Is mandating "smart meters" smart?

Thomas-Olivier Léautier*
Toulouse School of Economics (IAE, IDEI, University of Toulouse)

PRELIMINARY AND INCOMPLETE

June 8, 2011

Abstract

This article examines the marginal welfare increase (lower installed generation capacity requirement, energy savings, and emissions reductions) arising from consumers switching to real-time price usage of electricity. It first develops an analytically tractable framework, that produces closed form solutions while matching actual demand data. It then applies the methodology to data from the French power market, and finds that, for small residential and non residential users, the marginal value of switching is around 10 € per customer per year, much lower than the cost of the "smart meters" required to enable the switch estimated around 25 € per customer per year. This finding challenges the economic wisdom of mandating full deployment of "smart meters", a policy adopted in many regions.

1 Introduction

"Smart meters", which allow electric power users to react to real time wholesale prices, are expected to transform the electric power industry. First, price sensitive users will consume less electricity on peak, hence reduce installed generation capacity requirements and emissions of CO_2 and other pollutants. Second, increased elasticity of demand will reduce potential exercise of market power by producers. Finally, smart meters will allow grid operators to improve operations of the grid. These effects are expected to generate significant savings. For example, Faruqui et al. (2009), estimate the annual potential benefits from "smart meters" deployment in all of Europe at € 5.9 billions: € 4.8 billions from reduced capacity cost, € 600 millions from reduced electricity consumption, and € 500 millions from T&D savings. The Present Value of these benefits exceeds the deployment cost, estimated at € 53 billions. As a result, full deployment of "smart meters" is mandated by the European Union and underway in many US states.

Yet, the notion that full deployment is optimal is surprising, since the largest users are orders of magnitude larger than the smallest ones, while the cost of smart meters are not that different. For

*leautier@cict.fr. I am grateful to Claude Crampes and seminar participants at the Toulouse School of Economics for insightful comments, and for a grant from Electricité de France that supported this research. All errors and omissions are mine.

example, the 35 000 largest industrial sites in France represent 43% of electricity consumption, more than the 30 millions residential users (32% of electricity consumption). As a result, the marginal value of a user switching to real-time price usage (i.e., perfectly adapting its consumption to real-time prices) should be (roughly) 1000 times smaller for a residential than for a large industrial customer. Equivalently, the marginal cost of enabling 1% of load to switch to real-time price usage should be 1000 times higher for residential users than for large industrials. This sharp convexity suggests there may exist a deployment ratio after which the marginal cost of enabling customers to switch exceeds the marginal value: optimal deployment ratio may not necessarily be 100%.

Curiously, I have not found any analysis of the optimal deployment level, neither in the academic nor in the policy literature. Allcott (2010) is the only related article, that estimates the value of a 20% users switching to real-time price usage, using market data from the mid Atlantic region of the United States. This article aims to fill that gap, that proposes a rigorous evaluation of the marginal value of the reductions in electricity consumption and required installed generation capacity arising from a fraction of load switching to real-time price usage, and compares it to the cost of installing the necessary "smart" meters.

This article contribution's is twofold. First, it proposes an analytically tractable approximation of the solution to the optimal investment problem for a power system. The general principles of peak-load pricing have been developed in the late 1940s (Boiteux (1949)), and revisited recently (e.g., Borenstein and Holland (2005), Joskow and Tirole (2007)). However, the approximation developed here is the only one I am aware of that provides (almost) closed form solutions to the problem, while closely matching real data. This approximation may be used to examine other issues pertaining to power markets, but also more general issues of sizing and pricing of facilities when demand is uncertain and multiple technologies are available (e.g., infrastructure, cloud computing, etc.)

This article's second contribution is an estimate of the increase in net surplus of a customer switching from constant price to real-time price usage. Using data from the French market, this value is estimated at 9 to 14 €/customer per year for a small non residential customer and 4 to 10 €/customer per year for a small residential customer[1]. As a comparison point, this value is far below the cost of installing smart meters for small customers (residential and non residential), currently estimated around 25 €/meter per year. This result should also hold for other power markets, as long as its main driver does: the marginal impact of a single small customer on the total electricity consumption (and required installed capacity) is much smaller than the cost of a "smart meter".

This analysis does not constitute a full-blown marginal cost benefit analysis. First, it does not include other benefits of "smart meters", such as reduction in metering costs and other optimization for the distribution network owner/operator. Second, the cost of installing a meter is the marginal cost of *enabling* the switch to real-time price usage, hence the analysis does not include the cost of informing consumers and inducing them to switch, as well as data storage and processing costs. Nor

[1] The structure of the French power industry, namely EDF's dominant position, is not pertinent for this analysis.

does it factor in the fact that, for a variety of reasons, not all consumers equipped with "smart meters" will switch.

While not being a full-blown cost benefit analysis, the modest magnitude of the estimated value of a small customer switching (around 10 €/customer per year) has three practical implications. First, additional – and more thorough – cost-benefit analysis is required before policy makers commit to the tens of billions of euros of investment in "smart meters". It may be the case that equipping the largest 20% of consumers provides 80% of the benefits for a fraction of the cost. Second, the economic value from getting a customer to switch seems low compared to a customer's acquisition cost (retailers often mention 20 to 50 €/customer[2]). The business model for energy retailers offering innovative energy savings solutions is therefore unclear. Finally, the benefits of smart meters arise from a variety of sources, and accrue to a different stakeholders (distribution network owner/operator, supplier, consumers). Sharing the costs among these classes will prove complex.

This article is structured as follows. The model used in this article is the one that developed by Borenstein and Holland (2005) and Joskow and Tirole (2007), building on the earlier work by Boiteux (1949). For convenience, Section 2 summarizes its main features and results. The reader familiar with the model can proceed to Section 3, that presents general results on the impact of a marginal switch to real-time price usage. Section 4 then presents the approximation leading to the closed form solution, and the main analytical results. Section 5 discusses the development of numerical simulations for the French market, and presents the main empirical results. Section 6 concludes, that proposes avenues of future work.

2 The model

2.1 Model structure

2.1.1 Uncertainty

Uncertainty is an essential feature of power markets. In this work, demand uncertainty is explicitly modeled. Including production uncertainty does not modify the economic insights. The number of possible states of the world is infinite, and these are indexed by $t \in [0, +\infty)$. $f(t)$ and $F(t)$ are respectively the ex ante probability and cumulative density functions of state t. Since all market participants have the same information about future demand and supply conditions, it is realistic to assume that all participants share a common perspective on $f(t)$ and $F(t)$.

2.1.2 Demand, supply, and rationing

Demand

[2] Source: private communication with the author.

Assumption 1 *Customers are homogeneous, and all have the same underlying demand $D(p;t)$ in state t up to a scaling factor, non increasing in p, the electric power price.*

Assumption 1 greatly simplifies the derivations, while preserving the main economics insights. Inverse demand is $P(q;t)$ is defined by $D(P(q;t);t) = q$, and gross consumers surplus is $S[p;t] = \int_0^{D(p;t)} P(q;t)\,dq$.

Without loss of generality, states of the world are ordered by increasing demand:

$$\frac{\partial D}{\partial t}(p;t) \geq 0$$

Customers are split in two categories: a fraction α of consumers faces and react to real time wholesale price ("price reactive" consumers), and a fraction $(1-\alpha)$ of consumers faces a two-part pricing scheme, with price p^R per MWh, constant across all states of the world, and connection charge A per year ("constant price" consumers).

