

# Fuzzy Conventions

Marcin Peński

University of Toronto

March 7, 2022

- Social interactions, positive externalities.
  - wearing a mask,
  - engaging in criminal activity,
  - technology adoption.
- A typical result: emergence of a (homogeneous) convention.
- But, in reality, conventions are often fuzzy:
  - some, but not all, wear masks,
  - married couples that use both iPhone and Android.
- People care not only about their neighbors, but they differ wrt. tastes, preferences.

- Binary coordination games on networks with random utility,
- (Statistical) heterogeneous preferences: i.i.d payoff shocks,
- I am interested in the set of average (i.e., aggregate) behavior  $x \in [0, 1]$ 
  - in static,
  - complete information equilibria,
  - when each agent number of connection is large.
- **Q:** What can we say about equilibrium sets? How do they depend on the network?

- agents  $i, j$  live on a network with weights  $g_{ij} = g_{ji} \geq 0$ ,
  - $g_i = \sum_j g_{ij}$  is degree of agent  $i$ ,
- payoffs:  $\sum_{j \neq i} g_{ij} u(a_i, a_j, \varepsilon_i)$ ,
  - each  $i$  chooses  $a_i \in \{0, 1\}$ ,
  - i.i.d. payoff shocks  $\varepsilon_i \sim F$ ,
  - positive externalities:  $u(\cdot, \cdot, \varepsilon_i)$  has increasing differences for each  $\varepsilon$ ,
- average behavior  $Av(a) = \frac{1}{\sum_i g_i} \sum_i g_i a_i$ ,
- equilibrium set

$$\text{Eq}(g, \varepsilon) = \{Av(a) : a \text{ is a Nash equilibrium in game } G(g, \varepsilon)\},$$

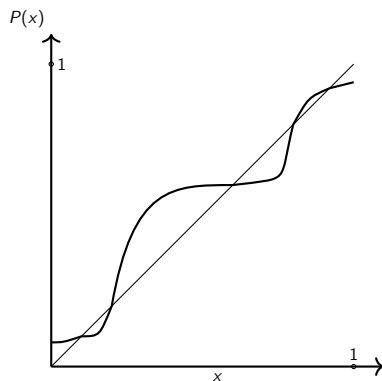
- Object of interest:  $\lim \text{Eq}(g, \cdot)$  as
  - $d(g) = \max_{i,j} \frac{g_{ij}}{g_i} \rightarrow 0$  - large degrees,
  - $w(g) = \max_{i,j} \frac{g_i}{g_j} < w_{\max} < \infty$  is bounded - not too much inequality.

- 4 theorems that characterize the largest and the smallest possible limit of equilibrium sets across all networks.

# Introduction

## Results

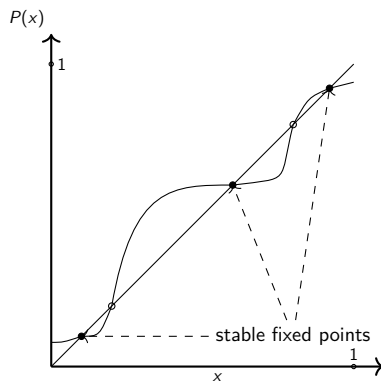
- Let  $P(x) = F\{\varepsilon : u(1, x, \varepsilon) \geq u(0, x, \varepsilon)\}$ ,
- continuum best response



# Introduction

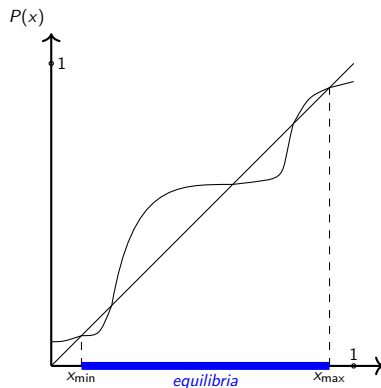
## Results

- Let  $P(x) = F\{\varepsilon : u(1, x, \varepsilon) \geq u(0, x, \varepsilon)\}$ ,
- continuum best response

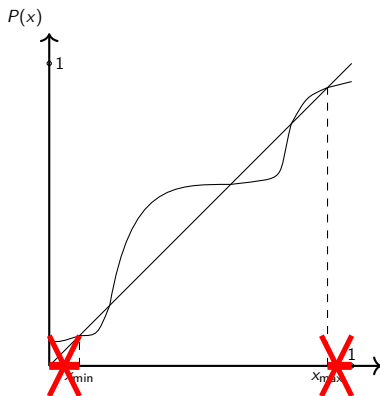




- **Theorem 1:** There exists a sequence of networks such that the limit equilibrium set is  $[x_{\min}, x_{\max}]$ .



- **Theorem 2:** All limit equilibrium sets are contained in  $[x_{\min}, x_{\max}]$ .

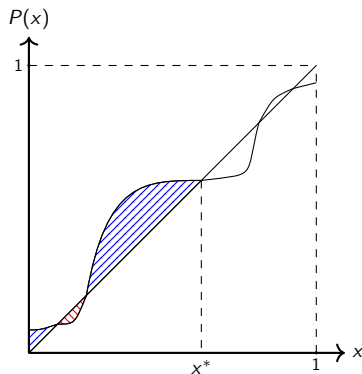


# Introduction

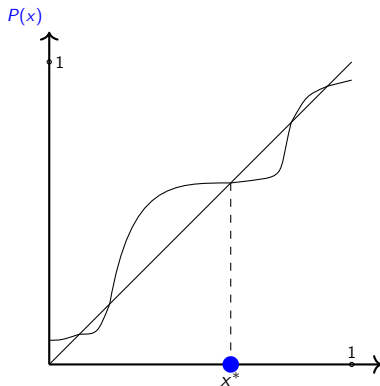
## Results

- Define *random utility (RU-) dominant* outcome

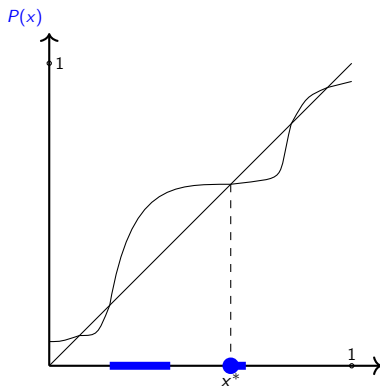
$$x^* \in \arg \max_x \int_0^x (y - P^{-1}(y)) dy.$$



- **Theorem 3:** There exists a sequence of networks such that the limit equilibrium set is  $\{x^*\}$ .



