Demographic Prisonner Dilemma as a Mean Field Games

René Carmona

Department of Operations Research & Financial Engineering
Program in Applied & Computational Mathematics
Princeton University

INRIA Short Talk 3/30/17

Turning a Two-Player One-Period Finite Game into an N-Player Game

- Start with your favorite **Two-Players, One Period, Finite States**
  - **Prisonner Dilemma** (or Rock-Paper-Scissor, or ....)
  - **Payoff**

    $$F(\alpha, \alpha') = \mathbf{1}_{\alpha=C}\big(R\mathbf{1}_{\alpha'=C} - S\mathbf{1}_{\alpha'=D}\big) + \mathbf{1}_{\alpha=D}\big(T\mathbf{1}_{\alpha'=C} - P\mathbf{1}_{\alpha'=D}\big)$$

    with $T > R > 0 > -S > -P$
- ***N*-player Game**
  - Fix an integer $N \geq 2$
  - A family $(\mathbf{N}_{(i,j)})_{(i,j)\in\mathcal{I}^*}$ of independent Poisson processes $\mathbf{N}_{(i,j)} = (N_{(i,j)}(t))_{t\geq 0}$ with rates $1/[N(N-1)]$.
- **Dynamics**
  - If one of the Poisson processes, say $\mathbf{N}_{(i,j)}$ jumps at time t,
  - we let players i and j play an instance of the one stage game.
  - If they use strategies α^i_t and α^j_t, their wealths Y^i_t and Y^j_t are updated:

$$Y^i_t = Y^i_{t-} + F(\alpha^i_t, \alpha^j), \qquad \text{and} \qquad Y^j_t = Y^j_{t-} + F(\alpha^j_t, \alpha^i).$$

Axelrod (1984), **Frank** (1993,1994)

- N agents **do not** consciously **optimize** over strategic alternatives.
- they inherit a **fixed strategy** (a phenotype) at birth
- **individuals are "hard-wired" to execute a fixed strategy**
- they repeat this inherited strategy over and over and over and over .....
- For the **Prisonner Dilemma** game individuals are "hard-wired" to execute a fixed strategy **C** or **D**
- Assume that the proportion of players wired with **C** is p
- Assume the players match at random, and when they do, they play a round of Prisonner Dilemma gate.
- In such a game their expected payoffs are given by:

$$\mathbb{E}[C] = pR - (1-p)S \qquad \text{and} \qquad \mathbb{E}[D] = pT - (1-p)P$$

J. Epstein

- Still **hard-wired** individuals of types **C** and **D**
- Epstein added a **spatial component**
  - m integer, $\mathbb{T}_m^2 = (\mathbb{Z}/m\mathbb{Z})^2$
  - For $i = 1, \cdots, N$, $(\mathbf{X}^i = (X_t^i)_{t\geq 0}$ i.i.d. standard random walks on $\mathbb{T}^2$
  - Players i and j are allowed to play at time t if
    - $N_{i,j}(t)$ jumps
    - $Y_t^i > 0$ and $Y_t^j > 0$
    - $X_t^i = X_t^j$
- The state of player i evolves as (X_t^i, Y_t^i, Z^i) where $Z^i = \mathbf{C}$ or $Z^i = \mathbf{D}$ does not change with time.
- Showed in Monte Carlo simulations ***zones of cooperation*** occur
- The Demographic Prisonner Dilemma game is **not really a (dynamic) game** since the control / strategy Z^i does not change with time !