Since all consumers have the same load profile up to a scaling factor by Assumption 1, α is constant across states of the world.

Assumption 2 *The SO has the technical ability to curtail "constant price" load while not curtailing "price reactive" load.*

Assumption 2 is unrealistic today, as the SO can only organize curtailment by zone, and cannot differentiate by type of customer. However, it will be fairly realistic when "smart meters" are being rolled out, which is precisely the situation considered.

Supply N generation technologies are available. c_n is the marginal cost, and r_n is the hourly investment cost (i.e., annual investment cost expressed in €$/MW/year$ divided by 8760 hours per year) of technology $n \in \{1,...,N\}$, both expressed in €$/MWh$. Generation technologies are ordered by increasing marginal cost: $c_n > c_m \ \forall\ n \geq m$.

As described in Section 4, not all available technologies are included in the optimal investment plan. However, to simplify the exposition, $n = 1$ (resp. $n = N$) denotes the first (resp. the last) technology used at the optimum.

There is a trade-off between investment and marginal costs: if a technology requires lower investment cost, it then produces at higher variable cost, i.e., $r_n < r_m \ \forall\ n \geq m$.

Rationing and Value of Lost Load Denote $\gamma \in [0,1]$ the *serving ratio*: $\gamma = 0$ means full curtailment, while $\gamma = 1$ means no curtailment. For state t, $\mathcal{D}(p,\gamma;t)$ is the demand for price p and serving ratio γ, and $\mathcal{P}(q,\gamma;t)$ is the inverse demand for a given serving ratio γ, defined by $\mathcal{D}(\mathcal{P}(q,\gamma;t),\gamma;t) = q$. Then $\mathcal{S}(p,\gamma;t) = \int_0^{\mathcal{D}(p,\gamma;t)} \mathcal{P}(q,\gamma;t)\,dq$ is the gross consumer surplus. We verify that: $\frac{\partial \mathcal{S}(\mathcal{D}(p,\gamma;t),\gamma;t)}{\partial p} = p\frac{\partial \mathcal{D}}{\partial p}$.

Any rationing technology satisfies: (1) $\mathcal{D}(p,0;t)=0$, (2) $\frac{\partial \mathcal{D}}{\partial \gamma}>0$ for $\gamma\in[0,1]$, and (3) $S(p;t)\equiv \mathcal{S}(p,1;t)$ and $D(p;t)\equiv \mathcal{D}(p,1;t)$.

The Value of Lost Load ($VoLL$) represents the value consumers would place on an extra unit of non delivered electricity. Formally, it is defined as

$$v(p,\gamma;t)=\frac{\frac{\partial \mathcal{S}}{\partial \gamma}}{\frac{\partial \mathcal{D}}{\partial \gamma}}(p,\gamma;t)$$

Assumption 3 *1. Rationing does not increase the net surplus:* $\forall p>0,\forall t\geq 0,\forall \gamma>0$

$$\mathcal{S}(p,\gamma;t)-p\mathcal{D}(p,\gamma;t)\leq S(p;t)-pD(p;t)$$

2. If the serving ratio is positive, the Value of Lost Load is always higher than the price of power, i.e., $\forall p>0,\forall t\geq 0,\forall \gamma>0$ *we have:*

$$v(p,\gamma;t)>p$$

Assumption 3 holds for example for anticipated and proportional rationing: $\mathcal{S}(p,\gamma)=\gamma S(p)$ and $\mathcal{D}(p,\gamma)=\gamma D(p)$, hence (1) $\mathcal{S}(p,\gamma;t)-p\mathcal{D}(p,\gamma;t)=\gamma(S(p;t)-pD(p;t))\leq S(p;t)-pD(p;t)$ for $\gamma\leq 1$; and (2) $v(p,\gamma;t)=\frac{S(p)}{D(p)}>p$.

Assumption 3 should hold for all possible rationing technologies: rationing does not increase net surplus, and consumers are always willing to pay at least as much for a MWh when curtailment is possible as they are for a MWh in normal circumstances.

2.2 Optimal dispatch and investment

2.2.1 First-order conditions

Under Assumption 1 to 3, Joskow and Tirole (2007) show that it is never optimal to ration "price reactive" customers. The total consumer surplus and demand in state t are therefore:

$$\begin{cases}\tilde{S}\left(p,p^R,\gamma,\alpha;t\right)=\alpha S(p;t)+(1-\alpha)\mathcal{S}\left(p^R,\gamma;t\right)\\ \tilde{D}\left(p,p^R,\gamma,\alpha;t\right)=\alpha D(p;t)+(1-\alpha)\mathcal{D}\left(p^R,\gamma;t\right)\end{cases}$$

The optimal program is then:

$$W(\alpha)=\begin{cases}\max\limits_{p(.),p^R,\gamma(.),u_n(.),k_n}\mathbb{E}\left\{\tilde{S}\left(p(t),p^R,\gamma(t),\alpha;t\right)-\sum\limits_{n\geq 1}c_n u_n(t)k_n\right\}-\sum\limits_{n\geq 1}r_n k_n\\ \quad st:\ \forall t\geq 0\quad \tilde{D}\left(p(t),p^R,\gamma(t),\alpha;t\right)\leq\sum\limits_{n\geq 1}u_n(t)k_n\quad(\lambda(t))\end{cases}$$

where $p(t)$ is the price faced by price reactive customers, $\gamma(t)\in[0,1]$ the serving ratio, $u_n(t)\in[0,1]$ the dispatch ratio of technology n, $\lambda(t)\geq 0$ the Lagrange multiplier in state t, p^R the optimal retail price, $k_n \geq 0$ the optimal investment in technology n.

The Lagrangian is:

$$\mathcal{L} = \mathbb{E}\left\{\tilde{S} - \sum_{n\geq 1} c_n u_n(t) k_n + \lambda(t) \left[\sum_{n\geq 1} u_n(t) k_n - \tilde{D}\right]\right\} - \sum_{n\geq 1} r_n k_n$$

and the first-order derivatives are:

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial p(t)} = \alpha \left(p(t) - \lambda(t)\right) D_p \\ \frac{\partial \mathcal{L}}{\partial u_n(t)} = \left(\lambda(t) - c_n\right) k_n \\ \frac{\partial \mathcal{L}}{\partial \gamma(t)} = (1-\alpha)\left(v(t) - \lambda(t)\right) \mathcal{D}\gamma \\ \frac{\partial \mathcal{L}}{\partial p^R} = (1-\alpha) \mathbb{E}\left[\left(p^R - \lambda(t)\right) \mathcal{D}_p\right] \\ \frac{\partial \mathcal{L}}{\partial k_n} = \mathbb{E}\left(\left[\lambda(t) - c_n\right] u_n(t)\right) - r_n \end{cases}$$

The first-order conditions yield familiar results (see for example Borenstein and Holland (2005) and Joskow and Tirole (2007), who also discuss sufficient conditions for the program to be concave).

First, $\frac{\partial \mathcal{L}}{\partial p(t)} = 0$ yields $p(t) = \lambda(t)$: price reactive customers pay the opportunity cost of electricity in each state.

Second, $\frac{\partial \mathcal{L}}{\partial u_n(t)}$ yields the dispatch rule:

$$u_n(t) = \begin{cases} 1 & if \quad c_n < p(t) \\ 0 & if \quad c_n > p(t) \\ \frac{\tilde{D} - \sum_{m<n} k_m}{k_n} & if \quad c_n = p(t) \end{cases}$$

Technology n produces at capacity (resp. does not produce) if its marginal cost is lower than the price (resp. exceeds the price) in state t. If technology n is marginal, i.e., price setting, energy balance sets the dispatch ratio. $p(t) = \lambda(t) > 0$ is therefore the real time wholesale electricity price.

