



**HAL**  
open science

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

Clémence Alasseur, Erhan Bayraktar, Roxana Dumitrescu, Quentin Jacquet

► **To cite this version:**

Clémence Alasseur, Erhan Bayraktar, Roxana Dumitrescu, Quentin Jacquet. A Rank-Based Reward between a Principal and a Field of Agents: Application to Energy Savings. 2023. hal-03770115v2

**HAL Id: hal-03770115**

**<https://hal.science/hal-03770115v2>**

Preprint submitted on 31 Jul 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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

Clémence Alasseur<sup>1</sup> Erhan Bayraktar<sup>2</sup> Roxana Dumitrescu<sup>3</sup> Quentin Jacquet<sup>1,4</sup>

<sup>1</sup> EDF Lab Saclay, Palaiseau, France

{clemence.alasseur,quentin.jacquet}@edf.fr

<sup>2</sup>Department of Mathematics, University of Michigan, USA

erhan@umich.edu

<sup>3</sup>Department of Mathematics, King's College London, United Kingdom

roxana.dumitrescu@kcl.ac.uk

<sup>4</sup>INRIA, CMAP, Ecole Polytechnique, Palaiseau, France

## Abstract

In this paper, we consider the problem of a Principal aiming at designing a reward function for a population of heterogeneous agents. We construct an incentive based on the ranking of the agents, so that a competition among the latter is initiated. We place ourselves in the limit setting of mean-field type interactions and prove the existence and uniqueness of the equilibrium distribution for a given reward, for which we can find an explicit representation. Focusing first on the homogeneous setting, we characterize the optimal reward function using a convex reformulation of the problem and provide an interpretation of its behaviour. We then show that this characterization still holds for a sub-class of heterogeneous populations. For the general case, we propose a convergent numerical method which fully exploits the characterization of the mean-field equilibrium. We develop a case study related to the French market of Energy Saving Certificates based on the use of realistic data, which shows that the ranking system allows to achieve the sobriety target imposed by the European commission.

**Keywords:** Ranking games, Principal-Agent problem, Mean-field games, Energy savings

## 1 Introduction

### 1.1 Motivation

In Europe, energy retailers have incentives to generate energy consumption savings at the scale of their customer portfolio. For example in France, since 2006, power retailers – called *Obligés* – have a target of a certain amount of Energy Saving Certificates<sup>1</sup> to hold at a predetermined future date (usually 3 or

---

E. Bayraktar is partially supported by the National Science Foundation under grant DMS-2106556 and by the Susan M. Smith chair.

<sup>1</sup><https://www.powernext.com/french-energy-saving-certificates>

4 years). If they fail to obtain this number of certificates, then they face financial penalties. Certificates can be acquired either by certifying energy savings at the customer or by buying certificates on the market. If a retailer holds more certificates than its target at the end of the period, the surplus can be sold on the Energy Saving Certificates market. The pluri-annual energy savings goal is determined by the government, and is function of the cumulative discounted amount of energy saved (thanks to thermal renovation for instance)<sup>2</sup>. Similar mechanisms – called *White certificates* – have been implemented in several countries in Europe (Great Britain, Italy or Denmark).

There is evidence from behavioral economics that energy consumption reductions can be motivated by providing a financial reward and/or information on social norms or comparison to customers, see e.g. see [AT14] or [DM15]. Especially, in [DM15], the authors find that social norms reduce consumption by around 6% (0.2 standard deviations). Secondly, they obtain that large financial rewards for targeted consumption reductions work very well in reducing consumption, with a 8% reduction (0.35 standard deviations) in energy consumption. For recent years, electricity providers are aware of this lever to make energy savings, and contracts offering bonus/rewards in compensation of reduction efforts appear, see e.g. the offers of “SimplyEnergy”<sup>3</sup>, “Plüm énergie”<sup>4</sup> or “OhmConnect”<sup>5</sup>. The interest of this kind of solutions is reinforced in the current situation of gas and power shortage where many countries intend to diminish their global energy consumption<sup>6</sup>.

## 1.2 Contributions

In this paper, we design a monetary reward based on the *rank* of each consumer. In our context, the rank measures the reduction effort of a consumer compared with the rest of the population (a rank  $r \in [0, 1]$  indicates that the consumer is among the  $r$  percent of the population with the highest consumption reduction). This new mechanism initiates a competition between similar consumers to be the best energy saver and unites the incentive potential of rankings with a financial reward.

We suppose that the interaction between the consumers is of mean-field type, i.e., the number of consumers is infinite. This choice is motivated by our application, where the game is played across a country (for e.g. around 30 millions of households in France). Given the reward, the problem reduces to a mean-field game. Our first main result is to characterize the (unique) mean-field Nash equilibrium of this game for rewards that linearly depends on the terminal consumption (Theorem 2.4).

We then study the Principal-Agent relation (Stackelberg game) between the provider and the population of consumers. We introduce the bi-level problem solved by the retailer, aiming at maximizing over reward functions the profit made on the whole time period, taking into account the consumption distribution achieved at the equilibrium. Our second main result is to derive a semi-explicit formula of the optimal reward in the homogeneous setting (Theorem 2.8), which follows by solving of a fixed-point equation. This relies on a convex reformulation of the problem, obtained by transforming the latter into an optimization over equilibrium distributions, and by expressing the sufficient optimality conditions for the reformulated problem. We show that the unique optimal reward can be approximated by a bounded

---

<sup>2</sup><https://www.ecologie.gouv.fr/dispositif-des-certificats-deconomies-denergie>

<sup>3</sup><https://www.simplyenergy.com.au/residential/energy-efficiency/reduce-and-reward>

<sup>4</sup><https://plum.fr/cagnotte/>

<sup>5</sup><https://www.ohmconnect.com/>

<sup>6</sup><https://www.politico.eu/article/eu-countries-save-energy-winter/>

function, where the sub-optimality of the latter is controlled and converges to zero for sufficiently large bounds (Corollary 2.9). In the general setting (heterogeneous population), we show that under uniform price elasticity and uniform relative volatility, the problem reduces to the previous case (Proposition 2.11). For the more general case, the reformulation in the distributions space does not apply, and we introduce a numerical algorithm (Algorithm 1) to optimize the shape of the reward. This black-box optimization procedure relies on a fast evaluation of the retailer objective function at each iteration, which is done by exploiting the characterization of the mean-field Nash equilibrium.

We then apply our approach to the French market of Energy Saving Certificates using realistic data (Section 4). We show that the numerical procedure exhibits a fast convergence, and successfully finds the optimal reward in the homogeneous setting, and provides significant consumption reduction in the general setting, while maintaining the satisfaction (utility) of the consumers. We also simulate some trajectories of consumers by using the reward found in the mean-field context to highlight the energy reduction capacity of this mechanism. In particular, we show that the ranking system allows to achieve the sobriety target imposed by the European commission.

Finally, we consider several extensions suitable to our context. First, we show that, for the class of reward functions considered here, the addition of common-noise in the consumption process only shifts the equilibrium distribution by a (random) constant. Besides, we focus on time-dependent costs of effort for the agents, reflecting the collective awareness of agents on the energy reduction’s necessity. We are also able to provide some invariance results, which show that the use of more sophisticated reward (a function that jointly depends on the rank and the consumption of the agent) is, at the equilibrium, equivalent to a reward that belongs to the class of purely rank-based rewards.

### 1.3 Related Works

Given the reward function provided by the retailer, the competition between agents is modeled by a mean-field game. These games have been introduced simultaneously by Lasry and Lions [LL06a; LL06b; LL07] and Huang, Caines and Malhame [HMC06; HCM07]. They refer to the study of differential games involving a large number of indistinguishable agents which interact through their empirical distribution. By looking at the limit case where a continuum of agents is involved, each of them asymptotically negligible, mean-field games provide efficient ways to compute approximations of Nash equilibria for stochastic games with large number of players (games which are otherwise rarely tractable). Among various techniques, the problem is often solved by a fixed-point method involving both a Hamilton-Jacobi-Bellman equation – characterizing the agents best response to a given population distribution – and a Fokker-Planck equation. Existence and uniqueness of a mean-field equilibrium are then analyzed through this system of coupled partial differential equations, see e.g. [Car+15; BCS17].

The design of a reward/incentive by the retailer is then modeled as a *Principal-Agent* problem, see e.g. the works of Sannikov [San08] and Capponi, Cvitanic and Yolcu [CCY12] in continuous-time settings. In such problems, the Principal (retailer) aims at designing a monetary reward that is offered to the agent, depending on the quantity of work achieved by the latter. In energy management, Aid, Possamaï and Touzi introduces an incentive mechanism to control both the average consumption and the volatility of the agents consumption. The additional difficulty in our context is the presence of a continuum of agents, and the interaction between them which is expressed in terms of a mean-field game. Such extensions

of the Principal-Agent problem have been considered by Elie, Mastolia and Possamaï [EMP19] – where an explicit contract has been found for a specific class of dynamics (encompassing the linear-quadratic setting) – and by Carmona and Wang [CW21] – focusing on the linear-quadratic setting and finite-state spaces. Shrivats, Firoozi and Jaimungal [SFJ21] introduce a Principal-Agent formulation to study the interaction between a regulator and a field of providers in the market of Renewable Energy Certificate (REC).

Our study is inspired by several works. We focus on rank-based interactions, previously introduced in [BZ16], where results of existence and uniqueness of the mean-field Nash equilibrium are provided for a general class of rewards. Extensions to principal-agent problem are then studied in [BCZ19; BZ21], deriving explicit expressions of optimal contract for several principal’s objectives (profit/effort/rank-performance maximization/distribution target). In comparison to these works, we provide new theoretical results for non purely rank-based reward in the case of a homogeneous population and general convex cost functions, and extends the latter to a sub-class of heterogeneous population, while keeping explicit characterizations of equilibria and optimal rewards. Finding such explicit expressions is rare in the literature, and is only possible by imposing a specific dynamics (as in [EMP19] and [CW21]). Another additional difficulty which arises from the application is to take into account the diversity of the agents: here, we consider that the overall population is clustered into a finite number of (infinite-size) independent sub-populations. This heterogeneous context (in absence of uniform elasticity) increases further the difficulty – both on analytic and numerical aspects – but is necessary on the application side for realism purposes, see e.g. [SFJ21; SFJ22] for applications of mean-field games to REC markets. In [Cam+21], Campbell et al. introduce deep learning algorithms to solve principal-agent mean field games under heterogeneity of agent types. Here, we propose an alternative method, which takes advantage of the specific structure of the problem (explicit solution of the underlying mean-field game and common rank-based reward across the sub-populations) to lower the numerical complexity and derive efficient computational methods.

The rest of the paper is organized as follows: in Section 2, we first define the model and characterize the equilibrium for the mean-field game between the agents. In Section 3, we propose a numerical approach to solve the problem in the heterogeneous setting, for which the convex reformulation seems not extendable. In Section 4, we apply the results to the French market of Energy Savings Certificates, and finally in Section 5, we tackle some extensions that naturally arise in the context of the application. *The proofs of the main results are given in the appendix.*

## 2 Model

### 2.1 Notation and Assumptions

In the sequel, we denote by  $\mathcal{P}(\mathbb{R})$  the set of distributions defined on  $\mathbb{R}$  and by  $\mathcal{P}^+(\mathbb{R})$  the set of distributions having strictly positive density. Moreover, for any  $\mu \in \mathcal{P}(\mathbb{R})$ ,  $F_\mu$  refers to the cumulative distribution function (cdf) of  $\mu$ , and when it exists,  $f_\mu$  (resp.  $q_\mu$ ) refers to the probability density function (pdf) (resp. the quantile function) of  $\mu$ . Moreover, we write  $X \sim \mu$  when  $X$  is distributed according to  $\mu \in \mathcal{P}(\mathbb{R})$ . The normal distribution centered in  $m$  with standard deviation  $\sigma$  is denoted by  $\mathcal{N}(m, \sigma)$  and its pdf is denoted by  $x \mapsto \varphi(x; m, \sigma)$ .

Let us successively introduce the different players involved in the Stackelberg game:

**Consumers.** We consider a *heterogeneous* population of consumers, and we suppose that a clustering algorithm can be applied as a preprocessing step in order to split the population into  $K$  sub-populations (or clusters), each of them composed of similar customers. Each cluster  $k \in [K] := \{1, \dots, K\}$  represents a proportion  $\rho_k$  of the overall population and corresponds to a given class of customers, categorized for example according to their usages, their heating system or the household composition. Here, we directly tackle mean-field interactions between the agents:

**Assumption 2.1.** *We assume that each sub-population is composed of an infinite number of indistinguishable agents, represented by a single consumer (representative agent).*

*Energy consumption.* Let  $(\Omega, \mathbb{F}, \mathbb{P})$  be a complete filtered probability space, which supports a family of  $K$  independent Brownian motions  $\{W_k\}_{1 \leq k \leq K}$ , and  $\mathbb{A}$  be the set of progressively measurable processes  $a$  satisfying the integrability condition  $\mathbb{E} \int_0^T |a(s)| ds < \infty$ . For a given control  $a_k \in \mathbb{A}$ , we denote by  $X_k^a(t)$  the forecasted energy consumption of an agent from cluster  $k$  (typically an household), forecast made at time  $t < T$  for the time period  $[0, T]$ . The controlled process  $X_k^a$  is described through the following stochastic differential equation:

$$\begin{cases} dX_k^a(t) = a_k(t)dt + \sigma_k dW_k(t), \\ X_k^a(0) = x_k^{\text{nom}}. \end{cases} \quad (1)$$

Here, we consider an arithmetic Brownian motion in the dynamics, expressing the uncertainty in the electricity needs. The use of such arithmetic noise (specific to Ornstein-Uhlenbeck processes) has been showed to be relevant for load modeling, see e.g. [RSM16]. Aïd, Possamai and Touzi considered in [APT22] the multidimensional version of this dynamics, to the same purpose of representing the electricity consumption. The process  $a_k$  in Equation (1) is then viewed as the consumer's effort to reduce his electricity consumption. Without any effort, customers are expected to have a *nominal* consumption of  $x_k^{\text{nom}}$ , and we define by  $f_k^{\text{nom}}$  the p.d.f. of  $X_k^a(T)$  under a zero effort ( $a_k$  is a constant process equals to 0):

$$f_k^{\text{nom}}(x) := \varphi\left(x; x_k^{\text{nom}}, \sigma_k \sqrt{T}\right). \quad (2)$$

Note that we do not explicitly impose bounds on the process  $X_k$  – typically non-negativity assumption – but this will be naturally enforced by the cost of effort and the volatility parameter  $\sigma_k$  so that the probability of negative consumption will be negligible.