- **Theorem 4:** All limit equilibrium sets contain  $x^*$ .



- Large models with random utility (matching: Dagsvik 00, Choo-Siow 06, Menzel (15), Peski (17, 22))
- Emergence of conventions: evolutionary approach
  - risk-dominance (Harsanyi-Selten 88),
  - complete networks (Kandori, Mailath Rob 93), (Young 93), line and some other networks (Ellison 93, Ellison 00), all networks (Peski 10).
- Global games and robustness to incomplete information
- Contagion (Morris 00):
  - some networks (lattices) admit contagion: a finite group of agents can spread risk-dominant behavior to the rest of the network,
  - contagion only works towards risk-dominant action.
- Here,
  - random utility instead of noise (or a perturbation),
  - static solution concept,
  - no aggregate uncertainty.

# Theorem 1

## Notation

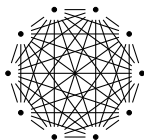
- Define a profile of neighborhood fractions  $\beta^a$ : for each  $i$

$$\beta_i^a = \frac{1}{g_i} \sum_{i \neq j} g_{ij} a_j,$$

- $A \subseteq_{\eta} B$  if for each  $a \in A$ , there is  $b \in B$  st.  $|a - b| \leq \eta$ ,  
 $A =_{\eta} B$  if  $A \subseteq_{\eta} B$  and  $B \subseteq_{\eta} A$ .

# Theorem 1

- Let  $g_{\text{complete}}^n$  be the complete graph with  $n$  nodes



## Theorem

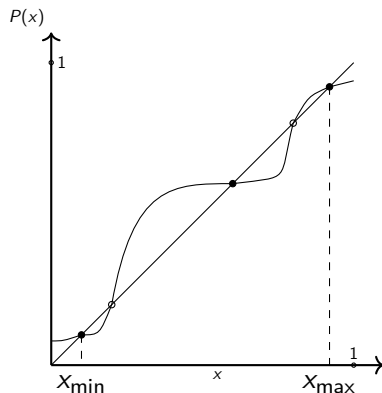
If  $x$  is a stable fixed point of  $P$ , then, for each  $\eta > 0$ ,

$$\lim_{n \rightarrow \infty} \text{Prob} \left( \{x\} \subseteq_{\eta} \text{Eq} \left( g_{\text{complete}}^n, \varepsilon \right) \right) \geq 1 - \eta.$$

- very simple proof,



# Theorem 1



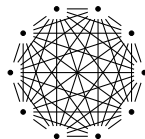
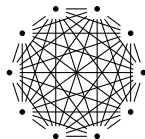
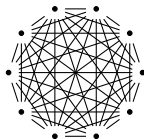
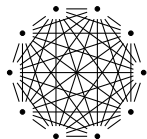
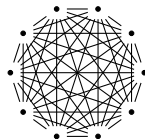
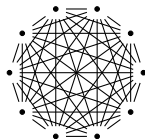
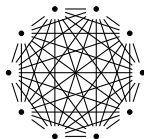
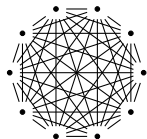
- Generically,  $x_{\min}$  and  $x_{\max}$  - the smallest and the largest fixed points - are stable.

# Theorem 1

## Corollary

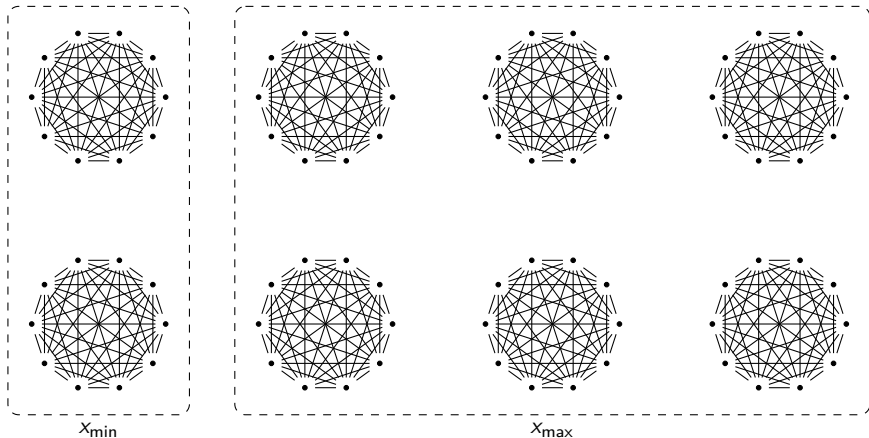
There exists a sequence of graphs  $g^n$  such that

$$\lim_{n \rightarrow \infty} \text{Prob}([x_{\min}, x_{\max}] \subseteq_{\eta} \text{Eq}(g^n, \varepsilon)) \geq 1 - \eta.$$



# Theorem 1

- Here,  $x = \frac{2}{8}x_{\min} + \frac{6}{8}x_{\max}$ .



## Theorem 2

- So far, we showed existence of networks  $g$  such that with a large probability,

$$[x_{\min}, x_{\max}] \subseteq_{\eta} \text{Eq}(g, \varepsilon).$$

- Next, we show that, for any  $g$  st.  $d(g) = \max_{i,j} \frac{g_{ij}}{g_i}$  is sufficiently small,

$$\text{Eq}(g, \varepsilon) \subseteq_{\eta} [x_{\min}, x_{\max}].$$

# Theorem 2

## Theorem

For any  $w_{\max} < \infty$ , any sequence of graphs  $g_n$ , if  $d(g_n) \rightarrow 0$  and  $w(g_n) \leq w_{\max}$ , then

$$\lim_{n \rightarrow \infty} \text{Prob}(Eq(g^n, \varepsilon) \subseteq_{\eta} [x_{\min}, x_{\max}]) = 1.$$

# Theorem 2

- Proof: surprisingly complicated.
- W.l.o.g., we want to show that, with a large probability, there is no profile  $a$  st  $Av(a) > x_{\max}$  and  $a$  is an equilibrium.

