

ROADEF 2023 (Rennes)

Optimality of Periodic Promotions for a Bilevel Problem with Customer Inertia: Application to Electricity Pricing

Quentin Jacquet, Wim van Ackooij,
Clémence Alasseur, Stéphane Gaubert

February 21, 2022



Section 1

Definition of the model

- 1 Definition of the model
 - Context
 - Bilevel pricing
 - Lifted MDP
- 2 Ergodic control
- 3 Algorithms
- 4 Application to electricity pricing

The consumer' decision *at time t*

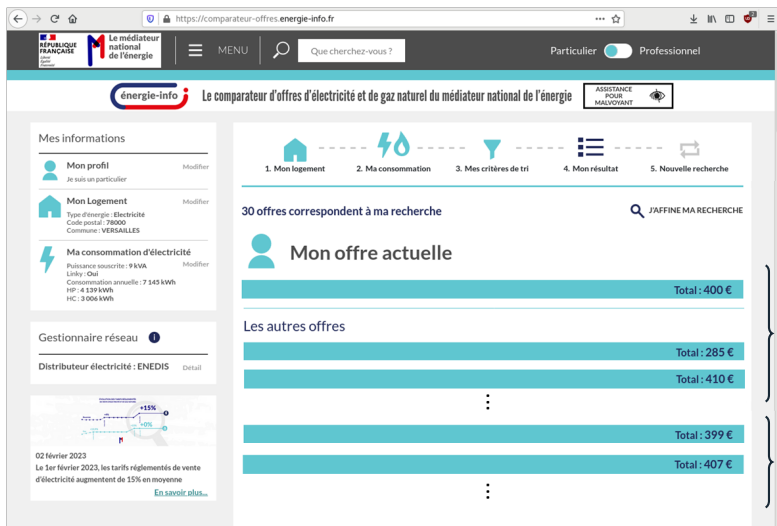


Figure: Comparison of existing offers (French electricity market)

Consumers *do not immediately* react

Intuition (Dubé et al., 2010; Horsky and Pavlidis, 2010)

“I switch to a new contract if there is a *sufficient* difference with my current offer.”

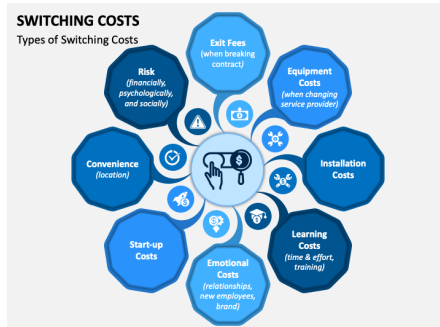


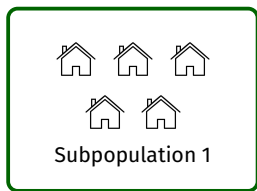
Image from <https://www.sketchbubble.com/en/presentation-switching-costs.html>

Bilevel pricing *at time t*

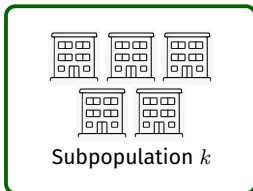
Provider



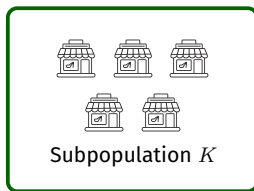
Competitors /
Regulated offers



...



...



Segmentation: Each subpopulation is composed of *similar* consumers (same behavior)

Static competition: Prices of competitors are *given* \Rightarrow *single-leader* problem

Bilevel pricing at time t

Upper level (leader)

Provider



Competitors /
Regulated offers



Fixed level

Lower level (followers)



Subpopulation 1

...



Subpopulation k

...



Subpopulation K

Segmentation: Each subpopulation is composed of *similar* consumers (same behavior)

Static competition: Prices of competitors are *given* \Rightarrow *single-leader* problem

Bilevel pricing at time t

Upper level (leader)

Decision/action:
prices of $N-1$ contracts $a_t \in \mathcal{A}$

Provider



Competitors /
Regulated offers



Fixed level

Lower level (followers)



Subpopulation 1

...



