Deep Principal-Agent Mean Field Games

Steven Campbell†

Yichao Chen†

Arvind Shrivats◇

Sebastian Jaimungal†

◇Princeton U., †U. Toronto & Oxford Man Institute

June 27, 2022

What problem do we address?

- We consider firms regulated by a REC market
- What is the set of **strategies** across **all** regulated firms that achieves a **Nash equilibrium** and an **endogenous price**?
  - Accounting for:
    - **Rental** costs
    - **Expansion** costs
    - **Trading** costs
- What is the **optimal penalty** a **Principal** should set?

Example setup...

- expansion of clean energy $\boldsymbol{\alpha_t}$
- rental of clean energy $\boldsymbol{g_t}$
- trading REC rate $\boldsymbol{\Gamma_t}$

State evolution:

$$
\begin{aligned}
dX_t^i &= (h_t^k + g_t^i + C_t^i + \Gamma_t^i)dt + \sigma^k\, dW_t^i, & X_0^i &= \xi^i \sim F(\Theta^k) \\
dC_t^i &= \alpha_t^i\, dt, & C_0^i &= 0.
\end{aligned}
$$

Performance criterion:

$$
\begin{aligned}
J^{A,i}(\alpha, g, \Gamma; \boldsymbol{\mu}) = \mathbb{E}\Bigg[\int_0^T \left\{ \frac{\zeta^k}{2}(g_u^i)^2 + \frac{\gamma^k}{2}(\Gamma_u^i)^2 + \frac{\beta^k}{2}(\alpha_u^i)^2 + S_u^{\boldsymbol{\mu}}\, \Gamma_u^i \right\} du \\
+ \phi_0 + \sum_{j=1}^{N} w_j (R_j - X_T^i)^+ \;\Bigg|\; \mathcal{G}_0^i \Bigg],
\end{aligned}
$$

Market clearing:

$$
\mathbb{E}[\,\Gamma_t^{\cdot}\,] = 0
$$

Example setup cont'd...

$$
\begin{aligned}
dX_t^{(k)} &= \left(h_t^k - \left(\tfrac{1}{\zeta^k} + \tfrac{1}{\gamma^k}\right) Y_t^{(k),X} - \tfrac{1}{\gamma^k} S_t^{\boldsymbol{\mu}} + C_t^i\right) dt + \sigma^k dW_t, && X_0^{(k)} = \xi^{(k)}, \\
dC_t^{(k)} &= -\tfrac{1}{\beta^k} Y_t^{(k),C} dt, && C_0^{(k)} = 0, \\
dY_t^{(k),X} &= Z_t^{(k),X} dW_t^{(k)}, && Y_T^{(k),X} = \partial_x g(X_T^{(k)}), \\
dY_t^{(k),C} &= -Y_t^{(k),X} dt + Z_t^{(k),C} dW_t, && Y_T^{(k),C} = 0,
\end{aligned}
$$

Optimal strategy:

$$
\alpha_t^{(k),\star} = -\tfrac{1}{\beta^k}\, Y_t^{(k),C}, \qquad g_t^{(k),\star} = -\tfrac{1}{\zeta^k}\, Y_t^{(k),X},
$$

$$
\Gamma_t^{(k),\star} = -\tfrac{1}{\gamma^k}\, \left(Y_t^{(k),X} - S_t^{\boldsymbol{\mu}}\right), \quad \text{and}
$$

$$
S_t^{\boldsymbol{\mu}} = -\sum_{k\in\mathcal{K}} \omega_k\, \mathbb{E}\left[Y_t^{(k),X}\right], \qquad \omega_k = \frac{\pi_k}{\gamma_k} \Bigg/ \sum_{k'\in\mathcal{K}} \frac{\pi_{k'}}{\gamma_{k'}}
$$

Principal-Agent MFG

Problem

Solve

$$\sup_{g \in \mathcal{G}_\beta[a,b]} J(g) \tag{P}$$

where

$$J(g) := \mathbb{E}\left[\mathcal{U}\left(g, (X_t^g)_{t\in[0,T]}, (Y_t^g)_{t\in[0,T]}, (\mathcal{L}(X_t^g, Y_t^g))_{t\in[0,T]}\right)\right],$$

$$\mathcal{G}_\beta[a,b] := \{g \in C^{1,1}[a,b] : g \text{ is convex, } ||g||_{C^{1,1}} \leq \beta\},$$

and (X^g, Y^g) *are solutions to the MV-FBSDE system*

$$\begin{aligned} dX_t^g &= \varphi(t, X_t^g, Y_t^g, \mathcal{L}(X_t^g, Y_t^g))dt + \sigma(t, X_t^g, Y_t^g)dW_t; \qquad X_0^g \sim \xi \\ dY_t^g &= -\rho(t, X_t^g, Y_t^g, \mathcal{L}(X_t^g, Y_t^g))dt + Z_t^g dW_t^i; \qquad Y_T^g = h(X_T^g) := \partial_x g(X_T^g). \end{aligned}$$

