SIAM FM2023

A Rank-Based Reward between a Principal and a Field of Agents: Application to Energy Savings

Clémence Alasseur, Erhan Bayraktar, Roxana Dumitrescu, Quentin Jacquet

June 6, 2023

Section 1

Introduction

Context

Obligations imposed by governments:

- In France: electricity providers ("*Obligés*") have a target of Energy Saving Certificates[1] to hold at a predetermined horizon ($\simeq$ 3 years).
  If they fail, they face financial penalties.

Existing incentives "Provider $\rightarrow$ customers":

- Comparison to similar customers
  - EDF, Total, Engie, . . .
- Reward/Bonus when reduction compared to past consumption
  - "SimplyEnergy"[2], "Plüm énergie"[3], "OhmConnect"[4]

[1] www.powernext.com/french-energy-saving-certificates
[2] www.simplyenergy.com.au/residential/energy-efficiency/reduce-and-reward
[3] www.plum.fr/cagnotte/
[4] www.ohmconnect.com/

Context

Obligations imposed by governments:

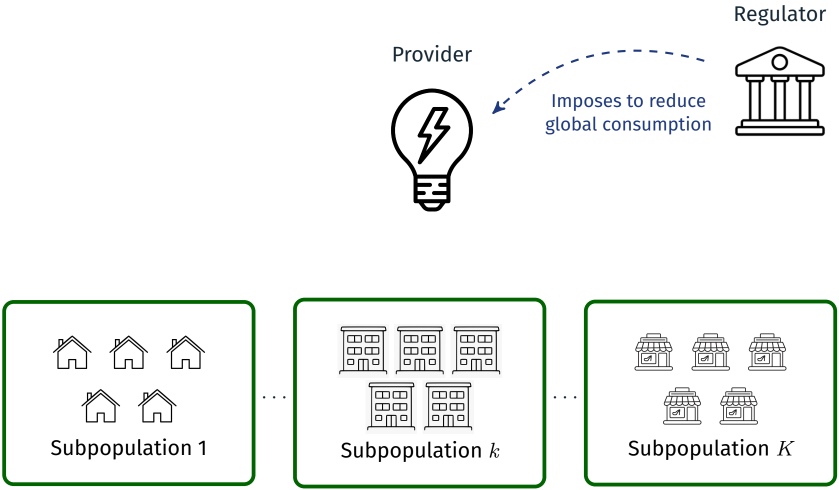
- In France: electricity providers ("*Obligés*") have a target of Energy Saving Certificates[1] to hold at a predetermined horizon ($\simeq$ 3 years).
  If they fail, they face financial penalties.

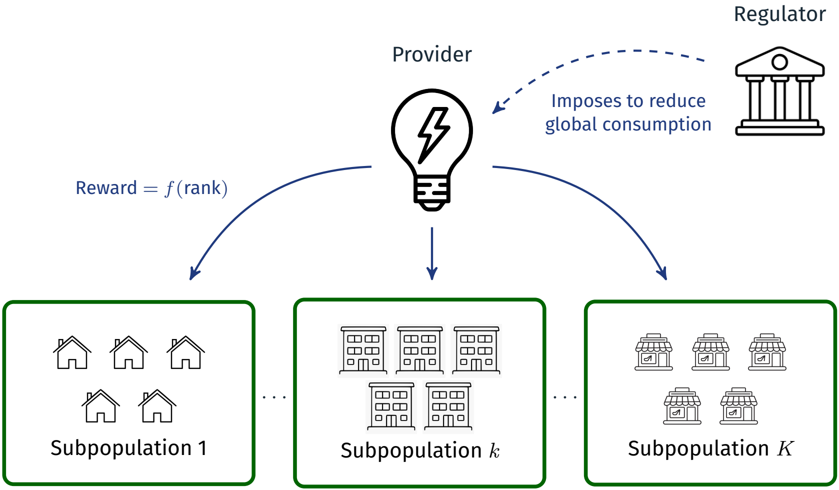
Existing incentives "Provider $\rightarrow$ customers":