Third, $\frac{\partial \mathcal{L}}{\partial \gamma(t)}$ yields the rationing rule:

$$\gamma(t) = \begin{cases} 1 & if \quad v(t) > p(t) \\ 0 & if \quad v(t) < p(t) \\ \frac{\tilde{D} - \sum_{m<n} k_m}{k_n} & if \quad v(t) = p(t) \end{cases}$$

Rationing occurs if only if the $VoLL$ is no higher than the real time price.

Fourth, $\frac{\partial \mathcal{L}}{\partial p^R} = 0$ yields:

$$p^R = \frac{\mathbb{E}\left[p(t) \mathcal{D}_p\right]}{\mathbb{E}\left[\mathcal{D}_p\right]}$$

As in Joskow and Tirole (2007), the optimal retail price is the weighted average wholesale price, where the weights are the marginal "rationed demand". Since $\mathcal{D}_p < 0$ and $\lambda(t) \geq 0$, $p^R \geq c$: the optimal retail price is always higher than the marginal cost. However, the retail price needs not cover the full production cost. The fixed part of two part retail price balances the retailers' profits.

Finally, $\frac{\partial \mathcal{L}}{\partial k_n} = 0$ yields:

$$\mathbb{E}\left[\left(p\left(t\right) - c_n\right) u_n\left(t\right)\right] = r_n$$

The optimal capacity is such that the marginal profit when the plant operates equals the investment (or capacity) cost.

2.2.2 Investment plan

If the program is concave (which we assume here), the first-order conditions determine a unique optimum. Denote $K_n = \sum_{m=1}^{n} k_m$ the total installed capacity up to and including technology n, t_n (resp. $\bar{t}_n$) the first state of the world where technology n is dispatched (resp. is at capacity). Adopt the convention $t_{N+1} \to +\infty$. From the previous first order conditions, real time price equals marginal cost of technology n when this one is marginal, but not yet constrained. When technology n is at capacity, but technology $(n+1)$ is not yet dispatched, the energy balance determines the price:

$$\alpha D\left(p;t\right) + \left(1-\alpha\right)\mathcal{D}\left(p^R,\gamma;t\right) = K_n$$

Hence:

$$p\left(t\right) = \begin{cases} c_n & for \quad t \in [t_n, \bar{t}_n] \\ \tilde{P}\left(K_n;t\right) = P\left(\frac{K_n - (1-\alpha)\mathcal{D}\left(p^R,\gamma;t\right)}{\alpha}\right) & for \quad t \in [\bar{t}_n, t_{n+1}] \end{cases}$$

for $n \in \{1, ..., , N\}$. By construction, we have:

$$\begin{cases} \tilde{P}\left(K_n;\bar{t}_n\right) = c_n \\ \tilde{P}\left(K_n;t_{n+1}\right) = c_{n+1} \end{cases}$$

This is illustrated on Figure 1.

As the price $p\left(t\right)$ is increasing, we have: $u_n(t) \geq 0 \; \forall \; t \geq t_n$ and $u_n(t) = 1 \; \forall \; t \geq \bar{t}_n$. The first-order condition determining k_n then becomes:

$$\mathbb{E}\left[\left(\tilde{P}\left(K_n;t\right) - c_n\right)\right]_{t \geq \bar{t}_n} = r_n \tag{1}$$

Equation (1) for $n = N$ yields:

$$\mathbb{E}\left[\left(\tilde{P}\left(K_N;t\right) - c_N\right)\right]_{t \geq \bar{t}_N} = r_N \tag{2}$$

Combined with $\tilde{P}\left(K_N;\bar{t}_N\right) = c_N$, we can solve for $\left(K_N, \bar{t}_N\right)$. As is well known, the total installed capacity is determined by the long-run marginal cost of the last invested technology.

Then, for $1 \leq n < N$, equation (1) yields:

$$\int_{\bar{t}_n}^{t_{n+1}} \left[\tilde{P}\left(K_n;t\right) - c_n\right] f\left(t\right) dt + \int_{t_{n+1}}^{\bar{t}_{n+1}} \left[c_{n+1} - c_n\right] f\left(t\right) dt + \int_{\bar{t}_{n+1}}^{+\infty} \left[\tilde{P}\left(t\right) - c_n\right] f\left(t\right) dt = r_n$$

$\Leftrightarrow$

$$\int_{\bar{t}_n}^{t_{n+1}} \left[\tilde{P}\left(K_n;t\right) - c_n\right] f\left(t\right) dt + \int_{t_{n+1}}^{+\infty} \left[c_{n+1} - c_n\right] f\left(t\right) dt = r_n - r_{n+1} \tag{3}$$

Combined with $\tilde{P}\left(K_n;t_{n+1}\right) = c_{n+1}$ and $\tilde{P}\left(K_n;\bar{t}_n\right) = c_n$ we solve for $\left(K_n, \bar{t}_n, t_{n+1}\right)$.

3 Increasing the proportion of price reactive customers

This section examines the impact of a marginal increase in α assuming all values, in particular, generation capacity and mix and retail prices, are optimal. This represents the long-term equilibrium. One could challenge this choice as being unrealistic: where are the installed generation mix and the retail price optimal? However, I believe this is the appropriate analysis, as it isolates the impact of switching to real-time price usage. Other analyses would run the risk of mixing this with the impact of moving closer to the optimum.

3.1 Impact on welfare

We first establish the following:

Result 1 *Increasing the proportion of price-reactive customers always increases the net surplus from consumption, since price-reactive customers are not rationed, and consume in each state according to the state-contingent price and not a fixed price (even optimally chosen).*
Proof. *The envelope theorem yields:*

$$\frac{dW}{d\alpha} = \frac{\partial W}{\partial \alpha} = \mathbb{E}\left\{\left[S\left(p\left(t\right);t\right) - p\left(t\right) D\left(p\left(t\right);t\right)\right] - \left[\mathcal{S}\left(p^R, \gamma\left(t\right);t\right) - p\left(t\right) \mathcal{D}\left(p^R, \gamma\left(t\right);t\right)\right]\right\}$$

then:

$$\mathcal{S}\left(p^R, \gamma\left(t\right);t\right) - p\left(t\right) \mathcal{D}\left(p^R, \gamma\left(t\right);t\right) < S\left(p^R;t\right) - p\left(t\right) D\left(p^R;t\right)$$

since rationing does not generate value; and

$$S\left(p^R;t\right) - p\left(t\right) D\left(p^R;t\right) < S\left(p\left(t\right);t\right) - p\left(t\right) D\left(p\left(t\right);t\right)$$

since $p = argmax_x \left\{S\left(x;t\right) - pD\left(x;t\right)\right\}$. ■

Result 1 differs from Borenstein and Holland (2005), who propose a counter-example, where increasing the share of price-reactive consumers reduces overall welfare. However, Borenstein and Holland (2005) operate in a slightly different setting, who assume that the retail price is set so that the retail profit is equal to zero. In this work, however, we consider, following Joskow and Tirole (2007), that retailers' budget balance can be achieved by a two-part tariff. In that case, the variable part of the retail price is chosen optimally, and the envelope theorem then applies.

