Riot Networks

Javier Barreda Lachlan Deer Chih-Sheng Hsieh
Michael D. König Gizem Korkmaz Fernando Vega-Redondo

June 12, 2019

Motivation

The emergence of collective action (“riots”), an important phenomenon

- Our approach emphasizes the role of social networks:
  - They underlie payoffs, by peer interaction
  - They channel crucial information, e.g. on the extent of support
  - They are endogenous, i.e. co-evolve with actions
- Our model sheds (theoretical) light on the following issues:
  - How does a large local population coordinate on collective action?
  - How do expectations form and adapt along the process?
  - What is the role of individual heterogeneity (e.g. in preferences)?
- Ensuing key objective:
  - Collection of “big” Twitter data on rioting events (Arab Spring)
  - Structural estimation of our model parameters

Related literature

A wide number of related literature strands

- **Coordination games in networks**:
  - *Fixed networks*: Blume (1993), Brock & Durlauf (2001), Morris (2000)
  - *Co-evolving networks*: Jackson & Watts (2002), Goyal & V-R (2005), König *et al.* (2014), Marsili & V-R (2017)
- **Learning**:
  DeMarzo *et al.* (2003), Golub & Jackson (2010), Acemoglu *et al.* (2014)
- **Collective action & threshold behavior**:
  Granovetter (1978), Chwe (2000), Barberà & Jackson (2016)

Our model integrates above features into a single framework leading to:

(i) a closed-form characterization of equilib. paths & full comparative analysis

(ii) a likelihood to be used in structural estimation of the model

Badev (2013) has model in similar vein, but equilibrium not characterized
— hence full-fledged theoretical analysis or structural estimation not possible.

Outline

1. The benchmark model: complete (global) information
    - The game-theoretic setup
    - The law of motion: action and link adjustment
    - The potential: characterization of the invariant distribution
    - Stochastically stable states
2. The belief-based model: learning and incomplete (local) information
    - The revised law of motion
    - Stochastic stability under belief formation
    - On beliefs and beachheads
    - External manipulation of beliefs
3. Empirical application (ongoing): Arab Spring on Twitter
4. Summary and conclusions

Benchmark model: the game-theoretic setup

- **Players**: (large) population $i \in \mathcal{N} = \{1, \ldots, n\}$.
- **Network**: $G = (g_{ij})_{i,j=1}^{n}$, with
  $g_{ij} \in \{1\texttt{(i \& j connected), 0 (i \& j not connected)}\}$,
- **actions**: $\mathbf{s} = (s_1, ..., s_n)$, $s_i \in \{-1\texttt{(safe)}, +1\texttt{(risky)}\}$
- **payoffs**:
  $\pi_i(\mathbf{s}, G) = (1-\theta-\rho)\gamma_i s_i + \theta \sum_{j=1}^{n} g_{ij} s_i s_j + \rho \sum_{j=1}^{n} s_j s_i - \kappa s_i - \zeta d_i$
  where
  - **idiosyncratic** characteristic of each agent i: $\gamma_i \in \{-1, +1\}$
  - **coordination**: local/peer effects θ ; global/population effects ρ
  - **costs**: κ on risky action; ζ on linking (d_i: i's degree)

The law of motion

Time is continuous. The dynamics induces a path of states $\{\omega_t\}_{t\geq 0} \subset \Omega$ where every $\omega = (\mathbf{s}, G) \in \Omega$ specifies an action profile and a network.

The dynamics involves three components:

(1) **Action adjustment**: At every t, each agent $i \in \mathcal{N}$ is selected at a rate $\chi > 0$ to revise his current action s_{it}, in which case he chooses the new action $s_i' \in \{-1, +1\}$ as a *noisy best response* to the prevailing state. Formally, for infinitesimal Δt, we posit:

$$\mathbb{P}\left(\omega_{t+\Delta t} = (s_i', \mathbf{s}_{-it}, G_t) | \omega_t = (s_{it}, \mathbf{s}_{-it}, G_t)\right) =$$
$$\chi\, \mathbb{P}\left(\pi_i(s_i', \mathbf{s}_{-it}, G_t) - \pi_i(s_{it}, \mathbf{s}_{-it}, G_t) + \varepsilon_{it} > 0\right) \Delta t + o(\Delta t).$$

for i.i.d. shocks ε_{it}, assumed logistically distributed with parameter η.

