Optimal Incentives to Mitigate Epidemics: A Stackelberg Mean Field Game Approach

Alexander Aurell[1]

joint work with René Carmona, Gökçe Dayanıklı & Mathieu Laurière

[1]**ORFE, Princeton University**
aaurell@princeton.edu

Stochastic Modeling and Financial Impacts of the Coronavirus Pandemic
SIAM Conference on Financial Mathematics and Engineering
June 1, 2021

Motivation

In the absence of a vaccine, how can the individuals be optimally **incentivized** to make the right **effort** in the fight against an epidemic?

Motivation

In the absence of a vaccine, how can the individuals be optimally **incentivized** to make the right **effort** in the fight against an epidemic?

A policy maker's problem: give incentives and penalties to the population that

1. the people accept and follow
2. yields a behavior that "controls" the epidemic

Setting

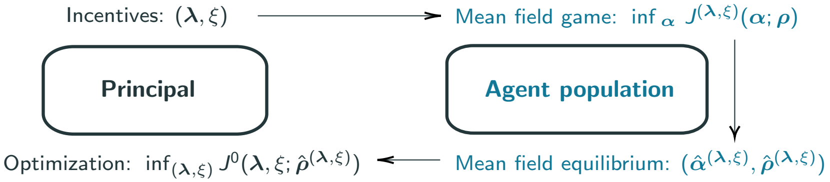
"Optimal incentives to mitigate epidemics: A Stackelberg mean field game approach"
A., Carmona, Dayanıklı, Lauriére, arXiv 2020.

→ Disease spreads depending on the **agents' efforts** to **slow the spread**.

→ The agents are not cooperating!

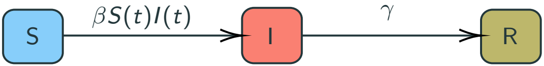
→ Principal **optimizes** a contract by taking into account the agents' response.

The principal and the population play a **Stackelberg game**[1]



[1] Holmström-Milgrom '87, Sannikov '08 '13, Djehiche-Helgesson '14, Cvitanić *e.a.* '18, Carmona-Wang '18, Elie *e.a.* '19

Epidemic modelling with the SIR model



→ Individuals are categorized as "**S**usceptible", "**I**nfected" or "**R**ecovered"

→ The system of equations that describes the evolution of the epidemic:

$$\begin{cases} \dot{S}(t) = -\beta S(t)I(t), & S(0) \geq 0 \\ \dot{I}(t) = \beta S(t)I(t) - \gamma I(t), & I(0) \geq 0 \\ \dot{R}(t) = \gamma I(t), & R(0) \geq 0 \\ S(0) + I(0) + R(0) = 1, & \end{cases}$$

SIR Modeling with Continuous Time Markov Chains

Consider N agents. Agent i has state $X_t^i \in \{\mathrm{S}, \mathrm{I}, \mathrm{R}\}$ at time t.

$\rightarrow$ **Pairwise meetings** at random with rate β.

$\rightarrow$ Susceptible agent meets an infected agent: Susceptible gets infected.

$$Q(p_t^N) = \begin{bmatrix} \dots & \beta p_t^N(I) & 0 \\ 0 & \dots & \gamma \\ 0 & 0 & 0 \end{bmatrix}$$

where $p_t^N(I)$ is the **proportion** of the population that is infected at time t.

SIR Modeling with Continuous Time Markov Chains

Consider N agents. Agent i has state $X_t^i \in \{\mathrm{S}, \mathrm{I}, \mathrm{R}\}$ at time t.

- → **Pairwise meetings** at random with rate β.
- → Susceptible agent meets an infected agent: Susceptible gets infected.
- → **The probability of infection is decreased by the action of two agents that meet in a multiplicative way.**

$$Q(p_t^N) = \begin{bmatrix} \dots & \beta p_t^N(I) & 0 \\ 0 & \dots & \gamma \\ 0 & 0 & 0 \end{bmatrix}$$

where $p_t^N(I)$ is the **proportion** of the population that is infected at time t.

Consider N agents. Agent i has state $X_t^i \in \{\mathrm{S}, \mathrm{I}, \mathrm{R}\}$ at time t.

- → **Pairwise meetings** at random with rate β.
- → Susceptible agent meets an infected agent: Susceptible gets infected.
- → **The probability of infection is decreased by the action of two agents that meet in a multiplicative way.**

Introduce "contact factor"

$$\begin{bmatrix} \dots & \beta\alpha_t^j \frac{1}{N}\sum_{k=1}^{N} \alpha_t^k \mathbb{1}_I(X_{t-}^k) & 0 \\ 0 & \dots & \gamma \\ 0 & 0 & 0 \end{bmatrix}$$

How to characterize the Stackelberg equilibrium?

