}\right)-P^*\left(\beta_i^a+\delta^{2/3}\right)\le 2\delta^{2/3}$ as P^* is Lipschitz with constant 1. The Hoeffding's inequality shows that

$$
\begin{aligned}
\mathbb{P}\left(\sum g_i^2X_i^2+\delta^{2/3}\sum g_i^2\ge\eta\sum g_i^2\right)&\le\mathbb{P}\left(\sum g_i^2\left(X_i-\mathbb{E}\,X_i\right)\ge\left(\eta-3\delta^{2/3}\right)\sum g_i^2\right)\\
&\le\exp\left(-\frac{\left(\sum g_i^2\right)^2}{2\sum g_i^4}\left(\eta-3\delta^{2/3}\right)^2\right).
\end{aligned}
$$

Finally, notice that $2\sum g_i^4\le\left(w\left(g\right)\right)^4g_{\min}^4N$ and $\left(\sum g_i^2\right)^2\ge g_{\min}^4N^2$. □

C.4. **Probability bound on the existence of an upper equilibrium.** This subsection finds a bound on the probability that, for any profile a_0, there exists a profile a with similar neighborhood behaviors as a_0, and such that a is an upper equilibrium.

Lemma 9. *For each $\xi>0$ and each $w<\infty$, there is $\delta>0$ so that, for each profile a_0 such that $Av\left(a_0\right)>x_{\max}+\xi$, and for each network g such that $d\left(g\right)\le\delta$ and $w\left(g\right)\le w$,*

$$\mathbb{P}\left(\textit{there exists } a \textit{ s.t. } a \textit{ is upper equilibrium and } d\left(\beta^a,\beta^{a_0}\right)\le\delta\right)\le 2\exp\left(-\delta N\right).$$

Proof. Choose $\eta,\delta>0$ such that

$$
\begin{aligned}
\left(1-\gamma\right)\xi&>\sqrt{w\left(g\right)}\left(2\delta+2\eta\right)+2\left(w\left(g\right)\right)^{\frac{1}{4}}\sqrt{2\eta}\text{ and}\\
\delta&\le\frac{c}{\left(w\left(g\right)\right)^4}\left(\eta^2-\delta\right)^2+\frac{1}{2\left(w\left(g\right)\right)^4}\left(\eta-3\delta^{2/3}\right)^2.
\end{aligned}
$$

Assume that $d(g) \leq \delta$.

Consider the following three events:

$$A = \left\{ d\left(\beta^{b(a_0,\varepsilon)}, \beta^{p^{a_0}}\right) \leq \eta \right\},$$
$$B = \left\{ \sup_{a: d(\beta^a, \beta^{a_0}) \leq \delta} d\left(\beta^{b(a_0,\varepsilon)}, \beta^{b(a,\varepsilon)}\right) \leq \eta \right\}.$$

Due to Lemmas 7 and 8, the probability that at least one of the two events does not hold is no larger than

$$\exp\left(-\frac{c}{(w(g))^4} N\left(\eta^2 - d(g)\right)^2\right) + \exp\left(-\frac{1}{2(w(g))^4} N\left(\eta - 3\delta^{2/3}\right)^2\right) \leq 2\exp(-\delta N).$$

Assume that the two events hold simultaneously. We will show that there exists no a such that $d(\beta^a, \beta^{a_0}) \leq \delta$ and such that a is an upper equilibrium.

On the contrary, suppose that such a exists. Then, $a = b(a, \varepsilon)$. Because d is a metric and events A and B hold,

$$d\left(\beta^a, \beta^{p^{a_0}}\right) = d\left(\beta^{b(a,\varepsilon)}, \beta^{p^{a_0}}\right) \leq d\left(\beta^{b(a,\varepsilon)}, \beta^{b(a_0,\varepsilon)}\right) + d\left(\beta^{b(a_0,\varepsilon)}, \beta^{p^{a_0}}\right) \leq 2\eta.$$

Because $\beta_i^{p^{a_o}} = \frac{1}{g_i}\sum_i g_{ij} P^*(a_{0,j}) \geq x_{\max}$ for each i, we can apply Lemma 6 to β^a instead of a and $\beta^{p^{a_0}}$ instead of b (notice that $p^a = P^*(\beta^a)$ by definition):

$$\mathrm{Av}(\beta^a) - \mathrm{Av}(p^a) \geq (1-\gamma)\left(\mathrm{Av}(\beta^a) - x_{\max}\right) - 2(w(g))^{\frac{1}{4}}\sqrt{2\eta}. \tag{9}$$

By Lemmas 2, 3, and 4, and because $d(\beta^a, \beta^{a_0}) \leq \delta$,

$$\left|\mathrm{Av}(p^a) - \mathrm{Av}(p^{a_0})\right| \leq \left|\mathrm{Av}(a) - \mathrm{Av}(a_0)\right| = \left|\mathrm{Av}(\beta^a) - \mathrm{Av}(\beta^{a_0})\right| \leq \sqrt{w(g)}\delta.$$

By Lemmas 2 and 4, and because event A holds,

$$\left|\mathrm{Av}(p^{a_0}) - \mathrm{Av}(b(a_0,\varepsilon))\right| = \left|\mathrm{Av}\left(\beta^{p^{a_0}}\right) - \mathrm{Av}\left(\beta^{b(a_0,\varepsilon)}\right)\right| \leq \sqrt{w(g)}\eta.$$

By Lemmas 2 and 4, because a is an upper equilibrium, and because event B holds,

$$\begin{aligned}\left|\mathrm{Av}\left(b\left(a_0,\varepsilon\right)\right)-\mathrm{Av}\left(\beta^a\right)\right| &= \left|\mathrm{Av}\left(b\left(a_0,\varepsilon\right)\right)-\mathrm{Av}\left(a\right)\right| \\ &= \left|\mathrm{Av}\left(b\left(a_0,\varepsilon\right)\right)-\mathrm{Av}\left(b\left(a,\varepsilon\right)\right)\right| \\ &= \left|\mathrm{Av}\left(\beta^{b(a_0,\varepsilon)}\right)-\mathrm{Av}\left(\beta^{b(a,\varepsilon)}\right)\right| \leq \sqrt{w\left(g\right)}\eta.\end{aligned}$$

Putting the three inequalities together, we obtain

$$\begin{aligned}&\left|\mathrm{Av}\left(\beta^a\right)-\mathrm{Av}\left(p^a\right)\right| \\ &\leq \left|\mathrm{Av}\left(b\left(a_0,\varepsilon\right)\right)-\mathrm{Av}\left(\beta^a\right)\right| + \left|\mathrm{Av}\left(b\left(a_0,\varepsilon\right)\right)-\mathrm{Av}\left(p^{a_a}\right)\right| + \left|\mathrm{Av}\left(p^a\right)-\mathrm{Av}\left(p^{a_0}\right)\right| \\ &\leq \sqrt{w\left(g\right)}\left(\delta+2\eta\right). \end{aligned} \tag{10}$$

Combining inequalities (9) and (10), we obtain

$$\left(\sqrt{w\left(g\right)}\left(\delta+2\eta\right)+2\left(w\left(g\right)\right)^{\frac{1}{4}}\sqrt{2\eta}\right) \geq \left(1-\gamma\right)\left(\mathrm{Av}\left(\beta^a\right)-x_{\max}\right).$$

By Lemmas 2, 4, and because $d\left(\beta^a,\beta^{a_0}\right)\leq\delta$,

$$\left|\mathrm{Av}\left(\beta^a\right)-\mathrm{Av}\left(a_0\right)\right| = \left|\mathrm{Av}\left(\beta^a\right)-\mathrm{Av}\left(\beta^{a_0}\right)\right| \leq \sqrt{w\left(g\right)}\delta.$$

Hence,

$$\left(\sqrt{w\left(g\right)}\left(2\delta+2\eta\right)+2\left(w\left(g\right)\right)^{\frac{1}{4}}\sqrt{2\eta}\right) \geq \left(1-\gamma\right)\left(\mathrm{Av}\left(a_0\right)-x_{\max}\right) \geq \left(1-\gamma\right)\xi.$$

However, this violates the choice of the parameters η and δ. □

C.5. **Metric entropy bound.** For each $\delta>0$, let $\mathcal{N}\left(\delta,\mathcal{B}\right)$ be the covering number of $\mathcal{B}$, i.e., the smallest cardinality n of a list of profiles $b^1,...,b^n\in\mathcal{B}$ such that for each $b\in\mathcal{B}$, there is $l\leq n$ so that $d\left(b,b^l\right)\leq\delta$.

Lemma 10. *There exists a constant $c<\infty$ such that, for each $\delta>0$, and each network g,*

$$\mathcal{N}\left(\delta,\mathcal{B}\right) \leq \exp\left(\frac{1}{\delta^2}c\left(w\left(g\right)\right)^2 d\left(g\right)N\right).$$

Proof. We will use the Sudakov's Minoration Inequality (Theorem 7.4.1 from [Vershynin(2018)]) which provides an upper bound on the covering number via the expectation of a certain

Gaussian process. For this, let Z_i for each agent i be an i.i.d. standard normal random variable. For each (possibly mixed) profile $a \in \mathcal{A}$, define

$$X_a = \frac{1}{\sqrt{\sum_i g_i^2}} \sum_i g_i a_i Z_i.$$

For any two profiles $a, b \in \mathcal{A}$,

$$\begin{aligned}\sqrt{\mathbb{E}\left(X_a - X_b\right)^2} &= \sqrt{\frac{1}{\sqrt{\sum g_i^2}}\, \mathbb{E}\left(\sum_i g_i\left(a_i - b_i\right) Z_i\right)^2} \\ &= \sqrt{\frac{1}{\sqrt{\sum g_i^2}} \sum_i g_i\left(a_i - b_i\right)^2} = d\left(a, b\right).\end{aligned}$$

Given the definition and the above property, the Sudakov's Minoration Inequality implies that, for some universal (i.e., independent of parameters and the current problem) constant $c_1 > 0$,

$$\log \mathcal{N}\left(\delta, \mathcal{B}\right) \leq c_1 \frac{\left(\mathbb{E}\sup_{b \in \mathcal{B}} X_b\right)^2}{\delta^2}.$$

We compute

$$\begin{aligned}\mathbb{E}\sup_{b \in \mathcal{B}} X_b = \mathbb{E}\sup_{a \in \mathcal{A}} X_{\beta^a} &= \mathbb{E}\left(\sup_{a \in \mathcal{A}} \frac{1}{\sqrt{\sum_i g_i^2}} \sum_i g_i Z_i \left(\frac{1}{g_i}\sum g_{ij} a_j\right)\right) \\ &= \frac{1}{\sqrt{\sum_i g_i^2}}\, \mathbb{E}\left(\sup_{a \in \mathcal{A}} \sum_i a_i \left(\sum_j g_{ij} Z_j\right)\right) \leq \frac{1}{\sqrt{\sum_i g_i^2}}\, \mathbb{E}\sum_i \left|\sum_j g_{ij} Z_j\right| \\ &\leq \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{\sum_i g_i^2}} \sum_i \sqrt{\sum_j g_{ij}^2},\end{aligned}$$

where the last inequality is due to a bound on the expectation of the absolute value of the normal variable $\sum g_{ij} Z_j$ via its standard deviation $\sigma_i = \sqrt{\sum_j g_{ij}^2}$. Because $\sum_j g_{ij}^2 \leq d\left(g\right) g_i^2$ and $\left(\sum_i g_i\right)^2 \leq N^2 \left(w\left(g\right)\right)^2 g_{\min}^2 \leq N\left(w\left(g\right)\right)^2 \sum g_i^2$, we have

$$\log \mathcal{N}\left(\delta, \mathcal{B}\right) \leq \sqrt{\frac{2}{\pi}} c_1 \frac{1}{\delta^2} \frac{1}{\sum_i g_i^2} \left(\sum_i \sqrt{d\left(g\right)} g_i\right)^2 d\left(g\right) \leq \frac{1}{\delta^2} \sqrt{\frac{2}{\pi}} c_1 \left(w\left(g\right)\right)^2 d\left(g\right) N.$$

□

C.6. **Proof of Theorem 2.** Fix $\eta > 0$ and $w < \infty$. Use Lemma 9 to find $\delta > 0$ and $\delta \leq \frac{1}{2\sqrt{w}}\eta$ such that, for each profile b, and each network g, if $\mathrm{Av}(b) \geq x_{\max} + \frac{1}{2}\eta$, $d(g) \leq \delta$, and $w(g) \leq w$, then

$$\mathbb{P}\left(\text{there exists } a \text{ s.t. } a \text{ is upper equilibrium and } d\left(\beta^a, \beta^b\right) \leq \delta\right) \leq 2\exp\left(-\delta N\right).$$