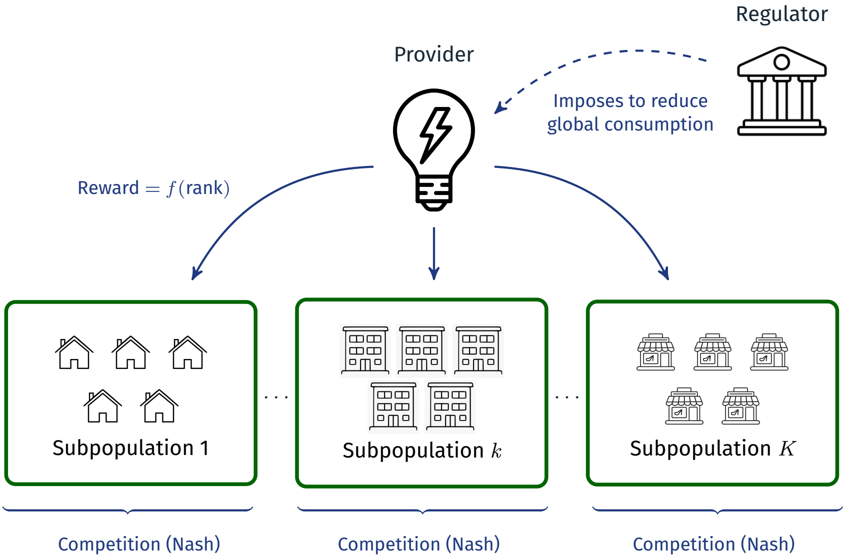
- Comparison to similar customers
  - EDF, Total, Engie, . . .
- Reward/Bonus when reduction compared to past consumption
  - "SimplyEnergy"[2], "Plüm énergie"[3], "OhmConnect"[4]

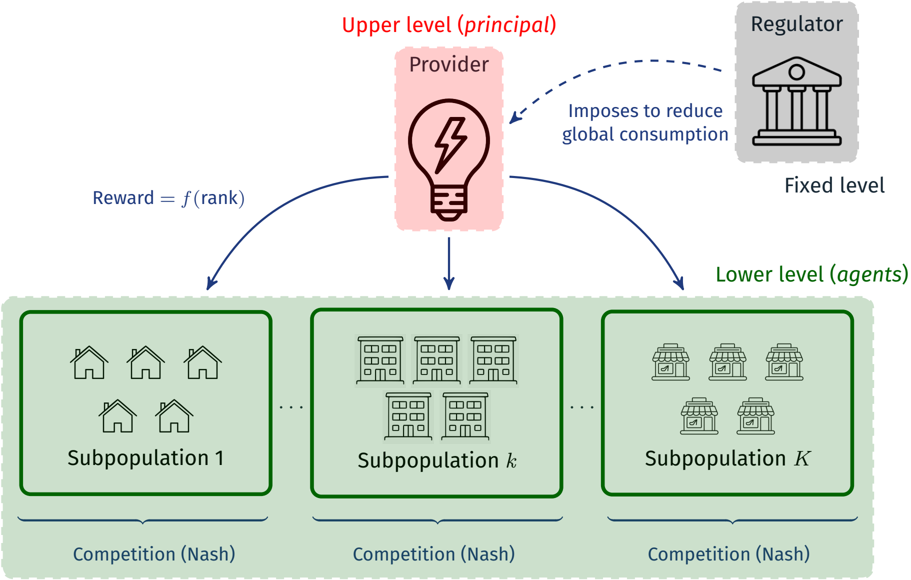
$\hookrightarrow$ Ranking games: A reward based on the comparison between similar customers

[1] www.powernext.com/french-energy-saving-certificates
[2] www.simplyenergy.com.au/residential/energy-efficiency/reduce-and-reward
[3] www.plum.fr/cagnotte/
[4] www.ohmconnect.com/


Regulator
Provider
Imposes to reduce
global consumption
Subpopulation 1
Subpopulation k
Subpopulation K


Regulator
Provider
Imposes to reduce
global consumption
Reward $= f(\text{rank})$
Subpopulation 1
Subpopulation k
Subpopulation K


Regulator
Provider
Imposes to reduce global consumption
Reward $= f(\text{rank})$
Subpopulation 1
Subpopulation k
Subpopulation K
$\cdots$
$\cdots$
Competition (Nash)
Competition (Nash)
Competition (Nash)



Mean-field assumption: Each subpopulation is composed of an *infinite* number of *indistinguishable* consumers

Section 2

Agents' problem

A field of agents at the lower level

- The population is divided into K clusters of *indistinguishable* consumers. Each cluster $k \in [K]$ represents a proportion ρ_k.
- $X_k^a(t)$ the *energy consumption* of a customer of k, forecasted at time t for consumption at $T > t$:

$$X_k^a(t) = X_k(0) + \int_0^t a_k(s)ds + \sigma_k \int_0^t dW_k(s), \quad X_k(0) = x_k^{\text{nom}} \quad , \tag{1}$$

  with
  - $\{W_k\}_{1 \le k \le K}$ a family of K independent Brownian motions
  - a_k a progressively measurable process satisfying $\mathbb{E} \int_0^T |a(s)|ds < \infty$

Interpretation:

- a_k is the consumer's *effort* to reduce his electricity consumption.
- Without effort ($a \equiv 0$), customers have a mean *nominal* consumption x_k^{nom}, and the terminal p.d.f. of $X_k^a(T)$ is:

$$f_k^{\text{nom}}(x) := \varphi\left(x;\, x_k^{\text{nom}}, \sigma_k\sqrt{T}\right) \quad ,$$

  where $\varphi(\,\cdot\,; \mu, \sigma)$ is the pdf for $\mathcal{N}(\mu, \sigma)$.