Result 1 matters for methodological reasons. Numerous analyses (e.g., Faruqui et al. (2009)) consider all demand reduction as a benefit. Result 1 shows this is incorrect, as one should also include the (lost) value of the foregone consumption in the analysis.

3.2 Impact on average price

Result 2 *Increasing the share of price reactive customers has no impact on the expected price.*
Proof. *By construction, we have:* $t_1 = 0$*. Equation (1) for* $n = 1$ *then yields:*

$$\mathbb{E}\left[(p(t) - c_1)\right]_{t \geq \bar{t}_1} = r_1$$

$\Leftrightarrow$

$$\mathbb{E}\left[p(t)\right]_{t \geq \bar{t}_1} = r_1 + c_1 \Pr(t \geq \bar{t}_1)$$

Then:

$$\mathbb{E}\left[p(t)\right] = c_1 \Pr(t \leq \bar{t}_1) + \mathbb{E}\left[p(t)\right]_{t \geq \bar{t}_1} = r_1 + c_1$$

hence

$$\frac{d\mathbb{E}\left[p(t)\right]}{d\alpha} = 0$$

■

Since the first technology dispatched produces in all states of the world, the zero-profit condition implies that the expected spot price is simply its long run marginal cost, i.e., the sum of its marginal and capacity cost. It is therefore independent of the share of price sensitive customers. This contradicts commonly held wisdom that real time pricing lowers average power price.

3.3 A specific case: linear demand, rationing linear and anticipated

In this Section, we assume demand is linear and is given by:

$$P(q, t) = a(t) - bq$$

We also assume that no rationing occurs at the optimum. This assumption is justified, since in most power markets, the largest customers already face real prices: demand elasticity is then high enough that no rationing be required to balance demand and supply in the high states of the world. It is verified in all numerical applications presented in Section 5.

We then establish the following:

Result 3 *When demand is linear and rationing does not occur at the optimum, the marginal net surplus is proportional to the spot price volatility.*
Proof. *Since demand is linear* $D(p,t) = \frac{a(t)-p}{b}$, *then*

$$\begin{aligned} S(p;t) &= \int_0^{D(p;t)} (a(t) - bq)\, dq = \left(a(t) - \frac{b}{2} D(p;t)\right) D(p;t) \\ &= \left(\frac{a(t)+p}{2}\right) D(p;t) \end{aligned}$$

Hence:

$$\begin{aligned} S(p;t) - pD(p;t) &= S(p;t) - pD(p;t) \\ &= \left(\frac{a(t)-p}{2}\right) D(p;t) = \frac{b}{2} D^2(p;t) \end{aligned}$$

Since there is no rationing at the optimum, we have: $\mathcal{D}_p = D_p = -\frac{1}{b}$ *hence*

$$p^R = \mathbb{E}[\lambda(t)] = c_1 + r_1$$

and

$$\begin{aligned} \mathcal{S}\left(p^R;t\right) - p\mathcal{D}\left(p^R;t\right) &= S\left(p^R;t\right) - pD\left(p^R;t\right) \\ &= \left(a(t) - \frac{bD\left(p^R;t\right)}{2} - (a(t) - bD(p;t))\right) D\left(p^R;t\right) \\ &= b\left(D(p;t) - \frac{D\left(p^R;t\right)}{2}\right) D\left(p^R;t\right) \end{aligned}$$

Then:

$$\begin{aligned} W'(\alpha) &= \frac{b}{2}\mathbb{E}\left\{\left(D(p(t);t) - D\left(p^R;t\right)\right)^2\right\} \\ &= \frac{1}{2b}\mathbb{E}\left\{\left(p^R - p(t)\right)^2\right\} = \frac{1}{2b} Var(p(t)) \end{aligned}$$

■

4 A closed form solution

A closed form solution is available if we further assume that $a(t) = a_0 - a_1 e^{-\lambda_2 t}$ and $f(t) = \lambda_1 e^{-\lambda_1 t}$. As will be shown in Section 5, for an optimal choice of the parameters $(a_0, a_1, \lambda_1, \lambda_2)$, this specification is consistent with observed load duration curves and estimated price elasticities, while leading to simple expressions for the values of interest. Richer specifications will be tested in further work. However, initial tests suggest that the results hold for changes in the parameters, hence the results are likely to be robust.

4.1 Optimal investment

Result 4 *The marginal technology N is the last technology such that*

$$c_N - c_{N-1} < \left(\frac{a_1}{\alpha}\right)^{\frac{\lambda}{1+\lambda}} \left([(1+\lambda) r_{N-1}]^{\frac{1}{1+\lambda}} - [(1+\lambda) r_N]^{\frac{1}{1+\lambda}} \right)$$

where $\lambda = \frac{\lambda_1}{\lambda_2}$. The optimal total capacity K_N is then the solution of:

$$\left[a_0 - bK_N - \left(\alpha c_N + (1-\alpha) p^R\right)\right]^{1+\lambda} = \alpha a_1^\lambda (1+\lambda) r_N \tag{4}$$

If

$$\left(a_0 - (1-\alpha) p^R - \alpha c_n\right)^{1+\lambda} - \left(a_0 - (1-\alpha) p^R - \alpha c_{n+1}\right)^{1+\lambda} > \alpha a_1^\lambda (1+\lambda)(r_n - r_{n+1})$$

K_n, the optimal capacity up to technology $n < N$, is the unique solution of:

$$\left(a_0 - bK_n - (1-\alpha) p^R - \alpha c_n\right)^{1+\lambda} - \left(a_0 - bK_n - (1-\alpha) p^R - \alpha c_{n+1}\right)^{1+\lambda} = \alpha a_1^\lambda (1+\lambda)(r_n - r_{n+1}) \tag{5}$$

otherwise, $K_n = 0$.
Proof. *The proof is presented in the appendix.* ■

The structure of result 4 is standard in the peak-load pricing literature. The demand and uncertainty specification selected here allows us to derive simple expressions, hence highlight the economic intuition. The first condition states that it is optimal to invest in a higher marginal cost technology as long as the marginal cost increase is lower than (a function of) the increase in investment cost. Then, equation (4) determines the optimal total capacity K_N, that depends only on the marginal and investment costs of the marginal technology N (and of course demand parameters and the fixed retail price) It is then optimal to invest in inframarginal technologies as long as (a function of) the reduction in marginal cost exceeds the investment cost increase. Equations (5) then determines the switching points between technologies n and $(n+1)$, that depend on the marginal and investment costs of both technologies.

4.2 No rationing conditions

No rationing is optimal as long as price is lower than the Value of Lost Load, which yields:

Result 5 *No rationing occurs at the optimum if and only if $R_n > 0$ where*

$$R_n = \begin{cases} \left[a_0 - 2c_{n+1} + p^R + \alpha (c_{n+1} - c_n)\right]^{1+\lambda} - \left[a_0 - 2c_{n+1} + p^R\right]^{1+\lambda} \\ \quad -\alpha a_1^\lambda (1+\lambda)(r_n - r_{n+1}) & n \in \{1, ..., (N-1)\} \\ \left[\frac{a_0 - 2c_N + p^R}{2}\right]^{1+\lambda} - \left(\frac{a_1}{\alpha}\right)^\lambda [(1+\lambda) r_N] & n = N \end{cases} \tag{6}$$

Proof. *The proof is presented in the appendix.* ■

The first conditions are for $t \in [\bar{t}_n, t_{n+1}]$ for $n \in \{0, ..., (N-1)\}$ and the last condition is for $t \geq \bar{t}_N$. From this last condition, rationing must occur for $\alpha = 0$, and by continuity, for low values of α.