The law of motion (cont.)

(2) **Link creation**: At every t, each pair of unconnected agents $i, j \in \mathcal{N}$ is selected at a rate $\lambda > 0$ to create a link. Then, they do so iff they both find it profitable given some noisy perceptions of the entailed payoffs. Formally, for infinitesimal $t + \Delta t$, we have:

$$\mathbb{P}\left[\omega_{t+\Delta t} = (\mathbf{s}_t, G_t + ij) | \omega_{t-1} = (\mathbf{s}, G_t)\right] =$$
$$\lambda\, \mathbb{P}\left[\{\pi_i(\mathbf{s}_t, G_t + ij) - \pi_i(\mathbf{s}_t, G_t) + \varepsilon_{ij,t} > 0\} \cap \right.$$
$$\left. \{\pi_j(\mathbf{s}_t, G_t + ij) - \pi_j(\mathbf{s}_t, G_t) + \varepsilon_{ij,t} > 0\}\right] \Delta t + o(\Delta t)$$

for i.i.d. shocks ε_{it}, assumed logistically distributed with parameter η.

The law of motion (cont.)

(3) **Link removal**: At every t, each pair of unconnected agents $i, j \in \mathcal{N}$ is selected at a rate $\xi > 0$ to remove a link. Then, they do so iff at least one of them finds it profitable, given some noisy perceptions of the entailed payoffs. Formally, for infinitesimal $t + \Delta t$, we postulate:

$$\mathbb{P}\left[\omega_{t+\Delta t} = (\mathbf{s}_t, G_t - ij) | \omega_{t-1} = (\mathbf{s}, G_t)\right] =$$
$$\lambda\, \mathbb{P}\left[\{\pi_i(\mathbf{s}_t, G_t - ij) - \pi_i(\mathbf{s}_t, G_t) + \varepsilon_{ij,t} > 0\} \cup\right.$$
$$\left.\{\pi_j(\mathbf{s}_t, G_t - ij) - \pi_j(\mathbf{s}_t, G_t) + \varepsilon_{ij,t} > 0\}\right] \Delta t + o(\Delta t)$$

for i.i.d. shocks ε_{it}, assumed logistically distributed with parameter η.

Potential and the invariant distribution

A first important result is that the game is a potential game (and hence *noiseless* best-response adjustment converges to some NE).

Proposition

The payoffs of the game admit a potential $\Phi : \Omega \to \mathbb{R}$ *defined by*

$$\Phi(\mathbf{s}, G) = (1-\theta-\rho)\sum_{i=1}^{n}\gamma_i s_i + \frac{\theta}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}s_i s_j + \frac{\rho}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} s_i s_j - \kappa\sum_{i=1}^{n} s_i - m\zeta,$$

where $m \equiv \sum_{i>j} a_{ij}$ *stands for the number of links in* G*.*

Potential and the invariant distribution (cont.)

For $\eta < \infty$, the process is ergodic. The potential function then leads to the following Gibbs-measure characterizing its (unique) invariant distribution.

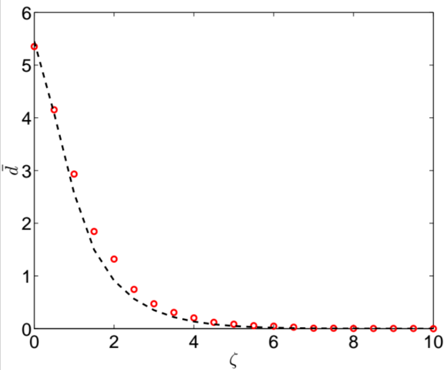
Proposition

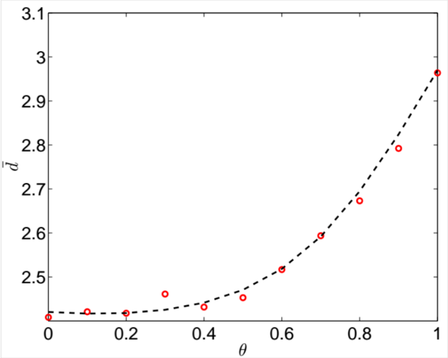
The unique stationary distribution μ^η defined on the measurable space $(\Omega, \mathcal{F})$ such that $\lim_{t\to\infty} \mathbb{P}(\omega_t = (\mathbf{s}, G)|\omega_0 = (\mathbf{s}_0, G_0)) = \mu^\eta(\mathbf{s}, G)$. The probability measure μ^η is given by