Use Lemma 10 to find a list of $n \leq \exp\left(\frac{1}{\delta^2} c\left(w(g)\right)^2 d(g) N\right)$ profiles $b^1, ..., b^n$ such that, for each profile $a \in \mathcal{A}$, there is $l \leq n$ such that $d\left(\beta^{b^l}, \beta^a\right) \leq \delta$. Observe that if a is such that $\mathrm{Av}(a) > x_{\max} + \eta$ and $d\left(\beta^{b^l}, \beta^a\right) \leq \delta$ for some l, then, by Lemmas 2 and 4,

$$\begin{aligned}\mathrm{Av}\left(b^l\right) - \left(x_{\max} + \frac{1}{2}\eta\right) &\geq \mathrm{Av}(a) - \left(x_{\max} + \eta\right) + \frac{1}{2}\eta - \left|\mathrm{Av}(a) - \mathrm{Av}\left(b^l\right)\right| \\ &\geq \frac{1}{2}\eta - \left|\mathrm{Av}\left(\beta^a\right) - \mathrm{Av}\left(\beta^{b^l}\right)\right| \geq \frac{1}{2}\eta - \sqrt{w} d\left(\beta^a, \beta^b\right) \geq 0.\end{aligned}$$

Putting the above observations together yields

$$\begin{aligned}&\mathbb{P}\left(\text{there exists } a \text{ st. } a \text{ is upper equilibrium and } \mathrm{Av}(a) \geq x_{\max} + \eta\right) \\ \leq &\sum_{l \leq n : \mathrm{Av}\left(b^l\right) \geq x_{\max} + \frac{1}{2}\eta} \mathbb{P}\left(\text{there exists } a \text{ st. } a \text{ is upper equilibrium and } d\left(\beta^a, \beta^{b^l}\right) \leq \delta\right) \\ \leq &2\exp\left(-\left(\delta - \frac{1}{\delta^2} c\left(w(g)\right)^2 d(g)\right) N\right)\end{aligned}$$

for some universal constant c. Because $N \geq \frac{1}{d(g)}$, if

$$d(g) \leq \min\left(\frac{1}{2}\delta^3 c^{-1}\left(w(g)\right)^{-2}, \frac{1}{2}\frac{1}{\log 2 - \log \eta}\delta \log \eta\right),$$

the above probability is smaller than η.

Appendix D. Proof of Theorem 3

D.1. **Proof description.** The proof is divided into five parts. Section D.2 is devoted to the existence of a contagion wave, i.e., the third step of the proof intuition from the main body of the paper.

Section D.3 introduces a two-dimensional lattice. In the limit, the neighborhoods converge to radius-1 balls in $\mathbb{R}^2$.

In Section D.4, we divide the lattice into areas, called *small cubes*, such that (a) there are many agents and the law of large numbers can be applied to describe the empirical distribution of payoff shocks inside each small cube, and (b) the cubes are sufficiently small so that agents from the same small cube have similar neighborhoods, which implies that their incentives are similar. The two properties imply that average behavior in a small cube is close to the behavior of a continuum of agents in the toy model.

Section D.5 studies the statistical distribution of bad small cubes, i.e. small cubes, where the empirical distribution of payoff shocks is not close to the distribution from which the shocks are drawn. We show that there are few of them and sufficiently sparse, so that the set of small cubes which are far away from the bad cubes contains a giant connected component.

The last section concludes the proof of the theorem.

D.2. **Contagion wave.** Consider a toy model, where agents are located on a line, each location has a continuum of agents, with a continuum best response function Q (not necessarily the same as P from the statement of the theorem), the connections depend only on the distance between agents, and and the cumulative weight of connections between agents x and agents in set $\{y' : y' \leq y\}$ is equal to $f(y-x)$, where $f : \mathbb{R} \to [0,1]$ is a function that is *balanced*: (a) $f(x)$ is strictly increasing for $x \in (-1,1)$, and (b) $f(-1) = 0$ and $f(x) + f(-x) = 1$ for each x. Given the interpretation of f stated above, condition $f(x) + f(-x) = 1$ is a consequence of the symmetry of the connection weights, and $f(-1) = 0$ means that agents separated by 1 or more are not connected. Notice that the weight of connections depends only on the distance between the agents.

Consider a strategy σ that is increasing in locations. For each location x, the average action of neighbors of agents in location x is equal to (assuming enough regularity, for intuition)

$$\int \sigma(y)\, df(y-x) = \lim_{a \to -\infty} \sigma(a) + \int (1 - f(y-x))\, d\sigma(y).$$

We say that σ is a contagion wave for Q if, at each location x, the best response of agents in such a location no higher than $\sigma(x)$ or, in other words, if the above average action is smaller than $Q^{-1}(\sigma(x))$.

This section contains two results: first, we show the existence of a contagion wave for a continuum best response function that can be represented by a step function, and next, we show the existence of a stronger version of a wave for the original best response function P.

We begin with a definition. An increasing function $q : \mathbb{R} \to [0,1]$ is a step function if the image $q(\mathbb{R})$ is finite. We refer to the elements of the image as steps. If q is a step function and $a \in q(\mathbb{R})$ is a step, then the most recent step before a is denoted as $a_- = \max\{b \in q(\mathbb{R}) : b < a\}$. For each $a \in [0,1]$, let $q^{-1}(a) = \min(v : q(v) \geq a)$ if the set is non-empty and $q^{-1}(a) = \infty$ if the set is empty. We have $q^{-1}(a_-) < q^{-1}(a)$ for each step a.

Lemma 11. *Let Q be a step function with steps $0 \leq a_0 < ... < a_{L+1} = 1$ and such that for each $a > a_0$, we have*

$$\int_{a_0}^{a} \left(Q^{-1}(x) - x\right) dx > 0. \tag{11}$$

Suppose that f is a continuous and balanced function. Then, there exist $0 = v_0 < v_1 ... < v_L \leq L$ such that, for each $l = 1, ..., L$,

$$Q^{-1}(a_{l+1}) \geq a_0 + \sum_{k \geq 0} (1 - f(v_k - v_l))(a_{k+1} - a_k).$$

We interpret each vector as a step strategy, where agents in locations $v_{l-1} < x \leq v_l$ play action a_l. Then, the right-hand side of the inequality is equal to the average action experiences in location v_l. The lemma says that, if Q is a step function, and it satisfies condition (11), then we can choose the step strategy such that the next step action a_{l+1} is a (Q-)best response for agents living on threshold v_l.

Proof. Let V be the set of all vectors $v = (v_0, ..., v_L)$ such that

$$0 = v_0 \leq ... \leq v_L \text{ and } v_{l+1} \leq v_l + 1 \text{ fov each } l = 0, ..., L-1.$$

(Abusing notation, we take $v_{-1} = -\infty$.) Define function $F : [-1, L+1] \times V \to \mathbb{R}$ so that

$$F(x|v) = a_0 + \sum_{k\geq 0} (1 - f(v_k - x))(a_{k+1} - a_k).$$

Then, for each strategy v, F is the (weighted) average action experienced by agents in location x.

Due to properties of function f, function F is continuous, strictly increasing in x for $x \in (v_0 - 1, v_{L-1} + 1)$ and decreasing in the lattice order on V^* (i.e., $F(x, v) \geq F(x, v')$ for any v, v' such that $\forall_k v_k \leq v'_k$.) For each $v \in V$ and $l = 1, ..., L$, define

$$b_l(v) = \inf\left\{x \geq 0 : F(x|v) \geq Q^{-1}(a_{l+1})\right\},$$

and we take $b_l(v) = \infty$ if the set is empty. $b_l(v)$ is the first location in which action a_{l+1} or higher is the best response given the strategy determined by v. The properties of F imply that b_l isweakly increasing in the lattice order on V, and, because $Q^{-1}(a_{l+1}) > Q^{-1}(a_l)$, we have $b_l(v) \leq b_{l+1}(v)$, with a strict inequality if either $b_l(v) \in (0, \infty)$ or $b_{l+1}(v) \in (0, \infty)$. It is also continuous for v such that $b_l(v) < \infty$. Let $b(v) = (b_l(v))_{l=1}^{L}$

Define function $b^* : V \to V$ so that

$$b^*_l(v) = \min(b_l(v), v_{l-1} + 1), \text{ for each } l = 1, ..., L-1.$$

Then, $b^*_l(v) \geq 0$, b^* is continuous and increasing in the lattice order. Moreover,

- if $b^*_l(v) = v_{l-1} + 1$, then $Q^{-1}(a_{l+1}) \geq F(b^*_l(v)|v)$,
- if $b^*_l(v) < v_{l-1} + 1$ and $b^*_l(v) > 0$, then $Q^{-1}(a_{l+1}) = F(b^*_l(v)|v)$, and
- if $b^*_l(v) = 0$ (which means that $b_l(v) = 0$), then $Q^{-1}(a_{l+1}) \leq F(0|v)$.

Consider a sequence $v^0 = (0, 0, ..., 0)$ and $v^n = b^*(v^{n-1})$ for $n > 0$. Because the sequence is bounded ($v^n \in V^*$ for each n) and b^* is continuous and increasing, it must converge to $v^* = b^*(v^*)$. The properties of b and b^* functions imply that if $v^*_l > 0$, then $b_l(v^*) \geq v^*_l$. (The reason is that if $n > 0$ is the first element of the sequence such that $v^n_l > 0$, then clearly $b_l(v^{n-1}) = v^n_l > 0 = v^{n-1}_l$, and by monotonicity, $b_l(v^{m-1}) \geq v^m_l$ for each m.)

Let $l_0 = \min(l = 1, ..., L \text{ st. } v_l = v_{l-1} + 1)$, where $l_0 = L + 1$ if the set is empty. We will show that $l_0 = 1$. On the contrary, suppose that $l_0 > 1$. Then, for each $l < l_0$,

$v_l^* = b_l^*(v) < v_{l-1}^* + 1$. The properties of b^* stated above imply that

$$\begin{aligned} Q^{-1}(a_{l+1}) \leq F(v_l^* | v^*) &= a_0 + \sum_{k \geq 0} (1 - f(v_k^* - v_l^*))(a_{k+1} - a_k) \\ &= a_0 + \sum_{k=0}^{l_0 - 1} (1 - f(v_k^* - v_l^*))(a_{k+1} - a_k), \end{aligned}$$

where the last equality follows from the fact that $v_k^* - v_l^* \geq 1$ for $l \leq l_0 - 1$ and $k \geq l_0$. Multiply both sides of the above inequality by $(a_{l+1} - a_l)$ and sum across all $l = 0, ..., l_0 - 1$ to obtain

$$\begin{aligned} &\sum_{l=0}^{l_0-1} Q^{-1}(a_{l+1})(a_{l+1} - a_l) \\ \geq & a_0 (a_{l_0} - a_0) + \sum_{k=0}^{l_0-1} \sum_{l=0}^{l_0-1} (1 - f(v_k^* - v_l^*))(a_{k+1} - a_k)(a_{l+1} - a_l) \\ = & a_0 (a_{l_0} - a_0) + \frac{1}{2} \sum_{k=0}^{l_0-1} \sum_{l=0}^{l_0-1} (1 - f(v_k^* - v_l^*) + 1 - f(v_l^* - v_k^*))(a_{k+1} - a_k)(a_{l+1} - a_l) \\ = & a_0 (a_{l_0} - a_0) + \frac{1}{2} \sum_{k=0}^{l_0-1} \sum_{l=0}^{l_0-1} (a_{k+1} - a_k)(a_{l+1} - a_l) \\ = & a_0 (a_{l_0} - a_0) + \frac{1}{2} (a_{l_0} - a_0)^2 = \frac{1}{2} \left(a_{l_0}^2 - a_0^2 \right) = \int_{a_0}^{a_{l_0}} x dx. \end{aligned}$$

(The first equality is obtained by exchanging indices k and l. The second one is due to f being balanced.) Because the LHS of the above inequality is equal to $\int_{a_0}^{a_{l_0}} Q^{-1}(x)\,dx$, we get a contradiction with (11). The contradiction shows that $l_0 = 1$.

Because $l_0 = 1$, $v_1^* = 1 > 0$, and we have $v_l^* \geq v_1^* > 0$ for each $l = 0, ..., L - 1$. The properties of the sequence v^n imply that $b_l(v^*) \geq v_l^* > 0$, which further implies that $Q^{-1}(a_{l+1}) \geq F(v_l^* | v^*)$, and, due to the definition of Q^{-1}, that $a_{l+1} \geq Q(F(v_l^* | v^*))$ for each l. Moreover, for each l, either $v_{l+1}^* = b_{l+1}(v^*) > b_l(v^*) \geq v_l$, or $v_{l+1}^* = v_l^* + 1$. In both cases, $v_{l+1}^* > v_l^*$. This establishes the existence of vector v with the required properties. $\square$

The next lemma strengthens the conclusion of Lemma 11.