Rank-based reward

In the N-players game setting:

- each subpopulation k contains N_k players
- the *terminal ranking* of a player i, consuming $X_k^i(T)$, is measured by

$$\frac{1}{N_k}\sum_{j=1}^{N_k} \mathbb{1}_{X_k^j(T)\leq X_k^i(T)} \qquad \left(\text{empirical cumulative distribution}\right)$$

$\Rightarrow$ The reward function should be decreasing (Low rank = good energy saver)

Rank-based reward

In the N-players game setting:

- each subpopulation k contains N_k players
- the *terminal ranking* of a player i, consuming $X_k^i(T)$, is measured by

$$\frac{1}{N_k}\sum_{j=1}^{N_k} \mathbb{1}_{X_k^j(T) \le X_k^i(T)} \qquad \left(\begin{array}{c}\textit{empirical cumulative}\\ \textit{distribution}\end{array}\right)$$

$\Rightarrow$ The reward function should be decreasing (Low rank = good energy saver)

With mean-field assumption:

- If $X_k(T) \sim \mu_k$, the *terminal ranking* of a player consuming x is $r = F_\mu(x)$

Rank-based reward

In the N-players game setting:

- each subpopulation k contains N_k players
- the *terminal ranking* of a player i, consuming $X_k^i(T)$, is measured by

$$\frac{1}{N_k}\sum_{j=1}^{N_k} \mathbb{1}_{X_k^j(T) \leq X_k^i(T)} \qquad \left(\text{\textit{empirical cumulative distribution}}\right)$$

$\Rightarrow$ The reward function should be decreasing (Low rank = good energy saver)

With mean-field assumption:

- If $X_k(T) \sim \mu_k$, the *terminal ranking* of a player consuming x is $r = F_\mu(x)$

Assumption: The reward R has the form

$$\mathbb{R} \times [0,1] \ni (x, r) \mapsto R(x, r) = B(r) - px \ , \tag{2}$$

- We call R the *total reward* and B the *additional reward*.
- $-px$ represents the *natural incentive* to reduce the consumption, coming from the price p to consume one unit of energy
- When $R(x, r)$ is independent of x, the reward is *purely ranked-based*

Mean-field game between consumers

Agents' problem:
Given the reward R and the terminal consumption distribution $\tilde{\mu}_k$,

$$V_k(R, \tilde{\mu}_k) := \sup_a \mathbb{E}\left[R_{\tilde{\mu}_k}(X_k^a(T)) - \underbrace{\int_0^T c_k a_k^2(t)\, dt}_{\text{cost of effort}} \right], \qquad (P^{\text{cons}})$$

where $R_\mu(x) = R(x, F_\mu(x))$.

Interpretation:

- The cost corresponds to the purchase of new equipment (new heating installation, isolation, ...).
- In exchange, the consumer receives $B(r)$, depending on his rank $r = F_{\tilde{\mu}_k}(x)$, where $\tilde{\mu}_k$ is the k-subpopulation's distribution.
- The quantity $V_k(R, \tilde{\mu}_k)$ is called the *optimal utility* of an agent of k.

Agents' best response

Theorem (Bayraktar and Zhang, 2021,Proposition 2.1)

Given $R \in \mathcal{R}$ and $\tilde{\mu}_k \in \mathcal{P}(\mathbb{R})$, let

$$\beta_k(\tilde{\mu}) = \int_{\mathbb{R}} f_k^{\text{nom}}(x) \exp\left(\frac{R_{\tilde{\mu}}(x)}{2c_k\sigma_k^2}\right) dx \quad (< \infty) \ . \tag{3}$$

Then, the *optimal terminal distribution* μ_k^* of cluster k has p.d.f.

$$f_{\mu_k^*}(x) = \frac{1}{\beta(\tilde{\mu}_k)} f_k^{\text{nom}}(x) \exp\left(\frac{R_{\tilde{\mu}_k}(x)}{2c_k\sigma_k^2}\right) \ , \tag{4}$$

and the optimal value is then $V_k(R, \tilde{\mu}_k) = 2c_k\sigma_k^2 \ln \beta_k(\tilde{\mu}_k)$.

Definition: $\mu_k \in \mathcal{P}(\mathbb{R})$ is an *equilibrium* if it is a fixed-point of the *best response* map

$$\Phi_k \ : \ \tilde{\mu}_k \mapsto \mu_k^* \ ,$$

with μ_k^* given by (4).