**Retailer.** In this model, an electricity provider, incentivised by a regulation agency, aims at designing a reward function based on the *terminal ranking* of the agents in order to lower the global consumption of the customers: considering that the terminal consumption of the agents in the  $k$ th population, i.e.  $X_k^a(T)$ , is distributed according to  $\mu_k$ , the ranking  $r$  of a player consuming the quantity  $x$ , is measured by the fraction of agents consuming less than  $x$ , i.e.,  $r = F_\mu(x)$ , where  $F_\mu$  denotes the cumulative distribution function on  $\mu$  (so that the worst performer/the highest consumption has rank one and the top performer has rank 0).

A reward function in our context is then a continuous real-valued function  $\mathbb{R} \times [0, 1] \ni (x, r) \mapsto R(x, r)$  that depends both on the terminal consumption  $x$  and the terminal ranking  $r$ . We consider only rewards that are non-increasing in both arguments, to favor low ranks. For any  $\mu \in \mathcal{P}(\mathbb{R})$ , we write  $R_\mu(x) =$

$R(x, F_\mu(x))$  and when  $R(x, r)$  is independent of  $x$ , we say that the reward is *purely rank-based*. In the sequel, we will consider the following decomposition assumption:

**Assumption 2.2.** *Each sub-population  $k \in [k]$  receives a reward  $R_k$  has the form*

$$R_k(x, r) = B_k(r) - px \quad , \quad (3)$$

where  $p \in \mathbb{R}$  and  $B_k \in \mathcal{B}$  with  $\mathcal{B}$  the set of purely rank-based (decreasing) functions. We then call  $R$  the total reward and its rank-dependent part  $B_k$  the additional reward (financial “bonus” for the consumer).

In the energy context, the second member “ $-px$ ” represents the classic invoice of the consumer, where  $p$  is the price to consume one unit of energy (e.g. in €/kWh). Here, this simple pricing strategy can be viewed as a regulated price (as this is the case in France for example<sup>7</sup>). The invoice is embedded in the reward function since it acts as a natural incentive to reduce the consumption. The first member  $B_k$  is then the additional financial reward offered to consumers based on their terminal ranking.

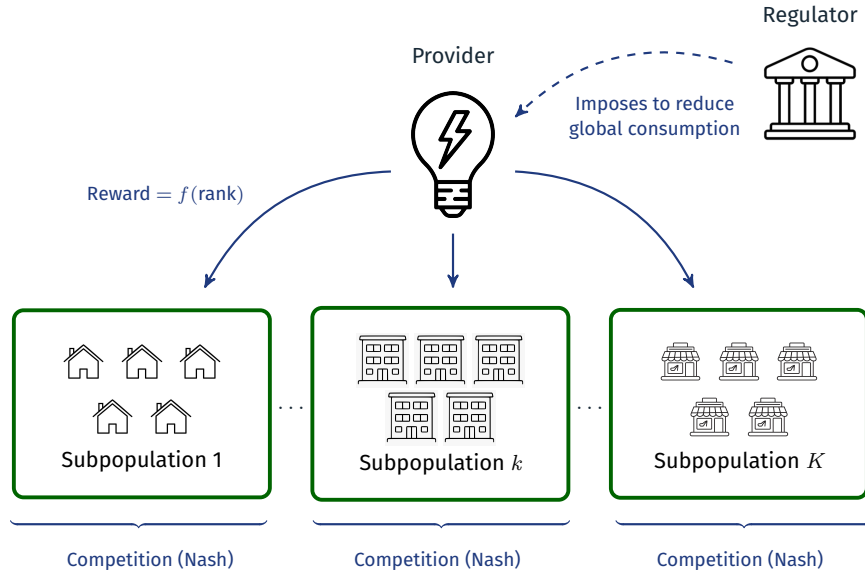


Figure 1: Relation between the Principal (provider) and the sub-populations composed of an infinite number of Agents (consumers).

In the modeling of energy consumption, a common-noise is often added (it can represent for example the outdoor temperature). However, we show that the insertion of such a noise only shifts the consumption distribution (by a random constant). This result was already mentioned for translation invariant functions (such as purely rank-based rewards), and we extend in Section 5 this property to the more general class of reward functions satisfying Assumption 2.2.

**Assumption 2.3** (Fair reward mechanism).

(i) *Each cluster is independent: the rank of an agent of cluster  $k \in [K]$  is only determined by the distribution of the cluster  $k$ .*

(ii) *The same unitary bonus is proposed to each cluster, i.e.,  $B_k(r) = x_k^{\text{nom}} \beta(r)$  for all  $k \in [K]$ .*

<sup>7</sup>“Tarif réglementé de vente” (TRV)

Assumption 2.3 imposes that the sub-populations evolve separately, but are linked through a common reward function. This assumption is taken for the sake of a fair reward mechanism: on one hand, consumers only compete with similar agents, i.e., with agents having the same characteristics (type of heating, household composition, ...) and on the other hand, the shape of the reward should be identical for each the sub-population to prevent from favoring one cluster compared to another. The function  $\beta$  is then the unitary bonus received by every customer (in €/kWh).

Figure 1 outlines the Principal-Agent relation between the retailer and the field of consumers. We then first focus on the competition among the agents before studying the principal problem.

## 2.2 Mean-field game between agents

In all this section, let us fix a cluster  $k \in [K]$ , as there is no interaction between clusters. We suppose here that the reward  $R_k(x, r)$  is given.

An agent of  $k$  is able to produce an effort  $a_k$  to reduce its consumption, but has to pay as a counterpart the quadratic cost  $c_k a_k^2(t)$  with  $c_k > 0$  a given positive constant. The convexity of the effort cost is natural in the context of our application. In particular, this cost either corresponds to the purchase of new equipment that is more efficient than the older one (new heating installation, isolation, ...) or corresponds to a change in the consumption pattern (sobriety). In the latter case, the convexity illustrates that small efforts (as for e.g. switching off the light when leaving a room) are easy to make while large consumption reduction (as for e.g. reducing heating or air conditioning) are more demanding. It is also possible to consider a more general convex cost, which is in non-quadratic form, since it would still lead to a tractable agent problem. However, quadratic costs are often considered in order to obtain explicit expression of the optimum, see e.g. [APT22] in the electricity context. In exchange of the effort, the consumer receives the reward  $R_k(x, r)$ , depending on his rank  $r = F_{\mu_k}(x)$  within the sub-population, where  $\mu_k$  is the  $k$ -sub-population distribution. His objective is then:

$$V_k(R_k, \mu_k) := \sup_{a \in \mathbb{A}} \mathbb{E} \left[ R_{k, \mu_k}(X_k^a(T)) - \int_0^T c_k a_k^2(t) dt \right] . \quad (P^{\text{cons}})$$

The quantity  $V_k(R_k, \mu_k)$  represents the *optimal expected utility* of an agent of class  $k$ , for a given provider's reward and population distribution.

We present below some results which will be used throughout the paper. The first theorem gives the explicit solution of the agent's best response to a population distribution  $\tilde{\mu}_k$ :

**Theorem 2.1** (Characterization of the best response). *Given a bounded total reward function  $R_k$  satisfying Assumption 2.2 and  $\tilde{\mu}_k \in \mathcal{P}(\mathbb{R})$ , let*

$$\gamma_k(\tilde{\mu}) = \int_{\mathbb{R}} f_k^{\text{nom}}(x) \exp\left(\frac{R_{k, \tilde{\mu}}(x)}{2c_k \sigma_k^2}\right) dx \quad (< \infty) . \quad (4)$$

*Then, the optimal terminal distribution  $\mu_k^*$  of a player from cluster  $k$  admits a pdf defined as*

$$f_{\mu_k^*}(x) = \frac{1}{\gamma(\tilde{\mu}_k)} f_k^{\text{nom}}(x) \exp\left(\frac{R_{k, \tilde{\mu}_k}(x)}{2c_k \sigma_k^2}\right) , \quad (5)$$



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

The above result corresponds to [BZ21, Proposition 2.1] and is obtained using the Schrödinger bridge approach, see [CGP15] for connections with optimal transport theory. The consumption process  $X_k$  under the optimal effort then satisfies the equation

$$dX_k(t) = a_k(t, X_k(t); \mu_k^*)dt + \sigma_k dW_k(t),$$

where the optimal effort  $a_k(\cdot, \cdot; \mu_k^*)$  is defined as

$$a_k(t, x, \mu) = \sigma_k^2 \partial_x \ln u_k(t, x, \mu), \quad u_k(t, x, \mu) = \mathbb{E} \left[ \exp \left( \frac{1}{2c_k\sigma_k^2} R_{k,\mu}(x + \sigma_k\sqrt{T-t}Z) \right) \right], \quad Z \sim \mathcal{N}(0, 1) . \quad (6)$$

We now introduce the notion of mean-field Nash equilibrium.

**Definition 2.2** (Mean-field Nash equilibrium). *We say that  $\mu_k \in \mathcal{P}(\mathbb{R})$  is an equilibrium (terminal distribution) if it is a fixed-point of the mapping  $\Phi_k : \tilde{\mu}_k \mapsto \mu_k^*$ , with  $\mu_k^*$  given by the solution of the equation (5).*

The existence of such an equilibrium has been proved in the general setting using Schauder's fixed point theorem (see [BZ16]). We give below a characterization of this equilibrium distribution, as well as an explicit expression for purely rank-based rewards:

**Theorem 2.3** (Characterization of the equilibrium distribution). *Given a bounded total reward function  $R_k : \mathbb{R} \times [0, 1] \rightarrow \mathbb{R}$ , the distribution  $\mu_k \in \mathcal{P}(\mathbb{R})$  is an equilibrium terminal distribution for cluster  $k$  if and only if its quantile function  $q_{\mu_k}$  satisfies*

$$N \left( \frac{q_{\mu_k}(r) - x_k^{\text{nom}}}{\sigma_k\sqrt{T}} \right) = \frac{\int_0^r \exp \left( -\frac{R_{k,\mu_k}(q_{\mu_k}(z))}{2c_k\sigma_k^2} \right) dz}{\int_0^1 \exp \left( -\frac{R_{k,\mu_k}(q_{\mu_k}(z))}{2c_k\sigma_k^2} \right) dz}, \quad (7)$$

where  $N$  is the standard normal c.d.f. In the specific case of a purely rank-based reward, we obtain that the equilibrium  $\nu_k$  is unique and the quantile is given by

$$q_{\nu_k}(r) = x_k^{\text{nom}} + \sigma_k\sqrt{T}N^{-1} \left( \frac{\int_0^r \exp \left( -\frac{B_k(z)}{2c_k\sigma_k^2} \right) dz}{\int_0^1 \exp \left( -\frac{B_k(z)}{2c_k\sigma_k^2} \right) dz} \right). \quad (8)$$

The mean consumption at the equilibrium is then  $m_{\mu_k} = \int_0^1 q_{\mu_k}(r)dr$ .

The above result is provided in [BZ21, Theorem 3.2], and below we extend the explicit characterization to the more general case of reward maps  $R$ , which not only depend on the rank, but also have a linear dependence on  $x$ .

**Theorem 2.4** (Explicit characterization for non purely rank-based rewards). *Suppose the reward is of the form defined in Assumption 2.2. Then, the equilibrium  $\mu_k$  is unique, and it satisfies*

$$q_{\mu_k}(r) = q_{\nu_k}(r) - \frac{pT}{2c_k}, \quad (9)$$

where  $\nu_k$  is the (unique) equilibrium distribution for the specific case  $p = 0$  (purely rank-based reward), defined in (8).

Theorem 2.4 shows that the addition of a linear part in the consumption acts as a shift on the probability density function. We emphasize that our uniqueness result of the equilibrium  $\mu$  generalizes the one established in [BZ21], the latter being obtained under the additional assumptions that the map  $r \mapsto R_k(x, r)$  is convex and  $r \mapsto \partial_x R_k(x, r)$  is non decreasing. Instead, we assume a linear dependence on the consumption for the reward, but no convexity requirement is made on its purely rank-based component  $B$ .

**Corollary 2.5** (Equilibrium without additional reward). *For  $R_k(x, r) = -px$ , the equilibrium follows the normal distribution  $\mathcal{N}(x_k^{\text{pi}}, \sigma_k \sqrt{T})$ , where  $x_k^{\text{pi}} = x_k^{\text{nom}} - \frac{pT}{2c_k}$  is the consumption under the natural incentive associated with the price  $p$ . Moreover, the optimal consumer's utility is*

$$V_k(R, \mu_k) = V_k^{\text{pi}} := -px_k^{\text{pi}} - \frac{p^2 T}{4c_k} . \quad (10)$$

*Proof.* For  $B_k \equiv 0$ , Eq. (8) gives us  $q_{\nu_k}(r) = x_k^{\text{nom}} + \sigma_k \sqrt{T} N^{-1}(r)$ , therefore  $\nu_k \sim \mathcal{N}(x_k^{\text{nom}}, \sigma_k \sqrt{T})$ . We then obtain by Theorem 2.4 the definition of the equilibrium  $\mu_k$ . Finally, using Lemma A.1, we get

$$2c_k \sigma_k^2 \ln \gamma_k(\tilde{\mu}_k) = \ln \left( \int_{\mathbb{R}} f_k^{\text{nom}}(x) \exp\left(\frac{-px}{2c_k \sigma_k^2}\right) dx \right) = -px_k^{\text{nom}} + \frac{p^2 T}{4c_k} .$$

□

Corollary 2.5 shows that the price of electricity constitutes a natural incentive, as the consumer already makes an effort to reduce his consumption from  $x^{\text{nom}}$  to  $x^{\text{pi}}$ . However, it induces a disutility for consumers ( $V^{\text{pi}} \leq 0$ ). An increase of the price would lead to a supplementary consumption reduction but would decrease further the utility of the agents, and is therefore a non-desirable energy saving strategy.