Principal-Agent MFG

Under usual Lipschitz continuity & square integrability...

Theorem (Existence and Uniqueness (MV-FBSDE) – Carmona & Delarue)

There exists a constant $c > 0$*, such that for* $T \leq c$ *and for any initial condition* $X_0 = \xi \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R})$*, the MV-FBSDE has a unique solution* $(X, Y, Z) \in \mathbb{S}^2 \times \mathbb{S}^2 \times \mathbb{H}^2$*.*

Principal-Agent MFG

Lemma (Stability of Solutions (MV-FBSDE))

There exist two constants $c, C \geqslant 0$, only depending on L, such that, with probability 1:

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{0\leqslant t\leqslant T}\left|X_t - X_t'\right|^2 + \sup_{0\leqslant t\leqslant T}\left|Y_t - Y_t'\right|^2 + \int_0^T \left|Z_t - Z_t'\right|^2 dt\right] \\
&\leq C\,\mathbb{E}\left[\left|\xi - \xi'\right|^2 + \left|\left(h - h'\right)(X_T)\right|^2 + \left(\int_0^T \left|\left(\varphi - \varphi', \rho - \rho'\right)(t, X_t, Y_t, Z_t, \mathcal{L}(X_t, Y_t))\right| dt\right)^2\right. \\
&\qquad\left. + \int_0^T \left|\left(\sigma - \sigma'\right)(t, X_t, Y_t, Z_t)\right|^2 dt\right]
\end{aligned}
$$

as long as $T \leqslant c$.

– extends Theorem 1.3 of Delarue (2002) to account for $\mathcal{L}(X_t, Y_t)$

Principal-Agent MFG

Lemma (Objective Function Continuity)

For sufficiently small T, the map $g \mapsto J(g)$ is continuous on $\mathcal{G}_\beta[a, b]$.

– proof relies on stability result, and joint continuity of $(g, X, Y, V) \mapsto \mathbb{E}[\mathcal{U}(g, X, Y, V)]$

Principal-Agent MFG

Theorem (Existence of Solution (PA-MFG))

There exists an optimizer $g^* \in \mathcal{G}_\beta[a, b]$ *to Problem* (P)

Moreover, for any sequence of closed subsets $(G_n)_{n\in\mathbb{N}}$ *of* $\mathcal{G}_\beta[a, b]$ *there exists a sequence of optimizers* $(g_n^*)_{n\in\mathbb{N}}$ *with* $g_n^* \in G_n$ *to Problem* (P) *when the feasible set* $\mathcal{G}_\beta[a, b]$ *is replaced by* G_n*.*

Lemma (Approximating Sequence of Sets of Penalties)

Let $G_n \subseteq G_{n+1}$ *be a sequence of subsets* $G_n \subseteq \mathcal{G}_\beta[a, b]$ *s.t.* $\bigcup_n G_n$ *is dense in* $\mathcal{G}_\beta[a, b]$*. If* $J(g)$ *is bounded from above and continuous on* $\mathcal{G}_\beta[a, b]$*, then*

$$\lim_{n\to\infty} \sup_{g_n \in G_n} J(g_n) = \sup_{g \in \mathcal{G}_\beta[a,b]} J(g).$$

Moreover, there exists a subsequence $(g_{n_k}^*)_{k\geq 0}$ *such that* $g^* := \lim_{k\to\infty} g_{n_k}^*$ *is a maximizer of* $J(g)$ *on* $\mathcal{G}_\beta[a, b]$*.*