Nash Equilibrium

For purely ranked-based reward (Bayraktar and Zhang, 2021, Theorem 3.2)

The equilibrium ν_k is *unique* and the quantile is given by

$$q_{\nu_k}(r) = x_k^{\text{nom}} + \sigma_k \sqrt{T} N^{-1} \left(\frac{\int_0^r \exp\left(-\frac{B(z)}{2c_k\sigma_k^2}\right) dz}{\int_0^1 \exp\left(-\frac{B(z)}{2c_k\sigma_k^2}\right) dz} \right) . \tag{5}$$

Theorem

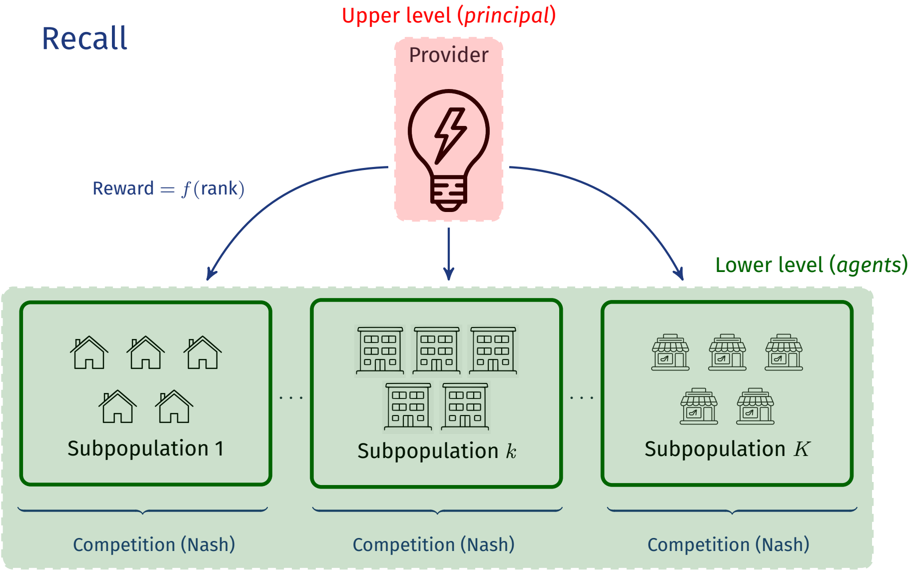
Let $R(x, r) = B(r) - px$. Then, the equilibrium μ_k is *unique*, and satisfies

$$q_{\mu_k}(r) = q_{\nu_k}(r) - \frac{pT}{2c_k} , \tag{6}$$

where ν_k is the (unique) equilibrium distribution for $p = 0$ (purely ranked-based reward), defined in (5).

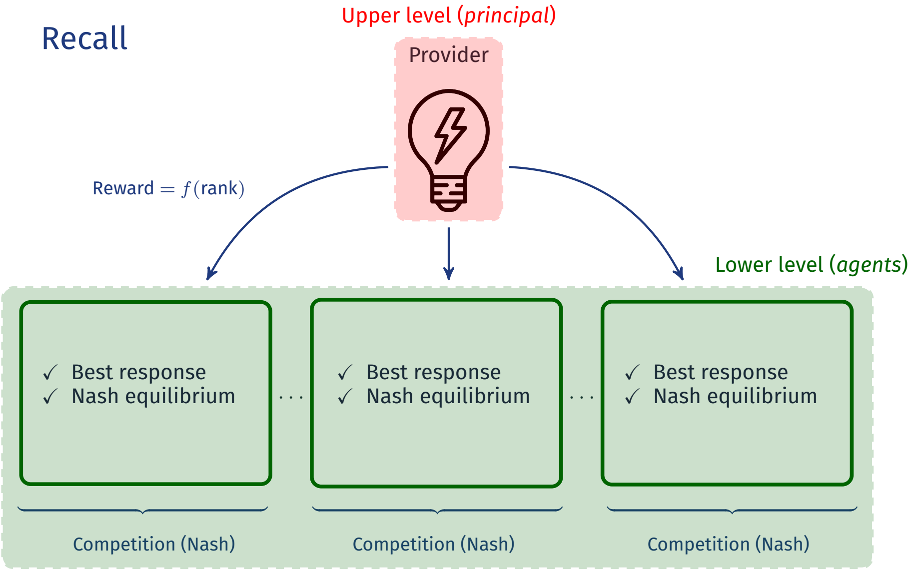
$\Rightarrow$ add of a linear part in “x” acts as a shift on the probability density function.