$$\mu^\eta(\mathbf{s}, G) = \frac{e^{\eta\Phi(\mathbf{s},G)}}{\sum_{G'\in\mathcal{G}^n} \sum_{\mathbf{s}'\in\{-1,+1\}^n} e^{\eta\Phi(\mathbf{s}',G')}}$$

where $\Phi(\cdot)$ is the potential.

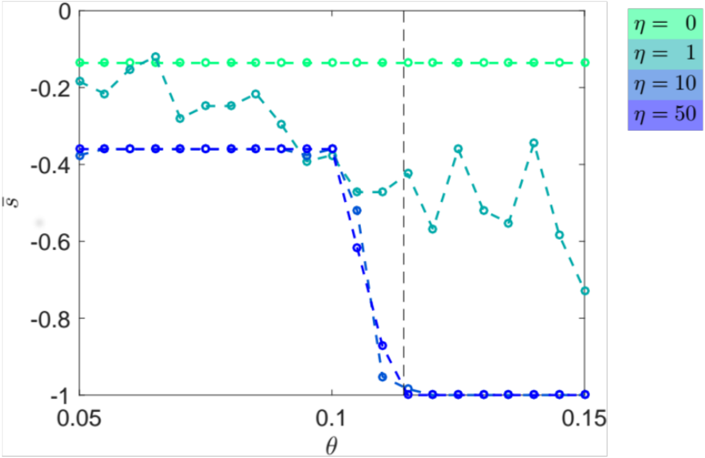
Checking the theory numerically





Actual & predicted avge degree for range of ζ and θ, $n = 10$, and equal number of $\gamma = \pm 1$

Checking the theory numerically



Segregated and Homogenous Societies

Stochastically stable states

Identify the states visited with positive prob. in the long run for **vanishing noise,** i.e. the **stochastically stable states** $\Omega^* \equiv \{\omega \in \Omega \text{ s.t. } \lim_{\eta\to\infty} \mu^\eta(\omega) > 0\}$.

Proposition

Assume $\zeta < \theta$ *and let* $n_+ \equiv \#(\{i \in \mathcal{N} : \gamma_i = +1\})$. *If*

$$\theta < \theta^* = \frac{(n - n_+)(\zeta - 2\rho) + 2(1 - \kappa - \rho)}{2 + n - n_+},$$

all states in Ω^* *involve a network with two cliques of sizes* n_+ *and* $n - n_+$*, with agents in them choosing* $s_i = \gamma_i = +1$ *and* $s_i = \gamma_i = -1$*, respectively. Instead, if* $\theta > \theta^*$*, all states in* Ω^* *involve a complete network* K_n *where all agents* $i \in \mathcal{N}$ *choose either* $s_i = +1$ *if* $n_+ > \frac{n}{2}$*, or* $s_i = -1$ *if* $n_+ < \frac{n}{2}$*.*

Observation: If pop. size n, global coordination effect are large, ρ, $\theta^* \le 0$.

The belief-based model: revised law of motion

Here, each agent i holds expectations $p_i \in [0,1]$ on the fraction of agents $j \in \mathcal{N}$ in the overall population choosing $s_j = +1$.

Then, the agent guides his behavior according to the **expected payoff**

$$\mathbb{E}_i\left(\pi_i | \mathbf{s}, \mathbf{p}, G\right) = (1-\theta-\rho)\gamma_i s_i + \theta \sum_{j=1}^{n} a_{ij} s_i s_j + \rho n p_i s_i - \kappa s_i - \zeta d_i,$$

in terms of which we posit **direct counterparts** of prior (1)-(3):

(1′) **Action adjustment**

(2′) **Link creation**

(3′) **Link removal**

Law of motion for beliefs under incomplete information

Agents revise beliefs by combining the following sources of local info.:

- the action frequencies of current partners, extrapolated globally;
- the beliefs of current partners, integrated à la DeGroot.