4.3 Marginal value of real-time pricing

Result 6 *We have:*

$$W'(\alpha) = \frac{8760}{b} \begin{cases} \sum_{n=1}^{N-1} \left[\left(\frac{a_1}{\alpha}\right)^2 \lambda_2 J_n + r_{n+1}(c_{n+1} - c_n) \right] + \left(\frac{a_1}{\alpha}\right)^{\frac{\lambda}{1+\lambda}} \frac{[(1+\lambda) r_N]^{\frac{2+\lambda}{1+\lambda}}}{(1+\lambda)(2+\lambda)} - \frac{r_1^2}{2} & N > 1 \\ r_1 \left[\left(\frac{a_1}{\alpha}\right)^{\frac{\lambda}{1+\lambda}} \frac{[(1+\lambda) r_1]^{\frac{1+\lambda}{1+\lambda}}}{2+\lambda} - \frac{r_1}{2} \right] & N = 1 \end{cases} \quad (7)$$

where $J_n = \int_{\bar{t}_n}^{t_{n+1}} \left(e^{-\lambda_2 \bar{t}_n} - e^{-\lambda_2 t} \right) e^{-(\lambda_1 + \lambda_2) t} dt.$

Proof. *The proof is presented in the appendix.* ■

Intuition for expression (7) for $N = 1$ is as follows:

$$\begin{aligned} Var[p(t)] &= Var[p(t) - c_1] = \mathbb{E}\left[(p(t) - c_1)^2\right] - (\mathbb{E}[p(t) - c_1])^2 \\ &= \int_{\bar{t}_1}^{+\infty} (p(t) - c_1)^2 f(t)\, dt - r_1^2 \end{aligned}$$

The specific shapes of $p(t)$ and $f(t)$ then yield $\int_{\bar{t}_1}^{+\infty} (p(t) - c_1)^2 f(t)\, dt = 2r_1 \left(\frac{a_1}{\alpha}\right)^{\frac{\lambda}{1+\lambda}} \frac{[(1+\lambda) r_1]^{\frac{1+\lambda}{1+\lambda}}}{2+\lambda}$, which then yields expression (7) for $N = 1$.

For $N > 1$, the same procedure is applied to each interval $\Omega_n = [t_n, t_{n+1}]$, for $n = 1, ..., N$:

$$Var[p(t)] = \sum_{n=1}^{N} \left\{ \int_{t_n}^{t_{n+1}} (p(t) - c_n)^2 f(t)\, dt - \left(\int_{t_n}^{t_{n+1}} (p(t) - c_n) f(t)\, dt \right)^2 \right\}$$

The computation is more complex, and not all terms cancel out, yielding expression (7) for $N > 1$.

For a given demand profile $(a_0, a_1, b, \lambda)$and available technology mix $\{c_n, r_n\}_n$, equation (7) can be used estimate the marginal value of real time pricing, as in Section 5 for the French market. Furthermore, it can be used to perform comparative statics on the marginal value of switching to real time pricing.

5 Optimal deployment of real time meters: application to the French market

To determine the optimal proportion of price reactive load α, estimates of (1) the demand curve, and (2) the marginal cost of increasing α are required.

5.1 Demand curve

The demand curve parameters are estimated in two steps: (i) an actual load duration curve, assuming price is constant, is used to estimate λ and derive a first set of relationships, and (ii) estimates of price elasticity are then used to derive the last relation among parameters. This approach is consistent with the reality of power markets: today, most customers pay a constant power price, denoted p_0. Observed demand fluctuations are due therefore to variations in the states of the world ($a(t)$ and $f(t)$). As demand becomes more price reactive, joint estimation of all parameters will become possible.

5.1.1 Estimation of λ and first set of relationships

Denote $G(.)$ the cumulative distribution of demand, i.e., $G(x)$ is the probability that demand is lower than x. If demand is linear:

$$\begin{aligned} G(x) &= \Pr\left(\frac{a(t)-p_0}{b} \leq x\right) = \Pr\left(a(t) \leq x + bp_0\right) \\ &= \Pr\left(t \leq a^{-1}(x+bp_0)\right) = F \circ a^{-1}(x+bp_0) \end{aligned}$$

Demand measured depends both on the state of the world t and demand conditional on that state of the world t. Estimating the distribution $G(.)$ allows us to identify $F \circ a^{-1}$. $F(.)$ and $a(.)$ cannot be identified separately.

If $a(t) = a_0 - a_1 e^{-\lambda_2 t}$ and $f(t) = \lambda_1 e^{-\lambda_1 t}$:

$$\begin{aligned} G(x) &= 1 - \exp\left[-\frac{\lambda_1}{\lambda_2} \ln \frac{a_0-(x+bp_0)}{a_1}\right] \\ &= 1 - \left[\frac{a_0-(x+bp_0)}{a_1}\right]^{\frac{\lambda_1}{\lambda_2}} \end{aligned}$$

Then, $1 - G(x) = \Pr(load \geq x) = \left[\frac{a_0-(x+bp_0)}{a_1}\right]^{\lambda}$ can be estimated from an actual load duration curve.

a_0 and a_1 cannot be estimated by Maximum Likelihood from the data. The minimum and maximum admissible values for load must be set exogenously. We choose these values to be the observed minimum and maximum values for load. Denote $\phi < 1$ the ratio of minimum to maximum demand for price p_0 and $Q^{\infty} = \lim_{t \to +\infty} Q(p_0, t) = \frac{a_0-p_0}{b}$ the maximum demand. We have:

$$\begin{cases} a_0 - bQ^{\infty} = p_0 \\ a_0 - a_1 - b\phi Q^{\infty} = p_0 \end{cases}$$

which yields:

$$\begin{cases} a_1 = bQ^{\infty}(1-\phi) \\ a_0 = p_0 + bQ^{\infty} \end{cases}$$

Estimation on 2009 demand in France (source: RTE website) leads to $Q^{\infty} = 92.4$ GW and $\phi =$

$\frac{31.5}{92.4} = 0.34$. Then, Maximum Likelihood estimation yields $\lambda = \frac{\lambda_1}{\lambda_2} = 1.78$.

Actual and fitted demand are presented on Figure 2.

5.1.2 Estimation of b and all other parameters

As of today, the empirical literature on price elasticity of electricity is inconclusive, which is not surprising, as most end-consumers pay a fixed price for electric power. Lijesen (2007) provides an up to date survey, as well as his own estimate.

Using the average elasticity of demand η for a given price δ and $\phi < 1$, we have:

$$\eta = -\frac{1}{b}\frac{\delta}{\mathbb{E}\left[Q\left(\delta, t\right)\right]} = -\frac{\delta}{\mathbb{E}\left[a\left(t\right)\right] - \delta} = -\frac{\delta}{\left(a_0 - \frac{a_1}{2}\right) - \delta}$$

hence

$$a_0 = \frac{a_1}{2} + \delta\left(1 - \frac{1}{\eta}\right)$$

We have $\delta = p_0 = 100$ €$/MWh$. This then leads to:

$$\begin{cases} bQ^{\infty} = -\frac{2\delta}{\eta(1+\phi)} \\ a_0 = \delta\left(1 - \frac{2}{\eta(1+\phi)}\right) \\ a_1 = -\frac{2\delta(1-\phi)}{\eta(1+\phi)} \end{cases}$$

From Lijesen (2007), we select as a base case $\eta = -0.05$ (at price $\delta = 100$ €$/MWh$), which corresponds to the upper estimate from Patrick and Wolak (1997) using UK data, and much higher than Lijesen (2007) own estimate on Dutch data. We also run a robustness check with $\eta = -0.1$. A higher elasticity of demand will render real-time pricing more attractive. This high elasticity case should therefore provide an upper bound for the optimal α.