Lemma 12. *Suppose that $P(1) < 1$ and x^* is strictly RU-dominant. For each $\eta > 0$, there exist $\delta > 0$, $a^* \leq x^* + \eta$, $L < \infty$, and a step function $\sigma : \mathbb{R} \to [0,1]$ such that $\sigma(0) = a^*$, $\sigma(L) = 1$, and, for each x,*

$$\sigma(x-\delta) \geq \delta + P\left(\delta + a^* + \sum_{a \in \sigma^{-1}(\mathbb{R})} \left(1 - f\left(\sigma^{-1}(a) - x\right)\right)(a - a_-)\right), \tag{12}$$

where the summation is over the consecutive steps of the step function σ.

We refer to σ as a δ-contagion wave for P.

Proof. Define $P^\delta(x) = P(x) + \delta$ for $\delta \in (0, 1 - P(1))$ and notice that for sufficiently small $\delta_1 > 0$, for each $\delta \leq \delta_1$, if a_0 is the highest maximizer

$$a_0 \in \sup \arg\max_a \int_0^a \left(\left(P^{\delta_1}\right)^{-1}(x) - x\right) dx,$$

then, $a_0 \leq x^* + \frac{1}{2}\eta$. Each P_{δ_1} can be approximated by a step function Q such that (a) $Q \geq P$ (hence $P^{-1} \geq Q^{-1}$), (b) each step is bounded by $a_l - a_{l-1} \leq \frac{1}{4}\delta_1$, for $l = 1, ..., L$, and (c) if a^* is the highest maximizer of

$$a^* \in \sup \arg\max_x \int_0^a \left(Q^{-1}(x) - x\right) dx,$$

then $a^* \leq x_0 + \frac{1}{2}\eta = x^* + \eta$. (We omit the details of finding such approximations.) Find $\delta_2 > 0$ s.t. $\delta_2 \leq \frac{1}{2}\delta_1$ and, for each $a > a^*$, we have

$$\int_{a^*}^a \left(Q^{-1}(x) - x - \delta_2\right) dx > 0.$$

Such δ_2 exists because Q is a step function and $\lim_{x \searrow a^*} Q^{-1}(x) > a^*$.

Let $Q_{\delta_2} = Q(x + \delta_2)$. Then, Q_{δ_2} is a step function that satisfies the hypothesis of Lemma 11. Let $0 = v_0 < v_1 ... < v_L \leq L$ be the thresholds from Lemma 11. Then, for each $l > 0$,

$$a_{l-1} \geq a_{l+1} - \frac{1}{2}\delta_1 \geq Q_{\delta_2}\left(a^* + \sum_{k\geq 0}\left(1 - f\left(v_k - v_l\right)\right)\left(a_{k+1} - a_k\right)\right) - \frac{1}{2}\delta_1$$
$$= Q\left(\delta_2 + a^* + \sum_{k\geq 0}\left(1 - f\left(v_k - v_l\right)\right)\left(a_{k+1} - a_k\right)\right) - \frac{1}{2}\delta_1$$
$$\geq P\left(\delta_2 + a^* + \sum_{k\geq 0}\left(1 - f\left(v_k - v_l\right)\right)\left(a_{k+1} - a_k\right)\right) + \frac{1}{2}\delta_1$$
$$\geq P\left(\delta_2 + a^* + \sum_{k\geq 0}\left(1 - f\left(v_k - v_l\right)\right)\left(a_{k+1} - a_k\right)\right) + \delta_2. \tag{13}$$

The first inequality follows from $a_{l+1} - a_{l-1} \leq a_l - a_{l-1} + a_{l+1} - a_l \leq \frac{1}{2}\delta_1$; the equality follows from $Q_{\delta_2}^{-1}(a) = Q^{-1}(a) - \delta_2$; the second inequality follows from $Q \geq P + \delta_1$; and the last inequality follows from $\delta_2 \leq \frac{1}{2}\delta_1$.

Define

$$\sigma(x) = \begin{cases} a^* & x < 0 \\ a_l & x \in [v_{l-1}, v_l) \text{ and } l = 1, ..., L \\ P(1) + \delta & x \geq v_L. \end{cases}$$

Find $\delta > 0$ such that $\delta \leq \delta_2$ and $\delta \leq v_{l+1} - v_l$ for each $l = 0, ..., L-1$. Because the right-hand side of inequality (12) is increasing in x, we have:

- If $x < \delta$, then $\sigma(x - \delta) = a^* = a_0$. Hence inequality (12) follows from inequality (13) for $l = 1$ and the fact that $x \leq 0 = v_0 < v_1$.
- If $v_{l-1} + \delta \leq x < v_l + \delta$ for $l = 1, ..., L$, then $\sigma(x - \delta) \geq a_{l-1}$. Hence inequality (12) follows from inequality (13) and $x \leq v_l$.
- If $x \geq v_L + \delta$, then $\sigma(x - \delta) \geq P^*(1) + \delta$. Hence inequality (12) is satisfied automatically.

□

D.3. **Lattice.** We start by describing the candidate network. For each $M \geq m$, the (M, m)-lattice is a network with

- $N = M^2$ nodes from the set $I_M = \{1, ..., M\}^2$. We define a distance on I_M by
$$d(i, j) = \frac{1}{m}\sqrt{\sum_l \left((i_l - j_l) \mod M\right)^2},$$
and a ball in this metric as $B(i, r) = \{y : d(x, y) \leq r\}$. The subtraction "mod$M$" turns the lattice into a subset of "Euclidean torus" $\left[0, \frac{M}{m}\right]^2$,
- connections $g_{i,j} = 1 \iff j \in B(i, 1)$.

In the course of the proof, we will assume that there exists values b and B such that $0 \ll b \ll m \ll B \ll M$ and such that B is divisible by b and M is divisible by B. This divisibility assumption simplifies the proof. The theorem remains valid without it, but the proof requires small modifications to take care of reminder items. We omit the details.

For each $i \in I_M$, and two sets $U, W \subseteq I_M$, let
$$d(i, W) = \min_{j \in W} d(i, j) \text{ and } d(U, W) = \min_{i \in U} \min_{j \in W} d(i, j). \tag{14}$$
For each set W, and each r, define the r-neighborhood of W:
$$B(W, r) = \{i : d(i, W) \leq r\} = \bigcup_{i \in W} B(i, r).$$

For large m, the neighborhoods of each agent behave in a similar way to open balls on a Euclidean plane. This is formalized as follows. Let $B_{\mathbb{R}^2}(x, r)$ be the ball on the plane with center $x \in \mathbb{R}^2$ and radius r. Let $|A|$ be a Lebesgue measure of a measurable set $A \subseteq \mathbb{R}^2$. Let
$$f_0(d, r_1, r_2) = \frac{1}{\pi}\left|B_{\mathbb{R}^2}((0, 0), r_1) \cap B_{\mathbb{R}^2}((d, 0), r_2)\right|$$
be the measure of the intersection of two balls, with radii r_1 and r_2 respectively, separated by distance d, and divided by the measure of the unit ball $B((0, 0), 1)$.

Lemma 13. *(1) For each $\rho > 0$, there exists $C_\rho < \infty$ such that if $m \geq C_\rho$, then for any two agents i, j, for any $r_1 \leq 1 \leq r_2$, we have*
$$\left|\frac{|B(i, r_1) \cap B(j, r_2)|}{|B(i, 1)|} - f_0(d(i, j), r_1, r_2)\right| \leq \rho.$$

(2) *Function f_0 has the following properties:*
  - *f_0 is Lipschitz over d and $r_1 \leq 1 \leq r_2$,*
  - *f_0 is decreasing in d, and*
  - *$f_0(d, r_1, r_2) = 0$ if $r_1 + r_2 \leq d$, and $f_0(d, r_1, r_2) = 1$ if $r_1 = 1$ and $d \leq r_2 - r_1$.*

(3) *Functions $f_1(x, r_1; r_2) = f_0(r_2 - x, r_1, r_2)$ for $r_1 \leq 1$ and $x \in \mathbb{R}$ converge uniformly to function $\lim_{r_2 \to \infty} f_1(x, r_1; r_2) = f_2(x, r_1)$. In particular, for each $\rho > 0$, there exists R_ρ such that, if $r_1 \leq 1$ and $r_2 \geq R_\rho$, then,*

$$\sup_{r_1 \leq 1, x} |f_2(x, r_1) - f_1(x, r_1; r_2)| \leq \rho.$$

Functions f_1 and f_2 are Lipschitz over d and $r_1 \leq 1$ and increasing in x.

(4) *Let $f(x) = f_2(x, 1)$. Function f is balanced (in the sense of the definition from Section D.2).*

Proof. The properties of f_0, f_1, f_2, and f follow from their geometric interpretations and the fact that the counting measure on I_M converges weakly to the Lebesgue measure on the torus. For example, $f_2(x, r_1)$ is a circle segment of a radius r_1 circle with height equal to $r_1 + x$ for $x \in (-r_1, r_1)$. $\square$

D.4. **Small cubes.** We divide the lattice into disjoint areas that we refer to as *small cubes*. Each cube is much smaller than the diameter of the neighborhood of each node so that the neighborhoods of nodes in the same cube are largely overlapping. At the same time, each small cube contains a sufficiently large number of nodes so that the distribution of payoff shocks within the cube can be probabilistically approximated by its expected distribution.

Let G be a (M, m)-lattice. Take any $b > 0$, where we intend $b \ll m$. For each real number x, let $\lfloor x \rfloor$ be the largest integer no larger than x. For each node i, the set of nodes

$$c^b(i) = \left\{ j \in \{1, ..., M\}^2 : \forall_l \lfloor i_l / b \rfloor = \lfloor j_l / b \rfloor \right\}$$

is referred to as a cube that contains i. Any two cubes are either disjoint or identical. Each cube c is uniquely identified by a pair of numbers $c_l = \lfloor i_l / b \rfloor$ for each $l = 1, 2$ and any $i \in c$. Due to the divisibility assumption, each cube contains exactly b^2 elements, and there are $\left(\frac{M}{b}\right)^2$ small cubes on the (M,m)-lattice.

Let $\mathcal{G}^b = \left\{c^b(i) : i \in G\right\}$ be the set of all cubes. We refer to the elements of $\mathcal{G}^b$ as *small cubes*, to distinguish them from the large cubes introduced in Section D.5. Sometimes, we treat $\mathcal{G}^b$ as a network with edges

$$g^b_{c,c'} = 1 \text{ iff } \sum_l \left|(c_l - c'_l) \operatorname{mod} \frac{M}{b}\right| = 1. \tag{15}$$

This way, each cube has four neighbors. We refer to $\left(\mathcal{G}^b, g^b\right)$ as a network of cubes.

For any $c, c' \in \mathcal{G}^b$, let $d^b(c, c')$ denote the length of the shortest path between c and c' in the network $\left(\mathcal{G}^b, g^b\right)$. For any $S \subseteq \mathcal{G}^b$, let $d^b(c, S) = \min_{c' \in S} d^b(c, c')$.

For each strategy profile $a = (a_i)_i$ and each small cube $c \in \mathcal{G}^b$, define

$$\begin{aligned}
a(c) &= \frac{1}{|c|} \sum_{i \in c} a_i, \\
\beta^a(i) &= \frac{1}{|B(i,1)|} \sum_{j \in B(i,1)} a_j = \frac{1}{|B(i,1)|} \sum_{j : d(i,j) \le 1} a_j, \text{ and} \\
\beta^a(c) &= \frac{1}{|c|} \sum_{j \in c} \beta^a(j) = \frac{1}{|c|} \sum_{i \in c} \frac{1}{|B(i,1)|} \sum_{j : d(i,j) \le 1} a_j,
\end{aligned}$$

where $a(c)$ is the average action within the cube, $\beta^a(i)$ is the fraction of neighbors of i who choose action 1, and $\beta(c)$ is the average fraction in cube c.

D.4.1. *Average fractions.* The next result shows that if the cube is sufficiently small, individual and average fractions are similar.