→ Nash equilibrium computation is notoriously hard in games with a large number of players, $N \gg 1$.

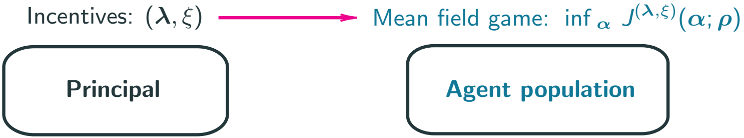
→ Approximation in the large-population limit $N \to \infty$: **Mean Field Games!**[2]

→ Can often be used when:

- Players are almost **identical**
- Interactions are of **mean-field** type

[2] Huang-Malhamé-Caines '06, Lasry-Lions '06

Agent Population



→ For very large N, approximate the game with **contact factor** control with **"extended finite-state MFG"**[3]

→ For a fixed **joint control-state distribution flow** ρ

- the cost for $\boldsymbol{\alpha} \in \mathbb{A}$ is

$$J^{\boldsymbol{\lambda},\xi}(\boldsymbol{\alpha}; \boldsymbol{\rho}) := \mathbb{E}\left[\int_0^T f(t, X_t, \alpha_t, \rho_t; \lambda_t)dt - U(\xi)\right],$$

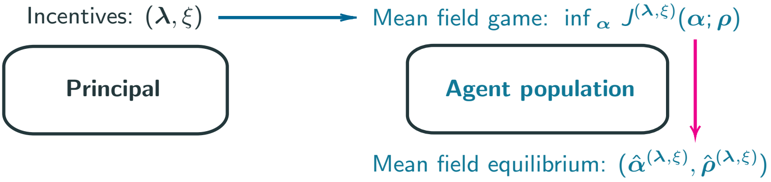
  where $(\boldsymbol{\lambda}, \xi)$ is principal's policy choice.

- The agent's state X_t jumps according to $Q(\alpha_t, \rho_t)$. In the SIR example:

$$Q(\alpha, \rho) = \begin{bmatrix} \cdots & \beta\alpha \int_A a\rho(da, I) & 0 \\ 0 & \cdots & \gamma \\ 0 & 0 & \cdots \end{bmatrix},$$

[3] Gomes *e.a.* '10 '13, Kolokoltsov '12, Carmona-Wang '16 '18, Cecchin-Fischer '18, Bayraktar-Cohen '18, Choutri *e.a.* '18 '19

Mean Field Nash Equilibrium



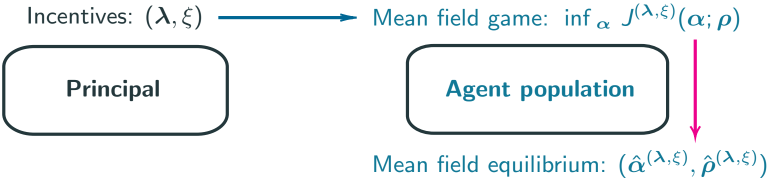
Definition: If the pair $(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\rho}})$ satisfies:

(i) $\hat{\boldsymbol{\alpha}}$ minimizes the cost of player given $\hat{\boldsymbol{\rho}}$;

(ii) $\forall t \in [0, T]$, $\hat{\rho}_t$ is the joint distribution of control $\hat{\alpha}_t$ and state X_t,

then $(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\rho}})$ is a **mean field Nash equilibrium** given the contract $(\boldsymbol{\lambda}, \xi)$.

Mean Field Nash Equilibrium



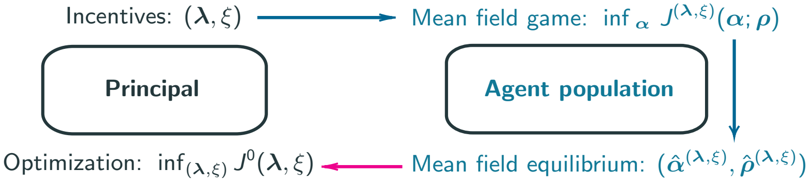
Definition: If the pair $(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\rho}})$ satisfies:

(i) $\hat{\boldsymbol{\alpha}}$ minimizes the cost of player given $\hat{\boldsymbol{\rho}}$;

(ii) $\forall t \in [0, T]$, $\hat{\rho}_t$ is the joint distribution of control $\hat{\alpha}_t$ and state X_t,

then $(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\rho}})$ is a **mean field Nash equilibrium** given the contract $(\boldsymbol{\lambda}, \xi)$.