## Theorem 2

- Proof: surprisingly complicated.
- W.l.o.g., we want to show that, with a large probability, there is no profile  $a$  st  $Av(a) > x_{\max}$  and  $a$  is an equilibrium.
- Bound

$$\begin{aligned} & \text{Prob}(\text{there exists } a \text{ st. } Av(a) \geq x \text{ and } a \text{ is equilibrium}) \\ & \leq \# \{a : Av(a) > x\} \cdot \text{Prob}(a \text{ is equilibrium}). \end{aligned}$$

## Theorem 2

- Proof: surprisingly complicated.
- W.l.o.g., we want to show that, with a large probability, there is no profile  $a$  st  $Av(a) > x_{\max}$  and  $a$  is an equilibrium.
- It is easy to show that  $a$  is unlikely to be an equilibrium: there exists  $\delta > 0$  st. for each  $a$ ,

$$\text{Prob}(a \text{ is equilibrium}) \leq \exp(-\delta N).$$

- But, there are many profiles  $a$ :

$$\#\{a : Av(a) > x\} \sim \exp((x \log x + (1-x) \log(1-x)) N).$$



## Theorem 2

- Proof: surprisingly complicated.
- W.l.o.g., we want to show that, with a large probability, there is no profile  $a$  st  $Av(a) > x_{\max}$  and  $a$  is an equilibrium.
- Problem: there are too many candidate profiles  $a$ .
- Observation I: the above proof treats events “ $a$  is equilibrium” for all  $a$ s as disjoint, whereas they are often correlated.
- Observation II: events “ $a$  is equilibrium” and “ $a'$  is equilibrium” are correlated more if  $\beta^a$  and  $\beta^{a'}$  are similar.
  - $\beta_i^a = \frac{1}{g_i} \sum_j g_{ij} a_j$ .
- Idea: divide all profiles  $a$  into “groups” with similar  $\beta^a$ .

## Theorem 2

- Proof: surprisingly complicated.
- W.l.o.g., we want to show that, with a large probability, there is no profile  $a$  st  $Av(a) > x_{\max}$  and  $a$  is an equilibrium.
- Problem: there are too many candidate profiles  $a$ .
- Observation I: the above proof treats events “ $a$  is equilibrium” for all  $a$ s as disjoint, whereas they are often correlated.
- Observation II: events “ $a$  is equilibrium” and “ $a'$  is equilibrium” are correlated more if  $\beta^a$  and  $\beta^{a'}$  are similar.
  - $\beta_i^a = \frac{1}{g_i} \sum_j g_{ij} a_j$ .
- Idea: divide all profiles  $a$  into “groups” with similar  $\beta^a$ .

## Theorem 2

- Proof: surprisingly complicated.
- W.l.o.g., we want to show that, with a large probability, there is no profile  $a$  st  $Av(a) > x_{\max}$  and  $a$  is an equilibrium.
- Problem: there are too many candidate profiles  $a$ .
- Observation I: the above proof treats events “ $a$  is equilibrium” for all  $a$ s as disjoint, whereas they are often correlated.
- Observation II: events “ $a$  is equilibrium” and “ $a'$  is equilibrium” are correlated more if  $\beta^a$  and  $\beta^{a'}$  are similar.
  - $\beta_i^a = \frac{1}{g_i} \sum_j g_{ij} a_j$ .
- Idea: divide all profiles  $a$  into “groups” with similar  $\beta^a$ .

# Theorem 2

- Proof: surprisingly complicated.
- W.l.o.g., we want to show that, with a large probability, there is no profile  $a$  st  $Av(a) > x_{\max}$  and  $a$  is an equilibrium.
- Problem: there are too many candidate profiles  $a$ .
- Observation I: the above proof treats events “ $a$  is equilibrium” for all  $a$ s as disjoint, whereas they are often correlated.
- Observation II: events “ $a$  is equilibrium” and “ $a'$  is equilibrium” are correlated more if  $\beta^a$  and  $\beta^{a'}$  are similar.
  - $\beta_i^a = \frac{1}{g_i} \sum_j g_{ij} a_j$ .
- Idea: divide all profiles  $a$  into “groups” with similar  $\beta^a$ .

- *Group division*: use “Euclidean” metric

$$d(\beta_i^a, \beta_i^{a_0}) = \sqrt{\frac{1}{\sum g_i^2} \sum g_i^2 (\beta_i^a - \beta_i^{a_0})^2}.$$

- *Uniform prob bound within group*: if  $\text{Av}(a_0) > x$ ,

$\text{Prob}(\{a : d(\beta^a, \beta^{a_0}) \leq d(g)\} \text{ contains an equilibrium}) \leq \exp(-cd(g))$

## Theorem 2

- *Number of groups*:  $\mathcal{N}(\delta)$  is the smallest  $n$  such that there exists  $a_1, \dots, a_n \in \mathcal{B}$  st. for each  $a$ , there is  $m$  st.  $d(\beta^a, \beta^{a_m}) \leq \delta$  (metric entropy):

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

- R. Vershynin “*High-dimensional probability*” (2018)

## Theorem 2

$$\begin{aligned} & \text{Prob}(\{a : d(\beta^a, \beta) \leq \delta\} \text{ contains an equilibrium}) \\ & \leq \mathcal{N}(\mathcal{B}, \delta) \cdot \sup_{a_0: \text{Av}(a_0) > x} \text{Prob}(\{a : d(\beta^a, \beta^{a_0}) \leq \delta\} \text{ contains an equilibrium}). \\ & \leq \exp\left(c \frac{1}{\delta^2} d(g) N - \delta N\right), \end{aligned}$$

which is small if  $d(g)$  is small enough.

# Theorem 3

## Random utility dominant outcome

- So far, we characterized a tight upper bound on the equilibrium set.
- Next, we turn to a lower bound.