Dynamic Game Version of DPD

- Allow players to change type (**C** or **D**) dynamically as a function of Y^i_t
  - Player i changes his/her status (control) dynamically ($\phi^i_t(Y^i_t) =$ **C** or **D**)
- Speed up the spacial walks $X^i_t \mapsto X^i_{\lambda t}$ for $\lambda \nearrow \infty$
- Homogenization (**Gibaud**) using T. Kurtz limit theorems for Markov processes
  - For each feedback (Markovian) $\phi = (\phi^1, \cdots, \phi^N)$, $\mathbf{Y} = (\mathbf{Y}^{\lambda,1}, \cdots, \mathbf{Y}^{\lambda,N})$ with with $\mathbf{Y}^{\lambda,i} = (Y^{\lambda,i}_t)_{t\geq 0}$ wealth of player i converges as $\lambda \nearrow \infty$
  - Essentially, homogenization in limit $\lambda \nearrow \infty$ brings the physical positions $(X^{\lambda,1}_t, \cdots, X^{\lambda,N}_t)$ to be picked according to their invariant measure

DYNAMICS OF THE DEMOGRAPHIC GAME

- If one of the Poisson processes, say $\mathbf{N}_{(i,j)}$ jumps at time t,
- one checks that the players i and j are still in the game,
- that their physical states X_t^i and X_t^j are neighbors,
- if so, they play an instance of the one stage game.
- If they use strategies α_t^i and α_t^j, their respective wealths are updated in the following way:

$$Y_t^i = Y_{t-}^i + F(\alpha_t^i, \alpha^j), \qquad \text{and} \qquad Y_t^j = Y_{t-}^j + F(\alpha_t^j, \alpha^i).$$

Credits & Starting Point

- Repeated **Prisonner Dilemma** Game exhibits same repeated one stage Nash equilibrium
- Infinite horizon, depending upon discount factor, *zones of cooperation* occur

J. Epstein

- Added a spatial component
- Showed in Monte Carlo simulations *zones of cooperation* occur

Gibaud revisited Epstein's **DPD** model

- proved
  - homogenization of the spatial component (T. Kurtz limit theorems for Markov processes)
  - propagation of chaos of the homogenized model (A.S. Sznitmann)

Dynamic DPD as an N-Player Game with Mean Field Interactions

$$Y_t = Y_0 + \int_{[0,t]} \sum_{(i,j)\in\mathcal{I}^*} \varphi_{(i,j)}(Y_{s-}, \phi(Y_{s-})N_{(i,j)}(ds)$$

with $\varphi_{(i,j)} = (\varphi^1_{(i,j)}, \cdots, \varphi^N_{(i,j)})$ are $\mathbb{R}^N$-valued

- If k fixed, $\varphi^k_{(i,j)} = 0$ unless $k = i$ or $k = j$
- **Goal**: rewrite the dynamics of Y^k_t to emphasize the role of Y^j_t with $j \neq k$

Fix k & rewrite the Poisson measures $\mathbf{N}_{ik}$ and $\mathbf{N}_{kj}$ as:

$$N_{ik}(dt) = \int_{[0,1]} \mathbf{1}_{[\frac{\sigma_t(i)-1}{N-1} < w \leq \frac{\sigma_t(i)}{N-1}]} \tilde{N}^1(dt, dw)$$

and

$$N_{kj}(dt) = \int_{[0,1]} \mathbf{1}_{[\frac{\sigma_t(j)-1}{N-1} < w \leq \frac{\sigma_t(j)}{N-1}]} \tilde{N}^2(dt, dw)$$

where

- $\tilde{\mathbf{N}}^1$ and $\tilde{\mathbf{N}}^2$ are independent Poisson random measures on $[0, \infty) \times [0, 1]$ with intensity $\frac{1}{2} Leb_2$
- $(\sigma_t)_{t\geq 0}$ is a predictable process with values in the set of one-to-one maps from $\{1, \cdots, N\} \setminus \{k\}$ onto $\{1, \cdots, N-1\}$ which we shall specify later on.