## 2.3 The Principal's problem

*In this section, we suppose that Assumption 2.2 is satisfied. Therefore, the equilibrium distribution is unique and is defined by (9). For a mean-field equilibrium  $(\mu_k)_{k \in [K]}$ , the mean consumption of the overall population is then  $m_\mu = \sum_{k \in [K]} \rho_k m_{\mu_k}$ .*

For a given  $k$ , we denote by  $\epsilon_k$  the mapping which associates to the total reward function the corresponding equilibrium distribution, i.e.  $\epsilon_k(R_k) = \mu_k$ , where  $\mu_k$  satisfies (9). The problem of the retailer can then be written as

$$\pi^* := \max_{\beta \in \mathcal{B}} \left\{ pm_\mu - \kappa(m_\mu) - \sum_{k \in [K]} \rho_k x_k^{\text{nom}} \int_0^1 \beta(r) dr \left| \begin{array}{l} R_k(x, r) = x_k^{\text{nom}} \beta(r) - px \\ \mu_k = \epsilon_k(R_k) \\ V_k(R_k, \mu_k) \geq V_k^{\text{pi}} + \tau x_k^{\text{nom}} \end{array} \right. \right\} \quad (P^{\text{ret}})$$

where  $\kappa(\cdot)$  denotes the mean selling cost function and  $m_\mu$  is the mean consumption at the equilibrium  $\mu$ . The optimal objective  $\pi^*$  then corresponds to the profit per agent (mean over the population) made on the interval  $[0, T]$  (in €). The inequality constraint on the utility ensures that consumers “play the game”,

as it procures a strictly better utility than without additional reward. Classically,  $\tau = 0$ , meaning that the effort achieved by consumers in order to save energy is compensated (in mean) by the reward offered by the retailer. Observe that with  $\tau = 0$ , some agents may have a negative reward, which is not always desirable. Therefore, for practical issue and acceptability, we allow for a positive  $\tau$  to take into account switching costs that appear when it comes to subscribing to a reward mechanism, see e.g. [MMV23].

**Assumption 2.4.** *The function  $\kappa : \mathbb{R} \rightarrow \mathbb{R}$  is increasing, convex and differentiable. Moreover,  $\kappa'(0) < p < \kappa'(x^{\text{pi}})$ .*

Assumption 2.4 is natural in the context of our application. In practice, the selling cost function is defined as  $\kappa : m \mapsto c_p(m) + s(m)$ , where

- ◇  $s(\cdot)$  denotes the penalty imposed by the regulator to favor a reduction in consumption,
- ◇ and  $c_p(\cdot)$  denotes the cost function, induced by the production of energy.

We assume here that the marginal cost  $\kappa'(\cdot)$  is lower than the marginal price  $p$  at 0 – meaning that it is always profitable to sell a positive quantity of energy – and conversely we assume that the marginal cost  $\kappa'(\cdot)$  is greater than the marginal price  $p$  at  $x^{\text{pi}}$  – meaning that it is not profitable to sell more electricity with the additional reward than without. The penalty function  $s$  is increasing and convex, since the regulator aims at encouraging consumption reduction by strongly penalizing huge consumption levels. Moreover, the retailer’s aggregated cost function is often considered as increasing and convex, due to a decreasing return to scale, see e.g. [Ale+19; ABM20]: the mechanism of day-ahead markets favors the “cheapest” (lowest marginal cost) power plants as the cheapest resource will participate to the electricity generation first, followed by the second cheapest option, and so on, until the demand is satisfied. In the case of non-convex aggregated cost, the convex hull of the aggregated cost function is often considered, see e.g. [Sch+16].

In the case of a homogeneous population and *linear dependence of the objective function with respect to the equilibrium distribution*, the results are obtained in [BZ21]. We extend them here to the more general case of *convex nonlinear dependencies*.

### 2.3.1 Homogeneous population

We consider in this section the specific case where there is a unique cluster of customers (homogeneous population). Therefore, we omit the dependence in  $k$ . Using Lemma A.2 (given in the Appendix), Problem ( $P^{\text{ret}}$ ) can be reformulated as a constrained maximization problem on the distribution space:

**Proposition 2.6.** *Let us consider the following minimization problem*

$$\begin{aligned}
 \min_{\mu \in \mathcal{P}^+(\mathbb{R})} \quad & \kappa \left( \int_{\mathbb{R}} y f_{\mu}(y) dy \right) + 2c\sigma^2 \int_{\mathbb{R}} \ln \left( \frac{f_{\mu}(y)}{f^{\text{nom}}(y)} \right) f_{\mu}(y) dy \\
 \text{s.t.} \quad & \int_{\mathbb{R}} f_{\mu}(y) dy = 1 \\
 & y \mapsto \ln \left( \frac{f_{\mu}(y)}{f^{\text{nom}}(y)} \right) + \frac{p}{2c\sigma^2} y \quad \text{decreasing}
 \end{aligned} \tag{11}$$

Then, the reward  $B_{\mu^*} \in \mathcal{B}$ , constructed from an optimal distribution  $\mu^* \in \mathcal{P}^+(\mathbb{R})$  of (11) as

$$B_{\mu^*}(r) = V^{\text{pi}} + \tau x^{\text{nom}} + 2c\sigma^2 \ln \left( \frac{f_{\mu^*}(q_{\mu^*}(r))}{f^{\text{nom}}(q_{\mu^*}(r))} \right) + pq_{\mu^*}(r) \quad (12)$$

is optimal for problem  $(P^{\text{ret}})$ .

*Proof.* From Lemma A.2,  $B_{\mu}$  defined in (12) is the reward that achieves a given equilibrium distribution  $\mu$  with the lowest cost while satisfying the utility condition in  $(P^{\text{ret}})$  (since  $V(R, \mu) = (1 + \tau)V^{\text{pi}}$  for any attainable equilibrium  $\mu$  and  $R(x, r) = B_{\mu}(r) - px$ ). The objective function is then rewritten as a function of the pdf  $f_{\mu}$  using the expression of the reward.  $\square$

We now relax (11) by ignoring the decreasingness of the additional reward in (11):

$$\min_{f: \mathbb{R} \rightarrow \mathbb{R}} \left\{ \kappa \left( \int_{\mathbb{R}} y f(y) dy \right) + 2c\sigma^2 \int_{\mathbb{R}} \ln \left( \frac{f(y)}{f^{\text{nom}}(y)} \right) f(y) dy \mid \int_{\mathbb{R}} f(y) dy = 1 \text{ and } f(x) \geq 0, x \in \mathbb{R} \right\} . \quad (\tilde{P}^{\text{ret}})$$

The discussion about the relation between the initial problem (11) and the relaxed one  $(\tilde{P}^{\text{ret}})$  is provided further. The optimal solution of this relaxed problem is then characterized by the following lemma:

**Lemma 2.7** (Characterization of the optimal distribution for the relaxed problem). *Let Assumption 2.4 holds. Then,  $(\tilde{P}^{\text{ret}})$  defines a convex problem. Moreover, if  $\mu^*$  admits a density  $f_{\mu^*}$  which minimizes  $(\tilde{P}^{\text{ret}})$ , then it satisfies the following optimality conditions: for  $\mu^*$ -almost every  $x$  in  $\mathbb{R}$ ,*

$$f_{\mu^*}(x) = \frac{1}{\alpha(\mu^*)} f^{\text{nom}}(x) \exp \left( -x \frac{\kappa'(m_{\mu^*})}{2c\sigma^2} \right) \quad (13)$$

where

$$\alpha(\mu) = \int_{\mathbb{R}} f^{\text{nom}}(y) \exp \left( -y \frac{\kappa'(m_{\mu})}{2c\sigma^2} \right) dy .$$

*Proof.* The convexity of the objective functional with respect to  $f$  comes from the convexity of  $\kappa$  (see Assumption 2.4) and the convexity of  $x \mapsto x \ln(x)$ . The first-order conditions for  $(\tilde{P}^{\text{ret}})$  are detailed in Appendix A. Furthermore, they are sufficient for this convex problem, see e.g. [LBD22, Theorem 3.3].  $\square$

In contrast with [BZ21], the optimal distribution is not explicit anymore due to the general function  $\kappa(\cdot)$ . Instead, the optimal distribution is implicitly known through the fixed-point equation (13). We simplify this condition in the following theorem to end up with one-dimension fixed-point equation on the mean consumption.

**Theorem 2.8.** *Let Assumption 2.4 holds, and let  $\delta : \mathbb{R} \rightarrow \mathbb{R}$  be a function given by*

$$\delta(m) = p - \kappa'(m) .$$

*Then, the distribution  $\mu^* = \mathcal{N}(m^*, \sigma\sqrt{T})$ , where  $m^*$  satisfies the fixed-point equation*

$$m - x^{\text{pi}} = \frac{T}{2c} \delta(m) , \quad (14)$$

is optimal for the problem  $(\tilde{P}^{\text{ret}})$ . Moreover, the associated reward  $B_{\mu^*}$  is

$$B_{\mu^*}(r) = \tau x^{\text{nom}} + \frac{c}{T} [(x^{\text{pi}})^2 - (m^*)^2] + q_{\mu^*}(r)\delta(m^*) , \quad (15)$$

and the associated retailer gain is

$$\pi^* = m^* \kappa'(m^*) - \kappa(m^*) + \left( \frac{m^* + x^{\text{pi}}}{2} \right) \delta(m^*) - \tau x^{\text{nom}} . \quad (16)$$

**Corollary 2.9.** *Let Assumption 2.4 holds. Then, the fixed-point equation (14) admits a unique solution  $m^* \in ]0, x^{\text{pi}}]$ . Moreover, the (unique) reward function is decreasing.*

*Proof.* The increasingness of  $\kappa'(\cdot)$  suffices to ensure that (14) admits a unique solution. Moreover, as  $\delta(0) \geq 0 \geq \delta(x^{\text{pi}})$ , the root of the equation  $\frac{T}{2c}\delta(m) - m + x^{\text{pi}} = 0$  must belong to  $]0, x^{\text{pi}}]$ . As a consequence,  $\delta(m^*) \leq 0$  and the reward function  $B_{\mu^*}$  is decreasing.  $\square$

The existence and uniqueness of a solution for (14) entails the existence and uniqueness of an optimal reward for  $(\tilde{P}^{\text{ret}})$ . The knowledge of the bounds for  $m^*$  along with the decreasingness of  $\delta(\cdot)$  allows to use for instance a binary search algorithm to numerically find the optimal mean consumption in logarithmic time.

*Remark.* For quadratic function  $s : m \mapsto \alpha_2 m^2 + \alpha_1 m + \alpha_0$ , the fixed point of (14) is analytically known:

$$m^* = \left( 1 + \frac{\alpha_2 T}{c} \right)^{-1} \left( x^{\text{nom}} - \frac{(\alpha_1 + c_p)T}{2c} \right) .$$

The function  $\delta(\cdot)$  is here interpreted as the *reduction desire* of the provider, as consumption reduction  $x^{\text{pi}} - m^*$  is proportional to  $|\delta(m^*)|$ , see (14). It expresses the marginal benefit coming from selling electricity (including the penalty function  $s$  provided by the regulator).

In the relaxed problem, we neglect that the reward is decreasing. However, this is directly ensured by Corollary 2.9: the reward provided in Theorem 2.8 is decreasing if and only if  $\delta$  is negative at the optimum. Therefore, it is also optimal for the original retailer problem  $(P^{\text{ret}})$ .

The optimal reward obtained in Eq. (15) is defined through the quantile of  $\mu^*$  and is therefore unbounded. From the application viewpoint (it is not realistic to give unbounded rewards to consumers) and for numerical issues, we now look at truncated reward. To this purpose, let us define for any  $M > 0$  the truncated optimal equilibrium distribution  $\mu_M$  through its p.d.f:

$$f_{\mu_M}(x) \propto h_M(x) := f^{\text{nom}}(x) \exp \left( \frac{-y\kappa'(m_{\mu^*}) \wedge M \vee (-M)}{2c\sigma^2} \right) . \quad (17)$$

**Theorem 2.10** (Bounded reward). *The total reward which leads to equilibrium  $\mu_M$  and gives to agents the utility  $V^{\text{pi}} + \tau x^{\text{nom}}$  is bounded for every consumption level and is defined as*

$$\forall x \in \mathbb{R}, R_{\mu_M}(x) = V^{\text{pi}} + \tau x^{\text{nom}} - 2c\sigma^2 \ln \int_{\mathbb{R}} h_M(y) dy + x\kappa'(m_{\mu^*}) \wedge M \vee (-M) . \quad (18)$$

Moreover, the mean consumption converges to the optimal one :

$$m_{\mu_M} = m_{\mu^*} + O\left(e^{-\frac{M}{2c\sigma^2}}\right) .$$

*Proof.* From Lemma A.2., the total reward associated to  $\mu_M$  is  $R_{\mu_M} = V^{\text{pi}} + \tau x^{\text{nom}} + 2c\sigma^2 \ln(f_{\mu_M}(y)/f^{\text{nom}}(y))$  and satisfies the utility constraint by construction. The result is then obtained using the definition of  $f_{\mu_M}$ . Besides, one can show (see [BZ21, Theorem 5.4]) that  $\int_{\mathbb{R}} h_M dx = \alpha(\mu^*) + O\left(e^{-\frac{M}{2c\sigma^2}}\right)$  and  $\int_{\mathbb{R}} x h_M(x) dx = \alpha(\mu^*) m_{\mu^*} + O\left(e^{-\frac{M}{2c\sigma^2}}\right)$ . As a consequence,  $m_{\mu_M} = m_{\mu^*} + O\left(e^{-\frac{M}{2c\sigma^2}}\right)$ .  $\square$

As the optimal (unbounded) total reward, its truncated analog obtained in (18) is linear in the terminal consumption (inside the bounds  $[-M, M]$ ). This means that the consumers are rewarded proportionally to their consumption reduction. Moreover, for both the theoretical bonus (15) and the bounded one (18),  $\tau$  only acts as a shift on the function in order to uplift or lower the bonus received by each agent. Consequently, it is possible to a posteriori choose  $\tau$  in such a way that the bonus of a given ranking corresponds to a certain amount.

### 2.3.2 Heterogeneous population

We consider here the more general setting of a heterogeneous population, not studied yet in the ranking games literature, which consists in a finite number of clusters  $K > 1$ . The transformation which leads to (11) still applies, but the additional constraint Assumption 2.3 has to be imposed to ensure the unitary reward is identical for every sub-population<sup>8</sup>.

As it will be seen below, we can recover explicitly solvable problems for a subclass of heterogeneous populations for which all agents of the overall population are similar up to a scaling factor.