Subpopulation k

...



Subpopulation K

Segmentation: Each subpopulation is composed of *similar* consumers (same behavior)

Static competition: Prices of competitors are *given* \Rightarrow *single-leader* problem

Bilevel pricing at time t

Upper level (leader)

Decision/action:
prices of $N-1$ contracts $a_t \in \mathcal{A}$

Provider



Competitors /
Regulated offers



Fixed level

Customer k' response: $\mu_{t-1}^k \in \Delta_N$ knowing μ_{t-1}^k

Lower level (followers)



Subpopulation 1

...



Subpopulation k

...

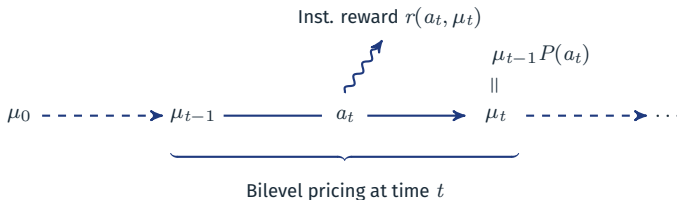


Subpopulation K

Segmentation: Each subpopulation is composed of *similar* consumers (same behavior)

Static competition: Prices of competitors are *given* \Rightarrow *single-leader* problem

High-level description as lifted MDP¹



1. *Heterogeneous population*: each cluster $k \in [K]$ represents a proportion ρ_k of the overall pop.
2. *Distribution*: $\mu_t^k \in \Delta_N$ the distribution of the population of cluster k over $[N]$.
3. *Instantaneous reward*:

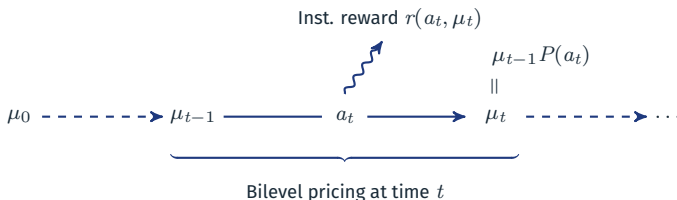
$$r : (a_t, \mu_t) \mapsto \sum_{k \in [K]} \rho_k \left\langle \theta^k(a_t), \mu_t^k \right\rangle_N$$

4. *(Linear) Transition*: $\mu_t^k = \mu_{t-1}^k P^k(a_t)$
5. *Leader's (global) objective* (average long-term reward):

$$g^*(\mu_0) = \sup_{\pi \in \Pi} \liminf_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T r(\pi_t(\mu_t), \mu_t) \quad (\text{AvR})$$

¹In a Mean-field context, see e.g. Motte and Pham, 2019

High-level description as lifted MDP¹



1. *Heterogeneous population*: each cluster $k \in [K]$ represents a proportion ρ_k of the overall pop.
2. *Distribution*: $\mu_t^k \in \Delta_N$ the distribution of the population of cluster k over $[N]$.
3. *Instantaneous reward*:

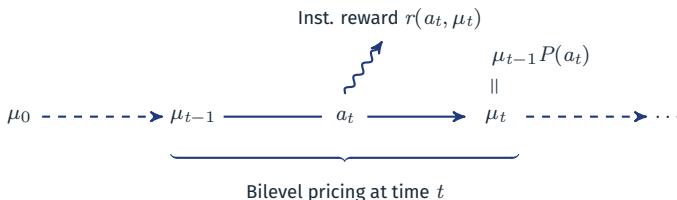
$$r : (a_t, \mu_t) \mapsto \sum_{k \in [K]} \rho_k \left\langle \theta^k(a_t), \mu_t^k \right\rangle_N \quad \leftarrow \text{upper level at time } t$$

4. *(Linear) Transition*: $\mu_t^k = \mu_{t-1}^k P^k(a_t) \quad \leftarrow \text{lower level at time } t$
5. *Leader's (global) objective* (average long-term reward):

$$g^*(\mu_0) = \sup_{\pi \in \Pi} \liminf_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T r(\pi_t(\mu_t), \mu_t) \quad (\text{AvR})$$