(4) **Belief adjustment**: At every t, each agent $i \in \mathcal{N}$ is selected at a rate $\tau > 0$ to revise his current belief p_i. In that case, he carries out a convex combination of the action frequencies and average beliefs of his current partners. Formally, for weight $\varphi \in (0,1]$, and

$$f_i(\mathbf{s}_t, \mathbf{p}_t, G_t) = \varphi \frac{1}{d_{it}} \sum_{j=1}^{n} a_{ij,t} s_{jt} + (1-\varphi) \frac{1}{d_{it}} \sum_{j=1}^{n} a_{ij,t} p_{jt}$$

Potential & invariant distribution: incomplete information

Again, we find that the game has a potential (a different one!) and hence the usual best-response adjustment converges to some NE.

Proposition

The payoffs of the game admit a potential $\tilde{\Phi} : \Omega \to \mathbb{R}$ *defined by*

$$\tilde{\Phi}i(\mathbf{s}, G) = (1-\theta-\rho)\sum_{i=1}^{n}\gamma_i s_i + \frac{\theta}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}s_i s_j + \rho p_i \sum_{i=1}^{n} p_i s_i - \kappa \sum_{i=1}^{n} s_i - m\zeta,$$

Potential and invariant distribution: incomplete info. (cont.)

Again, for $\eta < \infty$, the process is ergodic. Thus we obtain an identical characterization of its unique invariant distribution – in terms of the new potential $\tilde{\Phi}(\cdot)$, instead of the former $\Phi(\cdot)$.

Proposition

The unique stationary distribution $\tilde{\mu}^\eta$ defined on the measurable space $(\Omega, \mathcal{F})$ such that $\lim_{t\to\infty} \mathbb{P}(\omega_t = (\mathbf{s}, G) | \omega_0 = (\mathbf{s}_0, G_0)) = \tilde{\mu}^\eta(\mathbf{s}, G)$. The probability measure $\tilde{\mu}^\eta$ is given by

$$\tilde{\mu}^\eta(\mathbf{s}, G) = \frac{e^{\eta\tilde{\Phi}(\mathbf{s},G)}}{\sum_{G'\in\mathcal{G}^n}\sum_{\mathbf{s}'\in\{-1,+1\}^n} e^{\eta\tilde{\Phi}(\mathbf{s}',G')}}$$

where $\tilde{\Phi}(\cdot)$ is the potential.

Stochastically stable states: incomplete information

Identify the states visited with positive prob. in the long run for **vanishing noise,** i.e. the **stochastically stable states** $\Omega^* \equiv \{\omega \in \Omega \text{ s.t. } \lim_{\eta \to \infty} \mu^\eta(\omega) > 0\}$.

Proposition

Assume $\zeta < \theta$ *and let* $n_+ \equiv \#(\{i \in \mathcal{N} : \gamma_i = +1\})$*. If*

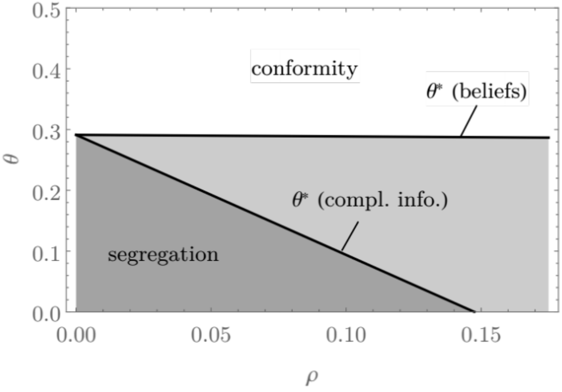
$$\theta < \tilde{\theta} = \frac{(n - n_+)\zeta + 2(1 - \kappa - \rho)}{2 + n - n_+},$$

all states in Ω^* *involve a network with two cliques of sizes* n_+ *and* $n - n_+$*, with agents in them choosing* $s_i = \gamma_i = +1$ *and* $s_i = \gamma_i = -1$*, respectively. Instead, if* $\theta > \theta^*$*, all states in* Ω^* *involve a complete network* K_n *where all agents* $i \in \mathcal{N}$ *choose either* $s_i = -1$ *if* $n_+ > \frac{n}{2}$*, or* $s_i = -1$ *if* $n_+ < \frac{n}{2}$*.*