Principal-Agent MFG: Discretisation

Discretize the MV-FBSDE

$$
\begin{aligned}
X_{t_{i+1}}^{g,\pi} &= X_{t_i}^{g,\pi} + \varphi(t_i, X_{t_i}^{g,\pi}, Y_{t_i}^{g,\pi}, \hat{\mathbb{P}}_{(X_{t_i}^{g,\pi}, Y_{t_i}^{g,\pi})})h + \sigma(t_i, , X_{t_i}^{g,\pi}, Y_{t_i}^{g,\pi})\Delta W_i, \\
X_0^{g,\pi} &\sim \xi, \\
Z_{t_i}^{g,\pi} &= \psi_i^g(X_{t_i}^{\pi}, Y_{t_i}^{\pi}), \\
Y_{t_{i+1}}^{\pi} &= \rho(t_i, X_{t_i}^{g,\pi}, Y_{t_i}^{g,\pi}, \hat{\mathbb{P}}_{(X_{t_i}^{g,\pi}, Y_{t_i}^{g,\pi})})h + Z_t^{g,\pi}\Delta W_i,
\end{aligned}
$$

where

- $Y_0^{g,\pi} = \mu_0^{g.\pi}(X_0^{g,\pi})$
- $(Z_{t_m}^{g,\pi})_{m=1}^M$, with $Z_{t_m}^{g,\pi} = \psi_m^{g.\pi}(X_m^{g,\pi}, Y_m^{g,\pi})$

are ensembles of neural-nets

For fixed g, solve

$$
\inf_{\mu_0 \in \mathcal{N}_0', \psi_i \in \mathcal{N}_i} \mathbb{E}\left|\partial_x g(X_T^{g,\pi}) - Y_T^{g,\pi}\right|^2
$$

Principal-Agent MFG: Discretisation

We assume monotonicity of σ, ρ, and φ and that they admit a modulus of continuity $\omega(\cdot)$ in time.

Theorem (MV-FBSDE error – Reisinger et.al. (2020))

There exists a constant C such that it holds for all sufficiently large N (given by some partition π) and for all $\theta^{g,\pi} \in \mathcal{N}_0' \times \prod_{i=0}^{N-1} \mathcal{N}_i$ that

$$\sup_{t\in[0,T]} \left(\mathbb{E}\left[\left| X_t - \hat{X}_t^{g,\pi,\theta} \right|^2 \right] + \mathbb{E}\left[\left| Y_t - \hat{Y}_t^{g,\pi,\theta} \right|^2 \right] \right) + \mathbb{E}\left[\int_0^T \left| Z_t - \hat{Z}_t^{g,\pi,\theta} \right|^2 \right] dt$$
$$\leq C \left(\omega\left(\tau_N\right)^2 + \mathcal{R}_\pi + \mathbb{E}\left[\left| \hat{Y}_T^{g,\pi,\theta} - h\left(\hat{X}_T^{g,\pi,\theta} \right) \right|^2 \right] \right)$$

where

$$\mathcal{R}_\pi := \max_{\substack{i\in\mathcal{N}_{<N} \\ t\in[t_i,t_{i+1}]}} \left(\mathbb{E}\left[|X_t - X_{t_i}|^2 \right] + \mathbb{E}\left[|Y_t - Y_{t_i}|^2 \right] \right) + \sum_{i=0}^{N-1} \mathbb{E}\left[\int_{t_i}^{t_{i+1}} \left| Z_t - \bar{Z}_{t_i} \right|^2 dt \right] \sim \Delta t_n$$

with $\bar{Z}_i := \frac{1}{\tau_N} \mathbb{E}_i \left[\int_{t_i}^{t_{i+1}} Z_s ds \right]$.

Principal-Agent MFG: Discretisation

Proposition (PA-MFG Approximation Bound)

Suppose there exist networks $\theta^{g,\pi} = (\mu_0^{g,\pi}, \{\psi_i^{g,\pi}\}_{i=0}^{N-1}) \in \mathcal{N}_0' \times \prod_{i=0}^{N-1} \mathcal{N}_i$ *s.t.*

$$\tau_N + \omega(\tau_N)^2 + \mathbb{E}\left|\partial_x g(\hat{X}_T^{g,\pi,\theta}) - \hat{Y}_T^{g,\pi,\theta}\right|^2 < \epsilon.$$

then

$$\Bigg|\sup_{g \in G_n} \mathbb{E}\left[\mathcal{U}(g, X^g, Y^g, \mathcal{L}(X^g, Y^g))\right] - \sup_{g \in G_n} \mathbb{E}\left[\mathcal{U}(g, \hat{X}^{g,\pi,\theta}, \hat{Y}^{g,\pi,\theta}, \mathcal{L}(\hat{X}^{g,\pi,\theta}, \hat{Y}^{g,\pi,\theta}))\right]\Bigg| \leq K\sqrt{\epsilon}$$