Lemma 14. *There exists an universal constant $D < \infty$ such that, if $\frac{b}{m} \le \rho$ and $m > C_\rho$, where C_ρ and is a constant from Lemma 13, then, for each profile a, each small cube, and each $i, j \in c$,*

$$|\beta^a(i) - \beta^a(c)| \le D\rho.$$

Proof. It is enough to show there exists $D < \infty$ such that $|\beta^a(i) - \beta^a(j)| \leq D\rho$ for each $i, j \in c$. Notice that

$$|\beta^a(i) - \beta^a(j)| \leq \frac{|B(i,1) \setminus B(j,1)|}{|B(i,1)|} + \frac{|B(j,1) \setminus B(i,1)|}{|B(j,1)|}.$$

By Lemma 13 and the fact that $d(i,j) \leq \sqrt{2}\rho$, the above is no larger than

$$\leq 2\rho + 2\left(1 - f_0\left(\sqrt{2}\rho, 1, 1\right)\right).$$

The claim follows from the Lipschitzness of function f_0 and the fact that $f_0(0,1,1) = 1$. ☐

D.4.2. *Average best response.* For each small cube $c \in \mathcal{G}^b$ and realization of payoff shocks, define the empirical cdf of best response thresholds:

$$P_c(x|\varepsilon) = \frac{1}{|c|} \sum_{i \in c} \mathbf{1}\{\beta(\varepsilon_i) < x\}.$$

(Recall that $\beta(\varepsilon_i)$ is the fraction of neighbors of individual i with payoff shock ε_i that would make her indifferent between the two actions.) For $\gamma > 0$, say that a small cube c is γ-bad, if there exists x such that $P_c(x|\varepsilon) > P(x) + \gamma$; otherwise, the cube is γ-good.

Next, we show that if a cube is good, then the average action can be approximated by a best response to average beliefs.

Lemma 15. *There exists a constant $D < \infty$ such that if $\frac{b}{m} \leq \rho$ and $m > C_\rho$, where C_ρ is a constant from Lemma 13, then, for each equilibrium profile a, if small cube c is γ-good, then*

$$a(c) \leq \gamma + P(\beta^a(c) + D\rho).$$

Proof. Notice that

$$\begin{aligned} a(c) &= \frac{1}{|c|} \sum_{i \in c} a_i \leq \frac{1}{|c|} \sum_{i \in c} \mathbf{1}(\beta(\varepsilon_i) \leq \beta_i^a) \leq \frac{1}{|c|} \sum_{i \in c} \mathbf{1}(\beta(\varepsilon_i) \leq \beta^a(c) + D\rho) \\ &= P_c(\beta^a(c) + D\rho|\varepsilon) \leq \gamma + P(\beta^a(c) + D\rho). \end{aligned}$$

The first inequality comes from the fact that if $a_i = 1$ is a best response, then $\beta(\varepsilon_i) \leq \beta_i^a$, and the second inequality is a consequence of Lemma 14. ☐

D.4.3. *Behavior dominance.* The next definition and result plays an important role in extending the contagion wave mechanics from a one-dimensional line to a two-dimensional lattice.

Let σ be an increasing step function (see Section D.2) for the definition. Let $a = (a_i)$ be a strategy profile. We say that profile a is (W, R, ρ)-*dominated* by σ given a set $W \subseteq \mathcal{G}^b$ of small cubes and $R > 0$ if for each small cube $c \in \mathcal{G}^b$, we have

$$a(c) \leq \sigma(d(c, W) - R) + \rho,$$

where distance between sets is defined in (14).

Lemma 16. *There is a constant $D < \infty$ with the following property: Fix $\rho > 0$. Suppose that $\frac{b}{m} < \rho$, $R > R_\rho$, and $m > C_\rho$, where C_ρ and R_ρ are the constants from Lemma 13. For each increasing step function $\sigma : \mathbb{R} \to [0, 1]$, and for each set of small cubes W, if strategy profile a is (W, R, ρ)-dominated by σ, then for each cube c,*

$$\beta^a(c) \leq a^* + \sum_{a \in \sigma^{-1}(\mathbb{R})} \left(1 - f\left(\sigma^{-1}(a) + R - d(c, W)\right)\right)(a - a_-) + D\rho.$$

Proof. By Lemma 14, there is a constant D_0 such that for any $i \in c$,

$$\beta^a(c) \leq \beta^a(i) + D_0\rho = a^* + \frac{1}{|B(i,1)|} \sum_{j \in B(i,1)} (a(j) - a^*) + D_0\rho$$

$$\leq a^* + \frac{1}{|B(i,1)|} \sum_{c' : d(i,c') \leq 1 - \sqrt{2}\rho} |c'| (a(c') - a^*) + \frac{\left|\left\{j : 1 - \sqrt{2}\rho < d(i,j) < 1\right\}\right|}{|B(i,1)|} + D_0\rho.$$

Lemma 13 implies that the third term is bounded by

$$\leq 1 - f_0\left(0, 1 - \sqrt{2}\rho, r_2\right) \leq D_1\rho$$

for some constant D_1 to the Lipschitzness of function f_0 and $f_0(0,1,1) = 1$. For the second term, we have

$$
\begin{aligned}
&\frac{1}{|B(i,1)|} \sum_{c': d(i,c') \leq 1-\sqrt{2}\rho} |c'| \left(a(c') - a^*\right) \\
\leq &\rho + \frac{1}{|B(i,1)|} \sum_{c': d(i,c') \leq 1-\sqrt{2}\rho} |c'| \left(\sigma\left(d(c,W) - R\right) - a^*\right) \\
\leq &\rho + \sum_{a\in\sigma(\mathbb{R})} (a - a_-) \frac{1}{|B(i,1)|} \sum_{c':\ d(i,c') \leq 1-\sqrt{2}\rho \text{ and } d(c', \bigcup W) \geq R + \sigma^{-1}(a)} |c'| \\
\leq &\rho + \sum_{a\in\sigma(\mathbb{R})} (a - a_-) \frac{1}{|B(i,1)|} \left|\left\{j : d(i,j) \leq 1, d\left(j, \bigcup W\right) \geq R + \sigma^{-1}(a)\right\}\right|. \qquad (16)
\end{aligned}
$$

(Recall that $\sigma(\mathbb{R})$ is the set of steps of the step function σ.) Let $i^* \in \arg\min_{j\in\bigcup W} d(i,j)$. Then, $d(i,i^*) = d(i, \bigcup W)$. Applied again, Lemma 13 implies that

$$
\begin{aligned}
&\frac{1}{|B(i,1)|} \left|\left\{j : d(i,j) \leq 1, d\left(j, \bigcup W\right) \geq R + \sigma^{-1}(a)\right\}\right| \\
\leq &\frac{1}{|B(i,1)|} \left|\left\{j : d(i,j) \leq 1, d(j, i^*) \geq R + \sigma^{-1}(a)\right\}\right| \\
= &1 - \frac{\left|B(i,1) \cap B\left(i^*, R + \sigma^{-1}(a) - \rho\right)\right|}{|B(i,1)|} \\
\leq &1 - f_0\left(d\left(i, \bigcup W\right), 1, R + \sigma^{-1}(a)\right) + \rho \\
= &1 - f_1\left(R + \sigma^{-1}(a) - \rho - d\left(i, \bigcup W\right), 1; R + \sigma^{-1}(a)\right) + \rho \\
\leq &1 - f_2\left(R + \sigma^{-1}(a) - \rho - d\left(i, \bigcup W\right), 1\right) + \rho \\
\leq &1 - f\left(R + \sigma^{-1}(a) - d\left(i, \bigcup W\right)\right) + (K+1)\rho,
\end{aligned}
$$

where K is a Lipschitz constant for f. Hence (16) is not larger than

$$
\leq \sum_{a\in\sigma(\mathbb{R})} (a - a_-) \left(1 - f\left(R + \sigma^{-1}(a) - d\left(i, \bigcup W\right)\right)\right) + D_2\rho
$$

for some constant $D_2 < \infty$ that may depend on the number of steps in the step function σ. The result follows from putting the estimates together. $\square$

D.5. **Good giant component of cubes.** We will show that if the lattice is sufficiently large then, with an arbitrarily large probability, we can find a set of small cubes that (a) contains almost all small cubes (we say that it is *giant)* (b) it is connected in the small cube network, (c) each cube in the set is far away from bad cubes, and (d) it contains a large set of agents for whom action 0 is dominant. Properties (b)-(c) will allow the contagion wave to spread across the entire set W, property (a) will ensure that spreading to set W means spreading almost everywhere, and property (d) will ensure that the set contains sufficiently many "initial infectors".

Formally, say that agent x is *extraordinary* if action 0 is strictly dominant for such an agent. A small cube $c \in \mathcal{G}^b$ is *extraordinary* if it only consists of extraordinary agents. In any equilibrium, $a(c) = 0$ for extraordinary cube c.

Say that set $W \subseteq \mathcal{G}^b$ of small cubes is (γ, R)-*good* if

(a) the union of all small cubes in W contains at least a fraction of $(1-\gamma)$ elements of the lattice, $|\bigcup W| \geq (1-\gamma) M^2$,

(b) it is connected as a subset of nodes on graph $\left(\mathcal{G}^b, g^b\right)$ (see the definition of a small cube network in (15)),

(c) if $c \in \mathcal{G}^b$ is γ-bad, then $d(c, c') \geq R$ for each $c' \in W$ (in particular, each cube in W is γ-good),

(d) it contains a cube c_0 such that each cube c s.t. $d(c, c_0) \leq R$ is extraordinary.

The goal of this subsection is to prove that large good sets of small cubes exists with a large probability:

Lemma 17. *For each $\gamma, \rho > 0$, and $R < \infty$, there exists constants $m_{\gamma,\rho,R}, A_{\gamma,\rho,R} > 0$ such that, if $m \geq m_{\gamma,\rho,R}$ and $M \geq (A_{\gamma,\rho,R})^{m^6}$, then there exists b so that $\frac{b}{m} \leq \rho$ and, if G is (M, m)-lattice with the associated small cube network $\mathcal{G}^b$, then*

$$\mathbb{P}\left(\textit{there exists } (\gamma, R)\textit{-good set } W \subseteq \mathcal{G}^b\right) \geq 1-\gamma.$$

D.5.1. *Large cubes.* In order to find a set W that is sufficiently far from bad small cubes, we are going to contain and separate bad small cubes in sufficiently large sets. Let B be a number that is divisible by b, $B = kb$, and such that M is divisible by B. Consider a network of cubes $\left(\mathcal{G}^B, g^B\right)$ defined in the same way as described in Section

D.4. We refer to elements of $\mathcal{G}^B$ as *large cubes* to distinguish from the elements of $\mathcal{G}^b$. Let $K = \frac{M}{B}$; then the number of large cubes is K^2.

For each set of large cubes $U \subseteq \mathcal{G}^B$, and for each R, define the small cube R-interior of U as the set of small cubes that are R-away from nodes that do not belong to U

$$W(U,R) = \left\{c \in \mathcal{G}^b : d\left(c, I_M \backslash \left(\bigcup U\right)\right) > R\right\}.$$

Here, $\bigcup U$ is the union of all large cubes in set U, and $I_M \backslash (\bigcup U)$ is the set of all nodes on (M,m)-lattice that do not belong to one of the large cubes in U. We have the following bound on the size of set $W(U,R)$.

Lemma 18. *Suppose that U is a subset of large cubes, $U \subseteq \mathcal{G}^B$. Then,*

$$\frac{1}{|\mathcal{G}|}\left|\bigcup W(U,R)\right| \geq \frac{|U|}{|\mathcal{G}^B|}\left(1 - 4\frac{1}{k}\left(\frac{Rm}{b}+1\right)\right).$$

.

Proof. Observe that

$$\frac{\left|\bigcup W(U,R)\right|}{|\mathcal{G}|} = \frac{\left|\bigcup \mathcal{G}^b\right|}{|\mathcal{G}|}\frac{\frac{\left|\bigcup W(U,R)\right|}{\left|\bigcup \mathcal{G}^b\right|}}{\frac{|W(U,R)|}{\left|\mathcal{G}^b\right|}}\frac{|W(U,R)|}{|W(U,0)|}\frac{|W(U,0)|}{|\mathcal{G}^b|}.$$

The bound is a consequence of the following observations:

- Because all small cubes have the same cardinality, we have $\left|\bigcup \mathcal{G}^b\right| = |\mathcal{G}|$ and $\frac{\left|\bigcup W(U,R)\right|}{\left|\bigcup \mathcal{G}^b\right|} = \frac{|W(U,R)|}{\left|\mathcal{G}^b\right|}$.
- For each regular large cube $C \in U$, $W(C,0)$ consists of k^2 small cubes, and $W(C,R)$ consists of at least $\left(k - 2\left(\frac{Rm}{b}+1\right)\right)^2$ small cubes. Hence $\frac{|W(U,R)|}{|W(U,0)|} \geq 1 - 4\frac{1}{k}\left(\frac{Rm}{b}+1\right)$.
- Finally, notice that $|W(U,0)| = k^2|U|$ and $\left|\mathcal{G}^b\right| = k^2\left|\mathcal{G}^B\right|$.

□

The next result shows that if U is a connected component of large cubes, then $W(U,R)$ is a connected component of small cubes.