With these values, for the base case, we have:

$$\begin{cases} bQ^{\infty} = 2\ 984 \text{ €}/MWh \\ a_0 = 3\ 084 \text{ €}/MWh \\ a_1 = 1\ 968 \text{ €}/MWh \end{cases}$$

Consider now the units. Equation (7) can be rewritten as $W'(\alpha) = 8760 Q^{\infty} \frac{L_N}{bQ^{\infty}}$, where $L_N = \frac{Var[p(t)]}{2}$. L_N is expressed in €$^2 \times MWh^{-2}$, which is denoted as $[L_N] =$€$^2 \times MWh^{-2}$. Then $\left[8760 Q^{\infty} \frac{L_N}{bQ^{\infty}}\right] = 10^3$€$/year$, since $[Q^{\infty}] = GW$, $[bQ^{\infty}] =$€$/MWh$, and $[8760] = h/year$. Finally, $\left[W'(\alpha) = 8.76 Q^{\infty} \frac{L_N}{bQ^{\infty}}\right] =$€ millions$/year$.

5.2 Production cost

French electricity is mostly produced from nuclear assets, with gas turbines providing peaking capacity. As a first approximation, this article ignores hydraulic assets and other thermal generation units. This reduces the marginal value of switching to real time price usage, as including these technologies would increase generation flexibility, hence reduce the value of demand flexibility.

IEA (2010) provides the following estimates for the cost of nuclear assets ($n = 1$) and gas turbines ($n = 2$):

	1	2
c_n	10.99	71.56
r_n	34.16	6.00

c_2 includes a 25 €/ton carbon price. The marginal value of switching therefore includes the environmental cost of emissions. r_2 is equivalent to 53 $€/kW/year$, slightly lower than most commonly used estimates of the annual fixed cost of peaking capacity (around 70 $€/kW/year$). The difference is attributable mostly to taxes. This is justified as this analysis examines the net total welfare, and taxes are internal transfers that do not affect it.

5.3 Cost of real time meters

Each real time meter is estimated to cost €250, a rather conservative estimate, as other estimates range around €500 per meter[3]. As a first approximation, this cost is assumed to be independent of the characteristics of the site where the meter is installed, in particular peak-demand. Assuming a cost of capital at 10%, the annualized cost of each meter is 25 €/meter/year.

Denote $C(\alpha)$ is the annualized cost of installing real-time meters for a proportion α of the total load. $C(\alpha) = 25\times$(number of sites required to reach the fraction α of the total load). We estimate $C(\alpha)$ for France, using data provided by the Commission de Régulation de l'Energie (CRE). The total cost for the 34.8 millions sites is around 8.7 € billions, which corresponds to $C(1) = 870$ millions € per year.

CRE provides the total number of sites and the total consumption (MWh) for four categories of customers:

1. large non residential: demand higher than 250 kW, representing 0.1% of the total number of sites, and 43% of total demand
2. medium non residential: peak-demand between 36 and 250 kW, representing 1% of the total number of sites, and 15% of total demand
3. small non residential: peak demand smaller than 36 kVA, 13% of the total number of sites, 10% of total demand
4. residential sites: peak demand lower than 36 kVA, 86% of the total number of sites, 32% of total demand.

[3]Sources: http://www.freenews.fr/spip.php?article8878?? , private communications with the author.

All customers in each class are assumed to have the same size. This is a conservative assumption, as allowing for differing sizes would increase the convexity of the cost function. After a few manipulations, we find:

$$C'(\alpha) = \begin{cases} 2 & if \quad \alpha \leq 43\% \\ 58 & if \quad 43\% \leq \alpha \leq 58\% \\ 1130 & if \quad 58\% \leq \alpha \leq 68\% \\ 2336 & if \quad 68\% \leq \alpha \end{cases}$$

where $C'(\alpha)$ is measured in € millions per percent per year. For example, the incremental cost of a 1% deployment for residential customers ($\alpha \geq 68\%$) is $2336 \times 10^{-2} = 23.36$ € million per year. The marginal cost increases rapidly, as the number of sites required to increase α by 1% increases significantly as sites become smaller.

It is also helpful to present the marginal value of one consumer switching to real-time pricing. Denote $\delta\alpha$ the incremental increase in α from a single consumer. Since all customers in each class are assumed to have the same size, $\delta\alpha$ is constant for each class, and given by:

$\alpha\,(\%)$	$(0,43)$	$(43,58)$	$(58,68)$	$(68,100)$
$\delta\alpha\,(\%/user)$	1.24×10^{-5}	4.31×10^{-7}	2.21×10^{-8}	1.07×10^{-8}

Then, δW (resp. δC), the incremental increase in net surplus (resp. incremental cost) from one consumer is estimated as $\delta W = W'(\alpha)\,\delta\alpha$ (resp. $\delta C = C'(\alpha)\,\delta\alpha$). Since $\delta\alpha$ is discontinuous at the boundaries between classes, so is δW ($\delta C = 25$ is continuous, while $C'(\alpha)$ is discontinuous). For these values, δW^- and δW^+(incremental values respectively for $\delta\alpha < 0$ and $\delta\alpha > 0$) are evaluated and reported.

5.4 Simulation results

K_1/Q^∞, K_2/Q^∞, $W'(\alpha)$ and δW are computed using equations (4), (5), and (7), and the previous values of the parameters.

For the base case $\eta = -0.05$:

α (%)	15	43	58	68	100
K_1/Q^∞ (%)	59.3	59.3	59.3	59.3	59.4
K_2/Q^∞ (%)	95.7	92.7	91.6	90.9	89.1
$W'(\alpha)$ (€ millions/year)	688	436	389	366	324
δW^- (€ /user/year)	8512	5386	168	8	3
δW^+ (€ /user/year)	8512	188	9	4	n/a

The no rationing conditions (6) are met for $\alpha > 15\%$.

Consider first the evolution of installed capacity generation (expressed here as a fraction of peak demand). The total installed capacity decreases as aggregate demand becomes more price reactive. For example, (1) increasing the share of price reactive demand from 15% to 100% reduces the installed capacity by 6.9%, slightly lower than Faruqui et al. (2009) who assumes a 10% decrease in peak demand, and (2) increasing α from 15% to 35% reduces the installed capacity by 2.4%, slightly higher than Allcott (2010), who estimates a 1.9% peak demand decrease from a 20% switch.

The nuclear installed capacity increases slightly.

Consider now the marginal value of switching. $W'(\alpha)$ decreases sharply as α increases, in particular for low values of α. The discontinuity at the boundaries between customer class is artificial. In reality, sites are continuously getting smaller. It may be that the optimal deployment should exclude the smallest medium non residential sites, or include the largest small non residential sites.

These numbers are lower from Allcott's (2010), who estimates the average welfare increase at 38.9 \$ per kW of (peak) demand switching per year (around 30 $€/kW/year$). Assuming the peak demand shares by class are equal to total demand shares by class, the marginal value per kW of peak demand ranges from 7 $€/kW/year$ for large industrials to 13 $€/kW/year$ for residential users. Further work will investigate this difference.