**Proposition 2.11** (Explicit characterization for a sub-class of heterogeneous population). *Let suppose that the following statement holds:*

$$\forall k \in [K], \quad \frac{x_k^{\text{nom}}}{x_1^{\text{nom}}} = \frac{\sigma_k}{\sigma_1} = \frac{c_1}{c_k} \quad (:= \theta_k) . \quad (19)$$

Then, any  $\mu_1, \dots, \mu_K$  equilibrium distributions associated to a common unitary reward  $\beta$  solution of  $(P^{\text{ret}})$  satisfies  $f_{\mu_k}(y) = \frac{1}{\theta_k} f_{\mu_1}\left(\frac{y}{\theta_k}\right)$  for all  $k \in [K]$ . Moreover, the retailer's profit problem simplifies to

$$\pi^* := \bar{\theta} \max_{\beta \in \mathcal{B}} \left\{ pm_{\mu_1} - \tilde{\kappa}(m_{\mu_1}) - x_1^{\text{nom}} \int_0^1 \beta(r) dr \left| \begin{array}{l} R_1(x, r) = x_1^{\text{nom}} \beta(r) - px \\ \mu_1 = \epsilon_1(R_1) \\ V_1(R_1, \mu_1) \geq V_1^{\text{pi}} + \tau x_1^{\text{nom}} \end{array} \right. \right\} , \quad (20)$$

with  $\tilde{\kappa}(m) = \bar{\theta}^{-1} \kappa(\bar{\theta} m)$  and  $\bar{\theta} = \sum_{k \in [K]} \rho_k \theta_k$ .

*Proof.* Using the characterization of the equilibrium in (9),  $q_{\mu_k}(r) = \theta_k q_{\mu_1}(r)$ . Therefore,  $F_{\mu_k}(y) =$

<sup>8</sup>Using Lemma A.2, there exists a common unitary reward leading to equilibrium  $\mu_1, \dots, \mu_K$  if and only if there exists for all  $k \in [K]$  a constant  $C_k$  such that  $\frac{c_k \sigma_1^2}{x_k^{\text{nom}}} \ln\left(\frac{f_{\mu_k}(x)}{f_k^{\text{nom}}(x)}\right) = \frac{c_1 \sigma_1^2}{x_1^{\text{nom}}} \ln\left(\frac{f_{\mu_1}(x)}{f_1^{\text{nom}}(x)}\right) + C_k$  for all  $x \in \mathbb{R}$ .

$F_{\mu_1} \left( \frac{y}{\theta_k} \right)$  and  $f_{\mu_k}(y) = \frac{1}{\theta_k} f_{\mu_1} \left( \frac{y}{\theta_k} \right)$ . Moreover,

$$\begin{aligned} \gamma(\mu_k) &= \int_{\mathbb{R}} f_k^{\text{nom}}(x) \exp \left( \frac{x_k^{\text{nom}} \beta(F_{\mu_k}(x)) - px}{2c_k \sigma_k^2} \right) dx \\ &= \int_{\mathbb{R}} \frac{1}{\theta_k} f_1^{\text{nom}} \left( \frac{x}{\theta_k} \right) \exp \left( \frac{x_1^{\text{nom}} \beta \left( F_{\mu_1} \left( \frac{x}{\theta_k} \right) \right) - p \frac{x}{\theta_k}}{2c_1 \sigma_1^2} \right) dx = \gamma(\mu_1) . \end{aligned}$$

Therefore,  $V_k(R_k, \mu_k) = \theta_k V_1(R_1, \mu_1)$ . As  $V_k^{\text{pi}} = \theta_k V_1^{\text{pi}}$ , the utility constraint is satisfied for every sub-population.  $\square$

Proposition 2.11 shows that in this specific case of heterogeneous population, the problem boils down to the homogeneous framework, up to a re-scaling of the cost function  $\kappa$ . Therefore, Theorems 2.8 and 2.10 and Corollary 2.9 still apply, and in particular, the optimal distribution is  $\mu_1^* = \mathcal{N}(m_1^*, \sigma_1 \sqrt{T})$  where  $\mu_1^*$  is uniquely determined by the equation  $m_1^* - x_1^{\text{pi}} = \frac{T}{2c_1} (p - \tilde{\kappa}'(m_1^*))$ . The condition (19) corresponds to the case where (i) the volatility of the noise is proportional to the nominal consumption and where (ii) the price elasticity is identical for all sub-populations (see Section 4.1 and (25) for the link between the cost of effort  $c_k$  and the elasticity). The second statement (ii) may be more debatable, as the elasticity of a consumer intuitively depends on the equipment of the housing (for instance the type of heating).

### 3 Numerical resolution in the non-uniform heterogeneous case

To deal with the general case of a heterogeneous population, we develop a numerical algorithm to compute the optimal reward from the original problem  $P^{\text{ret}}$ . For a given  $N \in \mathbb{N}$ , we denote by  $\Sigma_N$  the uniform discretization of the interval  $[0, 1]$  by  $N$  points, such that  $\Sigma_N := \{0 = \eta_1 < \eta_2 < \dots < \eta_N = 1\}$ . Let  $M \in \mathbb{R}_+$ , then we define the class of bounded piecewise linear rewards adapted to  $\Sigma_N$  as

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

The reward function obtained as a linear interpolation of a non-increasing vector  $b$  is denoted by  $\hat{\beta}[b]$ . For this special class of reward, the computation of some integrals can be simplified. The integral that appears in the equilibrium characterization (8) becomes

$$\int_0^1 \exp \left( -\frac{x_k^{\text{nom}} \hat{\beta}[b](r)}{2c_k \sigma_k^2} \right) dr = 2c_k \sigma_k^2 (x_k^{\text{nom}})^{-1} \sum_{i=1}^{N-1} \frac{\eta_{i+1} - \eta_i}{b_{i+1} - b_i} \left[ \exp \left( -\frac{x_k^{\text{nom}} b_{i+1}}{2c_k \sigma_k^2} \right) - \exp \left( -\frac{x_k^{\text{nom}} b_i}{2c_k \sigma_k^2} \right) \right]$$

and the integral of the bonus simplifies into

$$\int_0^1 \hat{\beta}[b](r) dr = \sum_{i=1}^{N-1} (\eta_{i+1} - \eta_i) \left( \frac{b_{i+1} + b_i}{2} \right) .$$

**Box maximization.** We define the following transformation:

$$\phi_M^N : [-1, 1]^N \rightarrow [-M, M]^N \quad \text{where} \quad \begin{cases} b_1 = Mz_1 \\ b_i = \frac{1}{2}(b_{i-1} - M) + \frac{1}{2}(b_{i-1} + M)z_i, \quad i > 1 \end{cases} . \quad (21)$$

For any  $M \in \mathbb{R}_+$  and  $N \in \mathbb{N}$ , the function  $\phi_M^N$  is invertible and  $(\phi_M^N)^{-1}$  is defined as:

$$(\phi_M^N)^{-1}(b) = \begin{cases} z_1 = \frac{1}{M}b_1 \\ z_i = \frac{2b_i - b_{i-1} + M}{b_{i-1} + M}, \quad i > 1 \end{cases}$$

As an example, Figure 2 displays  $(\eta_i, z_i)_{i \in [N]}$  and the corresponding bonus function  $\hat{\beta}[\phi_M^N(z)]$ .

We denote by  $\pi_\lambda : \mathcal{B} \rightarrow \mathbb{R}$  the Lagrangian function of  $(P^{\text{ret}})$ , defined as

$$\pi_\lambda(\beta) := \left\{ \begin{array}{l} pm_\mu - \kappa(m_\mu) - \sum_{k \in [K]} \rho_k x_k^{\text{nom}} \int_0^1 \beta(r) dr \\ -\lambda \sum_{k \in [K]} \rho_k \left( V_k^{\text{pi}} + \tau x_k^{\text{nom}} - V_k(R_k, \mu_k) \right)^+ \end{array} \left| \begin{array}{l} R_k(x, r) = x_k^{\text{nom}} \beta(r) - px \\ \mu_k = \epsilon_k(R_k) \end{array} \right. \right\}, \quad (22)$$

where  $(\cdot)^+ := \max(0, \cdot)$ . For fixed Lagrangian multiplier  $\lambda > 0$ ,  $\pi_\lambda$  constitutes a relaxed version of the initial problem  $(P^{\text{ret}})$ , where violations of the utility condition are not fully forbidden but rather strongly penalized in the objective for large values of  $\lambda$ .

**Proposition 3.1** (*Maximization with box constraints*).

$$\max_{z \in [-1, 1]^N} \pi_\lambda(\hat{\beta}[\phi_M^N(z)]) = \max_{\beta \in \mathcal{B}_M^N} \pi_\lambda(\beta) . \quad (23)$$

*Proof.* By definition of  $\mathcal{B}_M^N$ ,  $\max_{z \in [-1, 1]^N} \pi_\lambda(\hat{\beta}[\phi_M^N(z)]) \leq \max_{\beta \in \mathcal{B}_M^N} \pi_\lambda(\beta)$ . As the map  $\phi_M^N$  is invertible, for any reward  $\beta \in \hat{\mathcal{B}}_M^N$ , there exists  $z \in [-1, 1]^N$  such that  $\beta = \hat{\beta}[\phi_M^N(z)]$ , hence the reverse inequality. Optimizing on  $\hat{\mathcal{B}}_M^N$  is then equivalent to optimize on  $[-1, 1]^N$  via the transformation  $\phi_M^N$ .  $\square$

Algorithm 1 aims at maximizing the function  $\pi_\lambda$ . To this end, we do not directly search the optimal reward but, as described previously, we use the invertible map  $\phi_M^N$  to search in the space  $[-1, 1]^N$ , see Proposition 3.1. From a computational viewpoint, the search space is now independent of  $M$ , and the decreasingness of the bonus function is directly encoded in the transformation. Therefore, the only remaining constraints are the ones ensuring that the solution belongs to the unit box. The search is then achieved by black-box optimization, since the evaluation of  $\pi_\lambda$  can be explicitly done using (8)-(9). In the numerical results, we use CMA-ES ([Han06]) as optimization solver through the C++ interface ([Fab13]). Convergence properties are analyzed in [HO97], and we display in Section 4 the numerical convergence of the objective along the iterations.

*Remark.* (i) The evaluation of  $\pi_\lambda$  linearly depends on the number of sub-populations (i.e.,  $K$ ) since, given a reward, the problem boils down to the computation of the equilibrium distributions for the  $K$  sub-populations.



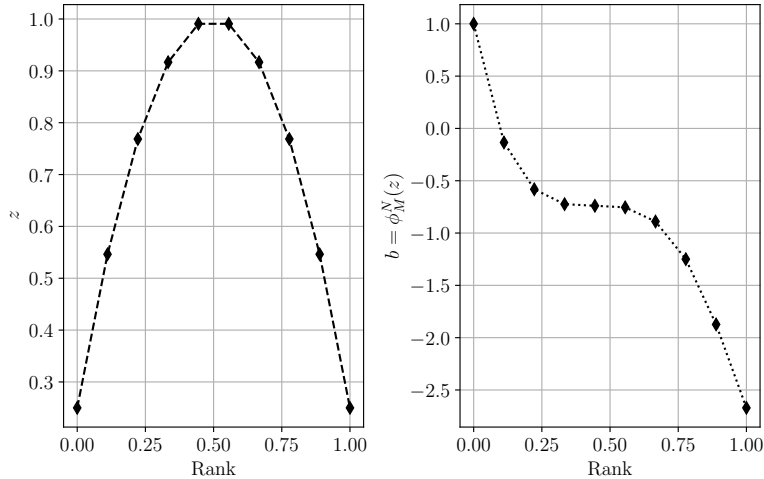


Figure 2: Example of transformation using function  $\phi_M^N$  for  $M = 4$  and  $N = 10$

- (ii) The reward function found by Algorithm 1 is bounded and decreasing, but might violate the utility constraint “ $V_k(R, \mu_k) \geq V_k^{\text{pi}} + \tau x_k^{\text{nom}}$ ” for small penalization values of  $\lambda$ . Note that if the optimizer for the discrete problem on a sufficiently precise grid is a global optimizer, then we get an  $\varepsilon$ -solution of the initial problem, see Theorem 2.10.

---

**Algorithm 1** OPTIMIZATION OF THE REWARD

---

**Require:**  $M, N, \lambda, \Sigma_N$ , solver  $\Pi$ , initial point  $z^0$ ,  
Construct  $\Theta$  as

$$\Theta : z \in ]-1, 1]^N \mapsto \pi_\lambda \left( \hat{\beta}[\phi_M^N(z)] \right) \quad (24)$$

Apply  $\Pi$  to maximize  $\Theta$  (starting from  $z^0$ ) and get the final state  $z^\Pi$ .  
**return**  $\beta^\Pi = \hat{\beta}[\phi_M^N(z^\Pi)]$ .

---

## 4 Application to Energy Savings

In this section, we develop a case study related to the French market of Energy Saving Certificates based on the use of realistic data. We compare the results with existing reward mechanisms, and analyze them in terms of consumption reduction (relatively to the target imposed by the European commission).

### 4.1 Instances

**Consumers.** We consider the case where the retailer aims at designing a reward for 4 types of consumers, listed in Table 1. Data on the average annual consumption correspond to the French case. The consumers are here distinguished according to the surface of the housing and the type of heating, which can represent up to 90% of the annual consumption. A more elaborated clustering might also take into account the location of the housing or the age of the occupants, but we focus here on the two main factors affecting the consumption. We suppose for simplicity that the overall population is composed of these four sub-populations, representing a total of 33 millions of households (current number of households in

	Distribution	Housing	Heating	Nb occupants	Consumption (mean/year)
Sub-pop. 1	26%	House 70 m <sup>2</sup>	Electric	3	9.9 MWh
Sub-pop. 2	49%	House 70 m <sup>2</sup>	Non-electric	3	1.5 MWh
Sub-pop. 3	9%	House 150 m <sup>2</sup>	Electric	4	20 MWh
Sub-pop. 4	16%	House 150 m <sup>2</sup>	Non-electric	4	2.2 MWh

Table 1: Annual electricity consumption by type of usage.  
The consumption data are extracted from “Agence France Electricité”<sup>9</sup>.

France). The distribution of the sub-populations is then computed by considering that there are thrice as many 70m<sup>2</sup>-houses as 150m<sup>2</sup>-houses (the mean surface in France<sup>10</sup> is around 90m<sup>2</sup>) and that a 35%<sup>11</sup> of the French households is equipped with electric heating. This gives us a mean annual consumption of 5.46MWh, or a total annual consumption of 180TWh. In comparison, the French annual consumption for residential households is around 155TWh. This slight over-estimation is due to the fact that we only consider here houses with three or four occupants.