Lemma 19. *Suppose that $R < \frac{b}{m}\left(\frac{1}{2}k - 1\right)$. If a set of large cubes $U \subseteq \mathcal{G}^B$ is a connected component in the network of large cubes, then the R-interior set of small cubes $W(U, R)$ is a connected component in the network of small cubes.*

Proof. For each large cube C, let $W(U,R) \cap \left\{c \in \mathcal{G}^b : c \subseteq C\right\}$ be a part of the R-interior that consists of small cubes which are contained in C. It is clear that $W(U,R) \cap \left\{c \in \mathcal{G}^b : c \subseteq C\right\}$ is connected in the network of small cubes. If C and C' are two neighboring large cubes, say $C_1 = C_1'$ and $C_2' = C_2' + 1$, then small cubes c and c' such that $c_1 = c_1' = B(C_1 - 1) + \left(\left\lceil \frac{Rm}{b} \right\rceil + 1\right)$ and $c_2' = c_2 + 1 = BC_2 + 1$ are neighbors and they both belong to $W(U,R)$. Hence, set $W(U,R) = \bigcup_{C \in U} W(U,R) \cap \left\{c \in \mathcal{G}^b : c \subseteq C\right\}$ is connected. □

D.5.2. *Percolation theory - deterministic bounds.* In order to establish the existence of a giant connected component of small cubes that are not too close to bad small cubes, we turn to the percolation theory. The percolation theory studies properties of graphs obtained by removal of some nodes. In this paper, we are especially interested in the size of the largest connected component of a so-obtained graph.

We divide the percolation theoretic arguments into two parts: deterministic and probabilistic.

Lemma 20. *For each connected $S \subseteq \mathcal{G}^B$ st. $|S| < K$, there are connected sets $\partial S \subseteq CS \subseteq \mathcal{G}^B \backslash S$ such that $\left|\mathcal{G}^B \backslash CS\right| \leq |S|^2$,*

$$\left\{c \in CS : d^B(c, S) \leq 1\right\} \subseteq \partial S \subseteq \left\{c \in CS : d^B(c, S) \leq 2\right\}.$$

Proof. Because $|S| < K$ is smaller than the length and width of the network of large cubes, set S can be contained in a cube of size $|S|^2$ in a way that the complement of the cube is connected and it contains at least $\left|\mathcal{G}^B\right| \backslash |S|^2$ elements. Let CS be the connected component of $\mathcal{G}^B \backslash S$ that contains the complement of the cube. Using Lemma 1 from [Bollobás *et al.*(2006)Bollobás, Riordan and Riordan], we can construct a finite path $c_0, ..., c_k$ of neighboring cubes in CS surrounding S in an intuitive way such that, if $\partial S = \{c_0, ..., c_k\}$, then ∂S satisfies the required inclusions. The construction implies that the path $c_0, ..., c_k$ may contain cubes that lie in distance 1 from set S (i.e., with a shared wall) and a distance of 2 from set S (that, intuitively, share a corner with a set S). $\square$

For each $S \subseteq \mathcal{G}^B$, say that a set $S \subseteq \mathcal{G}^B$ is k-connected if, for any subset $T \subseteq S$, $\min_{c\in T, c'\in S\setminus T} d^B(c, c') \leq 2$. In other words, a 2-connected set cannot be split into two parts that are more than 2 away from each other.

Lemma 21. *For each k-connected $S \subseteq \mathcal{G}^B$ st. $|S| < K$, there are connected sets $\partial S \subseteq CS \subseteq \mathcal{G}^B \setminus S$ such that $\left|\mathcal{G}^B \setminus CS\right| \leq 2^{4k} |S|^2$,*

$$\left\{c \in CS : d^B(c, S) \leq 1\right\} \subseteq \partial S \subseteq \left\{c \in CS : d^B(c, S) \leq 2\right\}.$$

Proof. If S is k-connected, then its k-neighborhood S^k, i.e., the set of large cubes that are within k-distance or less from S is connected. The cardinality of S^k is at most $4^k |S|$ as each cube has at most 4 neighbors. The claim follows from Lemma 20. $\square$

Lemma 22. *Suppose that $S_1, ..., S_J$ is a collection of 2-connected subsets of lattice $\mathcal{G}^B$ such that each $|S_j| < K$ and such that for any $i \neq j$, $\min_{c\in S_i, c'\in S_j} d^B(c, c') > 2$. Then, graph $\mathcal{G}^B \setminus \bigcup S_j$ contains a connected component of size not smaller than $\left|\mathcal{G}^B\right| \setminus \sum_j 2^8 |S_j|^2$.*

The above result says that that, for each collection of 2-connected sets $S_1, ..., S_J$ that are at least 3-distant from each other, the network of large cubes without those sets contains a connected component of the size that is not smaller than $\left|\mathcal{G}^B\right| \setminus \sum_j |S_j|^2$. It is worthwhile to make two observations:

- The reason why we require sets S_j to be at least 3-distant is that sets 2-distant can form an obstruction to connected paths: our definition look at 2-connected sets instead of 1-connected is because

Proof. Suppose that $S_1, ..., S_J$ is a collection of connected subsets as in the statement of the lemma. For each $j \leq J$, let $\partial S_j \subseteq CS_j$ be as in Lemma 20. Let $C = \bigcap_i CS_i$. Then, $|C| = |\bigcap_i CS_i| \geq \left|\mathcal{G}^B\right| \setminus \sum_j 2^8 |S_j|^2$.

For each $i \neq j$, suppose that $\partial S_i \cap CS_j \neq \emptyset$. Then, $\partial S_i \cap CS_j$ is connected. Because the distance between S_i and S_j is strictly greater than 2, $\partial S_i \cap S_j = \emptyset$. Hence, $\partial S_i \subseteq CS_j$. It follows that, if $\partial S_i \cap C \neq \emptyset$, then $\partial S_i \subseteq C$.

It is enough to show that C is connected. Take $a, b \in C$ and construct an arbitrary path from $a = a_0, ..., a_n = b$ of neighboring cubes in network $\mathcal{G}^B$. Such a path may go outside set C and, if so, let $l = \min\{m : a_m \notin C\}$. Suppose that $a_l \notin CS_i$ for some i. Then, $a_{l-1} \in \partial S_i \cap C$, and, by the above argument, $\partial S_i \subseteq C$. Let $k = \max\{m : a_m \notin CS_i\}$. Such k is well-defined and $k < n$ because $a_n = b \in C$. Hence $a_{k+1} \in \partial S_i \subseteq C$.

Because ∂S_i is connected, the segment of the path between a_{l-1} and a_{k+1} can be replaced by a path that lies completely within $\partial S_j \subseteq C$. We can repeat such a modification for any other segment of the path that lies outside of set C. After finitely many modifications, we obtain a path from a to b that is entirely within C. It follows that C is connected. □

The last result in this part provides an upper bound on the number of distinct 2-connected sets.

Lemma 23. *The number of distinct* 2*-connected subsets of* $\mathcal{G}^B$ *of cardinality* r *is no larger than* $K^2(12)^r$.

Proof. Each r-element 2-connected set S can be (not necessarily uniquely) encoded as a pair of a signature $(t_0, ..., t_{r-1})$ such that $\sum t_i = r - 1$ and a tuple

$$\left(c_0, c_1, ..., c_{t_0}, c_{t_0+1}, ..., c_{t_0+t_1+1}, ..., c_{t_0+...+t_{l-1}+1}, ..., c_{t_0+...+t_l}, ..., c_r\right),$$

where

- $c_1, ..., c_{t_0}$ is the list of all 2-neighbors (i..e, cubes that have d^B distance no larger than 2) of c_0,
- more generally, for each l, $c_{t_0+...+t_{l-1}+1}, ..., c_{t_0+...+t_l}$ is a list of all 2-neighbors of c_l that have not yet been listed.

The number of different signatures is no larger than 2^r. Given signature $(t_0, ..., t_r)$, notice that there are at most K^2 choices of c_0; given c_0, there are at most $(12)^{T_0}$ choices of $c_1, ..., c_{t_0}$ (there are 12 cubes that are at most 2-distant from a given cube), etc.. Thus, the number of encodings is no larger than

$$K^2 \cdot (12)^{t_0} \cdot ... \cdot (12)^{t_{r-1}} = K^2 (12)^{r-1}.$$

The result follows. □

D.5.3. *Percolation theory - probabilistic arguments.* Next, we consider a standard model of percolation theory, where nodes are removed i.i.d. with probability $p \in (0,1)$. Let $\mathcal{G}^B_{(p)}$ denote a random graph obtained from the lattice of large cubes $\mathcal{G}^B$ by removing i.i.d. nodes. The following two results provide the bounds on the probability of the existence of a giant component of $\mathcal{G}^B_{(p)}$.

Lemma 24. *There exists a universal constant $\xi < \infty$ such that, for each $\gamma \in (0,1)$, K, and p, if $p \leq \xi\gamma^2$ and $K^2 (12p)^K \leq \frac{1}{2}\gamma$, then*

$$\mathbb{P}\left(\mathcal{G}^B_{(p)} \text{ has a connected component of size not smaller than } (1-\gamma) K^2\right) \geq 1-\gamma.$$

Proof. Let $E \subseteq I_K$ be the (random) set of nodes removed to obtain graph $\mathcal{G}^B_{(p)}$. For each removed node $a \in E$, let $S(a) \subseteq E$ be the maximally 2-connected component of removed nodes that contains a. In other words, $S(a)$ is 2-connected, and if $c \in E$ is such that $d^B(c, S(a)) \leq 2$, then $c \in S(a)$. Let $\mathcal{S} = \{S(a) : a \in E\}$ be a collection of such components. The construction ensures that, for each $S, T \in \mathcal{S}$, if $S \neq T$, then $\min_{c \in S, c' \in T} d^B(c, c') > 2$.

Let $r_{\max} = \max_{S \in \mathcal{S}} |S|$. Let

$$X_r = |\{S \in \mathcal{S} : |S| \geq r\}| \text{ for each } r \geq 1,$$

$$X = 2^8 \sum_{S \in \mathcal{S}} |S|^2 = 2^8 \sum_r r^2 (X_r - X_{r+1}) = 2^8 \sum_r \left(r^2 - (r-1)^2\right) X_r = 2^8 \sum_r (2r-1) X_r.$$

We compute the expected value of X. By Lemma 23, the number of r-element 2-connected sets is bounded by $K^2 (12)^r$. The probability that all elements of a particular r-element tuple are removed is equal to p^r. The linearity of the expectation implies that $\mathbb{E}\, X_r \leq K^2 (12)^r p^r$ and

$$\mathbb{E}\, X = 2^8 \sum_r (2r-1) \mathbb{E}\, X_r \leq 2^8 K^2 \sum_r 2^r (12p)^r \leq 2^8 K^2 \frac{24p}{1-24p}$$

The probability that there exists a 2-connected component not smaller than K is not larger

$$\mathbb{P}(r_{\max} \geq K) \leq \mathbb{E}\, X_K \leq K^2 (12p)^K.$$

By Lemma 22, the probability that $\mathcal{G}^B_{(p)}$ does not have a connected component not smaller than $(1-\gamma)\left|\mathcal{G}^B\right|$ is not larger than

$$\begin{aligned}
&\mathbb{P}\left(\mathcal{G}^B_{(p)} \text{ has no connected component of size not smaller than } (1-\gamma)\,K^2\right)\\
\leq{}& \mathbb{P}\left(X \geq \gamma K^2\right) + \mathbb{P}\left(r_{\max} \geq K\right) \leq \frac{\mathbb{E}\,X}{\gamma K^2} + \mathbb{P}\left(r_{\max} \geq K\right) \leq \frac{1}{\gamma}2^8\frac{24p}{1-24p} + K^2\left(12p\right)^K.
\end{aligned}$$

(The second inequality is due to the Markov inequality.) Assuming that $\gamma < 1$, if $p \leq \frac{1}{2^{12}\cdot 24}\gamma^2$ and $K^2 2^{-K} \leq \frac{1}{2}$, then $1-24p \geq \frac{1}{2}$, and

$$\begin{aligned}
\frac{1}{\gamma}2^8\frac{24p}{1-24p} + K^2\left(12p\right)^K &\leq \frac{1}{\gamma}\cdot 2\cdot 2^8\cdot\frac{1}{2^{12}}\gamma^2 + K^2\left(12\sqrt{p}\right)^K p^{K/2}\\
&\leq \frac{1}{2}\gamma + K^2\left(12\frac{1}{2^6}\right)^K\frac{1}{2}\gamma \leq \gamma.
\end{aligned}$$

□

Next, we find a probability bound on the existence of a giant component of large cubes that do not have any bad small cubes. A large cube $C \in \mathcal{G}^B$ is γ-clean if it does not contain any γ-bad small cube. Let $\mathcal{G}^B_\gamma$ be the random subgraph of the network of large cubes that consists only of γ-clean cubes.