¹In a Mean-field context, see e.g. Motte and Pham, 2019

High-level description as lifted MDP¹



Assumptions:

- (A1) $a \mapsto P^k(a)$ is continuous,
- (A2) There exists L such that for any sequence of actions $(a_1, \dots, a_L) \in \mathcal{A}^L$,
 $\prod_{i=1}^L P(a_i) \gg 0$,
- (A2') For any action $a \in \mathcal{A}$, $P(a) \gg 0$,
- (A3) $\exists M_r$ such that, $|\theta^{kn}(a)| \leq M_r$ for every $k \in [K]$, $n \in [N]$ and $a \in \mathcal{A}$.

¹In a Mean-field context, see e.g. Motte and Pham, 2019

Specification to the Electricity Market context

Given k and an offer $n < N$, we know

- **Reservation price** R^{kn} : max. price that k want to spend on n ,
- **Energy consumption** E^{kn} : fixed consumption if k chooses n ,
- **Utility** $U^{kn}(a) := R^{kn} - E^{kn} a^n$, where a^n is the price for one unit of n .

The utility of the alternative option is normalized to 0, i.e., $U^{kN} = 0$.

The (linear) reward for the provider is then

$$\theta^{kn}(a) = \underbrace{E^{kn} a^n}_{\text{electricity invoice}} - \underbrace{C^{kn}}_{\text{cost}}, \quad n < N, \quad \theta^{kN} = 0.$$

Assumption: The transition probability follows a *logit response*, see e.g. Pavlidis and Ellickson, 2017:

$$[P^k(a)]_{n,m} = \frac{e^{\beta[U^{km}(a) + \gamma^{kn} \mathbb{1}_{m=n}]}}{\sum_{l \in [N]} e^{\beta[U^{kl}(a) + \gamma^{kn} \mathbb{1}_{l=n}]}} ,$$

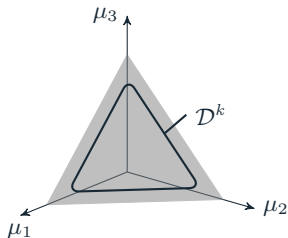
- γ^{kn} is the cost for segment k to *switch* from contract n to another one,
- β is the intensity of the choice (it can represent a “*rationality* parameter”).

Section 2

Ergodic control

- 1 Definition of the model
- 2 Ergodic control
 - Existence
 - Literature
 - Sketch of proof
- 3 Algorithms
- 4 Application to electricity pricing

Ergodic control



Let $\mathcal{D}^k := \text{vex}(\{\mu^k P_L^k(a) \mid a \in \mathcal{A}, \mu^k \in \Delta_N\})$,
and $\mathcal{D} = \times_{k \in [K]} \mathcal{D}^k$.

Lemma

Let (A1) – (A2) hold. Then $\mathcal{D}^k \subseteq \text{relint } \Delta_N^K$.
Moreover, for $t \geq 1$, $\mu_t \in \mathcal{D}$ for any policy $\pi \in \Pi$.

For $v : \Delta_N^K \rightarrow \mathbb{R}$, the *Bellman operator* \mathcal{B} is

$$\mathcal{B} v(\mu) = \max_{a \in \mathcal{A}} \{r(x, \mu) + v(\mu P(a))\}.$$

Theorem

Let (A1) – (A2) hold. Then, the *ergodic eigenproblem*

$$g \mathbf{1}_{\mathcal{D}} + h = \mathcal{B} h$$

admits a solution $g^* \in \mathbb{R}$ and h^* Lipschitz and convex on \mathcal{D} .

Moreover, g^* satisfies (AvR), and $a^*(\cdot) \in \arg \max \mathcal{B} h^*$ defines an *optimal policy*.