We suppose that the consumption levels displayed in Table 1 corresponds to customers having subscribed to a regulated offer, corresponding to a fixed price of electricity  $p$ . As showed in Corollary 2.5, nominal consumption ( $x^{\text{nom}}$ ) and consumption under price  $p$  ( $x^{\text{pi}}$ ) are linked by the relation  $x^{\text{pi}} = x^{\text{nom}} - \frac{p}{2c}$  (we consider annual consumption in Table 1, i.e.,  $T = 1$ ).

In [NYK20], the authors used several utility concave utility function to model the price elasticity of the electricity demand. In particular, they studied a quadratic utility function similar to the cost of effort we consider: with  $T = 1$  and constant effort,  $V_k^{\text{pi}} = \max_{x \in \mathbb{R}} \{-px - c(x - x^{\text{nom}})^2\}$ . This corresponds to the welfare maximization with quadratic utility, defined as  $U(x, x^{\text{nom}}) = -c(x - x^{\text{nom}})^2$ . For this type of utility function, the elasticity is defined as  $\eta = 1 - \frac{x^{\text{nom}}}{x^{\text{pi}}}$ , see e.g. [NYK20, Eq. 19]). As a consequence, using the relation between  $x^{\text{pi}}$  and  $x^{\text{nom}}$  and the definition of the elasticity, one can obtain the following relations:

$$c = \frac{-p}{2\eta x^{\text{pi}}} \quad , \quad x^{\text{nom}} = x^{\text{pi}}(1 - \eta) \quad . \quad (25)$$

Several values of price elasticity are reported in [NYK20; Cse20], and we use here  $\eta = -0.32$ , which corresponds to the estimation of the long-run residential price elasticity made by [Bön+15] on the EPEX spot market between 2012 and 2014. Price elasticity is always studied at the scale of a country (or even broader), and therefore we take an estimate which is identical for all the agents (*uniform* elasticity). In the numerical results, we will analyze the influence of a non-uniform elasticity, see Section 4.

Regarding the volatility, in the Low Carbon London pricing study, Carmichael et al. [Car+14] reported a deviation of  $\pm 200$  Watt for a demand of 1000 Watt. We take here a deviation  $\sigma\sqrt{T}$  equals to 10% of the total consumption  $X_T$  under zero effort for each of the four sub-populations. Finally, we consider here for  $p$  the price of the regulated offer (“Tarif Bleu”) in 2019, that is 145 €/MWh<sup>12</sup>.

<sup>10</sup><https://www.lamaisonsaintgobain.fr/blog/insolites/metre-carre-et-confort-connaissiez-vous-la-moyenne-francaise>

<sup>11</sup><https://www.voltalis.com/comprendre-electricite/les-types-de-chauffage-preferes-des-foyers-francais-1772>

<sup>12</sup><https://www.agence-france-electricite.fr/consommation-electrique/moyenne-par-jour/>

<sup>12</sup><https://prix-elec.com/tarifs/evolution/2019>

	$c_k$ (€/MWh <sup>2</sup> )	$\sigma_k$ (MWh)
Sub-pop. 1	24	0.57
Sub-pop. 2	156	0.09
Sub-pop. 3	12	4.15
Sub-pop. 4	107	0.13

Table 2: Cost of effort and volatility parameters.

**Retailer cost.** We consider here the year 2019 (just before the energy crisis). We display in Table 3 the marginal cost and the annual production for each type of power plants.

Power plant	Marginal cost (€/MWh)	Production (TWh)
Hydro/Wind/Solar	0 to 15	115
Nuclear	30	380
Gas	70	30
Coal	86	7
Fuel	162	5

Table 3: Marginal price and annual production. Source: *RTE Bilan électrique 2019* and *Ademe*

By aggregating the production capacities by increasing cost (as in merit order curves for day-ahead markets), we can obtain an estimate of the supply cost according the production, see Figure 3. The total cost is then obtained by dividing the supply cost by 0.35 as this approximately corresponds to the weight of supply in the total cost<sup>13</sup>. To fit with our situation where we only look at the residential part of the consumption, we shift the cost curve so that a residential consumption of 180TWh is “cleared” by a gas power plant (as it is often the case in the day-ahead market) and we regularize it to be differentiable.

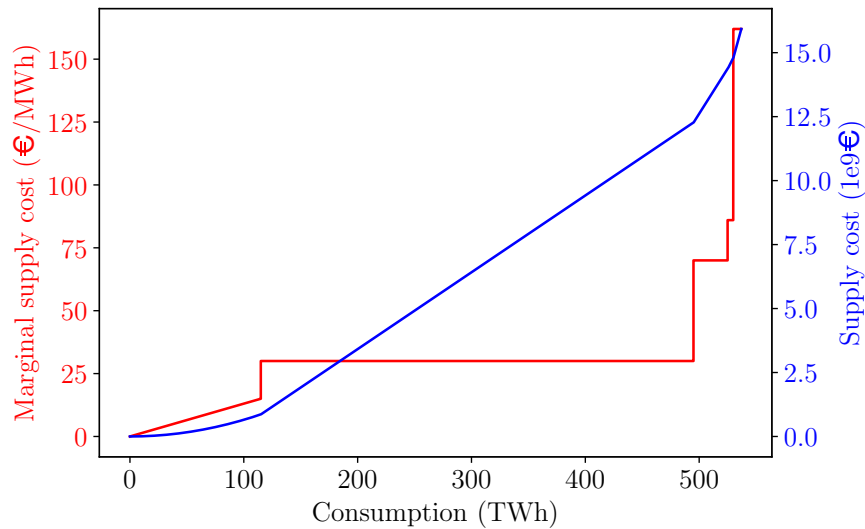


Figure 3: Estimation of supply cost through marginal costs

<sup>13</sup><https://www.ecologie.gouv.fr/commercialisation-lelectricite>

**Valuation of energy savings.** Electricity retailers are obliged by the French government<sup>14</sup> to reduce the global consumption of their customers, in the context of energy efficiency and sobriety. From 2024 to 2030, the European regulation will impose a reduction target of 1.49% of the annual consumption, and aspire to reach 1.9% by the end of 2030. If a retailer does not succeed in gathering a sufficient amount of Energy Saving Certificates, a penalty of 15€/MWh is applied (for “classic” certificates)<sup>15</sup>. In addition, each provider can buy (resp. sell) on a market a certain quantity of certificates if the quantity of energy consumption overshoots (resp. undershoots) the target. In 2023, the price of certificates is around 7.5€/MWh<sup>16</sup>. We consider here a target of 5% of consumption reduction over 3 years ( $T = 3$ ), corresponding to a mean consumption of 15.6MWh for the three years. The valuation function is then defined as  $s_{\theta}(m) = \text{softplus}_{\theta}(15(m - 15.6))$ , where  $\text{softplus}_{\theta} = \theta^{-1} \log(1 + \exp(\theta x))$ . Figure 4 shows the two extreme cases : a purely liquid market ( $\theta = 0$ ) and the absence of exchange ( $\theta = \infty$ ). We choose here  $\theta = 0.3$  to represent an intermediate case.

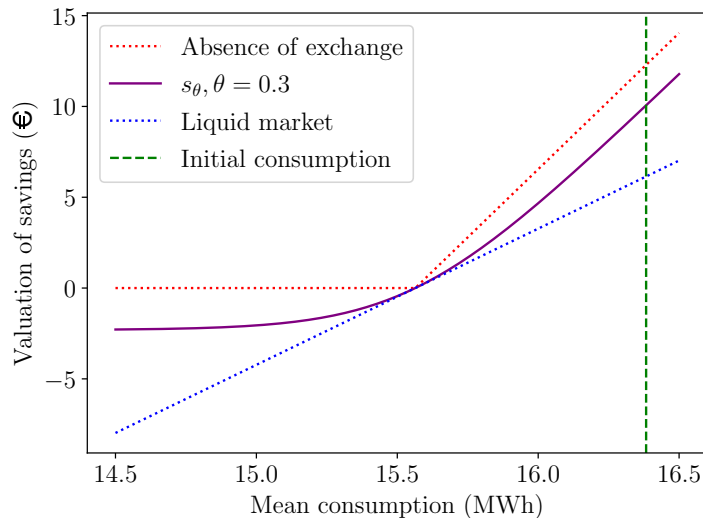


Figure 4: Penalty function  $s(\cdot)$  given by the regulator.

## 4.2 Numerical Results

We use  $N = 20$  discretization points for the bonus description and  $M = 0.1p$ . This means that the maximal unitary bonus given to an agent cannot exceed 10% of the electricity price. We take  $z^0 \equiv 1$  as initial guess. The main advantage of this initial guess is that it satisfies the utility constraint (if  $\tau < M$ ). The initial standard deviation parameter of CMA-ES<sup>17</sup> was set to 5%. The numerical results<sup>18</sup> – parallelized on 10 threads – were obtained on a laptop i7-1065G7 CPU@1.30GHz.

<sup>14</sup>Loi POPE, 2005 : <https://www.ecologie.gouv.fr/dispositif-des-certificats-deconomies-denergie>

<sup>15</sup><https://www.calculcee.fr/les-primas-cee.php>

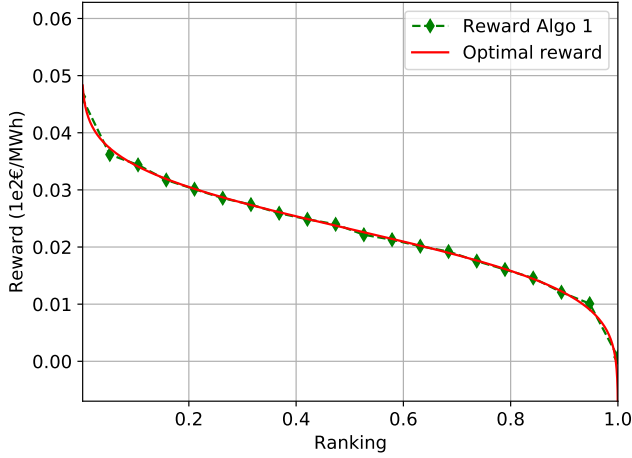
<sup>16</sup><https://c2emarket.com/>

<sup>17</sup>We use the C++ implementation of CMA-ES, available at <https://github.com/CMA-ES/libcmaes>.

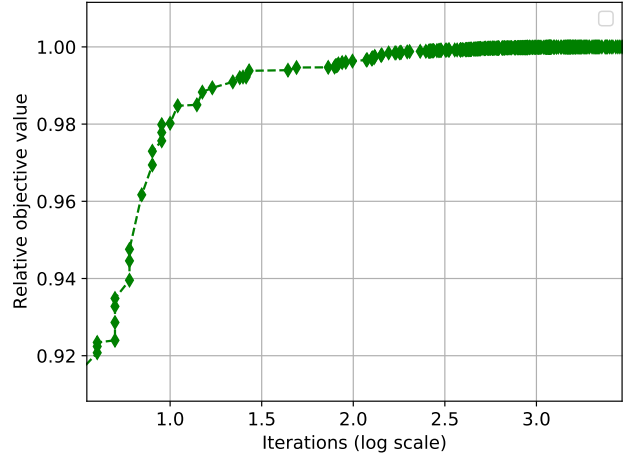
Practical hints are provided for the choice of the parameters.

<sup>18</sup>The whole code is available on the GitHub repository: [https://github.com/jacqu/rk\\_games\\_electricity](https://github.com/jacqu/rk_games_electricity).

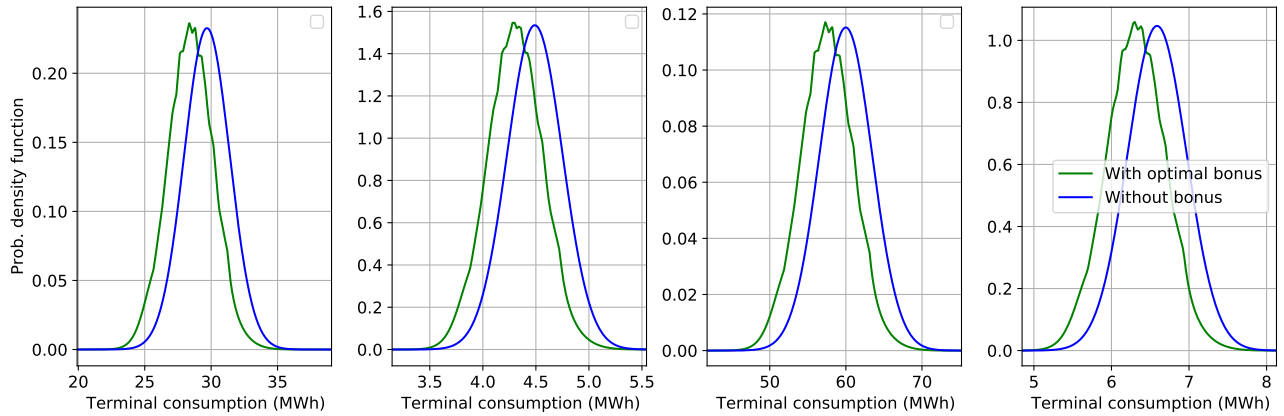
**Uniform elasticity.** Figure 5 shows the results for the test case described in Section 4.1, where the price elasticity is identical for all the sub-populations. As a consequence, Proposition 2.11 applies and we can analyze in this setting the performance of the numerical solving procedure: in Figure 5a, the reward found by Algorithm 1 is very close to the (theoretical) optimal reward, showing that the solver successfully finds the global optimum. About the computational cost, the algorithm converged in approximately 3000 iterations (around 400 seconds), but succeeded in reducing the optimality gap to less than 0.5% in 100 iterations.



(a) Analytic optimal reward in red, compared to the unitary bonus function found by Algorithm 1.



(b) Evolution of the relative objective value along the iterations.



(c) Terminal consumption distribution for the four sub-populations

Figure 5: Numerical results for the four populations described in Tables 1 and 2 (scalable case).

We depict in Figure 5 the distribution of the terminal consumption for the four sub-populations with and without the bonus. As shown in Corollary 2.5, the distribution without reward is a Gaussian process centered in  $x^{\text{pi}}$  (which corresponds to three times the annual consumption displayed in Table 1). The terminal distribution with the optimal reward is then a shift of this normal distribution – see Proposition 2.11. We observe that, as expected, the terminal distribution is also identical for the four sub-populations, up to a scaling ( $f_{\mu_k^*}(x) = \theta_k^{-1} f_{\mu_1^*}(\theta_k^{-1}x)$ ). Here, the mean pluri-annual consumption on the whole population decreased from 16.38MWh to 15.7MWh, giving a saving ratio of 4.1%. This has to be compared with the initial objective of the regulator (a reduction of 5% of the pluri-annual consumption): the retailer found

a compromise between the penalty imposed by the regulator, the cost to propose a reward mechanism and its natural willing to sell electricity.