It is also helpful to compare the marginal value of switching to the marginal cost of installing smart meters. $\delta W^{-}(43\%) = 5386\,€/kW/year > \delta C = 25\,€/kW/year$ and $\delta W^{-}(58\%) = 168\,€/kW/year > \delta C$: for large and medium non residential customers, the value of switching exceeds the cost of installing a smart meter. If a mechanism can be devised for them to appropriate a share of this surplus, these customers will accept to pay for installation of the meters, and switch to real time pricing.

However, $\delta W^{+}(58\%) = 9\ €/kW/year < \delta C$: the marginal value of the "first" small non residential customer switching is lower than the marginal cost of installing a "smart meter". As mentioned in the introduction, this finding does not constitute a full blown cost-benefit analysis. Including additional costs and benefits, installing smart meters for small customers may still be socially optimal.

Consider now the very elastic demand case, $\eta = -0.1$ at price $\delta = 100\ €/MWh$. Following the procedure described above:

α (%)	16	43	58	68	100
K_1/Q^∞ (%)	61.1	61.2	61.2	61.3	61.4
K_2/Q^∞ (%)	95.6	91.7	90.2	89.2	86.7
$W'(\alpha)$ (€ millions/year)	981	681	620	592	533
δW^- (€ /user/year)	12 129	8 419	268	13	6
δW^+ (€ /user/year)	12 129	294	14	6	n/a

If the underlying demand were more elastic, the no rationing conditions (6) would be met for $\alpha \geq 16\%$. The optimal total installed capacity would be lower, as expected. The maximum effect is for $\alpha = 1$, where installed capacity would be 2.7% lower. The optimal mix would also change: base generation assets would raise to 70.8% of total installed capacity compared to 67.2%.

The marginal value of switching to real time price usage would also increase, for example by 64.5% for $\alpha = 1$. Yet, this increase would be not be sufficient to balance the marginal cost of smart meters. The marginal cost of installing real-time would exceed the marginal benefit for small non residential users.

These observations are summarized in the following:

Result 7 *For small users (less than 36 kVA peak demand) in the French power market, the marginal value of switching to real-time price usage of electricity is lower than the marginal cost of the "smart meters" enabling that switch. This challenges the economic wisdom of mandating full deployment of "smart meters".*

6 Conclusion

This article has derived the marginal value of a share of demand (or a consumer) switching to real time price usage of power. Using data from the French power market, it has compared this marginal value to the marginal cost of installing smart meters, and found that the latter exceeds the former.

This analysis can be expanded in at least three directions. First, the methodology will be applied to other power markets. The main finding – that the marginal cost of installing smart meters for small users (residential and non residential) exceeds its marginal value – is expected to be confirmed in other power markets, as they exhibit higher supply flexibility than the French market: for example, most American markets have a large fraction of combined cycle gas turbines in their generation mix, with lower capital cost than the nuclear assets that constitute the core of the French generation fleet.

Second, the impact of aggregate demand elasticity on the exercise of generators' market power will be included in the analysis. Allcott (2010) finds a limited impact, but it is worth validating this finding.

Finally, alternative specification of demand can be used. For example, demand can be assumed to be log-linear, multiple classes of users and intertemporal substitution can be introduced. This would likely result in closed-form solutions no longer being available. Instead, numerical analysis will be required.

A Derivations of the closed-form solutions

A.1 Optimal investment

For $n \in \{1,...,N\}$, $\tilde{P}(K_n;t) = \frac{a(t)-bK_n-(1-\alpha)p^R}{\alpha}$, hence:

$$\begin{cases} \tilde{P}(K_n;\bar{t}_n) = \frac{a(\bar{t}_n)-bK_n-(1-\alpha)p^R}{\alpha} = c_n \\ \tilde{P}(K_n;t_{n+1}) = \frac{a(t_{n+1})-bK_n-(1-\alpha)p^R}{\alpha} = c_{n+1} \end{cases}$$

$\Leftrightarrow$

$$\begin{cases} a_0 - a_1 e^{-\lambda_2 \bar{t}_n} - bK_n - (1-\alpha)p^R = \alpha c_n \\ a_0 - a_1 e^{-\lambda_2 t_{n+1}} - bK_n - (1-\alpha)p^R = \alpha c_{n+1} \end{cases}$$

Then:

$$e^{-\lambda_2 \bar{t}_n} - e^{-\lambda_2 t_{n+1}} = \frac{\alpha}{a_1}(c_{n+1} - c_n)$$

For $n \in \{1,...,N\}$, define $I_n = \int_{\bar{t}_n}^{t_{n+1}} \left[\tilde{P}(K_n;t) - \tilde{P}(K_n;\bar{t}_n)\right] f(t)\,dt$. I_n is determined as:

$$\begin{aligned} I_n &= \frac{a_1}{\alpha}\int_{\bar{t}_n}^{t_{n+1}} \left(e^{-\lambda_2 \bar{t}_n} - e^{-\lambda_2 t}\right)\lambda_1 e^{-\lambda_1 t} dt \\ &= \frac{a_1}{\alpha}\left\{\left[-\left(e^{-\lambda_2 \bar{t}_n} - e^{-\lambda_2 t}\right)e^{-\lambda_1 t}\right]_{\bar{t}_n}^{t_{n+1}} + \lambda_2 \int_{\bar{t}_n}^{t_{n+1}} e^{-(\lambda_1+\lambda_2)t} dt\right\} \\ &= \frac{a_1}{\alpha}\left\{-\frac{\alpha}{a_1}(c_{n+1}-c_n)e^{-\lambda_1 t_{n+1}} + \frac{\lambda_2}{\lambda_1+\lambda_2}\left(e^{-(\lambda_1+\lambda_2)\bar{t}_n} - e^{-(\lambda_1+\lambda_2)t_{n+1}}\right)\right\} \\ &= -(c_{n+1}-c_n)e^{-\lambda_1 t_{n+1}} + \frac{a_1}{\alpha}\frac{\lambda_2}{\lambda_1+\lambda_2}\left(e^{-(\lambda_1+\lambda_2)\bar{t}_n} - e^{-(\lambda_1+\lambda_2)t_{n+1}}\right) \end{aligned}$$

with the convention $t_{N+1} \to +\infty$.

Equation (2) is:

$$I_N = r_N$$

As $t_{N+1} \to +\infty$, this yields:

$$\frac{a_1}{\alpha}\frac{\lambda_2}{\lambda_1+\lambda_2}e^{-(\lambda_1+\lambda_2)\bar{t}_N} = r_N$$

$\Leftrightarrow$

$$\begin{aligned} \Pr\left(\tilde{D}(t) = K_N\right) &= e^{-\lambda_1 \bar{t}_N} = \left(\frac{\alpha r_N}{a_1}\frac{\lambda_2+\lambda_1}{\lambda_2}\right)^{\frac{\lambda_1}{(\lambda_2+\lambda_1)}} \\ &= \left(\frac{\alpha r_N}{a_1}(1+\lambda)\right)^{\frac{\lambda}{1+\lambda}} \end{aligned}$$

The optimal total capacity is the solution of:

$$\left[a_0 - bK_N - \left(\alpha c_N + (1-\alpha) p^R\right)\right]^{1+\lambda} = \alpha a_1^{\lambda}(1+\lambda) r_{N-1}$$