# Theorem 3

## Random utility dominant outcome

- Define *random utility (RU-) dominant* outcome

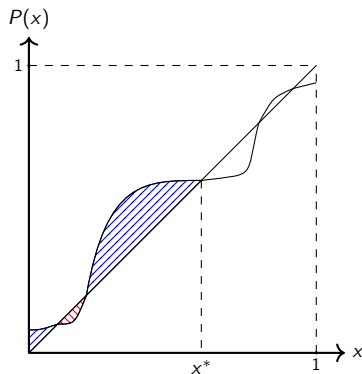
$$x^* \in \arg \max_x \int_0^x (y - P^{-1}(y)) dy.$$

# Theorem 3

## Random utility dominant outcome

- Define *random utility (RU-) dominant* outcome

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



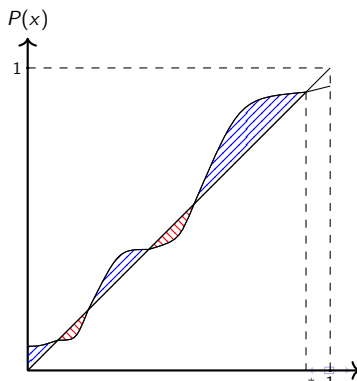
# Theorem 3

## Random utility dominant outcome

- Define *random utility (RU-) dominant* outcome

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

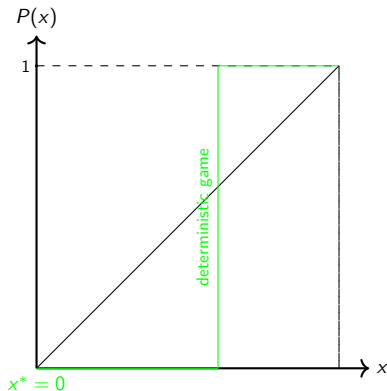
- RU-outcome can be  $x_{\min}$  or  $x_{\max}$ .



# Theorem 3

## Random utility dominant outcome

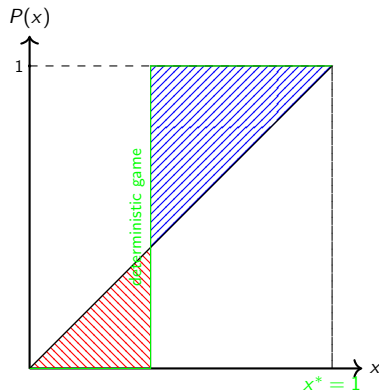
- When game is deterministic, RU-dominance is equivalent to Harsanyi-Selten risk-dominance



# Theorem 3

## Random utility dominant outcome

- When game is deterministic, RU-dominance is equivalent to Harsanyi-Selten risk-dominance



# Theorem 3

## Random utility dominant outcome

- Formula

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

appears in Morris and Shin (06).

- continuum toy model,
- observe that the coordination game has a potential,
- the above outcome maximizes potential,
- hence it is robust to incomplete information.

# Theorem 3

## Random utility dominant selection

### Theorem

Assume  $0 < P(0) < P(1) < 1$ . There exists a sequence of networks  $g^n$  st. for each  $\eta > 0$ ,

$$\lim_{n \rightarrow \infty} \text{Prob}(Eq(g^n, \varepsilon) =_{\eta} \{x^*\}) \geq 1 - \eta.$$

- For some networks,  $x^*$  is the unique average equilibrium behavior.
- The assumption ensures that, for each action, there is a positive probability that the action is dominant.

# Theorem 3

## Proof

- Networks: 2-dimensional lattices
  - line (1-dimensional lattice) is not enough
- Static result, but proof based on best response dynamics.
  - review of contagion arguments (Ellison 93, Blume 93, Morris 00),
  - contagion wave on “toy” line,
  - why line is not enough and why 2-dimensional lattice is.



# Theorem 3

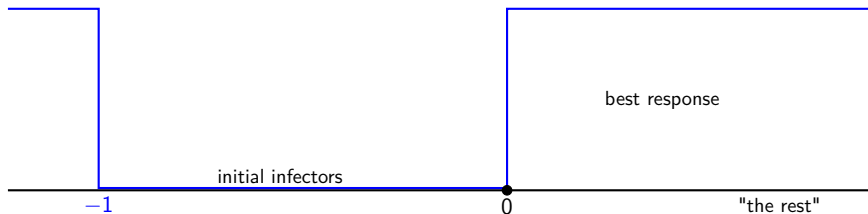
Proof: Review of contagion arguments

- Start with deterministic case, but with small group of initial infectors.
- Assume 0 is risk-dominant.
- We want to show that 0 is the only equilibrium.
- $\rightarrow$  contagion.

# Theorem 3

Proof: Review of contagion arguments

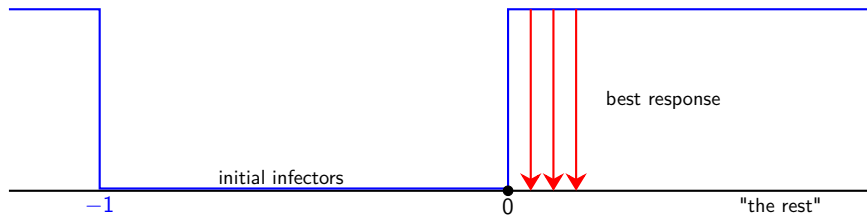
- Ellison 93: suppose that action 0 is risk-dominant,
- initial infectors  $-1 \leq i \leq 0$  play 0; the rests play 1,



# Theorem 3

Proof: Review of contagion arguments

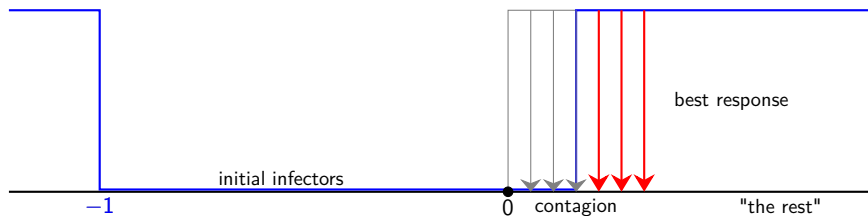
- Ellison 93: suppose that action 0 is risk-dominant,
- initial infectors  $-1 \leq i \leq 0$  play 0; the rests play 1,
- best response dynamics