Massage the Formula

We have:

$$
\begin{aligned}
Y_t^k = Y_0^k &+ \sum_{i=1,\, i\neq k}^{N} \int_{[0,t]\times[0,1]} \varphi_{ik}^k(Y_{s-}, \phi(Y_{s-}))\, \mathbf{1}_{\left[\frac{\sigma_s(i)-1}{N-1} < w \leq \frac{\sigma_s(i)}{N-1}\right]} \tilde{N}^1(ds, dw) \\
&+ \sum_{j=1,\, j\neq k}^{N} \int_{[0,t]\times[0,1]} \varphi_{kj}^k(Y_{s-}, \phi(Y_{s-}))\, \mathbf{1}_{\left[\frac{\sigma_s(j)-1}{N-1} < w \leq \frac{\sigma_s(j)}{N-1}\right]} \tilde{N}^2(ds, dw) \qquad (1) \\
= Y_0^k &+ \int_{[0,t]\times[0,1]} \left(\sum_{i=1,\, i\neq k}^{N} \varphi_{ik}^k(Y_{s-}, \phi(Y_{s-}))\, \mathbf{1}_{\left[\frac{\sigma_s(i)-1}{N-1} < w \leq \frac{\sigma_s(i)}{N-1}\right]} \right) \tilde{N}(ds, dw)
\end{aligned}
$$

where $\tilde{N}(ds, dw) = \tilde{N}^1(ds, dw) + \tilde{N}^2(ds, dw)$ is Poisson random measure on $[0, \infty) \times [0, 1]$ with intensity Leb_2.

Search for Best Response of Player k

Assume all players $j \neq k$ use same strategy $\tilde{\phi}^k$, and find best response ϕ^k by player k

$$Y_t^k = Y_0^k + \int_{[0,t]\times[0,1]} \mathbf{1}_{Y_{s-}^k > 0} \sum_{i=1,\, i\neq k}^{N} \mathbf{1}_{Y_{s-}^i > 0} F(\phi^k(Y_{s-}^k), \tilde{\phi}^k(Y_{s-}^i)) \mathbf{1}_{[\frac{\sigma_s(i)-1}{N-1} < w \leq \frac{\sigma_s(i)}{N-1}]} \tilde{N}(ds, dw)$$

Choose the Predictable $\sigma_t(i)$ Appropriately

- Order the wealths $Y_t^{(1),-k} \leq Y_t^{(2),-k} \leq \cdots \leq Y_t^{(N-1),-k}$
- Choose $\sigma_t : \{1, \cdots, N\} \setminus \{k\} \mapsto \{1, \cdots, N-1\}$ so that

$$Y_t^{(i),-k} = Y_t^{\sigma_t^{-1}(i)}, \qquad i = 1, \cdots, N-1.$$

- Denote by $\tilde{\mu}_t^{-k}$ tje empirical measure $\tilde{\mu}_t^{-k} = \frac{1}{N-1} \sum_{i=1,\, i \neq k}^{N} \delta_{Y_t^i}$
- Denote $[0,1] \ni w \mapsto \tilde{Q}_t^{-k}(w) \in [0, \infty)$ its quantile function

From

$$Y_t^k = Y_0^k + \int_{[0,t]\times[0,1]} \mathbf{1}_{Y_{s-}^k > 0} \Big(\sum_{i=1}^{N-1} \mathbf{1}_{Y_{s-}^{(i),-k} > 0} F(\phi^k(Y_{s-}^k, \tilde{\phi}^k(Y_{s-}^{(i),-k})) \, \mathbf{1}_{[\frac{i-1}{N-1} < w \leq \frac{i}{N-1}]} \Big) \tilde{N}(ds, dw$$

and

$$\frac{i-1}{N-1} < w \leq \frac{i}{N-1} \iff \tilde{Q}_t^{-k}(w) = Y_t^{(i),-k}, \qquad t \geq 0$$

we get

$$Y_t^k = Y_0^k + \int_{[0,t]\times[0,1]} F\Big(Y_{s-}^k, \tilde{Q}_{s-}^{-k}(w), \phi^k(Y_{s-}^k), \tilde{\phi}^k(\tilde{Q}_{s-}^{-k}(w)) \Big) \tilde{N}(ds, dw)$$