$\rightarrow$ For given the contract $(\boldsymbol{\lambda}, \xi)$, mean-field Nash equilibria are characterized with a **forward-backward SDE** (FBSDE).

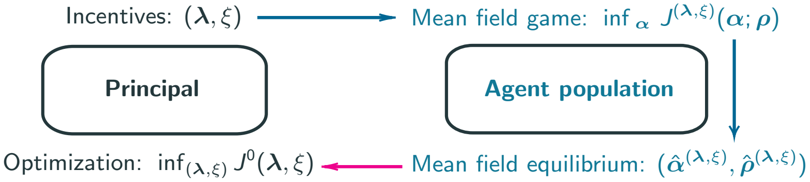


→ The principal's cost for policy $(\boldsymbol{\lambda}, \xi)$ is

$$J^0(\boldsymbol{\lambda}, \xi) := \mathbb{E}\left[\int_0^T \Big(c_0(t, \hat{p}_t^{\boldsymbol{\lambda},\xi}) + f_0(t, \lambda_t)\Big)\, dt + C_0(\hat{p}_T^{\boldsymbol{\lambda},\xi}) + \xi\right]$$

where $\hat{\boldsymbol{p}}^{(\boldsymbol{\lambda},\xi)}$ is the state-marginal of $\hat{\boldsymbol{\rho}}^{(\boldsymbol{\lambda},\xi)}$.

Principal



→ The principal's cost for policy $(\boldsymbol{\lambda}, \xi)$ is

$$J^0(\boldsymbol{\lambda}, \xi) := \mathbb{E}\left[\int_0^T \Big(c_0(t, \hat{p}_t^{\boldsymbol{\lambda},\xi}) + f_0(t, \lambda_t)\Big)\, dt + C_0(\hat{p}_T^{\boldsymbol{\lambda},\xi}) + \xi\right]$$

where $\hat{\boldsymbol{p}}^{(\boldsymbol{\lambda},\xi)}$ is the state-marginal of $\hat{\boldsymbol{\rho}}^{(\boldsymbol{\lambda},\xi)}$.

→ The principal's **optimization problem** is

$$\inf_{(\boldsymbol{\lambda},\xi)\in\mathcal{C}} \inf_{\substack{(\boldsymbol{\alpha},\boldsymbol{\rho})\in\mathcal{N}(\boldsymbol{\lambda},\xi) \\ J^{\boldsymbol{\lambda},\xi}(\boldsymbol{\alpha};\boldsymbol{\rho})\le\kappa}} J^0(\boldsymbol{\lambda}, \xi).$$

How to solve the principle's optimization problem?

Numerical Approach

Rewriting → **Discretization** → **Parameterization**

→ Reposing the FBSDE as optimal control of a forward-in-time control problem[4]

→ Time-discretization and Monte Carlo-approximation

→ Parametrizing the optimization variables (principle contract + FBSDE) with **neural networks**[5]

[4] Sannikov '08, '13
[5] Carmona-Lauriére '19

An example...

Example: Evaluation of Numerical Approach

Agent Population: Set $U(\xi) = \xi$ and

$$f(t, x, \alpha, \rho; \lambda) = \frac{c_\lambda}{2}\left(\lambda^{(S)} - \alpha\right)^2 \mathbb{1}_S(x) + \left(\frac{1}{2}\left(\lambda^{(I)} - \alpha\right)^2 + c_I\right)\mathbb{1}_I(x) + \frac{1}{2}\left(\lambda^{(R)} - \alpha\right)^2 \mathbb{1}_R(x),$$

where $c_\lambda, c_I \in \mathbb{R}_+$ are constants.

Principal: Set $C_0(p) = 0$ and

$$c_0(t, p) = c_{\mathrm{Inf}}\, p(I)^2, \quad f_0(t, \lambda) = \sum_{i \in \{S,I,R\}} \frac{\bar{\beta}^{(i)}}{2}\left(\lambda^{(i)} - \bar{\lambda}^{(i)}\right)^2$$

for constant $\bar{\lambda}, \bar{\beta} \in \mathbb{R}_+^m$ and $c_{\mathrm{Inf}} > 0$.