Lemma 25. *There exists a universal constant $\xi < \infty$ such that, if $b \geq \frac{1}{2\gamma}\left(\log\frac{\xi k^2}{\gamma^2}\right)^{1/2}$ and $K^2 2^{-K} \leq \frac{1}{2}\gamma$, then*

$$\mathbb{P}\left(\mathcal{G}^B_\gamma \textit{ has a connected component of } \gamma\textit{-clean large cubes and size at least} (1-\gamma)\left|\mathcal{G}^B\right|\right) \geq 1-\gamma.$$

The giant component from the lemma is obviously uniquely defined. We refer to it as U_γ.

Proof. Due to the Dvoretzky–Kiefer–Wolfowitz–Massart inequality, the probability that a small cube c is γ-bad is bounded by

$$\mathbb{P}\left(c \text{ is } \gamma\text{-bad}\right) \leq \mathrm{e}^{-2b^2\gamma^2}.$$

The probability that a large cube C is not γ-clean is bounded by

$$\mathbb{P}\left(C \text{ is not } \gamma\text{-clean}\right) \leq k^2\mathrm{e}^{-2b^2\gamma^2}.$$

By Lemma 24 and some algebra, the claim holds if $K^2 2^{-K} \leq \frac{1}{2}\gamma$ and $k^2 \mathrm{e}^{-2b^2\gamma^2} \leq \frac{1}{\xi}\gamma^2$ for some universal constant $\xi < \infty$. □

D.5.4. *Extraordinary set.* A large cube $C \in \mathcal{G}^B$ is extraordinary if it only consists of extraordinary agents. The next result bounds the probability that the large component identified in the previous section contains an extraordinary large cube.

Lemma 26. *There exists a universal constant $\xi < \infty$ such that, if $\mathrm{e}^{-(1-\gamma)K^2 P(0)^{k^2 b^2}} \leq \frac{1}{2}\gamma$, $b \geq \frac{2}{\gamma}\left(\log \frac{\xi k^2}{\gamma^2}\right)^{1/2}$, and $K^2 2^{-K} \leq \frac{1}{4}\gamma$, then*

$$\mathbb{P}\left(|U_\gamma| \geq (1-\gamma) K^2 \text{ and } U_\gamma \text{ contains an extraordinary large cube}\right) \geq 1-\gamma.$$

Proof. The probability that a single agent is extraordinary is $P(0) = \mathbb{P}(\beta(\varepsilon_i) \leq 0)$. The probability that a cube $C \in \mathcal{G}^B$ is extraordinary is $P(0)^{(kb)^2}$. Because each extraordinary cube is also γ-clean, the probability that C is extraordinary conditionally on C being part of the giant component U_γ and on an arbitrary realization of payoff shocks outside of C is no smaller than $P(0)^{(kb)^2}$. Conditionally on $|U_\gamma| \geq (1-\gamma) K^2$, the probability that the giant component has no extraordinary cube is bounded by

$$\begin{aligned}
&\mathbb{P}\left(U_\gamma \text{ has no extraordinary cube} \middle| |U_\gamma| \geq (1-\gamma) K^2\right) \\
&\leq \left(1 - P(0)^{(kb)^2}\right)^{(1-\gamma)K^2} \leq \mathrm{e}^{-(1-\gamma)K^2 P(0)^{k^2 b^2}}.
\end{aligned}$$

The claim follows from the above bound and Lemma 25. □

D.5.5. *Proof of Lemma 17.* Assume w.l.o.g. that $R \geq 1$ and $\gamma, \rho < 1$. Let $k_m = \left\lceil \frac{100}{\gamma} Rm \right\rceil$ and $b_m = \left\lceil \frac{20}{\gamma}\left(\log \frac{100\xi k_m^2}{\gamma^2}\right)^{1/2} \right\rceil$, where ξ is the constant from Lemma 26. Then, $k_m, b_m \geq 1$ and there is a constant $m_{\gamma,\rho,R}$ such that, if $m \geq m_{\gamma,\rho,R}$, then $\frac{b_m}{m} \leq \rho$. Moreover, the assumptions of Lemma 19 are satisfied:

$$\frac{b_m}{m}\left(\frac{1}{2}k_m - 1\right) \geq \frac{k_m}{2m} - \frac{b_m}{m} \geq \frac{50}{\gamma}R - \rho > R.$$

Find constant $A_{\gamma,\rho,R} < \infty$ such that for each $m \geq m_{\gamma,\rho,R}$,

$$(A_{\gamma,\rho,R})^{m^6} \geq k_m b_m \max\left(20, 2\log 2\left(-\log\left(\frac{1}{40}\gamma\right)\right), \frac{2}{1-\gamma}\left(-\log\left(\frac{1}{20}\gamma\right)\right)(P(0))^{-k_m^2 b_m^2}\right).$$

(Such a constant exists because $k_m \leq \frac{200}{\gamma} Rm$ and $b_m \leq m$.) Take $K \geq K_m = \left\lceil \frac{1}{k_m b_m} \left(A_{\gamma,\rho,R}\right)^{m^6} \right\rceil$ and let $M = K k_m b_m$. Then, the assumptions of Lemma 26 are satisfied with $\frac{1}{10}\gamma$ instead of γ:

$$\mathrm{e}^{-(1-\gamma)K^2 P(0)^{k_m^2 b_m^2}} \leq \frac{1}{20}\gamma \text{ and } K^2 2^{-K} \leq 2^{-\frac{1}{2}K} \leq \frac{1}{40}\gamma.$$

Finally,

$$2\frac{b_m}{M} + 4\frac{1}{k_m}\left(\frac{Rm}{b_m} + 1\right) \leq 2\frac{1}{k_m} + 4\frac{Rm}{k_m} + \frac{4}{100}\gamma \leq \gamma,$$

which implies that the bound in the brackets of Lemma 18 is larger than $1 - 4\frac{1}{k_m}\left(\frac{Rm}{b_m} + 1\right) \geq 1 - \gamma$.

Lemma 26 implies that

$$\mathbb{P}\left(|U_\gamma| \geq \left(1 - \frac{1}{10}\gamma\right) K^2 \text{ and } U_\gamma \text{ contains a extraordinary large cube}\right) \geq 1 - \frac{1}{10}\gamma.$$

If $|U_\gamma| \geq \left(1 - \frac{1}{10}\gamma\right) K^2$, Lemma 18 implies that $\left|\bigcup W\left(U_\gamma, R\right)\right| \geq (1-\gamma) M^2$, and Lemma 19 implies that $W\left(U_\gamma, R\right)$ is connected in the network of small cubes. The definition of $W\left(U_\gamma, R\right)$ implies that each small cube that is not γ-good, and hence not contained in U, is at least R-distant from each small cube contained in $W\left(U_\gamma, R\right)$. Finally, because $R < \frac{b_m}{m}\left(\frac{1}{10}k_m - 1\right)$, if $C_0 \in U_\gamma$ is an extraordinary large cube, then $W\left(C_0, R\right)$ is non-empty and it contains a small cube $c_0 \in W\left(C_0, R\right) \subseteq W\left(U_\gamma, R\right)$ such that for any c, if $d\left(c, c_0\right) \leq R$, then $c \in C_0$ and c is extraordinary. Therefore set $W\left(U_\gamma, R\right)$ is (γ, R)-good.

D.6. **Proof of Theorem .** Fix $\eta > 0$. We are going to show that, for each $\eta > 0$, there exist constants $A, m_0 > 0$ such that, if $m \geq m_0$ and $M \geq A^{m^6}$, and G is (M, m)-lattice, then the probability that there is an equilibrium a on the (M, m)-lattice such that $\mathrm{Av}\left(a\right) = \frac{1}{M^2}\sum a \geq x^* + \eta$ is smaller than η. The argument for the lack of equilibria with average action below $x^* - \eta$ is analogous (and it follows from exchanging the roles for binary actions 0 and 1). Combining the two bounds (and taking maximum of respective constants A and m_0) delivers the result.

Apply Lemma 12 to $\frac{1}{2}\eta$ and find $\delta > 0$, $a^* < x + \frac{1}{2}\eta$, $L < \infty$, and a δ-contagion wave σ for P.

Let $D \geq 1$ be a constant that is larger than the sum of constants from Lemmas 15 and 16. Choose $\rho \leq \frac{1}{D}\delta$ and $\gamma \leq \min\left(\delta, \frac{1}{4}\eta\right)$. Let R_ρ be the constant from Lemma 13. Let $R = R_\rho + L$. Let $m_0 = m_{\gamma,\rho,R}$ and $A = A_{\gamma,\rho,R}$. Choose $m \geq m_0$, $M \geq A^{m^6}$, and b be as in Lemma 17.

Let W denote a (γ, R)-good set of cubes in the network of small cubes $\mathcal{G}^b$ if such a set exists. Let $c_0 \in W$ be the cube such that for each c, if $d(c, c_0) \leq R$, then c is extraordinary.

Let a be any equilibrium on the lattice. Let $W_d \subseteq W$ be a maximal subset of small cubes such that the equilibrium a is (W_d, γ, R_ρ)-dominated by σ. If W exists, then $c_0 \in W_d$ and W_d is non-empty. (To see why, notice that $a(c) = 0 \leq \sigma(d(c, c_0) - R_\rho)$ for each extraordinary cube, including all cubes c st. $d(c, c_0) \leq R$. Additionally, $\sigma(d(c, c_0) - R_\rho) \geq \sigma(L) = 1 \geq a(c)$ for each cube c such that $d(c, c_0) > R$.) By Lemmas 15 and 16, for each γ-good small cube c,

$$\begin{aligned} a(c) &\leq \gamma + P\left(a^* + \sum_{a \in \sigma(\mathbb{R})} \left(1 - f\left(\sigma^{-1}(a) + R_\rho - d(c, W_d)\right)\right)(a - a_-) + D\rho\right) \\ &\leq \delta + P\left(a^* + \sum_{a \in \sigma(\mathbb{R})} \left(1 - f\left(\sigma^{-1}(a) + R_\rho - d(c, W_d)\right)\right)(a - a_-) + \delta\right). \end{aligned}$$

Because σ is a δ-contagion wave (see Lemma 12), the above is no larger than

$$\leq \sigma(d(c, W_d) - R_\rho - \delta).$$

Suppose that $W_d \neq W$. Because W is connected, there is a cube $c_d \in W \backslash W_d$ such that c_d is a neighbor of $c_d' \in W_d$ in the network of small cubes. Then, $d(c_d, c_d') \leq \rho$, and, by the triangle inequality, $d(c, W_d \cup \{c_d\}) \geq d(c, W_d) - \rho$ for any cube c. We have:

- for each γ-good cube c, because $\rho \leq \delta$,

$$a(c) \leq \sigma(d(c, W_d) - R - \delta) \leq \sigma(d(c, W_d \cup \{c_d\}) - R).$$

- for each cube c that is not γ-good, we have $d(c, W_d \cup \{c_d\}) \geq R \geq R_\rho + L$ due to $W_d \cup \{c_d\} \subseteq W$. But then, $a(c) \leq 1 = \sigma(L) = \sigma(d(c, W_d \cup \{c_d\}) - R)$.

It follows that equilibrium a is $(W_d \cup \{c_d\}, \gamma, R_\rho)$-dominated by σ. But this is a contradiction with the choice of W_d as a maximal set.

Therefore, $W_d = W$, a is (W, γ, R_ρ)-dominated by σ, and for each $c \in W$,

$$a(c) \leq \sigma(d(c, W) - R) + \rho = \sigma(-R) + \rho \leq a^* + \frac{1}{4}\eta.$$

Hence

$$\begin{aligned}\mathrm{Av}(a) &= \frac{1}{M^2}\sum a_i = a^* + \frac{1}{|\mathcal{G}^b|}\sum_{c\in W}(a(c) - a^*) + \frac{|I_M \backslash \bigcup W|}{M^2}\sum_{i\notin \bigcup W}(a_i - a^*)\\ &\leq a^* + \frac{1}{4}\eta + \gamma \leq x^* + \eta.\end{aligned}$$

Because the probability that (γ, R)-good set of small cubes exists is at least $1-\gamma \geq 1-\eta$, the above inequality demonstrates our claim.

Appendix E. Proof of Theorem 4

E.1. **Proof overview.** We formally describe the best response dynamics: initial profile and the updating process. Next, we compute capacity-type bounds on the dynamics, i.e., calculations (8) from the main body of the paper. We show that the reminder terms are small. We use this to show that the average payoffs at the end of the dynamics cannot be significantly different from x^* and conclude the proof of the theorem.

E.2. **Initial profile.** In this part of the Appendix, we define the initial profile for the dynamics and its properties. Let x^* be the RU-dominant outcome. For each relation $r \in \{=, <, >\}$, let $E_r = \{\varepsilon_i : u(0, x^*, \varepsilon_i)\ r\ u(1, x^*, \varepsilon_i)\}$. Then, $E_=$ is the set of payoff shocks that make player indifferent if exactly fraction $x^*$of their neighbors plays action 1. Then, because x^* is an RU-dominant outcome, $F(E_<) \leq x^* \leq F(E_<) + F(E_=)$. If $F(E_=) \neq 0$, define $p = \frac{F(E_<)+F(E_=)-x^*}{F(E_=)}$. For each player i, let Y_i be the binomial i.i.d. variable equal to 1 with probability p and equal to 0 otherwise.