The optimal bonus offered to customers takes negative values for the 1% consuming the most (we choose  $\tau$  a posteriori in this sense) and goes up to more than 4€ per MWh, which corresponds to a bonus of 66€ in average over the three years. This should be compared for instance with the “Bonus Conso” proposed by TotalEnergies<sup>19</sup>, where 30€ are proposed for a reduction of 5% over one year.

**The  $N$ -players game.** We now numerically illustrate the behavior of several individual consumers incentivized by the optimal bonus found in Figure 5a. The simulation of the trajectories is done using a Euler-Maruyama scheme, see e.g. [NT15] for details on the discretization, as for convergence rates.

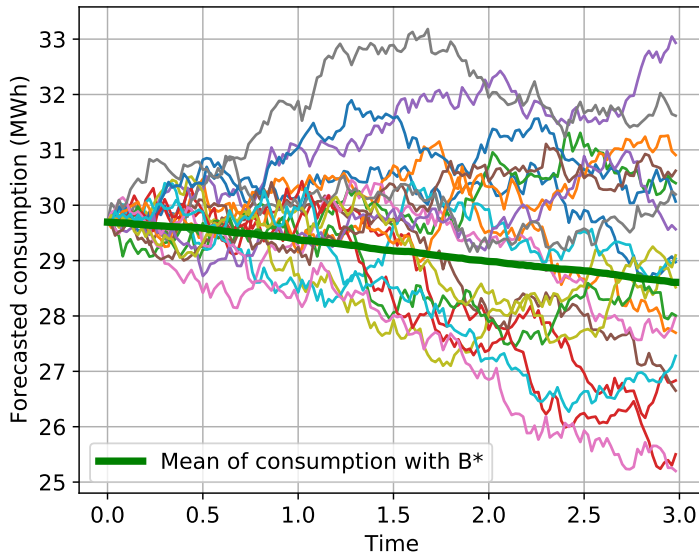


Figure 6: Deviation of the consumption from the no-bonus case  
Trajectories for 20 consumers from sub-population 1.

Figure 6 displays the evolution of the forecasted consumption  $X_1^{a_1^*}$ , from which we subtracted the deviation coming from price in order to clearly distinguish the supplementary effort made through the influence of the bonus. This corresponds to the quantity

$$Y_1(t) = X_1^{a_1^*}(t) + \frac{p(t-T)}{2c_1} ,$$

where  $a_1^*$  is the optimal effort in the presence of the bonus. We observe the same consumption decrease as in Figure 5c, and this reduction has a linear behavior. Indeed, we showed in (18) that the optimal total reward is linear in  $x$ , and for any reward  $R_{k,\mu} = \alpha_0 - \alpha_1 x$ , the corresponding effort is  $a_k^*(t) = -\frac{\alpha_1}{2c_k}$  – see (6) – and the consumption reduction is then  $\frac{\alpha_1}{2c_k} t$ . This has a strong implication on the behavior of the model: the effort made at time  $0 \leq t \leq T$  by a consumer is independent from his current situation, i.e., is not influenced by the hazard  $W_t$ . This means that a consumer will not stop/reduce his effort even if he is undergoing an adverse hazard.

<sup>19</sup><https://www.totalenergies.fr/bonus-conso>

**Non-uniform price elasticity.** We now slightly change the previous test case by considering that the price elasticity is not constant across the population, but rather depends on the characteristics of each agent. In particular, we consider here that the price elasticity of a consumer with electric heating is greater than someone with another heating technology. This greater specific adaptability is for instance exploited by some energy providers<sup>20</sup>. To see the influence of non uniform elasticity, we divide by two the elasticity of sub-populations 2 and 4 – as they do not have electric heating – and multiply by 1.5 the elasticity of sub-populations 1 and 3. In this setting, the scaling condition (19) is no longer satisfied, and so, contrary to the previous case, we are not able to find the theoretical optimal bonus function, but only able to perform a numerical optimization using Algorithm 1.

Figure 7 shows the results for the test case with modified elasticity parameters. We use here  $N = 40$  discretization points and let the algorithm runs up to 5000 iterations. The convergence of Algorithm 1 is still fast since the gap between the solution at iteration 100 was already close to the final solution to less than 1%. About the terminal consumption distribution, we observe that the mean consumption for sub-populations 1 and 3 is reduced by 5.3% whereas the mean consumption for sub-populations 2 and 4 is reduced by 2.3%. Indeed, it reflects the increase (resp. decrease) of price-elasticity for 1 and 3 (resp. 2 and 4). This should be compared with the uniform consumption reduction of 4.1% in the previous setting.

The unitary bonus found by Algorithm 1 is lower than in Figure 5: for example, in the uniform-elasticity case, every agent with a ranking lower than 0.6 received a unitary bonus greater than 2€ per MWh, while in the non-uniform case, only consumers with ranking lower than 0.2 can claim this level of reward. This highlights the fact that the retailers does not need to propose a reward as huge as in the previous case since the reduction effort is mostly endorsed by users with electric heating, now more compliant to lower their consumption.

## 5 Extensions

We propose in this section several extensions to fit with more general settings.

**Energy consumption with common-noise.** The add of common-noise is not rare in the modeling of electricity consumption. But in this present case, it does not impact the retailer problem. Intuitively, as the reward is determined by the ranking of the agents, an identical perturbation of the consumption will not modify the rankings, and so the effort made by the agents is independent of the common-noise.

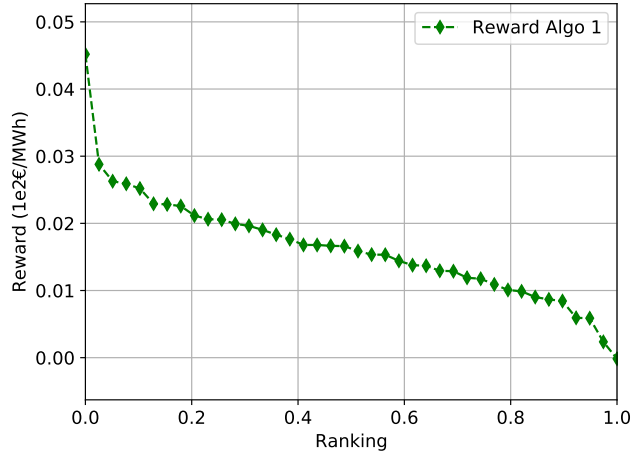
Let us prove this intuitive behavior. To this purpose, we fix a sub-population  $k \in [K]$ , and suppose that the dynamics is now described as:

$$dX_k^a(t) = a_k(t)dt + \sigma_k dW_k(t) + \sigma^0 dW^0(t), \quad X_k(0) = x_k^{\text{nom}} . \quad (26)$$

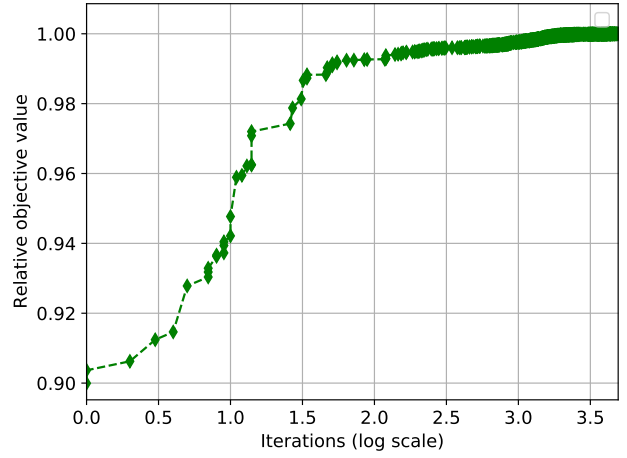
**Proposition 5.1** (Translation invariance of the effort). *Let  $R_k$  be the total reward for sub-population  $k$  (satisfying Assumption 2.2) and  $\mu_k$  be the equilibrium distribution under  $R_k$  and without common-noise*

---

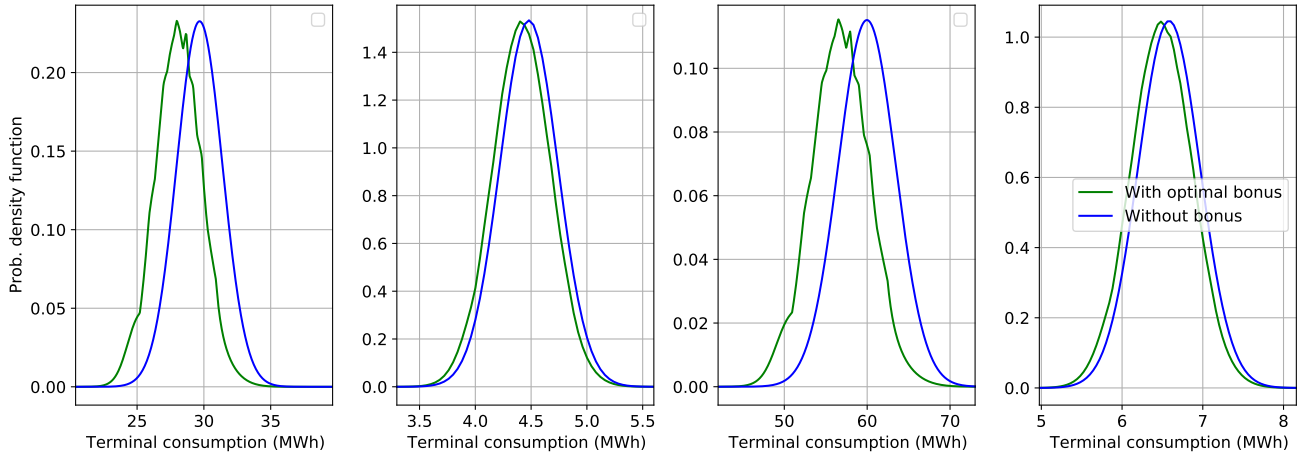
<sup>20</sup><https://www.sowee.fr/>



(a) Unitary bonus function found by Algorithm 1.



(b) Evolution of the relative objective value along the iterations.



(c) Terminal consumption distribution for the four sub-populations

Figure 7: Numerical results for the four populations with different price elasticity.

(given by (8)-(9)). Then

$$\mu_k^0 := x \mapsto \mu_k(x - \sigma^0 W^0(T)) \quad (27)$$

is a (random) equilibrium distribution under  $R_k$  and dynamics (26).

*Proof.* For all  $x, q \in \mathbb{R}$  and  $\mu \in \mathcal{P}(\mathbb{R})$ , we have:

$$R_{k, \mu_k}(x + q) = B_k(F_{\mu_k}(x + q)) - p(x + q) = B_k(F_{\mu_k(\cdot + q)}(x)) - p(x + q) = R_{k, \mu_k(\cdot + q)}(x) - pq .$$

Therefore, according to the expression of the optimal effort in (6)

$$u_k(t, x + q, \mu) = q^{-1} u_k(t, x, \mu(\cdot + q))$$

and  $a_k(t, x + q; \mu_k) = a_k(t, x; \mu(\cdot + q))$ . Therefore, the drift is translation invariant, and the results of [LW15] apply:  $\mu^0$  defined in (27) is an equilibrium distribution for the dynamics with common-noise.  $\square$



In contrast with the purely rank-based case, total rewards satisfying Assumption 2.2 are not translation invariant. Nonetheless, the drift obtained through the optimal effort is translation invariant, enabling to use the results of [LW15]. For a common-noise  $W^0$  such that  $\mathbb{E}[W^0(\cdot)] = 0$ , maximizing the (expected version of the) profit, defined in  $(P^{\text{ret}})$ , will boil down to the same problem, and so will lead to the same optimal unitary reward.

**General reward  $R(x, r)$ .** We consider here a more general form of reward, coupling the terminal consumption and the ranking. Therefore, Assumption 2.2 is no longer satisfied and the equilibrium cannot be explicitly computed with Theorem 2.8. Instead, one can use fixed-point resolution techniques to compute the equilibrium. To this purpose, let us denote by  $W_1(f_1, f_2)$  the 1-Wasserstein metric for distribution  $f_1, f_2 \in \mathcal{P}_1(\mathbb{R}) = \{\mu \in \mathcal{P}(\mathbb{R}) : \int_{\mathbb{R}} |x| d\mu(x) < \infty\}$ . Algorithm 2 follows the standard way to numerically compute mean-field Nash equilibria – see [AL20] – by iteratively updating the distribution using the best response operator. Here, the operator is explicitly given by (5), which still applies for general forms of reward function, see [BZ21].

---

**Algorithm 2** FIXED-POINT RESOLUTION

---

**Require:**

- initial p.d.f.  $f_{\mu_k^{(0)}}$  of cluster  $k$ ,
- error tolerance  $\varepsilon$ ,
- iteration maximum  $n_{max}$ ,
- sequence of damping coefficients  $\{l_i\}_{i \in \mathbb{N}}$ .

$d, i \leftarrow 2\varepsilon, 0$

**while**  $d \geq \varepsilon$  or  $n \leq n_{max}$  **do**

$f_{\mu_k^{(i+1/2)}} \leftarrow \Phi_k(f_{\mu_k^{(i)}})$  ▷ Best-response map defined in Definition 2.2

$f_{\mu_k^{(i+1)}} \leftarrow l_i f_{\mu_k^{(i+1/2)}} + (1 - l_i) f_{\mu_k^{(i)}}$  ▷ damping  $l_i$

$d \leftarrow W_1(f_{\mu_k^{(i)}}, f_{\mu_k^{(i+1)}})$  ▷ distance between two iterates

$i \leftarrow i + 1$

**end while**

---

Instead of Picard iterates ( $l_i = 1$ ), a decreasing damping  $l_i = \left(\frac{1}{i+1}\right)^p$ ,  $p \in \mathcal{N}$  can be used. The latter sequence of inertial parameters defines iterates of Krasnoselskii-Mann type, which has been proved to converge for pseudo-contractive map in Hilbert space, see [Raf07]. Such a damping has been used for example to solve Linear-Quadratic mean-field control problems in [Gra+16].