# Theorem 3

Proof: Review of contagion arguments

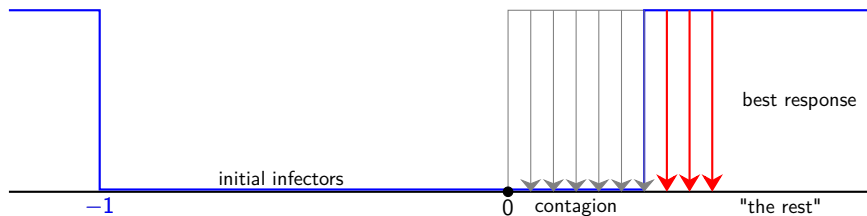
- Ellison 93: suppose that action 0 is risk-dominant,
- initial infectors  $-1 \leq i \leq 0$  play 0; the rests play 1,
- best response dynamics  $\rightarrow$  contagion



# Theorem 3

Proof: Review of contagion arguments

- Ellison 93: suppose that action 0 is risk-dominant,
- initial infectors  $-1 \leq i \leq 0$  play 0; the rests play 1,
- best response dynamics  $\rightarrow$  contagion



# Theorem 3

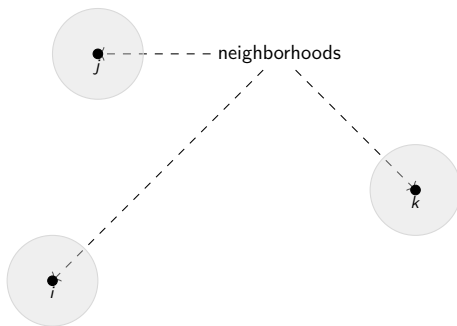
Proof: Review of contagion arguments

- Blume 93, Morris 00 - the same mechanics works on other networks, like 2 (or higher)-dimensional lattices.
- Key step: half of neighbors of “threshold agents” must be infected to spread contagion.

# Theorem 3

Proof: Review of contagion arguments

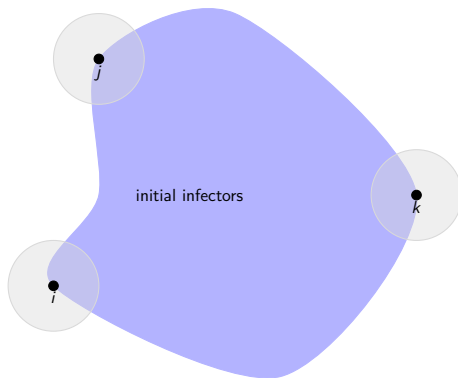
- Blume 93, Morris 00 - the same mechanics works on other networks, like 2 (or higher)-dimensional lattices.
- Key step: half of neighbors of “threshold agents” must be infected to spread contagion.



# Theorem 3

Proof: Review of contagion arguments

- Blume 93, Morris 00 - the same mechanics works on other networks, like 2 (or higher)-dimensional lattices.
- Key step: half of neighbors of “threshold agents” must be infected to spread contagion.

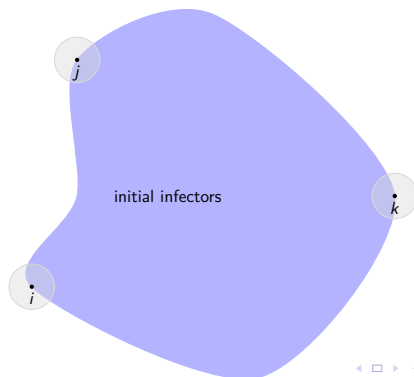




# Theorem 3

## Proof: Review of contagion arguments

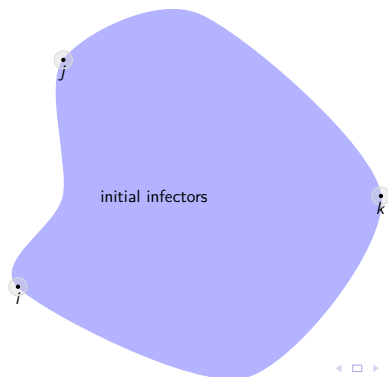
- Blume 93, Morris 00 - the same mechanics works on other networks, like 2 (or higher)-dimensional lattices.
- Key step: half of neighbors of “threshold agents” must be infected to spread contagion
- - > initial infectors must be large enough relative to neighborhoods.



# Theorem 3

## Proof: Review of contagion arguments

- Blume 93, Morris 00 - the same mechanics works on other networks, like 2 (or higher)-dimensional lattices.
- Key step: half of neighbors of “threshold agents” must be infected to spread contagion
- -  $\rightarrow$  initial infectors must be large enough relative to neighborhoods.



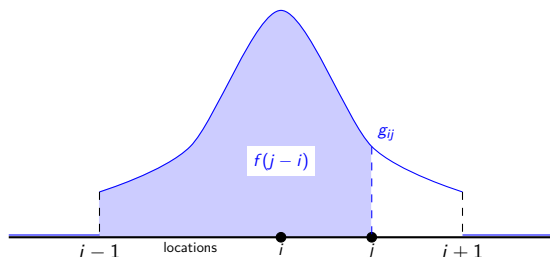
# Theorem 3

Proof: Contagion wave on toy line

- Random utility payoffs (so, not deterministic)
- Toy line: Continuum of agents in each location.

# Theorem 3

Proof: Contagion wave on line, RU case



- Toy line: agents in location  $i$  are connected with agents in location  $j$ 
  - connection density  $g_{ij} = g_{ji} = g_{i+l,j+l}$  for any  $l$ ,
  - $g_{ij} = 0$  for  $j > i + 1$ ,
  - $f(j-i) = \frac{1}{g_i} \int_{i-1}^j g_{il} dl$ ,
  - $f(x) + f(1-x) = 1$ .

# Theorem 3

Proof: Contagion wave on line, RU case

- For simplicity, assume that  $x^* = 0$  is *RU*-dominant, i.e.

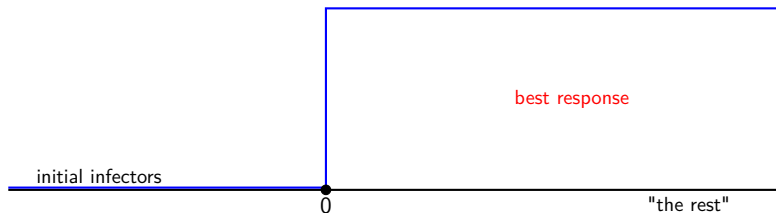
$$\int_0^x (y - P^{-1}(y)) dy < 0 \text{ for each } x > 0.$$

- Starting from arbitrary profile with a group of initial infectors playing  $x^*$ , best response dynamics will spread  $x^*$  to the whole line.