Deterministic MDP without controllability – the most degenerate case

	Time	Transitions	Assumption
Schweitzer, 1985	discrete	stochastic	unichain ²
Biswas, 2015	discrete	stochastic	Doebelin / minorization ³
Mallet-Paret and Nussbaum, 2002	discrete	deterministic	quasi-compactness
Fathi, 2010	continuous	deterministic	controlability ⁴
Zavidovique, 2012	discrete	deterministic	controlability
Calvez et al., 2014	continuous	deterministic	contraction of the dynamics (A2)
<i>This work</i>	discrete	deterministic	contraction of the dynamics (A2)

Standard unichain/Doebelin type conditions entail that the eigenvector is *unique*, up to an additive constant, this is *no longer true* in our case.

²the Markov Chain induced by any deterministic stationary policy consists of a single recurrent class plus a –possibly empty– set of transient states (i.e., there exists a subset of states that are visited infinitely often with probability 1 independently of the starting state)

³for all state s , action a and measurable subset B of the state space, $P(B|x, a) \geq \epsilon \mu(B)$

⁴for every pair of states (s, s') , there exists an action a making s' accessible from s

Ergodic control – Sketch of the proof (existence)

We use a contraction argument directly on the dynamics (*not on* the Bellman Operator):

Let d_H be the Hilbert's projective metric defined as

$$d_H(u, v) = \max_{1 \leq i, j \leq n} \log \left(\frac{u_i}{v_i} \frac{v_j}{u_j} \right) .$$

Under (A1) – (A2), (\mathcal{D}, d_H) is a complete metric space.

Birkhoff theorem

Every matrix $Q \gg 0$ is a contraction in Hilbert's projective metric, i.e.,

$$\forall \mu, \nu \in (\mathbb{R}_{>0}^N), \quad d_H(\mu Q, \nu Q) \leq \kappa_Q d_H(\mu, \nu) ,$$

where $\kappa_Q := \tanh(\text{Diam}_H(Q) / 4) < 1$.

We then use the method of *vanishing discount approach* (Lions et al., 1987):

→ the family of α -discounted objective function $(V_\alpha(\cdot))_\alpha$ is *equi-Lipschitz*, which entails the existence of the eigenvector by a *compactness* argument.

Section 3

Algorithms

- 1 Definition of the model
- 2 Ergodic control
- 3 Algorithms**
 - Relative Value Iteration
 - Policy Iteration
 - Numerical results
- 4 Application to electricity pricing

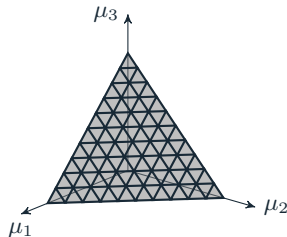
Relative Value Iteration with Krasnoselskii-Mann damping

- ◇ Regular grid Σ of the simplex Δ_N^K ,
- ◇ Bellman Operator \mathcal{B}^Σ using Freudenthal triangulation (Lovejoy, 1991).

Algorithm RVI with Mann-type iterates

Require: $\Sigma, \mathcal{B}^\Sigma, \hat{h}_0$

- 1: $v_{max} \leftarrow -\infty$
 - 2: Initialize $\hat{h} = \hat{h}_0, \hat{h}'(\mu) = \mathcal{B}^\Sigma \hat{h}$
 - 3: **while** $\text{sp}(\hat{h}' - \hat{h}) > \epsilon$ **do**
 - 4: $\hat{h} \leftarrow (\hat{h}' - \max\{\hat{h}'\}e + \hat{h})/2$
 - 5: $\hat{h}'(\hat{\mu}) \leftarrow (\mathcal{B}^\Sigma \hat{h})(\hat{\mu})$ for all $\hat{\mu} \in \Sigma$
 - 6: **end while**
 - 7: $\hat{g} \leftarrow (\max(\hat{h}' - \hat{h}) + \min(\hat{h}' - \hat{h}))/2$
 - 8: **return** \hat{g}, \hat{h}
-



Proposition (Gaubert and Stott, 2020)

Convergence time of RVI = $O(\epsilon^{-2})$

Policy Iteration

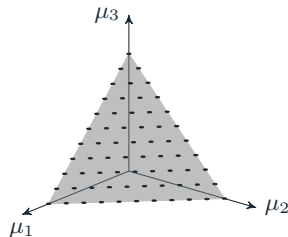
- ◇ Regular grid Σ of the simplex Δ_N^K ,
- ◇ Bellman Operator \mathcal{B}^Σ using semi-lagrangian discretization.