We then show that the uniqueness of the reward function is no longer true in the general setting, and there exists a family of equivalent reward function, going from purely rank-based rewards to purely consumption-based reward ones:

**Proposition 5.2** (Invariance). *Let  $R^*(x, r)$  be an optimal reward function for the following problem*

$$\max_{R(x,r)} \left\{ -\kappa(m_\mu) - \int_{\mathbb{R}} R_\mu(x) f_\mu(x) dx \mid \begin{array}{l} \mu = \epsilon(R) \\ V(R, \mu) \geq V^{\text{Pi}} \end{array} \right\} \quad (28)$$

*This equilibrium distribution obtained with  $R^*$  is denoted by  $\mu^*$ . Then,*

(i) the purely rank-based reward function  $\hat{B} : r \mapsto R^*(q_{\mu^*}(r), r)$  is also an optimal reward,

(ii) the reward function  $\hat{R} : x \mapsto R^*(x, F_{\mu^*}(x))$  is also an optimal reward.

*Proof.* By definition, the two reward functions  $\hat{B}$  and  $\hat{R}$  also satisfy the characterization of the equilibrium (7) with  $\mu_k = \mu^*$ . Therefore, under these rewards, agents reach the same equilibrium as with  $R$ , and their utility is identical. Moreover, the objective in (28).  $\square$

In practice, Proposition 5.2 has very useful implications. It states that complicated reward policies simplify into simple rules. The first item shows that we can construct a purely *competitive* game in the sense that the consumers receives incentives only through their rank. The second item shows that we can construct a *decentralized* reward since the incentive of each customer only depends on their own consumption. Note that this notion of invariance applies at the equilibrium, and the equivalence of the reward is no longer true outside the equilibrium.

**Time-dependent effort cost.** In the context of the ecological transition, the consumers are more willing to contribute to the energy reduction, and therefore the effort cost  $c$  can be viewed as a time dependent parameter, modeling the change of customers' behavior.

In this case, with a cost profile  $c_k(t)$ ,  $t \in [0, T]$  for each cluster  $k$ , the consumer's problem becomes

$$V_k(R, \mu_k) := \sup_a \mathbb{E} \left[ R_{\mu_k}(X_k^a(T)) - \int_0^T c_k(t) a_k^2(t) dt \right] . \quad (29)$$

As a direct extension of [BZ16], we have the following existence result:

**Theorem 5.3.** *Assume that the cost profiles are bounded such that there exist  $(\underline{c}_k, \bar{c}_k)$  verifying for all  $t \leq T$*

$$0 < \underline{c}_k \leq c_k(t) \leq \bar{c}_k .$$

*Then, there exists at least one equilibrium.*

Nonetheless, there is no more explicit formula (even for the best response of the agents) in presence of time-varying cost of effort, as the Schrödinger bridge method requires a quadratic cost of effort that is constant over time. To illustrate the behavior of the agents with a time-dependent cost of effort, we draw in Figure 8 the trajectories of the same 20 consumers as in Figure 6 obtained with the incentive depicted in Figure 5a and a cost of effort  $c_k(t) = 24 - 1.5t$  €/MWh. As expected, the energy savings are greater than the previous case (the terminal consumption is now around 27.6MWh whereas it was around 28.5MWh with  $c_k(t) = 24$ €/MWh).

## 6 Conclusion

In this work, we study a Principal-Agent mean-field game where the incentive designed by the principal is based on the ranking of each agent, initiating a competition between them. This specific framework allows us to derive explicit formula for the (unique) mean-field Nash equilibrium for the agents' problem. Incorporating this characterization in the principal profit maximization problem, we prove in the homogeneous setting that the optimal reward can be obtained by solving a convex reformulation of the problem

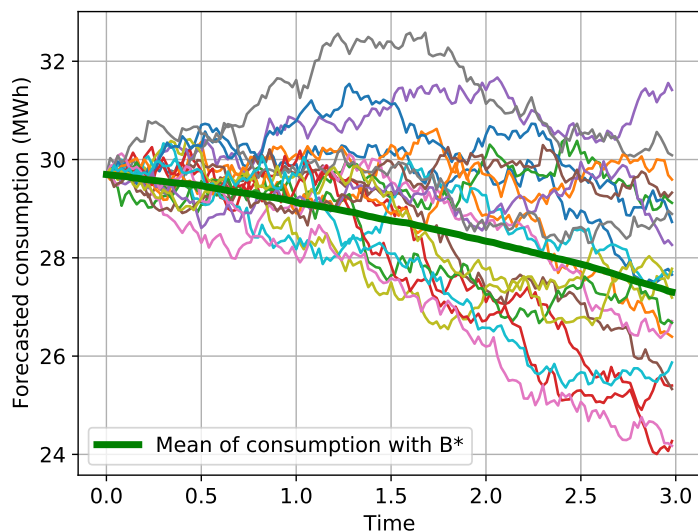


Figure 8: Deviation of the consumption from the no-bnous case. Trajectories of 20 consumers from sub-population 1 obtained with the optimal control from the mean-field approximation and a time-dependent cost of effort.

in the distribution space. We exploit the optimality conditions of the latter to then get the optimal reward through a fixed-point equation. In the general case, we show that the problem can be recast as a finite-dimensional maximization over a box, which can be efficiently solved by numerical algorithms.

We apply the results to electricity markets where a provider aims at designing a reward for its consumers portfolio in order to incentivise them to energy sobriety. We construct realistic instances for the French market of Energy Saving Certificates, and numerically observe that the rank-based rewards can constitute efficient mechanisms to make substantial energy reduction, while staying sufficiently simple to be easily grasped by the consumers.

## Acknowledgments

We thank the reviewers for their comments, which contributed to an improved version of the paper.

## References

- [AT14] Hunt Allcott and Rogers Todd. “The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation”. In: *American Economic Review* 104.10 (2014), pp. 3003–37.
- [DM15] Paul Dolan and Robert Metcalfe. “Neighbors, knowledge, and nuggets: two natural field experiments on the role of incentives on energy conservation”. In: *Becker Friedman Institute for Research in Economics Working Paper* (2015).
- [LL06a] Jean-Michel Lasry and Pierre-Louis Lions. “Jeux à champ moyen. I—Le cas stationnaire”. In: *Comptes Rendus Mathématique* 343.9 (2006), pp. 619–625.
- [LL06b] Jean-Michel Lasry and Pierre-Louis Lions. “Jeux à champ moyen. II—Horizon fini et contrôle optimal”. In: *Comptes Rendus Mathématique* 343.10 (2006), pp. 679–684.

- [LL07] Jean-Michel Lasry and Pierre-Louis Lions. “Mean field games”. In: *Japanese Journal of Mathematics* 2.1 (2007), pp. 229–260.
- [HMC06] Minyi Huang, Roland P. Malhame, and Peter E. Caines. “Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle”. In: *Commun. Inf. Syst.* 6.3 (2006), pp. 221–252.
- [HCM07] M. Huang, P. E. Caines, and R. P. Malhame. “Large-Population Cost-Coupled LQG Problems With Nonuniform Agents: Individual-Mass Behavior and Decentralized Nash Equilibria”. In: *IEEE Transactions on Automatic Control* 52.9 (2007), pp. 1560–1571.
- [Car+15] Pierre Cardaliaguet, P Jameson Graber, Alessio Porretta, and Daniela Tonon. “Second order mean field games with degenerate diffusion and local coupling”. In: *Nonlinear Differential Equations and Applications NoDEA* 22 (2015), pp. 1287–1317.
- [BCS17] Jean-David Benamou, Guillaume Carlier, and Filippo Santambrogio. “Variational mean field games”. In: *Active Particles, Volume 1: Advances in Theory, Models, and Applications* (2017), pp. 141–171.
- [San08] Yuliy Sannikov. “A Continuous-Time Version of the Principal-Agent Problem”. In: *Review of Economic Studies* 75.3 (July 2008), pp. 957–984. URL: <https://doi.org/10.1111/j.1467-937x.2008.00486.x>.
- [CCY12] Agostino Capponi, Jakša Cvitanić, and Türkay Yolcu. “Optimal contracting with effort and misvaluation”. In: *Mathematics and Financial Economics* 7.1 (Sept. 2012), pp. 93–128. URL: <https://doi.org/10.1007/s11579-012-0088-z>.
- [EMP19] Romuald Elie, Thibaut Mastrolia, and Dylan Possamaï. “A Tale of a Principal and Many, Many Agents”. In: *Mathematics of Operations Research* 44.2 (May 2019), pp. 440–467. URL: <https://doi.org/10.1287/moor.2018.0931>.
- [CW21] René Carmona and Peiqi Wang. “Finite-State Contract Theory with a Principal and a Field of Agents”. In: *Management Science* 67.8 (Aug. 2021), pp. 4725–4741. URL: <https://doi.org/10.1287/mnsc.2020.3760>.
- [SFJ21] Arvind Shrivats, Dena Firoozi, and Sebastian Jaimungal. *Principal agent mean field games in REC markets*. 2021. URL: <https://arxiv.org/abs/2112.11963>.
- [BZ16] Erhan Bayraktar and Yuchong Zhang. “A rank-based mean field game in the strong formulation”. In: *Electronic Communications in Probability* 21 (2016), pp. 1–12.
- [BCZ19] Erhan Bayraktar, Jakša Cvitanić, and Yuchong Zhang. “Large tournament games”. In: *The Annals of Applied Probability* 29.6 (Dec. 2019). URL: <https://doi.org/10.1214/2F19-aap1490>.
- [BZ21] Erhan Bayraktar and Yuchong Zhang. “Terminal Ranking Games”. In: *Mathematics of Operations Research* 46.4 (Nov. 2021), pp. 1349–1365. URL: <https://doi.org/10.1287/moor.2020.1107>.
- [SFJ22] Arvind V. Shrivats, Dena Firoozi, and Sebastian Jaimungal. “A mean-field game approach to equilibrium pricing in solar renewable energy certificate markets”. In: *Mathematical Finance* 32.3 (Feb. 2022), pp. 779–824. URL: <https://doi.org/10.1111/2Fmfi.12345>.
- [Cam+21] Steven Campbell, Yichao Chen, Arvind Shrivats, and Sebastian Jaimungal. *Deep Learning for Principal-Agent Mean Field Games*. 2021.
- [RSM16] Ciaran Roberts, Emma M Stewart, and Federico Milano. “Validation of the Ornstein-Uhlenbeck process for load modeling based on  $\mu$ PMU measurements”. In: *2016 Power Systems Computation Conference (PSCC)*. IEEE. 2016, pp. 1–7.
- [APT22] René Aid, Dylan Possamaï, and Nizar Touzi. “Optimal Electricity Demand Response Contracting with Responsiveness Incentives”. In: *Mathematics of Operations Research* 47.3 (Aug. 2022), pp. 2112–2137. URL: <https://doi.org/10.1287/moor.2021.1201>.

- [CGP15] Yongxin Chen, Tryphon T. Georgiou, and Michele Pavon. “On the Relation Between Optimal Transport and Schrödinger Bridges: A Stochastic Control Viewpoint”. In: *Journal of Optimization Theory and Applications* 169.2 (Sept. 2015), pp. 671–691. URL: <https://doi.org/10.1007/s10957-015-0803-z>.
- [MMV23] Marco Magnani, Fabio M. Manenti, and Paola Valbonesi. “Measuring Switching Costs in the Italian Residential Electricity Market”. In: *SSRN Electronic Journal* (2023). URL: <https://doi.org/10.2139/ssrn.4350592>.
- [Ale+19] Ekaterina Alekseeva, Luce Brotcorne, Sébastien Lepaul, and Alain Montmeat. “A bilevel approach to optimize electricity prices”. In: *Yugoslav Journal of Operations Research* 29.1 (2019), pp. 9–30. URL: <https://doi.org/10.2298/yjor171115002a>.
- [ABM20] Clémence Alasseur, Imen Ben Tahar, and Anis Matoussi. “An extended mean field game for storage in smart grids”. In: *Journal of Optimization Theory and Applications* 184.2 (2020), pp. 644–670.
- [Sch+16] Dane A. Schiro, Tongxin Zheng, Feng Zhao, and Eugene Litvinov. “Convex Hull Pricing in Electricity Markets: Formulation, Analysis, and Implementation Challenges”. In: *IEEE Transactions on Power Systems* 31.5 (Sept. 2016), pp. 4068–4075. URL: <https://doi.org/10.1109/tpwrs.2015.2486380>.
- [LBD22] Nicolas Lanzetti, Saverio Bolognani, and Florian Dörfler. “First-order conditions for optimization in the Wasserstein space”. In: *arXiv preprint arXiv:2209.12197* (2022).
- [Han06] Nikolaus Hansen. “The CMA Evolution Strategy: A Comparing Review”. In: *Towards a New Evolutionary Computation*. Springer Berlin Heidelberg, 2006, pp. 75–102. URL: [https://doi.org/10.1007/3-540-32494-1\\_4](https://doi.org/10.1007/3-540-32494-1_4).
- [Fab13] Alexander Fabisch. *CMA-ESpp*. <https://github.com/AlexanderFabisch/CMA-ESpp>. 2013.
- [HO97] Nikolaus Hansen and Andreas Ostermeier. “Convergence properties of evolution strategies with the derandomized covariance matrix adaptation: The (/I,)-ES”. In: *Eufit* 97 (1997), pp. 650–654.
- [NYK20] Amir Niromandfam, Ahmad Sadeghi Yazdankhah, and Rasool Kazemzadeh. “Modeling demand response based on utility function considering wind profit maximization in the day-ahead market”. In: *Journal of Cleaner Production* 251 (Apr. 2020), p. 119317. URL: <https://doi.org/10.1016/j.jclepro.2019.119317>.
- [Cse20] Zsuzsanna Cserekyei. “Price and income elasticities of residential and industrial electricity demand in the European Union”. In: *Energy Policy* 137 (Feb. 2020), p. 111079. URL: <https://doi.org/10.1016/j.enpol.2019.111079>.
- [Bön+15] Werner Bönnte, Sebastian Nielen, Niyaz Valitov, and Torben Engelmeyer. “Price elasticity of demand in the EPEX spot market for electricity—New empirical evidence”. In: *Economics Letters* 135 (Oct. 2015), pp. 5–8. URL: <https://doi.org/10.1016/j.econlet.2015.07.007>.
- [Car+14] Richard Carmichael, James Schofield, Matt Woolf, Mark Bilton, Ritsuko Ozaki, and G. Strbac. “Residential consumer attitudes to time varying pricing”. In: (Sept. 2014).
- [NT15] Hoang-Long Ngo and Dai Taguchi. “Strong rate of convergence for the Euler-Maruyama approximation of stochastic differential equations with irregular coefficients”. In: *Mathematics of Computation* 85.300 (Oct. 2015), pp. 1793–1819. URL: <https://doi.org/10.1090/Fmcom3042>.
- [LW15] Daniel Lacker and Kevin Webster. “Translation invariant mean field games with common noise”. In: *Electronic Communications in Probability* 20 (Jan. 2015). URL: <https://doi.org/10.1214/2015-ecp.v20-3822>.