Define an initial strategy profile as a function of the payoff shocks:

$$a_i^0 = \begin{cases} BR_i(a_{-i}; \varepsilon_i) & \text{if } |BR_i(a_{-i}; \varepsilon_i)| = 1 \\ Y_i & \text{otherwise.}\end{cases} \tag{17}$$

For each player i, let $\beta_i^0 = \frac{1}{g_i}\sum g_{ij}a_j^0$ be the fraction of neighbors of agent i who play action 1 under profile a_i^0. The next result derives a probabilistic bound on the average distance of neighborhood behaviors from the RU-dominant outcome.

Lemma 27. *For each $\eta > 0$, there exists $d > 0$ such that if $d(g) \leq d$, then*

$$\mathbb{P}\left(\sum g_i \left|\beta_i^0 - x^*\right| > \eta\left(\sum g_i\right)\right) < \eta.$$

Proof. Variables a_j^0 are independent of each other and $\mathbb{E}\, a_j^0 = x^*$. Hence, for each i,

$$\mathbb{E}\left(\beta_i^0 - x^*\right)^2 = \sum_j \frac{g_{ij}^2}{g_i^2}\,\mathbb{E}\left(a_j^0 - x^*\right)^2 \leq \sum_j d(g)\,\frac{g_{ij}}{g_i} = d(g)\,.$$

By the Cauchy-Schwartz inequality, we get $\mathbb{E}\left|\beta_i^0 - x^*\right| \leq 2\sqrt{d(g)}$. Let $d(g) \leq d = \frac{1}{4}\eta^4$. Then, by the Markov's equality, for each η,

$$\mathbb{P}\left(\sum g_i \left|\beta_i^0 - x^*\right| > \eta\left(\sum g_i\right)\right) \leq \frac{\mathbb{E}\left(\sum g_i \left|\beta_i^0 - x^*\right|\right)}{\eta\left(\sum g_i\right)} \leq \frac{2\sqrt{d(g)}}{\eta} \leq \eta.$$

□

E.3. **Best response process.** In this subsection, we formally define best response dynamics: starting from the initial profile a^0, agents who play 0 but have 1 as a best response revise their actions to 1, in an arbitrary (but fixed) order. Assume that all players are labeled with numbers $i \in \{1, ..., N\}$. For all $t \geq 0$, and for each i, let

$$\begin{aligned}
\beta_i^t &= \frac{1}{g_i}\sum g_{ij}a_j^t, \qquad (18)\\
p_i^t &= P\left(\beta_i^t\right),\\
i_t &= \min\left\{i : a_i^t = 0 \text{ and } u\left(1, \beta_i^t, \varepsilon_i\right) \geq u\left(0, \beta_i^t, \varepsilon_i\right)\right\},\\
a_i^{t+1} &= \begin{cases} 1 & \text{if } i = i_t \\ a_i^t & \text{otherwise.}\end{cases}
\end{aligned}$$

We refer to p_i^t as the expected action of agent i in period t. Because at most one player changes actions at each step, we have $\left|\beta_i^t - \beta_i^{t+1}\right| \leq d(g)$ for each i. The stochastic process $\left(a^t, \beta^t, p^t\right)_t$ depends on the realization of payoff shocks ε.

If the set in the third line is empty, the process stops. Because there are finitely many players, the dynamics must stop in a finite time. We denote the final outcome of the process as $\left(a_i^U, \beta_i^U, p_i^U\right)$.

E.4. **Main step.** For each profile of expected actions p, define the functional

$$\mathcal{F}(p) = \frac{1}{2}\sum_{i,j} g_{ij}\left(p_i - p_j\right)^2.$$

Clearly, $\mathcal{F}(p^t) \geq 0$ for each t. Also, define function

$$L(x) = \int_{x^*}^{x} \left(P^{-1}(y) - y\right) dy.$$

Because x^* is RU-dominant, it is the unique minimizer of $L(x)$. Hence $L(x^*) = 0$ and $L(x) > 0$ for each $x \neq x^*$.

The next lemma fills calculations behind formula (8) in the main body of the paper.

Lemma 28. *For each* t*,*

$$2\sum_i g_i L\left(p_i^{T+1}\right) \leq \mathcal{F}\left(p^0\right) + A + 2\sum_i g_i \left|\beta_i^0 - x^*\right| + 2d(g)\sum g_i, \tag{19}$$

where A *is defined as*

$$A = \sum_{t\leq T}\sum_i \left(p_i^{t+1} - p_i^t\right)\sum_j g_{ij}\sum_{s=t,t+1}\left(a_j^s - p_j^s\right).$$

Proof. Observe that for each t,

$$\begin{aligned}
&\mathcal{F}\left(p^{t+1}\right) - \mathcal{F}\left(p^t\right)\\
=&\sum_i g_i\left(p_i^{t+1}\right)^2 - \sum_i g_i\left(p_i^t\right)^2 - \sum_{i,j} g_{ij}\left(p_i^{t+1}p_j^{t+1} - p_i^t p_j^t\right)\\
=&\sum_i g_i\left(p_i^{t+1}\right)^2 - \sum_i g_i\left(p_i^t\right)^2 - \sum_{i,j} g_{ij}\left(\left(p_i^{t+1} - p_i^t\right)p_j^{t+1} + p_i^t\left(p_j^{t+1} - p_j^t\right)\right)\\
=&\sum_i g_i\left(p_i^{t+1}\right)^2 - \sum_i g_i\left(p_i^t\right)^2 - \sum_i \left(p_i^{t+1} - p_i^t\right)\sum_j g_{ij}\sum_{s=t,t+1} p_j^s\\
=&\sum_i g_i\left(p_i^{t+1}\right)^2 - \sum_i g_i\left(p_i^t\right)^2 - \sum_i g_i\left(p_i^{t+1} - p_i^t\right)\sum_{s=t,t+1}\beta_i^s + \sum_i \left(p_i^{t+1} - p_i^t\right)\sum_j g_{ij}\sum_{s=t,t+1}\left(a_j^s - p_j^s\right),
\end{aligned}$$

where, in the last line, we used $g_i\beta_i^s = \sum_j g_{ij}a_j^s$. Summing up across $t \leq T$, we obtain

$$\mathcal{F}\left(p^{T+1}\right) - \mathcal{F}\left(p^0\right) = \sum_{t\leq T}\left(\mathcal{F}\left(p^{t+1}\right) - \mathcal{F}\left(p^t\right)\right)$$

$$= \sum_i g_i\left(p_i^{T+1}\right)^2 - \sum_i g_i\left(p_i^0\right)^2 - \sum_{t\leq T}\sum_i g_i\left(p_i^{t+1} - p_i^t\right)\sum_{s=t,t+1}\beta_i^s + A$$

$$= A + \sum_i g_i\left[\left(p_i^{T+1}\right)^2 - \left(p_i^0\right)^2 - 2\int_{p_i^0}^{p_i^{T+1}} P^{-1}(y)\,dy\right]$$

$$+ \sum_{t\leq T}\left[2\int_{p_i^t}^{p_i^{t+1}} P^{-1}(y)\,dy - \left(p_i^{t+1} - p_i^t\right)\sum_{s=t,t+1}\beta_i^s\right].$$

The second term of the above is equal to

$$\sum_i g_i\left[\left(p_i^{T+1}\right)^2 - \left(p_i^0\right)^2 - 2\int_{p_i^0}^{p_i^{T+1}} P^{-1}(y)\,dy\right]$$

$$= 2\sum_i g_i\left[\int_{p_i^0}^{p_i^{T+1}} y\,dy - \int_{p_i^0}^{p_i^{T+1}} P^{-1}(y)\,dy\right] = 2\sum_i g_i\left(L\left(p_i^0\right) - L\left(p_i^{T+1}\right)\right).$$

Notice that $L(x^*) = L(P(x^*)) = 0$ and $L(P(\beta_i^0))$ is Lipschitz with constant 1. Hence the above is no larger than

$$\leq -2\sum_i g_i L\left(p_i^{T+1}\right) + \sum g_i\left|\beta_i^0 - x^*\right|.$$

Recall that $\sup_{t\leq T}\left(\beta_i^{t+1} - \beta_i^t\right) \leq d(g)$. By definition of the Lebesgue integral,

$$\sum_{t\leq T}\beta_i^t\lambda\left(y : \beta_i^t \leq P^{-1}(y) < \beta_i^{t+1}\right) \leq \int_{p_i^0}^{p_i^{T+1}} P^{-1}(y)\,dy$$

$$\leq \sum_{t\leq T}\left(\beta_i^t + d(g)\right)\lambda\left(y : \beta_i^t \leq P^{-1}(y) < \beta_i^{t+1}\right),$$

where λ is the Lebesgue measure on the interval $[0,1]$. The definition of inverse function P^{-1} as well as $p_i^t = P\left(\beta_i^t\right)$ for each t imply that

$$\lambda\left(y : \beta_i^t \leq P^{-1}\left(y\right) < \beta_i^{t+1}\right) = p_i^{t+1} - p_i^t.$$

Hence

$$\sum_i g_i \sum_{t\leq T}\left[2\int_{p_i^t}^{p_i^{t+1}} P^{-1}\left(y\right)dy - \left(p_i^{t+1}-p_i^t\right)\sum_{s=t,t+1}\beta_i^s\right] \leq 2d\left(g\right)\sum g_i.$$

The result follows from putting the estimates together and the fact that $\mathcal{F}\left(p_i^{T+1}\right) \geq 0$. □

E.5. **Estimates.** In this section, we provide estimates of the terms on the right-hand side of (19).

Lemma 29. *For each $\eta > 0$, there exists $d_\eta^F > 0$ such that, if $d\left(g\right) \leq d_\eta^F$, then*

$$\mathbb{P}\left(\mathcal{F}\left(p^0\right) > \eta\left(\sum g_i\right)\right) < \eta.$$

Proof. Note that

$$\mathcal{F}\left(p^0\right) \leq \sum_i g_i\left(p_i - x^*\right)^2 \leq \sum_i g_i\delta\left(\beta_i^0\right),$$

where $\delta\left(x\right) = \left(P\left(x\right) - x^*\right)^2$. Note that $\delta\left(x^*\right) = 0$. Choose $\xi > 0$ such that $\delta\left(\sqrt{\xi}\right) + \sqrt{\xi} < \frac{1}{2}\eta$. Let $d < \eta$ be small enough so that Lemma 27 holds for ξ. Then,

$$\mathbb{P}\left(\sum_{i:\beta_i^0\geq\sqrt{\xi}} g_i > \sqrt{\xi}\sum g_i\right) \leq \xi,$$

and, if the event in the brackets does not hold, we have

$$\sum_i g_i\delta\left(\beta_i^0\right) \leq \sum_i g_i\left(\delta\left(\sqrt{\xi}\right) + \sqrt{\xi}\right) \leq \eta\left(\sum_i g_i\right).$$

□

To gain estimates on term A, we need a preliminary lemma:

Lemma 30. *For each j and s,*

$$\mathbb{E}\left(a_j^s - \max\left(x^*, p_j^s\right)|\varepsilon_{-j}\right) \leq 0.$$

Proof. Fix player j. The stochastic process $\left(a^t, \beta^t, p^t\right)_t$ can be defined on the probability space $\Omega = E^N$ composed of the realizations of the payoff shock for each individual. Consider an auxiliary stochastic processes $\left(a'^t, \beta'^t, p'^t\right)_t$ defined on the same probability space with the same equations (17)-(18) as the original process, but with setting $a_j'^t \equiv a_j^0$ for each t. Additionally, define

$$a_j^{*t+1} = 1 \text{ iff } u\left(1, \max\left(x^*, \beta_j'^t\right), \varepsilon_j\right) \geq u\left(0, \max\left(x^*, \beta_j'^t\right), \varepsilon_j\right).$$

So defined a_j^{*t} depends on ε_{-j} only through process β'. Hence, for each ε_{-j},

$$\begin{aligned}\mathbb{P}\left(a_j^{*t+1} = 1|\varepsilon_{-j}\right) &= \mathbb{P}\left(u\left(1, \max\left(x^*, \beta_j'^t\right), \varepsilon_j\right) \geq u\left(0, \max\left(x^*, \beta_j'^t\right), \varepsilon_j\right)|\varepsilon_{-j}\right) \\ &= P\left(\max\left(x^*, \beta_j'^t\right)\right).\end{aligned}$$

Notice that $a_j^{*t} \geq a_j^t$ for each t. Indeed, let $t_0 = \inf\left\{t : a_j^t = 1\right\}$ and equal ∞ if the set is empty. Then, $\beta_i^t = \beta_i'^t$ for each i and $t < t_0$. Moreover, $a_j^{t_0} = 1$ implies $u\left(1, \beta_j^{t_0-1}, \varepsilon_j\right) \geq u\left(0, \beta_j^{t_0-1}, \varepsilon_j\right)$, which implies that $a_j^{*t_0} = 1$.