On-the-fly generation of transitions, refining (C.-Terrasson et al., 1998).

→ solve the spectral problem

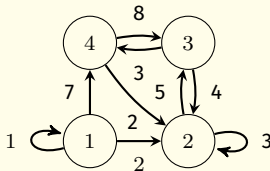
$$\max_{1 \leq j \leq n} (A_{ij} + x_j) = \lambda + x_i.$$

→ the transition is *decomposed* on each segment



Example

$$A = \begin{bmatrix} 1 & 2 & 0 & 7 \\ 0 & 3 & 5 & 0 \\ 0 & 4 & 0 & 3 \\ 0 & 2 & 8 & 0 \end{bmatrix}$$



Proposition

PI has *finite* time convergence

Numerical results

Instance ⁵	(node, arcs)	RVI-KM	PI ⁶	This work ⁷
$K = 2, N = 2$ $\delta_\mu = 1/50$	$(7.4 \cdot 10^5, 6.9 \cdot 10^8)$	7h 15Mo	390s 13Go	70s 103Mo

⁵ K : segments, N : contracts, δ_μ : discretization's precision (for each dimension)

⁶ Cochet-Terrasson et al., 1998

⁷ Each method ran on a 10 threads on a laptop i7-1065G7 CPU@1.30GHz.

Section 4

Application to electricity pricing

- 1 Definition of the model
- 2 Ergodic control
- 3 Algorithms
- 4 Application to electricity pricing**
 - Steady-states
 - Impact of switching costs

Steady-states

Theorem

Given a *constant* action a , the distribution sequence $(\mu_t^k)_t$ converges to $\bar{\mu}^k(a)$, defined as

$$\bar{\mu}^{kn}(a) = \frac{\eta^{kn}(a)\mu_L^{kn}(a)}{\sum_{l \in [N]} \eta^{kl}(a)\mu_L^{kl}(a)} . \quad (1)$$

where $\eta^{kn}(a) := 1 + \left[e^{\beta \gamma^{kn}} - 1 \right] \mu_L^{kn}(a)$, and

$$\mu_L^{kn} = e^{\beta U^{kn}(a)} / \sum_{l \in [N]} e^{\beta U^{kl}(a)} . \quad (2)$$

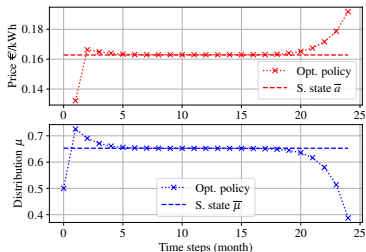
As a consequence, the optimal steady-state can be found by solving the *static* problem

$$\bar{g} = \max_{a \in \mathcal{A}} r(a, \bar{\mu}(a)) . \quad (3)$$

Impact of switching costs γ on toy model

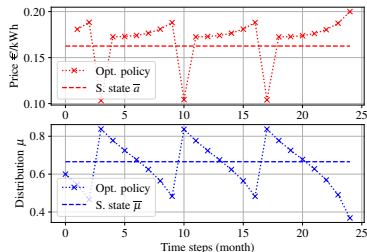


“Turnpike” like strategy:
Attraction to a steady-state



(a) Optimal finite horizon trajectory (provider action and customer distribution) for *low* switching cost.

Cyclic strategy:
A promotion is periodically applied



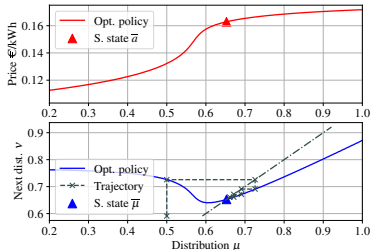
(b) Optimal finite horizon trajectory (provider action and customer distribution) for *high* switching cost.

↪ Confirms *optimality of periodic promotions*, already observed in Economics, see e.g. Horsky and Pavlidis, 2010.