Principal-Agent MFG: Discretisation

Algorithm 1: Principal-Agent Optimization

Initialize forward network parameters θ. Initialize principal parameters Υ. Initialize memory buffer $\mathfrak{M}$. Initialize $\mathbf{u}^{(0)}$

Define update rule for sampling radius ϵ, network parameters $\boldsymbol{\theta}, \Upsilon$ and values $\mathbf{u}^{(\cdot)}$;

foreach n in $[N_O]$ **do**

    Initialize local samples $\mathcal{U} := (\mathbf{u}_i^{(j)})_{i \in [N_S]} \subset B_\epsilon(\mathbf{u}^{(j)})$;

    foreach $\mathbf{u}' \in \mathcal{U}$ **do**

        foreach k in $[N_F]$ **do**

            Sample paths of the discretized MV-FBSDE using the ensemble network $\boldsymbol{\theta}$;

            Compute the MV-FBSDE loss $\mathcal{L}_F(\boldsymbol{\theta})$;

            Update $\boldsymbol{\theta}$;

            if $\mathcal{L}_F(\boldsymbol{\theta}) < TOL_F$ **then**

                break;

            end

        end

        Estimate the principal loss via $\hat{\mathcal{L}}_P(\psi(\mathbf{u}'))$;

        Store $(\mathbf{u}', \hat{\mathcal{L}}_P(\psi(\mathbf{u}')))$ in $\mathfrak{M}$;

    end

Principal-Agent MFG: Discretisation

Algorithm 2: Principal-Agent Optimization

    foreach k in $[N_A]$ **do**
        Sample from $\mathfrak{M}$ a batch $\mathbf{U} := (\mathbf{u}_m)_{m\in[N_B]}$ and associated principal losses $\hat{\mathcal{L}}_P(\psi(\mathbf{U})) := (\hat{\mathcal{L}}_P(\psi(\mathbf{u}_m)))_{m\in[N_B]}$;
        Compute the MSE loss $||\mathcal{L}_P^{\Upsilon}(\psi(\mathbf{U})) - \hat{\mathcal{L}}_P(\psi(\mathbf{U}))||_2^2$;
        Update Υ;
    foreach k in $[N_P]$ **do**
        Compute $\mathcal{L}_P^{\Upsilon}(\psi(\mathbf{u}^{(j)}))$;
        Update $\mathbf{u}^{(j)} \mapsto \mathbf{u}^{(j+1)}$;
        Increment $j \mapsto j+1$;
    Update ϵ;
    if $||\mathbf{u}^{(j)} - \mathbf{u}^{(j-N_P)}||_2 < TOL$ **then**
        break;

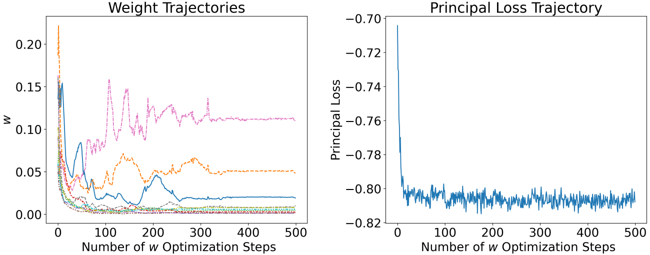


Figure: Trajectory of optimal weights (left) and principal loss (right).

MOVIE

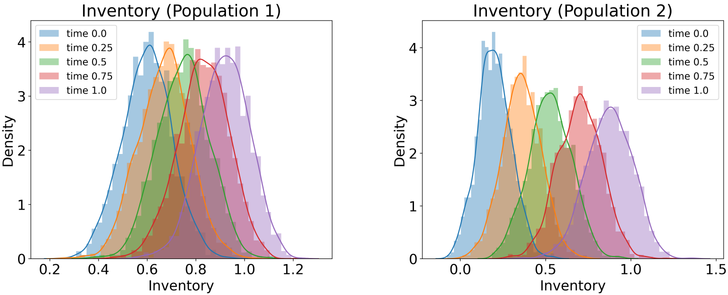


Figure: Inventories for population 1 (left) and population 2 (right) throughout time.