On peer pressure, segregation, and information

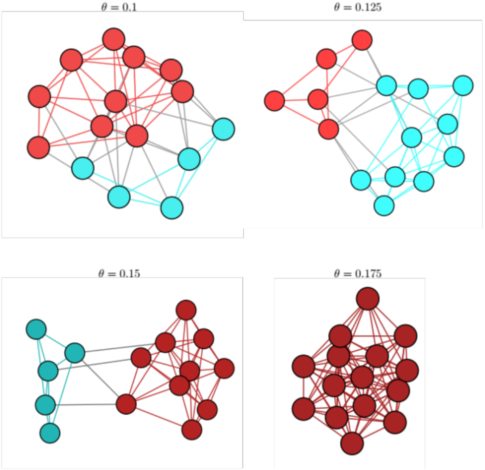
Comparing *segregation thresholds* under complete & incomplete info.:

$$\frac{(n-n_+)(\zeta-2\rho)+2(1-\kappa-\rho)}{2+n-n_+}=\theta^* < \tilde{\theta} = \frac{(n-n_+)\zeta+2(1-\kappa-\rho)}{2+n-n_+}$$



For large n and ρ, $\theta^* = 0$ while $\hat{\theta}$ almost constant

Sample Network Structures



Role of homophily on network structure

On segregation and beliefs (cont.)

Why is the comparison of thresholds θ^* and $\tilde{\theta}$ important?

- It bears on whether a relatively small group of $n_- \equiv n - n_+$ revolutionaries can gain a stable beachhead.
- Thereafter, through "drift" (or gradual change in socio-political conditions), it can take over.

Explicit modeling of beliefs also allows one to study **belief manipulation** (by, say, a government), modifying the belief formation rule as

$$f_i(\mathbf{s}_t, \mathbf{p}_t, G_t) = \underbrace{\nu g}_{propaganda} + \varphi \frac{1}{d_{it}} \sum_{j=1}^{n} a_{ij,t} s_{jt} + (1 - \varphi - \nu) \frac{1}{d_{it}} \sum_{j=1}^{n} a_{ij,t} p_{jt}$$

$\nu \in (0, 1]$ parametrizes manipulation strength, $g = -1$ is preferred action

Naturally, manipulation seen to reduce the "segregation parameter region"

Empirical application I: Social unrest in Latin America

A wide and diverse range of social unrest in Latin America (LA)

- Context: a large set of cross-continental different riot instances, approx 15K
- Over almost two years - November 1st, 2012 - August 31st, 2014.
- Our data:
  - Network/communication: Twitter, with 600M tweets
  - 10% sample of all tweets in LA for the period
  - Complementary data: list of events including location, dates, and type of protest, population involved, reported evidence on the event, etc.

Country	Events
Mexico	4,454
Venezuela	3,072
Brazil	3,051
Paraguay	1,800
Argentina	1,227
Colombia	1,069
Chile	638
El Salvador	608
Uruguay	570
Ecuador	410

Operationalization, Objectives and Examples

Operationalization of theory:

- Generate social network: (i) retweets (ii) @-mentions
- NLP analysis of Twitter data to extract message content & intrinsic features of users (gender, location, etc.) for econometric analysis

Objectives of Empirical Analysis:

- Testing the model and structural estimation of the parameters
- Comparative analysis and policy prediction

Examples:

1. Mexican Riots, 'Yo soy 132'
   - June 2012, biased media coverage, students
2. 'Brazillian Spring'
   - June 2013, increases in bus fares, general population

Structural estimation of the model: the essential idea

Our explicit (closed-form) solution of the model underlies the econometric maximum-likelihood strategy, applied either statically or dynamically:

- **Static approach**: For any observed strategy profile and network $(\mathbf{s}, G)$, we use as likelihood the invariant distribution:

$$\tilde{\mu}^{\eta}(\mathbf{s}, G; \varpi) = \frac{e^{\eta \tilde{\Phi}(\mathbf{s}, G)}}{\sum_{G' \in \mathcal{G}^n} \sum_{\mathbf{s}' \in \{-1,+1\}^n} e^{\eta \tilde{\Phi}(\mathbf{s}', G')}}$$

  where $\varpi = (\gamma, \theta, \rho, \kappa, \zeta)$ are the parameters to be used as likelihood maximizers.