# Theorem 3

Proof: Contagion wave on line, RU case

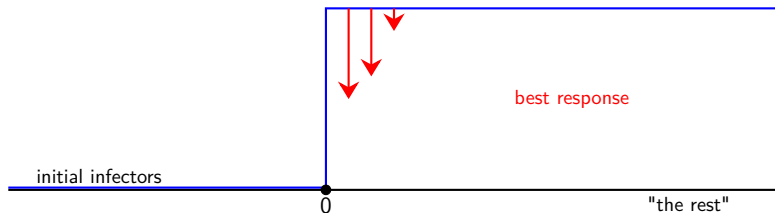
- Initial infectors play  $x^* = 0$ .



# Theorem 3

Proof: Contagion wave on line, RU case

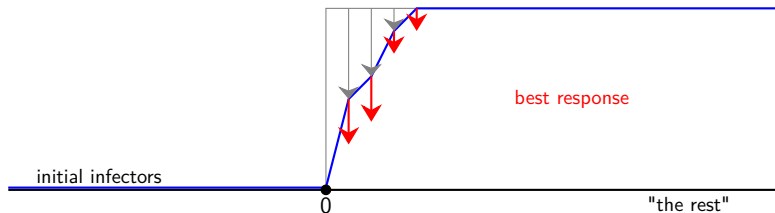
- Initial infectors play  $x^* = 0$ .



# Theorem 3

Proof: Contagion wave on line, RU case

- Initial infectors play  $x^* = 0$ .

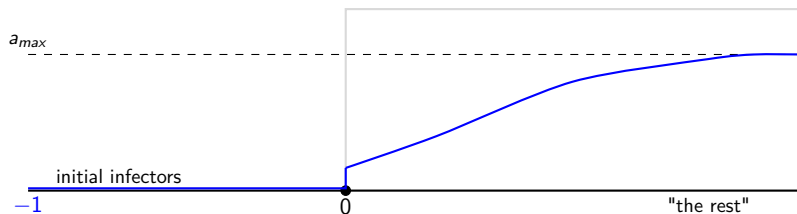




# Theorem 3

Proof: Contagion wave on line, RU case

- Suppose that stops before spreading everywhere.



# Theorem 3

Proof: Contagion wave on line, RU case

- If the contagion stops, then at each location  $i > 0$ ,

$$a_i \leq P \left( \int a_{i+k} df(k) \right).$$

- Taking inverse and integrating by parts

$$P^{-1}(a_i) \leq \int a_{i+k} df(k) = \int_0^{a_{\max}} f(i-j) da_j.$$

- Integrate over  $a_i \in [0, a_{\max}]$ ,

$$\int_0^{a_{\max}} P^{-1}(a_i) da_i \leq \int_0^{a_{\max}} \int_0^{a_{\max}} f(i-j) da_j da_i.$$

# Theorem 3

Proof: Contagion wave on line, RU case

- If the contagion stops, then at each location  $i > 0$ ,

$$a_i \leq P \left( \int a_{i+k} df(k) \right).$$

- Taking inverse and integrating by parts

$$P^{-1}(a_i) \leq \int a_{i+k} df(k) = \int_0^{a_{\max}} f(i-j) da_j.$$

- Integrate over  $a_i \in [0, a_{\max}]$ ,

$$\int_0^{a_{\max}} P^{-1}(a_i) da_i \leq \int_0^{a_{\max}} \int_0^{a_{\max}} f(i-j) da_j da_i.$$

# Theorem 3

Proof: Contagion wave on line, RU case

- If the contagion stops, then at each location  $i > 0$ ,

$$a_i \leq P \left( \int a_{i+k} df(k) \right).$$

- Taking inverse and integrating by parts

$$P^{-1}(a_i) \leq \int a_{i+k} df(k) = \int_0^{a_{\max}} f(i-j) da_j.$$

- Integrate over  $a_i \in [0, a_{\max}]$ ,

$$\int_0^{a_{\max}} P^{-1}(a_i) da_i \leq \int_0^{a_{\max}} \int_0^{a_{\max}} f(i-j) da_j da_i.$$

# Theorem 3

Proof: Contagion wave on line, RU case

- Integrate over  $a_i \in [0, a_{\max}]$ ,

$$\begin{aligned} & \int_0^{a_{\max}} P^{-1}(a_i) da_i \\ & \leq \int_0^{a_{\max}} \int_0^{a_{\max}} f(i-j) da_j da_i \end{aligned}$$

# Theorem 3

Proof: Contagion wave on line, RU case

- Integrate over  $a_i \in [0, a_{\max}]$ ,

$$\begin{aligned} & \int_0^{a_{\max}} P^{-1}(a_i) da_i \\ & \leq \int_0^{a_{\max}} \int_0^{a_{\max}} f(i-j) da_j da_i \\ & = \frac{1}{2} \int_0^{a_{\max}} \int_0^{a_{\max}} f(i-j) da_j da_i + \frac{1}{2} \int_0^{a_{\max}} \int_0^{a_{\max}} f(j-i) da_j da_i \end{aligned}$$

# Theorem 3

Proof: Contagion wave on line, RU case

- Integrate over  $a_i \in [0, a_{\max}]$ ,

$$\begin{aligned} & \int_0^{a_{\max}} P^{-1}(a_i) da_i \\ & \leq \int_0^{a_{\max}} \int_0^{a_{\max}} f(i-j) da_j da_i \\ & = \frac{1}{2} \int_0^{a_{\max}} \int_0^{a_{\max}} f(i-j) da_j da_i + \frac{1}{2} \int_0^{a_{\max}} \int_0^{a_{\max}} f(j-i) da_j da_i \\ & = \frac{1}{2} \int_0^{a_{\max}} \int_0^{a_{\max}} [f(i-j) + f(j-i)] da_j da_i \end{aligned}$$

# Theorem 3

Proof: Contagion wave on line, RU case

- Integrate over  $a_i \in [0, a_{\max}]$ ,

$$\begin{aligned} & \int_0^{a_{\max}} P^{-1}(a_i) da_i \\ & \leq \int_0^{a_{\max}} \int_0^{a_{\max}} f(i-j) da_j da_i \\ & = \frac{1}{2} \int_0^{a_{\max}} \int_0^{a_{\max}} f(i-j) da_j da_i + \frac{1}{2} \int_0^{a_{\max}} \int_0^{a_{\max}} f(j-i) da_j da_i \\ & = \frac{1}{2} \int_0^{a_{\max}} \int_0^{a_{\max}} [f(i-j) + f(j-i)] da_j da_i \end{aligned}$$

- Recall that  $f(i-j) + f(j-i) = 1$ .