MEAN FIELD GAME PROBLEM

◇ ***Best Response Step:***

- Fix a distributed feedback function $\tilde{\phi} : [0,\infty) \times [0,\infty) \ni (t,y) \mapsto \tilde{\phi}_t(y) \in A$
- Fix a flow $(\tilde{\mu}_t)_{t\geq 0}$ of probability measures
- Solve:

$$\sup_{\phi} \mathbb{E}\Big[g(Y_T, \tilde{\mu}_T) + \int_0^T f(t, Y_t, \phi_t, \tilde{\phi}_t, \tilde{\mu}_t)dt\Big] \tag{2}$$

under the dynamic constraint:

$$Y_t = Y_0 + \int_{[0,t]\times[0,1]} F\Big(Y_{s-}, \tilde{Q}_{s-}(w), \phi(Y_{s-}), \tilde{\phi}(\tilde{Q}_{s-}(w))\Big)\tilde{N}(ds, dw) \tag{3}$$

where $\tilde{Q}_t$ denotes the quantile function of the probability measure $\tilde{\mu}_t$.

,

◇ ***Fixed Point Step:*** If $(\hat{\phi}, \hat{\mathbf{Y}} = (\hat{Y}_t)_{t\geq 0})$ solves the above problem, demand:

$$\tilde{\phi} = \hat{\phi}, \qquad \text{and} \qquad \tilde{\mu}_t = \mathcal{L}(\hat{Y}_t), \quad \text{for } \ 0 \leq t \leq T. \tag{4}$$

Example with Wealth in $\{0, 1, \cdots, 100\}$, $\mu_0 = \delta_{50}$

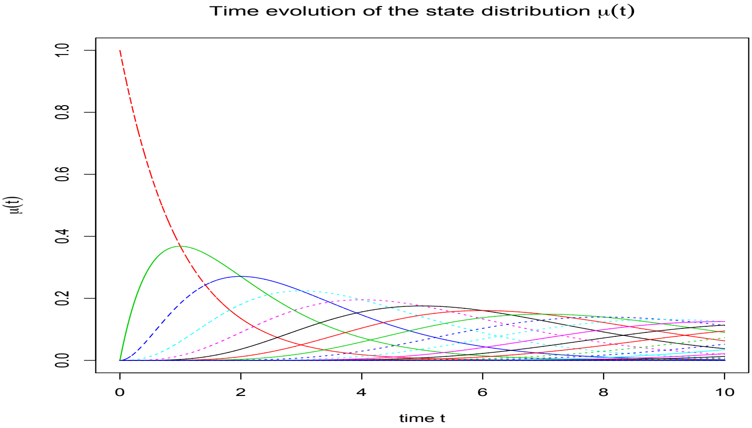

FIGURE: Time evolution of the total mass of the distribution μ_t. Killing has no effect.

Example with Wealth in $\{0, 1, \cdots, 50\}$

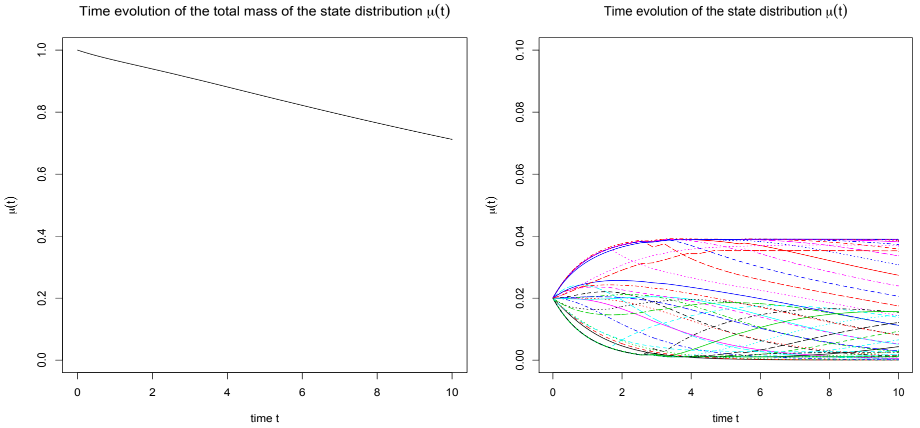

FIGURE: Left: Time evolution of total mass (killing if $Y_t < 0$ or $Y_t > 50$. Right: Time evolution of the distribution $\mu_t(y)$ for $y = 0, 1, \cdots, 50$.