- **Dynamic approach**: For any observed sequence of networks $(G_t)_{t=t_1}^{t_2}$ and action profiles $(\mathbf{s}_t)_{t=t_1}^{t_2}$ we use as likelihood:

$$\mathbb{P}((G_t, \mathbf{s}_t)_{t=t_1+1}^{t_2} | \mathbf{s}_{t_1}, G_{t_1}; \varpi) = \prod_{t=t_1+1}^{t_2} \mathbb{P}((G_t, \mathbf{s}_t)_{t=t_1+1}^{t_2} | G_{t-1}, \mathbf{s}_{t-1}; \varpi).$$

ML estimates actions-network (@ mention) Mexican "Yo soy 132"

	(I)	(II)	(III)
Local spillover (θ)	0.3123***	0.0244	0.1742***
	(0.0140)	(0.0219)	(0.0122)
Global conformity (ρ)	1.75E-05***		1.79E-05***
	(5.27E-07)		(4.36E-07)
Individual heterogeneity (γ)			
female	-0.3523***		-0.3400***
	(0.0200)		(0.0170)
capital	0.1525***		0.1700***
	(0.0097)		(0.0095)
follower count	0.0006		0.0019
	(0.0048)		(0.0049)
kloutscore	-0.0028***		-0.0029***
	(0.0007)		(0.0009)
Linking cost (ζ)			
constant		12.7970***	12.9531***
		(0.0391)	(0.0118)
$\lvert \text{female}_i - \text{female}_j \rvert$		0.0959	0.0661**
		(0.0720)	(0.0288)
$I(\text{capital}_i = \text{capital}_j)$		-0.9417***	-1.00064***
		(0.0557)	(0.0065)
$\lvert \text{follower}_i - \text{follower}_j \rvert$		0.2397***	0.2597***
		(0.0179)	(0.0043)
$\lvert \text{kloutscore}_i - \text{kloutscore}_j \rvert$		0.0242***	0.0212***
		(0.0024)	(0.0016)
Sample size		60,837	

ML estimation of actions-network (retweeting) Brazilian "bus riots"

	(I)	(II)	(III)
Local spillover (θ)	0.1850***	0.1671***	0.1826***
	(0.0089)	(0.0116)	(0.0051)
Global conformity (ρ)	-2.08E-05***		-2.01E-05***
	(3.41E-07)		(3.49E-07)
Individual heterogeneity (γ)			
female	-1.0408***		-0.8223***
	(0.0644)		(0.0408)
capital	0.4104***		0.3822***
	(0.0228)		(0.0138)
follower count	-0.2865***		-0.2939***
	(0.0111)		(0.0098)
kloutscore	0.0277***		0.0281***
	(0.0015)		(0.0012)
Linking cost (ζ)			
constant		12.7474***	12.8553***
		(0.0231)	(0.0134)
$\lvert \mathrm{female}_i - \mathrm{female}_j \rvert$		0.3688***	0.2734***
		(0.0419)	(0.0165)
$I(\mathrm{capital}_i = \mathrm{capital}_j)$		-1.8217***	-1.6673***
		(0.0608)	(0.0306)
$\lvert \mathrm{follower}_i - \mathrm{follower}_j \rvert$		0.1971***	0.1913***
		(0.0125)	(0.0095)
$\lvert \mathrm{kloutscore}_i - \mathrm{kloutscore}_j \rvert$		0.0486***	0.0417***
		(0.0023)	(0.0018)
Sample size		96,566	

Main points from the preliminary case-specific analysis:

- Both local and global effects (θ and ρ) are as predicted, positive and significant
- Importance of global effetc provides strong motivation for the belief-based formulation under incomplete information
- Heterogeneity in γ_i important, intuitive

Summary

- The rise of collective action in **large populations** is not well understood – cannot be modeled as a standard coordination game
- We propose a rich model where actions and links co-evolve, as dictated by (myopic) payoff considerations
- Motivated by the large-population context, incomplete information and locally-formed expectations alternative to benchmark setup
- The model can be fully solved analytically, which permits its structural estimation, both statically and dynamically
- This approach can be applied to a wide range of collective-action problems, for many of which extensive data are becoming available.
  - Example provided: Unrest in Latin America
  - To come: Arab Spring.