Recall

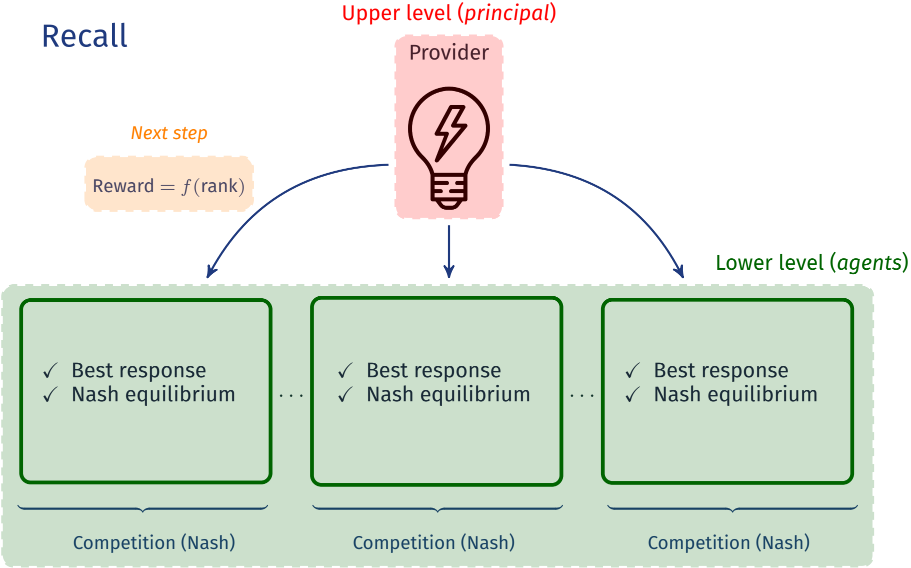


Upper level (*principal*)
Provider
Reward $= f(\text{rank})$
Lower level (*agents*)
Subpopulation 1
. . .
Subpopulation k
. . .
Subpopulation K
Competition (Nash)
Competition (Nash)
Competition (Nash)

Recall


Upper level (*principal*)
Provider
Reward = f(rank)
Lower level (*agents*)
✓ Best response
✓ Nash equilibrium
Competition (Nash)
✓ Best response
✓ Nash equilibrium
Competition (Nash)
✓ Best response
✓ Nash equilibrium
Competition (Nash)

Recall


Upper level (*principal*)
Provider
Next step
Reward $= f(\text{rank})$
Lower level (*agents*)
✓ Best response
✓ Nash equilibrium
· · ·
✓ Best response
✓ Nash equilibrium
· · ·
✓ Best response
✓ Nash equilibrium
Competition (Nash)
Competition (Nash)
Competition (Nash)

Section 3

Principal's problem

Retailer's problem

For an equilibrium $(\mu_k)_{k\in[K]}$, the mean consumption is $m_{\mu_k} = \int_0^1 q_{\mu_k}(r)\,dr$, and the overall mean consumption is $m_\mu = \sum_{k\in[K]} \rho_k m_{\mu_k}$.

Principal's problem:

$$\max_{B\in\mathcal{R}_b^r} \left\{ s\left(m_\mu\right) + (p - c_r)m_\mu - \int_0^1 B(r)\,dr \;\middle|\; \begin{array}{l} \mu_k = \epsilon_k(B) \\ V_k(B) \geq V_k^{\mathrm{pi}} \end{array} \right\} \qquad (P^{\mathrm{ret}})$$

where

- $\mathcal{R}_b^r$ is the set of *bounded* and *decreasing* rewards,
- $\mu_k = \epsilon_k(B)$ the *agents' equilibrium* given additional reward $B(\cdot)$,
- $s(\cdot)$ denotes the *valuation of the energy savings* (given by regulator),
- c_r denotes the *production cost* of energy,
- V^{pi} is the *reservation utility* (utility when $B \equiv 0$)

In the sequel, we denote by $g(\cdot)$ the function $g : m \mapsto s(m) - c_r m$.

Optimal reward – Homogeneous population ($K = 1$)

Principal's problem:

$$\max_{B \in \mathcal{R}_b^r} \left\{ s\,(m_\mu) + (p - c_r) m_\mu - \int_0^1 B(r)\,dr \;\middle|\; \begin{array}{l} \mu = \epsilon(B) \\ V(B) \geq V^{\mathrm{pi}} \end{array} \right\} \qquad (P^{\mathrm{ret}})$$

Optimal reward – Homogeneous population ($K = 1$)

Principal's problem:

$B = \epsilon^{-1}(\mu)$

Idea: $$\max_{\substack{\cancel{B\in\mathcal{R}^r_b} \\ \mu \text{ distrib.}}} \left\{ s\left(m_\mu\right) + (p - c_r)m_\mu - \int_0^1 B(r)\,dr \;\middle|\; \begin{array}{l} \cancel{\mu = \epsilon(B)} \\ V(B) \geq V^{\mathrm{pi}} \end{array} \right\} \qquad (P^{\mathrm{ret}})$$

$+B$ bounded and decreasing

Optimal reward – Homogeneous population ($K = 1$)

Principal's problem:

$$\textit{Idea:} \quad \max_{\substack{\cancel{B \in \mathcal{R}_b^r} \\ \mu \text{ distrib.}}} \left\{ s(m_\mu) + (p - c_r) m_\mu - \int_0^1 B(r)\, dr \;\middle|\; \begin{array}{l} \overset{B = \epsilon^{-1}(\mu)}{\cancel{\mu = \epsilon(B)}} \\ V(B) \geq V^{\text{pi}} \end{array} \right\} \qquad (P^{\text{ret}})$$

$+B$ bounded and decreasing

Using the characterization of the equilibrium,

$$B_\mu(r) = V^{\text{pi}} + 2c\sigma^2 \ln\left(\zeta_\mu(q_\mu(r))\right) + p q_\mu(r) \qquad \left(= \epsilon^{-1}(\mu)\right) \quad ,$$

with $\zeta_\mu := f_\mu / f^{\text{nom}}$.