Solutions: Inactive Principal

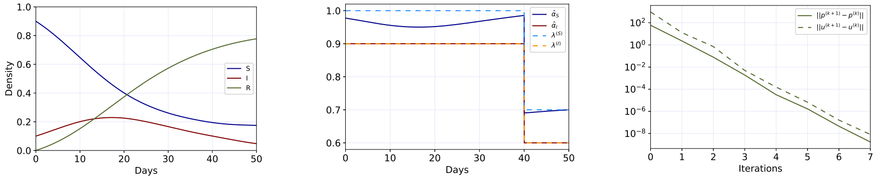


Figure 1: Late lockdown, ODE solution. Evolution of the population state distribution (left), evolution of the controls (middle), convergence of the solver (right).

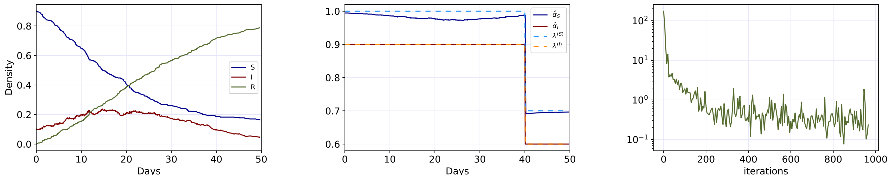


Figure 2: Late lockdown, numerical solution. Evolution of the population state distribution (left), evolution of the controls (middle), convergence of the loss value (right).

Solutions: Active Principal

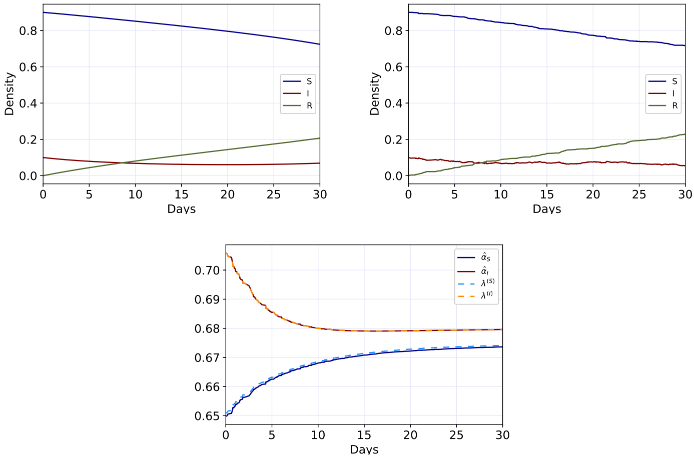


Figure 3: SIR Stackelberg. Evolution of the population state distribution in ODE solution (top left), evolution of the population state distribution in numerical solution (top right), evolution of the controls in numerical solution (bottom).

T	p^0	c_λ	c_I	c_{Inf}	$\bar{\beta}$	$\bar{\lambda}$	β	γ	κ
30	[0.9, 0.1, 0]	10	0.5	1	[0.2, 1, 0]	[1, 0.7, 0]	0.25	0.1	0

Epidemic disease spread in a heterogeneous population

"Finite State Graphon Games with Applications to Epidemics"

A., Carmona, Dayanıklı, Lauriére, on arXiv very soon!

→ Disease spreads depending on the agents' efforts to slow the spread

→ The agents are not cooperating

→ Agents are **heterogeneous** and have individual rates for infection, recovery, etc.

The population plays a **graphon game**[6]

[6] Delarue '17, Parise-Ozdaglar '19, Carmona *e.a.* '19, Caines e.a. '18,'19,'20, A.-Carmona-Lauriére '21

Epidemic disease spread in a heterogeneous population

→ A **continuum of players**, labeled by $x \in [0,1]$

→ Players see a weighted aggregate: player characteristics (like state, control, etc) weighted by a **graphon** w.

- $w : [0,1] \times [0,1] \to [0,1]$ measurable and symmetric
- the contact factor aggregate $\int_A a\rho_t(da, I)$ now becomes (for player x)

$$Z_t^x = \int_I w(x,y) \left(\int_A a\rho_t^y(da, \mathrm{I}) \right) dy$$

- In the SIR model with contact factor control

$$Q^x(\alpha_t^x, Z_t^x) = \begin{bmatrix} \dots & \beta(x)\alpha_t^x Z_t^x & 0 \\ 0 & \dots & \gamma(x) \\ 0 & 0 & \dots \end{bmatrix}$$

- Individual costs

→ Leads to a **graphon game** between the players

Thank you!