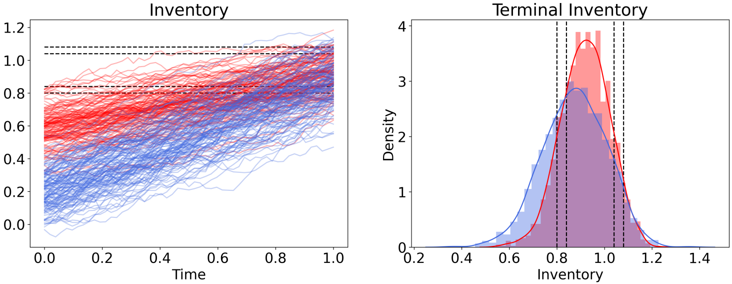


Figure: Inventory paths (left) and terminal inventory (right). Population 1 indicated in red, population 2 indicated in blue. The lines denote the non-negligible nodes in the terminal penalty g.

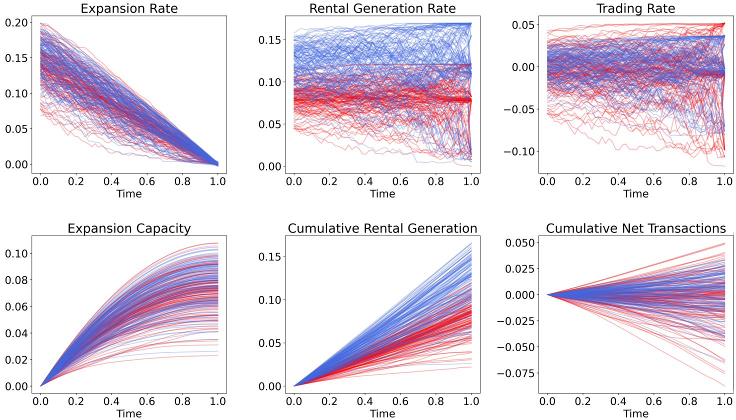


Figure: Expansion rate and total expansion (left column), rental generation rate and cumulative rental generation (middle column), trading rate and net trading position (right column) across sampled paths of representative agent from each sub-population. Population 1 indicated by red, population 2 indicated by blue.

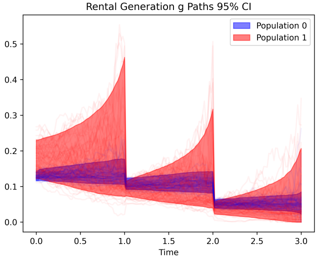
Finite Banking Case

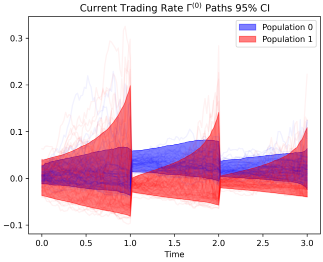
We consider multi-period REC problem where the firms have N compliance periods and the inventories expire after their second compliance periods since generation.

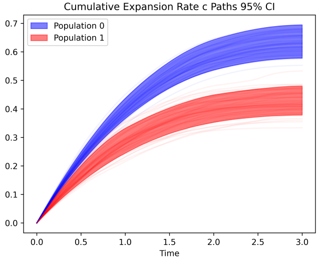
$$dX_t^{i,(k)} = (h_t^{(k)} + g_t^i + \Gamma_t^i + C_t^i)dt + \sigma^{(k)} dW_t^{i,(k)}, \qquad X_0 \sim \xi$$

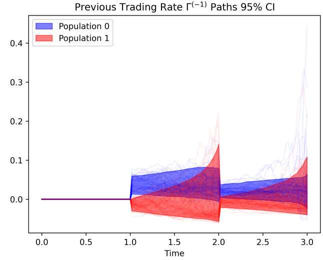
$$dV_t^{i,(k)} = \Gamma_t^{V,i} dt \qquad V_0 = 0$$

Finite Banking Case









Main Contributions

- Main Contributions:
  - Deep learning MV-FBSDE approach for solving PA-MFGs
  - Convergence of deep PA-MFG algorithm
  - Applied to single period REC market problem/ multi-period finite-banking REC market
- On-going Work:
  - Theoretical foundations of multi-period REC market problems (indefinite banking/finite-banking/infinite-horizon) under PA-MFGs framework

Thank you for your attention!

https://sebastian.statistics.utoronto.ca

based on

Deep Learning for Principal-Agent Mean Field Games

https://arxiv.org/abs/2110.01127