Reformulation in the distribution space:

$$(P^{\text{ret}}) \begin{cases} \max\limits_{\mu} \quad g\left(\displaystyle\int_{-\infty}^{+\infty} y f_\mu(y)\, dy\right) - V^{\text{pi}} - 2c\sigma^2 \displaystyle\int_{-\infty}^{+\infty} \ln\left(\frac{f_\mu(y)}{f^{\text{nom}}(y)}\right) f_\mu(y)\, dy \\ \text{s.t.} \quad \displaystyle\int_{-\infty}^{+\infty} f_\mu(y)\, dy = 1 \\ \qquad y \mapsto \ln\left(\dfrac{f_\mu(y)}{f^{\text{nom}}(y)}\right) + \dfrac{p}{2c\sigma^2} y \text{ bounded and decreasing} \end{cases}$$

Optimal reward – Homogeneous population ($K = 1$)

Principal's problem:

$B = \epsilon^{-1}(\mu)$

Idea: $$\max_{\cancel{B \in \mathcal{R}_b^r}} \left\{ s\left(m_\mu\right) + (p - c_r) m_\mu - \int_0^1 B(r)\, dr \;\middle|\; \begin{array}{l} \cancel{\mu = \epsilon(B)} \\ V(B) \geq V^{\text{pi}} \end{array} \right\} \qquad (P^{\text{ret}})$$

μ distrib.

$+B$ bounded and decreasing

Using the characterization of the equilibrium,

$$B_\mu(r) = V^{\text{pi}} + 2c\sigma^2 \ln\left(\zeta_\mu(q_\mu(r))\right) + p q_\mu(r) \qquad \left(= \epsilon^{-1}(\mu)\right) \; ,$$

with $\zeta_\mu := f_\mu / f^{\text{nom}}$.

~~*Reformulation*~~ *in the distribution space*:

Relaxation

$$\cancel{(P^{\text{ret}})} \; (\widetilde{P}^{\text{ret}}) \begin{cases} \max\limits_{\mu} & g\left(\displaystyle\int_{-\infty}^{+\infty} y f_\mu(y)\, dy\right) - V^{\text{pi}} - 2c\sigma^2 \displaystyle\int_{-\infty}^{+\infty} \ln\left(\frac{f_\mu(y)}{f^{\text{nom}}(y)}\right) f_\mu(y)\, dy \\ \text{s.t.} & \displaystyle\int_{-\infty}^{+\infty} f_\mu(y)\, dy = 1 \\ & \cancel{y \mapsto \ln\left(\frac{f_\mu(y)}{f^{\text{nom}}(y)}\right) + \frac{p}{2c\sigma^2} y \text{ bounded and decreasing}} \end{cases}$$

Optimal reward – Homogeneous population ($K = 1$)

Assumption: The function $s : \mathbb{R} \to \mathbb{R}$ is supposed to be decreasing, concave and differentiable with $\|s'(m)\| \leq M_s$.

Lemma

The optimal distribution μ^* for ($\widetilde{P}^{\text{ret}}$) satisfies the following equation:

$$f_\mu(y) \propto f^{\text{nom}}(y) \exp\left(y \frac{g'(m_\mu)}{2c\sigma^2} \right) \tag{7}$$

Sketch of proof: Use optimality conditions, sufficient for ($\widetilde{P}^{\text{ret}}$)

Theorem – Analytic formula of the optimal reward

Let $\delta(m) = p - c_r + s'(m)$. The distribution $\mu^* \hookrightarrow \mathcal{N}(m^*, \sigma\sqrt{T})$, where m^* satisfies

$$m^* = x^{\text{pi}} + \frac{T}{2c}\delta(m^*) \ , \tag{8}$$

is optimal for ($\widetilde{P}^{\text{ret}}$) . Moreover, the associated reward B^* is

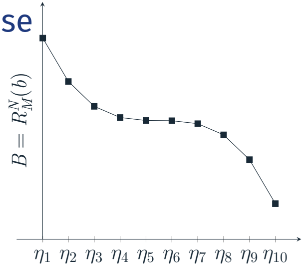
$$B^*(r) = \frac{c}{T}\left[(x^{\text{pi}})^2 - (m^*)^2 \right] + q_{\mu^*}(r)\delta(m^*) \ . \tag{9}$$

Remark: The function $\delta(\cdot)$ is viewed as the *reduction desire* of the provider.