# Theorem 3

Proof: Contagion wave on line, RU case

- Integrate over  $a_i \in [0, a_{\max}]$ ,

$$\begin{aligned} & \int_0^{a_{\max}} P^{-1}(a_i) da_i \leq \int_0^{a_{\max}} \int_0^{a_{\max}} f(i-j) da_j da_i \\ &= \frac{1}{2} \int_0^{a_{\max}} \int_0^{a_{\max}} f(i-j) da_j da_i + \frac{1}{2} \int_0^{a_{\max}} \int_0^{a_{\max}} f(j-i) da_j da_i \\ &= \frac{1}{2} \int_0^{a_{\max}} \int_0^{a_{\max}} [f(i-j) + f(j-i)] da_j da_i \\ &= \frac{1}{2} \int_0^{a_{\max}} \int_0^{a_{\max}} da_j da_i = \int_0^{a_{\max}} a da. \end{aligned}$$

- We get contradiction with  $\int_0^{a_{\max}} (y - P^{-1}(y)) dy < 0$ .

# Theorem 3

Proof: Contagion wave on line, RU case

- Integrate over  $a_i \in [0, a_{\max}]$ ,

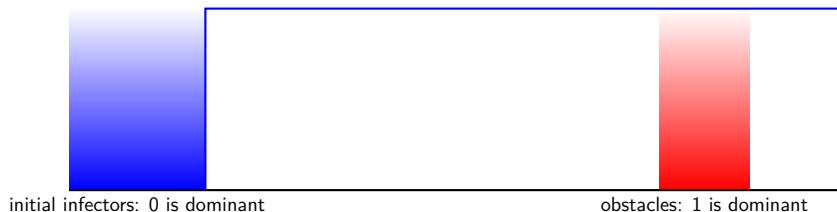
$$\begin{aligned} & \int_0^{a_{\max}} P^{-1}(a_i) da_i \leq \int_0^{a_{\max}} \int_0^{a_{\max}} f(i-j) da_j da_i \\ &= \frac{1}{2} \int_0^{a_{\max}} \int_0^{a_{\max}} f(i-j) da_j da_i + \frac{1}{2} \int_0^{a_{\max}} \int_0^{a_{\max}} f(j-i) da_j da_i \\ &= \frac{1}{2} \int_0^{a_{\max}} \int_0^{a_{\max}} [f(i-j) + f(j-i)] da_j da_i \\ &= \frac{1}{2} \int_0^{a_{\max}} \int_0^{a_{\max}} da_j da_i = \int_0^{a_{\max}} a da. \end{aligned}$$

- We get contradiction with  $\int_0^{a_{\max}} (y - P^{-1}(y)) dy < 0$ .

# Theorem 3

Proof: Contagion wave on line, RU case

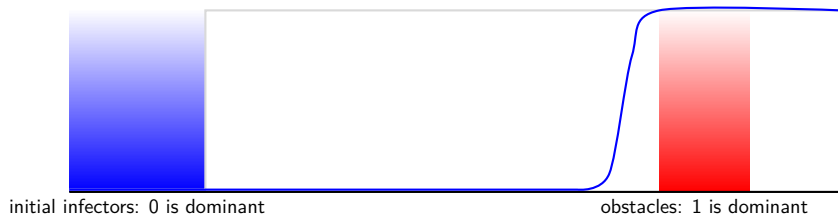
- Hence the contagion has to spread to the entire line.
- But! - so far we assumed that locations contain continuum.
- Contagion can be also stopped by unusual payoff shocks, like those that make 1 dominant.



# Theorem 3

Proof: Contagion wave on line, RU case

- Hence the contagion has to spread to the entire line.
- But! - so far we assumed that locations contain continuum.
- Contagion can be also stopped by unusual payoff shocks, like those that make 1 dominant.



# Theorem 3

Proof: Contagion wave on line, RU case

- We can compare the relative likelihood of infectors vs obstacles.
- On line, the latter can be more frequent.
- But not on 2-dimensional lattices.

# Theorem 3

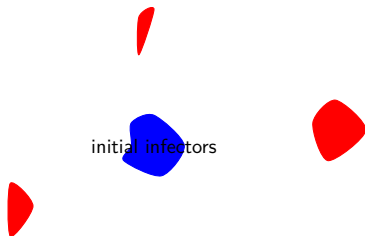
Proof: Contagion wave on line, RU case

- We can compare the relative likelihood of infectors vs obstacles.
- On line, the latter can be more frequent.
- But not on 2-dimensional lattices.

# Theorem 3

Proof: Contagion wave on line, RU case

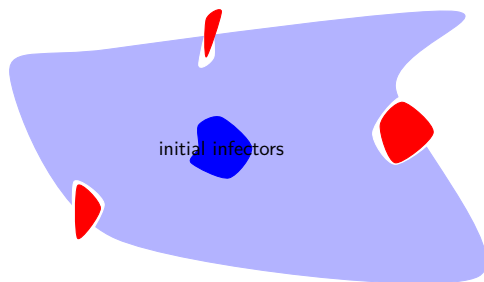
- We can compare the relative likelihood of infectors vs obstacles.
- On line, the latter can be more frequent.
- But not on 2-dimensional lattices.



# Theorem 3

Proof: Contagion wave on line, RU case

- We can compare the relative likelihood of infectors vs obstacles.
- On line, the latter can be more frequent.
- But not on 2-dimensional lattices.





# Theorem 4

- So far, we showed that there are networks  $g$  such that  $\text{Eq}(g, \varepsilon) \subseteq_{\eta} \{x^*\}$  with a large probability.
- Next, we show that if  $d(g) = \max_{i,j} \frac{g_{ij}}{g_i}$  is sufficiently small, then  $\{x^*\} \subseteq_n \text{Eq}(g, \varepsilon)$ .