Impact of switching costs γ on toy model

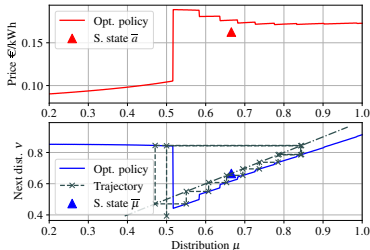


“Turnpike” like strategy:
Attraction to a steady-state



(a) Optimal decision for the long-run average reward (provider action and next customer distribution)

Cyclic strategy:
A promotion is periodically applied



(b) Optimal decision for the long-run average reward (provider action and next customer distribution)

↪ Confirms *optimality of periodic promotions*, already observed in Economics, see e.g. Horsky and Pavlidis, 2010.

Conclusion and Perspectives

Conclusion

- ◇ Resolution of deterministic lifted MDP using a eigenproblem representation
- ◇ Refinement of Policy Iteration for Heterogeneous populations
- ◇ Application to electricity pricing, and highlight of the switching cost's impact

Perspectives








- ◇ Conditions for the convergence to a steady-state
- ◇ Links between dissipativity condition (control theory) and strict subsolutions (weak-KAM theory)
- ◇ Study of other transitions (non logit-based)

↪ All this work was published here: Jacquet et al., 2022.

Thank you for your attention !



References

-  Schweitzer, P. J. (1985). On undiscounted markovian decision processes with compact action spaces. [RAIRO - Operations Research - Recherche Opérationnelle](#), 19(1), 71–86.
-  Lions, P.-L., Papanicolaou, G., & Varadhan, S. (1987). [Homogenization of hamilton-jacobi equation](#).
-  Lovejoy, W. S. (1991). Computationally feasible bounds for partially observed markov decision processes. [Operations Research](#), 39(1), 162–175.
-  Cochet-Terrasson, J., Cohen, G., Gaubert, S., McGettrick, M., & Quadrat, J.-P. (1998). Numerical computation of spectral elements in max-plus algebra. [IFAC Proceedings Volumes](#), 31(18), 667–674.
-  Mallet-Paret, J., & Nussbaum, R. (2002). Eigenvalues for a class of homogeneous cone maps arising from max-plus operators. [Discrete and Continuous Dynamical Systems](#), 8(3), 519–562.
-  Dubé, J.-P., Hitsch, G. J., & Rossi, P. E. (2010). State dependence and alternative explanations for consumer inertia. [The RAND Journal of Economics](#), 41(3), 417–445.
-  Fathi, A. (2010). [The weak-KAM theorem in lagrangian dynamics](#) [Book to appear].

References



Horsky, D., & Pavlidis, P. (2010). Brand loyalty induced price promotions: An empirical investigation. [SSRN Electronic Journal](#).



Zavidovique, M. (2012). Strict sub-solutions and mañé potential in discrete weak KAM theory. [Commentarii Mathematici Helvetici](#), 1–39.



Calvez, V., Gabriel, P., & Gaubert, S. (2014). Non-linear eigenvalue problems arising from growth maximization of positive linear dynamical systems. [Proceedings of the 53rd IEEE Annual Conference on Decision and Control \(CDC\), Los](#) 1600–1607.



Biswas, A. (2015). Mean field games with ergodic cost for discrete time markov processes.



Pavlidis, P., & Ellickson, P. B. (2017). Implications of parent brand inertia for multiproduct pricing. [Quantitative Marketing and Economics](#), 15(4), 369–407.



Motte, M., & Pham, H. (2019). Mean-field markov decision processes with common noise and open-loop controls.



Gaubert, S., & Stott, N. (2020). A convergent hierarchy of non-linear eigenproblems to compute the joint spectral radius of nonnegative matrices. [Mathematical Control & Related Fields](#), 10(3), 573–590.

References



Jacquet, Q., van Ackooij, W., Alasseur, C., & Gaubert, S. (2022). Ergodic control of a heterogeneous population and application to electricity pricing. [2022 IEEE 61st Conference on Decision and Control \(CDC\)](#).