- [AL20] Yves Achdou and Mathieu Laurière. “Mean Field Games and Applications: Numerical Aspects”. In: *Lecture Notes in Mathematics*. Springer International Publishing, 2020, pp. 249–307. URL: [https://doi.org/10.1007%2F978-3-030-59837-2\\_4](https://doi.org/10.1007%2F978-3-030-59837-2_4).
- [Raf07] Arif Rafiq. “On Mann iteration in Hilbert spaces”. In: *Nonlinear Analysis: Theory, Methods and Applications* 66.10 (May 2007), pp. 2230–2236. URL: <https://doi.org/10.1016%2Fj.na.2006.03.012>.
- [Gra+16] Sergio Grammatico, Francesca Parise, Marcello Colombino, and John Lygeros. “Decentralized Convergence to Nash Equilibria in Constrained Deterministic Mean Field Control”. In: *IEEE Transactions on Automatic Control* 61.11 (Nov. 2016), pp. 3315–3329. URL: <https://doi.org/10.1109%2Ftac.2015.2513368>.

## A Proofs

In this section, we collect several results and proofs.

**Lemma A.1.**

$$f_k^{\text{nom}}(x) \exp(\tau x) = \exp\left(\tau x_k^{\text{nom}} + \frac{1}{2}\tau^2\sigma_k^2 T\right) \varphi\left(x; x_k^{\text{nom}} + \tau\sigma_k^2 T, \sigma_k\sqrt{T}\right). \quad (30)$$

*Proof.*

$$\begin{aligned} f_k^{\text{nom}}(x) \exp(\tau x) &= \frac{1}{\sigma\sqrt{T}\sqrt{2\pi}} \exp\left(-\frac{(x - x_k^{\text{nom}})^2 - 2\tau\sigma^2 T x}{2\sigma^2 T}\right) \\ &= \frac{1}{\sigma\sqrt{T}\sqrt{2\pi}} \exp\left(-\frac{(x - [x_k^{\text{nom}} + \tau\sigma^2 T])^2}{2\sigma^2 T} + \tau x_k^{\text{nom}} + \frac{1}{2}\tau^2\sigma^2 T\right) \end{aligned}$$

□

**Lemma A.2** (Set of attainable equilibria). *(i) For a given cluster  $k$ , the set of equilibria attainable by an additional reward function  $B$  is given by*

$$\mathcal{E}_k = \{\mu \in \mathcal{P}^+(\mathbb{R}) : 2c_k\sigma_k^2 \ln \zeta_{k,\mu_k}(q_{\mu_k}(r)) + pq_{\mu_k}(r) \text{ is bounded and decreasing}\},$$

with  $\zeta_{k,\mu} := f_\mu / f_k^{\text{nom}}$ .

*(ii) If  $\mu_k \in \mathcal{E}_k$ , then*

$$\epsilon_k^{-1}(\mu_k) = \{2c_k\sigma_k^2 \ln \zeta_{k,\mu_k}(q_{\mu_k}(r)) + pq_{\mu_k}(r) + C_k : C_k \in \mathbb{R}\}$$

*(iii) Suppose that additional reservation “utility” constraint  $V_k(R, \mu_k) \geq V_k^{\text{pi}} + \tau x_k^{\text{nom}}$  and budget constraint  $\int_0^1 B(r)dr \leq K$ , then the constant  $C_k$  in (ii) is restricted to*

$$V_k^{\text{pi}} + \tau x_k^{\text{nom}} \leq C_k \leq K - 2c_k\sigma_k^2 \int_0^1 \ln \zeta_{k,\mu_k}(q_{\mu_k}(r))dr - pm_{\mu_k}.$$

In particular, such a  $C_k$  exists if and only if

$$2c_k\sigma_k^2 \int_0^1 \ln \zeta_{k,\mu_k}(q_{\mu_k}(r)) dr - pm_{\mu_k} \leq K - V_k^{\text{pi}} - \tau x_k^{\text{nom}} .$$

*Proof.* Items (i) and (iii) directly comes from [BZ21]. For (ii), the condition of Theorem 2.3 is verified:

$$\int_0^r \exp\left(-\frac{R_\mu(q_\mu(z))}{2c\sigma^2}\right) dz = \int_0^r (\zeta_\mu(q_\mu(r)))^{-1} dz = \int_{-\infty}^{q_\mu(r)} f^{\text{nom}}(z) dz .$$

As the uniqueness is concerned, suppose that  $B$  and  $B'$  lead to the same distribution  $\mu$  with  $p \neq 0$ . Then,  $B$  and  $B'$  lead to the same distribution  $\nu$  with  $p = 0$ , see Theorem 2.4. Therefore, as shown in [BZ21],  $B$  and  $B'$  are equal up to a constant.  $\square$

## Proof of Theorem 2.4

We give here the proof for a given class and, for simplicity, we omit the dependence in  $k$ .

*Characterization of an equilibrium.* First, suppose that  $\nu$  is an equilibrium distribution for the case  $p = 0$ . Let  $\gamma \in \mathbb{R}$  whose value will be determined later. By definition of  $f_\nu$  (see (5)), we get

$$\begin{aligned} \int_0^r \exp\left(-\frac{B(z) - p(q_\nu(z) + \gamma)}{2c\sigma^2}\right) dz &= \int_{-\infty}^{q_\nu(r)} \exp\left(-\frac{B(F_\nu(x))}{2c\sigma^2} + \frac{p}{2c\sigma^2}(x + \gamma)\right) f_\nu(x) dx \\ &= \frac{e^{\frac{p}{2c\sigma^2}\gamma}}{\gamma(\nu)} \int_{-\infty}^{q_\nu(r)} \exp\left(-\frac{B(F_\nu(x))}{2c\sigma^2} + \frac{p}{2c\sigma^2}x\right) f^{\text{nom}}(x) \exp\left(\frac{B(F_\nu(x))}{2c\sigma^2}\right) dx. \end{aligned}$$

Using (30) with  $\tau = \frac{p}{2c\sigma^2}$  and the change of variables  $u = \frac{x - (x^{\text{nom}} + \frac{pT}{2c})}{\sigma\sqrt{T}}$ , we deduce

$$\begin{aligned} \int_0^r \exp\left(-\frac{B(z) - p(q_\nu(z) + \gamma)}{2c\sigma^2}\right) dz &= \frac{1}{\gamma(\nu)} e^{\frac{1}{2c\sigma^2}\left(\gamma + px^{\text{nom}} + \frac{Tp^2}{4c}\right)} \int_{-\infty}^{q_\nu(r)} \varphi\left(x; x^{\text{nom}} + \frac{pT}{2c}, \sigma\sqrt{T}\right) dx \\ &= \frac{1}{\gamma(\nu)\sqrt{2\pi}} e^{\frac{1}{2c\sigma^2}\left(\gamma + px^{\text{nom}} + \frac{Tp^2}{4c}\right)} \int_{-\infty}^{\frac{q_\nu(r) - (x^{\text{nom}} + \frac{pT}{2c})}{\sigma\sqrt{T}}} \exp\left(-\frac{u^2}{2}\right) du \\ &= \frac{1}{\gamma(\nu)} e^{\frac{1}{2c\sigma^2}\left(\gamma + px^{\text{nom}} + \frac{Tp^2}{4c}\right)} N\left(\frac{q_\nu(r) - (x^{\text{nom}} + \frac{pT}{2c})}{\sigma\sqrt{T}}\right) . \end{aligned}$$

Therefore, taking  $\gamma = -\frac{pT}{2c}$ , we end up with

$$N\left(\frac{\left[q_\nu(r) - \frac{pT}{2c}\right] - x^{\text{nom}}}{\sigma\sqrt{T}}\right) = \frac{\int_0^r \exp\left(-\frac{B(z) - p\left[q_\nu(z) - \frac{pT}{2c}\right]}{2c\sigma^2}\right) dz}{\int_0^1 \exp\left(-\frac{B(z) - p\left[q_\nu(z) - \frac{pT}{2c}\right]}{2c\sigma^2}\right) dz} .$$

By setting  $q_\mu(r) = q_\nu(r) - \frac{pT}{2c}$ , we recover the characterization of an equilibrium (see Theorem 2.3).

Conversely, suppose now that  $\mu$  is the equilibrium for  $p \in \mathbb{R}$ . Then, following the same steps,

$$N \left( \frac{\left[ q_\mu(r) + \frac{pT}{2c} \right] - x^{\text{nom}}}{\sigma\sqrt{T}} \right) = \frac{\int_0^r \exp\left(-\frac{B(z)}{2c\sigma^2}\right) dz}{\int_0^1 \exp\left(-\frac{B(z)}{2c\sigma^2}\right) dz}.$$

The distribution  $\nu$  defined as  $q_\nu(r) = q_\mu(r) + \frac{pT}{2c}$  is a valid equilibrium.

*Uniqueness of the equilibrium.* Suppose that there exist two distinct equilibrium distributions  $\mu$  and  $\mu'$  such that  $q_\mu \neq q_{\mu'}$ . Then by the above proof, we derive the existence of two distinct equilibrium distributions  $\nu$  and  $\nu'$  for the case  $p = 0$  satisfying  $q_\nu \neq q_{\nu'}$ . We get a contradiction by the uniqueness of the equilibrium for purely rank-based rewards.

### Proof of Lemma 2.7

We apply the KKT conditions on  $(\tilde{P}^{\text{ret}})$  (relaxing the positivity assumption on  $f$ ): for  $\mu^*$ -almost every  $x$  in  $\mathbb{R}$ ,

$$\begin{cases} 0 = x\kappa'(m_{\mu^*}) + 2c\sigma^2 \ln\left(\frac{f_{\mu^*}(x)}{f^{\text{nom}}(x)}\right) + \lambda, \\ \int_{-\infty}^{+\infty} f_{\mu^*}(y) dy = 1 \end{cases}, \lambda \in \mathbb{R}$$

From which we can deduce that  $f_{\mu^*}(x) = f^{\text{nom}}(x) \exp\left(-\frac{x\kappa'(m_{\mu^*}) + \lambda}{2c\sigma^2}\right)$ , which is positive for all  $x$ . The Lagrange multiplier  $\lambda$  is then computed using the normalization condition on  $f_{\mu^*}$ .

### Proof of Theorem 2.8

Integrating (13) gives us

$$\begin{aligned} m_\mu &= \int_{-\infty}^{+\infty} y f_\mu(y) dy = \frac{1}{\alpha(\mu)} \int_{-\infty}^{+\infty} y f^{\text{nom}}(y) \exp\left(-y \frac{\kappa'(m_\mu)}{2c\sigma^2}\right) dy \\ &= \int_{-\infty}^{+\infty} y \phi\left(y; x^{\text{nom}} - \frac{T\kappa'(m_\mu)}{2c}, \sigma\sqrt{T}\right) dy \\ &= x^{\text{nom}} - \frac{T\kappa'(m_\mu)}{2c} = x^{\text{pi}} + \frac{T}{2c} \delta(m_\mu), \end{aligned}$$

where we use Lemma A.1 between the two first lines in order to recover a gaussian process.

We can now recover the reward:

$$\begin{aligned} B^*(r) &= V^{\text{pi}} + \tau x^{\text{nom}} + 2c\sigma^2 \ln(\zeta_{\mu^*}(q_{\mu^*}(r))) + p q_{\mu^*}(r) \\ &= V^{\text{pi}} + \tau x^{\text{nom}} + q_{\mu^*}(r) [p - \kappa'(m_{\mu^*})] - 2c\sigma^2 \ln\left(\int_{-\infty}^{+\infty} f^{\text{nom}}(y) \exp\left(-y \frac{\kappa'(m_{\mu^*})}{2c\sigma^2}\right) dy\right) \\ &= V^{\text{pi}} + \tau x^{\text{nom}} + \frac{c}{T} [(x^{\text{nom}})^2 - m^2] + q_{\mu^*}(r) \delta(m_{\mu^*}) \\ &= \tau x^{\text{nom}} + \frac{c}{T} [(x^{\text{pi}})^2 - m^2] + q_{\mu^*}(r) \delta(m_{\mu^*}), \end{aligned}$$



where we use Lemma A.1 to get the value of the integral. From the definition of the provider objective,

$$\begin{aligned}
\pi &= pm - \kappa(m) - \int_0^1 B^*(r) dr \\
&= pm - \kappa(m) - \frac{c}{T} [(x^{\text{pi}})^2 - m^2] - m [p - \kappa'(m)] - \tau x^{\text{nom}} \\
&= m\kappa'(m) - \kappa(m) + \left(\frac{x^{\text{pi}} + m}{2}\right) \frac{2c}{T} (m - x^{\text{pi}}) - \tau x^{\text{nom}} \\
&= m\kappa'(m) - \kappa(m) + \left(\frac{x^{\text{pi}} + m}{2}\right) \delta(m) - \tau x^{\text{nom}} .
\end{aligned}$$

### Proof of Proposition 5.2

- (i) By construction, the reward  $\tilde{B}$  is also bounded and decreasing. Then, the cost induced by the additional reward is the same with  $R^*$  and  $\hat{B}$ :

$$\int_{-\infty}^{+\infty} R_{\mu^*}^*(x) f_{\mu^*}(x) dx = \int_0^1 \hat{B}(r) dr .$$

Finally,  $\mu^*$  is also an equilibrium for the reward  $\hat{B}$ :

$$\frac{1}{\hat{\gamma}(\mu^*)} f^{\text{nom}}(x) \exp\left(\frac{\hat{B}(F_{\mu^*}(x))}{2c\sigma^2}\right) = \frac{1}{\gamma^*(\mu^*)} f^{\text{nom}}(x) \exp\left(\frac{R_{\mu^*}^*(x)}{2c\sigma^2}\right) = f_{\mu^*} ,$$

where  $\hat{\gamma}$  and  $\gamma^*$  are computed respectively with  $\hat{B}$  and  $R^*$ . The last equality comes from the characterization of an equilibrium. Therefore, the reward function  $\hat{B}$  satisfies the constraints and produces the same objective value as  $R^*$ . It is also optimal.

- (ii) The proof follows the same ideas as at the previous item.