# Theorem 4

## Theorem

*For any sequence of graphs  $g_n$ , if  $d(g_n) \rightarrow 0$ , then*

$$\lim_n \text{Prob}(\{x^*\} \subseteq_{\eta} \text{Eq}(g_n)) = 1.$$

# Theorem 4

- Hence  $\{x^*\}$  is the smallest equilibrium set.
- Equilibrium selection theory: no matter what network, there is an equilibrium with aggregate behavior,
  - the proof tries to make this idea more precise.
- Analog of a result from Morris “Contagion”: if all but finitely many agents play risk-dominant action, the best response dynamics won’t move towards risk-dominated action.

# Theorem 4

Proof: Morris “Contagion”

- Morris: “Contagion”:
- Initial profile  $a^0$ : all but finitely many play risk-dominant action 0
- Consider a best response dynamics  $a^0 < a^1 < a^2 < \dots$ 
  - each “round” only one agent changes action
- For each profile  $a$ , define *capacity to infect*:

$$\mathcal{F}_0(a) = \sum_{i,j:a_i=1,a_j=0} g_{ij}.$$

# Theorem 4

Proof: Morris “Contagion”

- Capacity must go down “significantly” at each round:
  - if  $i$  changes action from 0 to 1 as a best response, the capacity changes by

$$\sum_{j:a_j=0} g_{ij} - \sum_{j:a_j=1} g_{ij}.$$

- but, if 1 is a best response and 0 is risk-dominant, then

$$\sum_{j:a_j=0} g_{ij} < \frac{1}{2} < \sum_{j:a_j=1} g_{ij}!$$

- Because the capacity cannot be negative, contagion has to stop.
- If the initial profile was close to 0, the capacity was small and the contagion will stop very soon, with most agents not changing their actions.

# Theorem 4

Proof: Morris “Contagion”

- Key feature of a good definition of capacity
  - it decreases along best response dynamics,
  - it is small,
  - cannot be negative.
- The number of stages until the dynamics stops is related to the initial capacity.

# Theorem 4

Proof: RU case

- Our proof follows a similar idea.
- Let  $x^*$  be  $RU$ -dominant outcome.
- Construct *initial profile*  $a^0$  st. for each  $i$ ,

$$a_i^0 \in \arg \max_a u_i(a, x^*, \varepsilon_i)$$

- many people play 0 and many play 1
- Consider best response dynamics  $a^0 < a^1 < \dots < a^T$ .
- We show that  $\frac{T}{N} \sim O(d(g))$ .
- Hence  $a_i^0 \in \arg \max_a u_i(a, x^*, \varepsilon_i)$  is a pretty safe action to take, whatever is the true network.

# Theorem 4

Proof: RU case

- Definition of capacity: original

$$\sum_{i,j:a_i=1,a_j=0} g_{ij}.$$

- The problem is that is is too big for the initial profile  $a^0$  .



# Theorem 4

Proof: RU case

- Definition of capacity: notice that

$$\sum_{i,j:a_i=1,a_j=0} g_{ij} = \frac{1}{2} \sum_{i,j} g_{ij} (a_i - a_j)^2.$$

# Theorem 4

Proof: RU case

- Definition of capacity:

$$\frac{1}{2} \sum_{i,j} g_{ij} (a_i - a_j)^2.$$

# Theorem 4

Proof: RU case

- Definition of capacity:

$$\mathcal{F}(a) = \frac{1}{2} \sum_{i,j} g_{ij} \left( P(\beta_i^a) - P(\beta_j^a) \right)^2.$$

- replace  $a_i$  by the “continuum best response” to the neighborhood profile  $\beta^a$ .

# Theorem 4

Proof: RU case

- Definition of capacity:

$$\mathcal{F}(a) = \frac{1}{2} \sum_{i,j} g_{ij} \left( P(\beta_i^a) - P(\beta_j^a) \right)^2.$$

- Capacity is small at the initial profile  $a^0$  (probabilistically),
  - $\beta_i^{a^0} \sim x^*$  for most  $i$  because  $x^* = P(x^*)$  and  $d(g) \sim 0$ , and
  - $P(\beta_i^{a^0}) \sim P(\beta_j^{a^0})$  for most  $i$  and  $j$ .

# Theorem 4

Proof: RU case

- Turns out that this is a good definition
  - capacity is small (probabilistically)
  - it is a sum of a martingale and a decreasing process,
  - ignoring (probabilistically) small terms, we show that, for each  $T$

$$\sum_i g_i \left[ \begin{array}{c} P(\beta_i^{a^T}) \\ \int_{x^*} (P^{-1}(y) - y) dy \end{array} \right] \leq \mathcal{F}(a^0)$$

- Potential game
- Evolutionary literature
- Small degree

- Binary coordination games have potential

$$\begin{aligned}V(\mathbf{a}; \boldsymbol{\varepsilon}) &= \frac{1}{2} \sum_{i,j} g_{ij} a_i a_j - \sum_i \epsilon_i a_i \\ &= \mathbf{a}^T \mathbf{G} \mathbf{a} - \mathbf{a}^T \boldsymbol{\varepsilon}.\end{aligned}$$

- Potential maximizers are
  - equilibria,
  - selected by evolutionary logistic dynamics (Blume)
  - robust to incomplete information.

- Binary coordination games have potential

$$\begin{aligned}V(a; \varepsilon) &= \frac{1}{2} \sum_{i,j} g_{ij} a_i a_j - \sum_i \epsilon_i a_i \\ &= a^T G a - a^T \varepsilon.\end{aligned}$$

- One way to prove Theorem 4 is to show that  $V$  always has a local (or global) maximum close  $a^0$ 
  - quadratic binary form with a random linear term,
  - maximize not minimize.
- But how?



# Conclusion

- Heterogeneous payoffs in coordination games on network.
- We characterized the largest and the smallest possible set of equilibrium average behaviors across all networks.
- Results:
  - The largest set achieved on a collection of complete graphs,
  - partial identification theory,
  - The smallest set achieved on 2-dimensional (but not necessarily 1-dimensional) lattice,
  - equilibrium selection theory.
- Main assumptions:
  - independent payoff shocks,
  - large degree,
  - both assumptions are important.