Section 4

Numerical results

Numerical computation for general case



Restriction to piecewise linear reward:

- For $N \in \mathbb{N}$, $\Sigma_N := \{0 = \eta_1 < \eta_2 < \ldots < \eta_N = 1\}$.
- For $M \in \mathbb{R}_+$, we define the class of bounded piece-wise linear rewards adapted to Σ_N as

$$\widehat{\mathcal{R}}_M^N := \left\{ r \in [0,1] \mapsto \sum_{i=1}^{N-1} \mathbb{1}_{r \in [\eta_i, \eta_{i+1}[} \left[b_i + \frac{b_{i+1} - b_i}{\eta_{i+1} - \eta_i}(r - \eta_i) \right] \;\middle|\; \begin{array}{l} b \in [-M, M]^N \\ b_1 \geq \ldots \geq b_N \end{array} \right\} .$$

- $R_M^N(b)$ is the reward function obtained as a linear interpolation of b.

Optimization by a black-box solver:

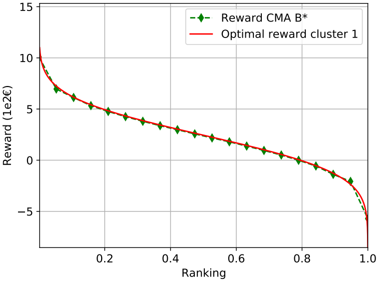
- We construct an oracle $b \in \mathbb{R}^N \mapsto \pi^{\text{ret}}(b)$, where $\pi^{\text{ret}}(b)$ is the retailer objective.
- We use a black-box solver, here CMA-ES (Hansen, 2006).

Instance

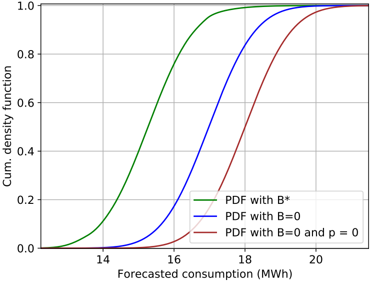
Parameter	Segment 1	Segment 2	Unit
T	3		years
p	0.17		€/kWh
c_r	0.15		€/kWh
$X(0)$	18	12	MWh
σ	0.6	0.3	MWh
c	2.5	5	€ [MWh]$^{-2}$ [years]^2
s	$m \mapsto 0.1m^2$		€
ρ	0.5	0.5	-

Table: Parameters of the instance

Results – $K = 1$



(a) Analytic optimal reward in red, compared to the reward function found by CMA



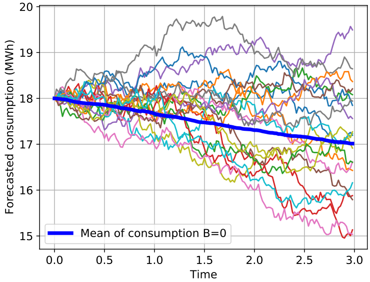
(b) Comparison of the three CDF: nominal, price incentive and with the optimal reward

Figure: Optimization in the homogeneous case

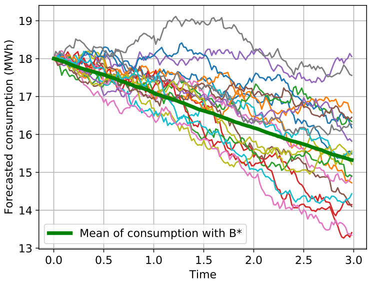
Consumption reduction:

- Nominal consumption: $x^{\text{nom}} = 18$ MWh
- With only price incentive: $x^{\text{pi}} = 17$ MWh
- With optimal reward B^*: $m = 15.4$ MWh

Results – $K = 1$



(a) Trajectories without additional reward



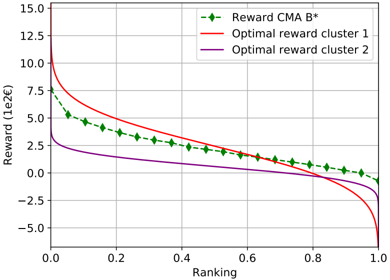
(b) Trajectories with optimal control from mean-field approximation

Figure: Trajectories for 20 consumers (homogeneous case)

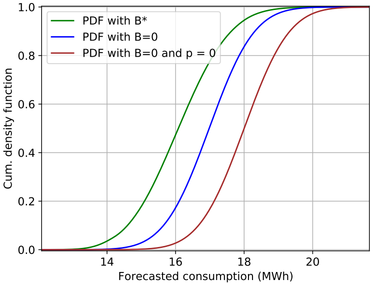
Consumption reduction:

- ◇ Nominal consumption: $x^{\text{nom}} = 18$ MWh
- ◇ With only price incentive: $x^{\text{pi}} = 17$ MWh
- ◇ With optimal reward B^*: $m = 15.4$ MWh

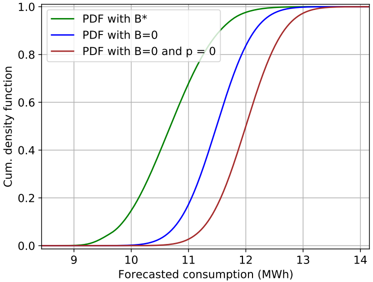
Results – $K > 1$



(a) Red and purple rewards are the optimal reward in the homogeneous case. The reward function found by CMA is displayed in green.