Further, payoff complementarities imply that, for each s, $\beta'^s \leq \beta^s$, and hence $p'^s \leq p^s$. Additionally, $p'^{s-1} \leq p'^s$. Thus,

$$\begin{aligned}\mathbb{E}\left(a_j^s - \max\left(x^*, p_j^s\right)|\varepsilon_{-j}\right) &= \mathbb{P}\left(a_j^s = 1|\varepsilon_{-j}\right) - \max\left(x^*, p_j'^s\right) \\ &\leq \mathbb{P}\left(a_j^{*s} = 1|\varepsilon_{-j}\right) - \max\left(x^*, p_j'^s\right) \\ &= P\left(\max\left(x^*, \beta_j'^{s-1}\right)\right) - \max\left(x^*, p_j'^s\right) \\ &= \max\left(x^*, p_j'^{s-1}\right) - \max\left(x^*, p_j'^s\right) \leq 0,\end{aligned}$$

where the first equality is due to the fact that $p_j'^{s-1}$ and $\beta_j'^{s-1}$ are measurable wrt. ε_{-i}. □

Lemma 31. *For each $\eta > 0$, there exists $d_\eta > 0$ such that, if $d(g) \leq d_\eta^1$, then*

$$\mathbb{P}\left(\frac{1}{g_i}\sum g_{ij}\left(a_j^s - \max\left(x^*, p_j^s\right)\right) \geq \eta\right) \leq \eta.$$

Proof. By Lemma 30, finite stochastic process $X_j = \frac{1}{g_i}\sum_{j'\leq j} g_{ij'}a^s_{j'}$ is a supermartingale. Take $d_\eta = -\frac{\eta}{\ln \eta}$. Then, the Azuma-Hoeffding's Inequality implies that

$$\mathbb{P}\left(\frac{1}{g_i}\sum g_{ij}a^s_j - p^s_j \geq \eta\right) \leq \exp\left(-\frac{\eta}{\sum \frac{g_{ij}^2}{g_i^2}}\right) \leq \exp\left(-\frac{1}{d\left(g\right)}\eta\right) \leq \exp\left(\ln \eta\right) = \eta.$$

□

Lemma 32. *For each $\eta > 0$, there exists $d^A_\eta > 0$ such that if $d\left(g\right) \leq d^A_\eta$, then for each i and s,*

$$\mathbb{P}\left(A \geq \eta \sum_i g_i\right) \leq \eta.$$

Proof. Because $p_i^{t+1} > p_i^t$ for each i,

$$\begin{aligned}
A =& \sum_{t\leq T}\sum_i \left(p_i^{t+1} - p_i^t\right)\sum_j g_{ij}\sum_{s=t,t+1}\left(a^s_j - \max\left(x^*, p^s_j\right)\right) \\
&+ \sum_{t\leq T}\sum_i \left(p_i^{t+1} - p_i^t\right)\sum_j g_{ij}\sum_{s=t,t+1}\left(\max\left(x^*, p^s_j\right) - p^s_j\right) \\
\leq& \sum_{t\leq T}\sum_i \left(p_i^{t+1} - p_i^t\right)\sum_j g_{ij}\sum_{s=t,t+1}\left(a^s_j - \max\left(x^*, p^s_j\right)\right) + 2\sum_j g_j\left|p^0_j - x^*\right| \\
=& A_1 + A_2.
\end{aligned}$$

We are going to bound each of the two terms separately.

Let d^2_η be the constant from Lemma 32. Then, if $d\left(g\right) \leq d^{A1}_\eta = d^2_{\frac{1}{8}\sqrt{\eta}}$,

$$\mathbb{E}\left(\sum_{t\leq T}\sum_i \left(p_i^{t+1} - p_i^t\right)\sum_j g_{ij}\sum_{s=t,t+1}\left(a^s_j - \max\left(x^*, p^s_j\right)\right)\right) \leq \frac{1}{4}\sqrt{\eta}\sum_i g_i.$$

By Markov's inequality,

$$\mathbb{P}\left(A_1 \geq \frac{1}{2}\eta\sum g_i\right) \leq \frac{\frac{1}{4}\sqrt{\eta}\sum_i g_i}{\frac{1}{2}\eta\sum g_i} \leq \frac{1}{2}\eta.$$

Take $\delta(x) = |P(x) - x^*|$. Note that $\delta(x^*) = 0$. Choose $\xi > 0$ such that $\max\left(\xi, 4\left(\delta\left(\sqrt{\xi}\right) + \sqrt{\xi}\right)\right) < \frac{1}{2}\eta$. Let $d_\eta^{A2} < \eta$ be sufficiently small so that Lemma 27 holds for ξ. Then,

$$\mathbb{P}\left(\sum_{i:\beta_i^0 \geq \sqrt{\xi}} g_i > \sqrt{\xi}\sum g_i\right) \leq \xi,$$

and, if the event in the brackets does not hold, we have

$$2\sum_j g_j \left|P\left(\beta_j^0\right) - x^*\right| \leq 2\left(\delta\left(\sqrt{\xi}\right) + \sqrt{\xi}\right)\sum_i g_i \leq \frac{1}{2}\eta\left(\sum_i g_i\right).$$

Take $d_\eta^A = \min\left(d_\eta^{A1}, d_\eta^{A2}\right)$. Then,

$$\mathbb{P}\left(A \geq \eta\sum_i g_i\right) \leq \mathbb{P}\left(A_1 \geq \frac{1}{2}\eta\sum_i g_i\right) + \mathbb{P}\left(A_2 \geq \frac{1}{2}\eta\sum_i g_i\right) \geq \eta.$$

□

E.6. **Average payoffs at the end of dynamics.** We show that the average payoffs when the upper best response dynamics stop are not much higher than x^*.

Lemma 33. *For each $\eta > 0$, there exists $d_\eta^U > 0$ such that, if $d(g) \leq d_\eta^U$, then*

$$\mathbb{P}\left(Av\left(a^U\right) \geq (\eta + x^*)\sum_i g_i\right) \leq \eta.$$

Proof. By definition, x^* is the unique maximizer of $L(x)$. Fix $\eta > 0$ and find $\xi > 0$ such that $\sqrt{\xi} \leq \eta$ and if $L(x) \leq \sqrt{\xi}$, then $x \leq x^* + \frac{1}{2}\eta$.

Let$(a^t, \beta^t, p^t)_t$ be the upper best response dynamics defined in Section E.3. By Lemmas 19, 29, and 32, if $d \leq d_\xi^U = \max\left(d_\xi^F, d_\xi^A\right)$, then

$$\sum_i g_i L\left(p_i^U\right) \leq \xi\sum_i g_i$$

with a probability of at least $1 - \xi$. It follows that $\sum_{i:L\left(p_i^U\right)\geq\sqrt{\xi}} g_i \leq \sqrt{\xi}$, which implies that $\sum_{i:\beta_i^U \geq x^* + \frac{1}{2}\eta} g_i \leq \sqrt{\xi}$. Hence,

$$\sum g_i\beta_i^U \leq \sum_{i:\beta_i^U \leq x^* + \frac{1}{2}\eta} g_i\left(x^* + \frac{1}{2}\eta\right) + \sqrt{\xi}\sum g_i \leq (x^* + \eta)\sum_i g_i.$$

Finally, notice that

$$\operatorname{Av}\left(a^U\right) = \sum_i g_i a_i^U = \sum_i \sum_j g_{ij} a_i^U = \sum_i \sum_j g_{ij} a_i^U = \sum_i g_i \beta_i^U.$$

The result follows from the above inequality. □

E.7. **Proof of Theorem 4.** Lemma 33 shows that the best response dynamics, where players only revise their actions upwards, stop with a profile a^U with average payoffs close to x^*. An analogous result shows that a lower version of the best response dynamics, initiated from the same profile a^0 and where players only revise their actions downwards, stop with a profile a^L with average payoffs also close to x^*.

Due to payoff complementarities, the lower best response dynamics initiated from profile a^U will stop at equilibrium profile a^{UL} that lies in between a^U and a^L. The latter implies that the average payoffs must lie in between the average payoffs $\operatorname{Av}\left(a^U\right)$ and $\operatorname{Av}\left(a^L\right)$. The claim follows.

E.8. **Extension to unweighted average.** The argument remains identical except for the following modification of Lemma 33: For each $\eta > 0$ and $w < \infty$, there exists $d_\eta^U > 0$ such that, if $d(g) \leq d_\eta^U$, and $w(g) \leq w$ then

$$\mathbb{P}\left(\operatorname{Av}_{\text{unweighted}}\left(Ua^0\right) \geq (\eta + x^*)\right) \leq \eta.$$

To see the above claim, recall that $a_i^U \geq a_i^0$. Hence

$$\begin{aligned}
&\operatorname{Av}_{\text{unweighted}}\left(a^U\right) - \operatorname{Av}_{\text{unweighted}}(a) \\
=&\frac{1}{N}\sum_i \left(a_i^U - a_i^0\right) = \frac{1}{\min_i g_i}\frac{1}{N}\sum_i \left(\min_j g_j\right)\left(a_i^U - a_i^0\right) \\
\leq&\frac{1}{\min_i g_i}\frac{1}{N}\sum_i g_i \left(a_i^U - a_i^0\right) \leq \frac{1}{\min_i g_i}\frac{\sum g_i}{N}\frac{1}{\sum g_i}\sum_i g_i\left(a_i^U - a_i^0\right) \\
\leq&\frac{\max_i g_i}{\min_i g_i}\left(\operatorname{Av}(Ua) - \operatorname{Av}\left(a^0\right)\right) = w(g)\left(\operatorname{Av}(Ua) - \operatorname{Av}\left(a^0\right)\right).
\end{aligned}$$

An application of Lemma 33 established the claim.

References

[Blume(1993)] Blume, L. E. (1993). The statistical mechanics of strategic interaction. *Games and economic behavior*, **5** (3), 387–424, publisher: Elsevier.

[Bollobás *et al.*(2006)Bollobás, Riordan and Riordan] Bollobás, B., Riordan, O. and Riordan, O. (2006). *Percolation*. Cambridge University Press.

[Ellison(1993)] Ellison, G. (1993). Learning, local interaction, and coordination. *Econometrica: Journal of the Econometric Society*, pp. 1047–1071, publisher: JSTOR.

[Ellison(2000)] — (2000). Basins of attraction, long-run stochastic stability, and the speed of step-by-step evolution. *The Review of Economic Studies*, **67** (1), 17–45, publisher: Wiley-Blackwell.

[Friedman(1991)] Friedman, D. (1991). Evolutionary games in economics. *Econometrica: Journal of the Econometric Society*, pp. 637–666.

[Harsanyi and Selten(1988)] Harsanyi, J. C. and Selten, R. (1988). A general theory of equilibrium selection in games. *MIT Press Books*, **1**, publisher: The MIT Press.

[Jackson and Zenou(2015)] Jackson, M. O. and Zenou, Y. (2015). Games on networks. In *Handbook of game theory with economic applications*, vol. 4, Elsevier, pp. 95–163.

[Kandori *et al.*(1993)Kandori, Mailath and Rob] Kandori, M., Mailath, G. J. and Rob, R. (1993). Learning, mutation, and long run equilibria in games. *Econometrica: Journal of the Econometric Society*, pp. 29–56, publisher: JSTOR.

[Morris(2000)] Morris, S. (2000). Contagion. *The Review of Economic Studies*, **67** (1), 57–78, publisher: Wiley-Blackwell.

[Neary(2012)] Neary, P. R. (2012). Competing conventions. *Games and Economic Behavior*, **76** (1), 301–328.

[Neary and Newton(2017)] — and Newton, J. (2017). Heterogeneity in preferences and behavior in threshold models. *Available at SSRN 3035289*.

[Newton(2021)] Newton, J. (2021). Conventions under heterogeneous behavioural rules. *The Review of Economic Studies*, **88** (4), 2094–2118.

[Peski(2010)] Peski, M. (2010). Generalized risk-dominance and asymmetric dynamics. *Journal of Economic Theory*, **145** (1), 216–248, publisher: Elsevier.

[Vershynin(2018)] Vershynin, R. (2018). *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge: Cambridge University Press.

[Young(1993)] Young, H. P. (1993). The evolution of conventions. *Econometrica: Journal of the Econometric Society*, pp. 57–84, publisher: JSTOR.