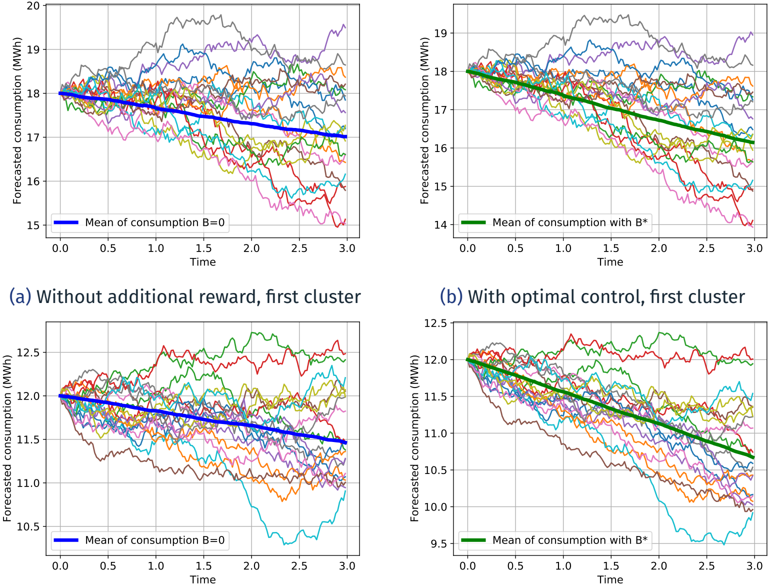


(b) Comparison of the three CDF (first cluster)



(c) Comparison of the three CDF (second cluster)

Figure: Optimization in the heterogeneous case



(a) Without additional reward, first cluster

(b) With optimal control, first cluster

(c) Without additional reward, second cluster

(d) With optimal control, second cluster

Figure: Trajectories for 20 consumers (heterogeneous case)

Section 5

Conclusion

Conclusion

Study of a *specific* framework where it is possible to

- characterize the mean-field equilibrium
- explicitly find the optimal reward ($K = 1$)
- numerically determine good reward functions ($K > 1$)

Perspectives:

- And if we can't (or don't want to) ensure *Utility* $\geq$ *Reservation utility* for all the agents ?
- More complex reward functions ?

Thank you for your attention !

SCAN ME

References

Hansen, N. (2006). The CMA evolution strategy: A comparing review. In Towards a new evolutionary computation (pp. 75–102). Springer Berlin Heidelberg.

Sannikov, Y. (2008). A continuous-time version of the principal–agent problem. Review of Economic Studies, 75(3), 957–984.

Capponi, A., Cvitanić, J., & Yolcu, T. (2012). Optimal contracting with effort and misvaluation. Mathematics and Financial Economics, 7(1), 93–128.

Adlakha, S., & Johari, R. (2013). Mean field equilibrium in dynamic games with strategic complementarities. Operations Research, 61(4), 971–989.

Fabisch, A. (2013). Cma-espp.

Chen, Y., Georgiou, T. T., & Pavon, M. (2015). On the relation between optimal transport and schrödinger bridges: A stochastic control viewpoint. Journal of Optimization Theory and Applications, 169(2), 671–691.

Ngo, H.-L., & Taguchi, D. (2015). Strong rate of convergence for the euler-maruyama approximation of stochastic differential equations with irregular coefficients. Mathematics of Computation, 85(300), 1793–1819.

Bayraktar, E., & Zhang, Y. (2016). A rank-based mean field game in the strong formulation. Electronic Communications in Probability, 21, 1–12.

References

Bayraktar, E., Cvitanić, J., & Zhang, Y. (2019). Large tournament games. The Annals of Applied Probability, 29(6).

Elie, R., Mastrolia, T., & Possamaï, D. (2019). A tale of a principal and many, many agents. Mathematics of Operations Research, 44(2), 440–467.

Bayraktar, E., & Zhang, Y. (2021). Terminal ranking games. Mathematics of Operations Research, 46(4), 1349–1365.

Carmona, R., & Wang, P. (2021). Finite-state contract theory with a principal and a field of agents. Management Science, 67(8), 4725–4741.

Gobet, E., & Grangereau, M. (2021). Extended mckean-vlasov optimal stochastic control applied to smart grid management.

Shrivats, A., Firoozi, D., & Jaimungal, S. (2021). Principal agent mean field games in rec markets.