Then, for $1 \leq n < N$:

$$I_n + [c_{n+1} - c_n]_1 e^{-\lambda_1 t_{n+1}} = r_n - r_{n+1}$$

Hence:

$$e^{-(\lambda_1+\lambda_2)\bar{t}_n} - e^{-(\lambda_1+\lambda_2)t_{n+1}} = \frac{\lambda_1+\lambda_2}{\lambda_2}\frac{\alpha}{a_1}(r_n - r_{n+1})$$

For $n \in \{1, ..., (N-1)\}$, define:

$$\theta_n(x) = \left(a_0 - bx - (1-\alpha) p^R - \alpha c_n\right)^{1+\lambda} - \left(a_0 - bx - (1-\alpha) p^R - \alpha c_{n+1}\right)^{1+\lambda} - \alpha a_1^{\lambda}(1+\lambda)(r_n - r_{n+1})$$

Then:

$$\theta_n'(x) = -b(1+\lambda)\left[\left(a_0 - bx - (1-\alpha) p^R - \alpha c_n\right)^{\lambda} - \left(a_0 - bx - (1-\alpha) p^R - \alpha c_{n+1}\right)^{\lambda}\right] < 0$$

$K_n > 0$ if and only if $\theta_n(0) > 0 \Leftrightarrow$

$$\left(a_0 - (1-\alpha) p^R - \alpha c_n\right)^{1+\lambda} - \left(a_0 - (1-\alpha) p^R - \alpha c_{n+1}\right)^{1+\lambda} > \alpha a_1^{\lambda}(1+\lambda)(r_n - r_{n+1})$$

Then, K_n is defined by:

$$\left(a_0 - bK_n - (1-\alpha) p^R - \alpha c_n\right)^{1+\lambda} - \left(a_0 - bK_n - (1-\alpha) p^R - \alpha c_{n+1}\right)^{1+\lambda} = \alpha a_1^{\lambda}(1+\lambda)(r_n - r_{n+1})$$

If $\theta_n(0) \leq 0$, $K_n = 0$: it is optimal not to invest in technologies "lower" than $(n+1)$, which is then the baseload technology.

Similarly, it is optimal to invest in technology N if and only if $K_{N-1} < K_N \Leftrightarrow \theta_{N-1}(K_N) < 0 \Leftrightarrow$

$$\left(\alpha^{\frac{1}{1+\lambda}} a_1^{\frac{\lambda}{1+\lambda}} \left[(1+\lambda) r_N\right]^{\frac{1}{1+\lambda}} + \alpha(c_N - c_{N-1})\right)^{1+\lambda} - \left(\alpha^{\frac{1}{1+\lambda}} a_1^{\frac{\lambda}{1+\lambda}} \left[(1+\lambda) r_N\right]^{\frac{1}{1+\lambda}}\right)^{1+\lambda} < \alpha a_1^{\lambda}(1+\lambda)(r_{N-1} - r_N)$$

$\Leftrightarrow$

$$\left(\alpha^{\frac{1}{1+\lambda}} a_1^{\frac{\lambda}{1+\lambda}} \left[(1+\lambda) r_N\right]^{\frac{1}{1+\lambda}} + \alpha(c_N - c_{N-1})\right)^{1+\lambda} < \alpha a_1^{\lambda}(1+\lambda) r_{N-1}$$

⇔

$$c_N - c_{N-1} < \left(\frac{a_1}{\alpha}\right)^{\frac{\lambda}{1+\lambda}} \left([(1+\lambda) r_{N-1}]^{\frac{1}{1+\lambda}} - [(1+\lambda) r_N]^{\frac{1}{1+\lambda}}\right)$$

A.2 No rationing condition

No rationing is optimum as long as the $VoLL$ exceeds than price faced by price reactive load. If rationing is proportional and anticipated:

$$v(p,\gamma;t) = \frac{S(p;t)}{D(p;t)}$$

Hence:

$$v(p,\gamma;t) = \frac{a(t)+p}{2}$$

Then, no rationing is optimal if and only if, for all $t \geq 0$:

$$\frac{a(t)+p^R}{2} \geq p(t)$$

⇔

$$\forall 1 \leq n \leq N : \begin{cases} \frac{a(t)+p^R}{2} \geq c_n & \forall t \in [t_n, \bar{t}_n] \\ \frac{a(t)+p^R}{2} \geq \tilde{P}(K_n;t) & \forall t \in [\bar{t}_n, t_{n+1}] \end{cases}$$

⇔

$$n \in \{1,...,N\} : \begin{cases} a(t_n) \geq 2c_n - p^R \\ a(t_{n+1}) \leq \frac{2}{2-\alpha} bK_n + p^R \end{cases}$$

Since $a(.)$ is increasing, we examine each condition in turns. For $n \in \{1,...,N\}$:

$$a(t_n) \geq 2c_n - p^R$$

⇔

$$bK_{n-1} + \alpha c_n + (1-\alpha) p^R \geq 2c_n - p^R$$

⇔

$$bK_{n-1} \geq (2-\alpha)\left(c_n - p^R\right)$$

Then, for $n \in \{1,...,(N-1)\}$:

$$a(t_{n+1}) \leq \frac{2}{2-\alpha} bK_n + p^R$$

⇔

$$(2-\alpha)\left(bK_n + \alpha c_{n+1} + (1-\alpha) p^R\right) \leq 2bK_n + (2-\alpha) p^R$$

⇔

$$\alpha bK_n \geq \alpha (2-\alpha)\left(c_{n+1} - p^R\right)$$

Since $\theta_n(.)$ is decreasing and $\theta_n(bK_n) = 0$, this is equivalent to $\theta_n\left[(2-\alpha)\left(c_{n+1} - p^R\right)\right] \geq 0 \Leftrightarrow$

$$\left[a_0 - 2c_{n+1} + p^R + \alpha\left(c_{n+1} - c_n\right)\right]^{1+\lambda} - \left[a_0 - 2c_{n+1} + p^R\right]^{1+\lambda} \geq \alpha a_1^\lambda (1+\lambda)(r_n - r_{n+1})$$

Finally, for $n = N$:

$$a(t_{N+1}) \leq \frac{2}{2-\alpha} bK_N + p^R$$

$\Leftrightarrow$

$$bK_N \geq (2-\alpha)\frac{a_0 - p^R}{2}$$

$\Leftrightarrow$

$$a_0 - a_1\left(\frac{\alpha r_N}{a_1}(1+\lambda)\right)^{\frac{1}{1+\lambda}} - \left(\alpha c_N + (1-\alpha)p^R\right) \geq (2-\alpha)\frac{a_0 - p^R}{2}$$

$\Leftrightarrow$

$$a_0 - \left(2c_N - p^R\right) \geq 2\left((1+\lambda) r_N\right)^{\frac{1}{1+\lambda}}\left(\frac{a_1}{\alpha}\right)^{\frac{\lambda}{1+\lambda}}$$

$\Leftrightarrow$

$$\left[\frac{a_0 - \left(2c_N - p^R\right)}{2}\right]^{1+\lambda} \geq \left(\frac{a_1}{\alpha}\right)^\lambda \left[(1+\lambda) r_N\right]$$

A.3 Marginal value of α

$W'(\alpha) = \frac{1}{2b}\mathbb{E}\left\{\left(p^R - p(t)\right)^2\right\} = \frac{8760}{b} L_N$, where $L_N = \frac{Var[p(t)]}{2}$. Then:

$$\begin{aligned}
2L_N &= \left\{\sum_{n=1}^{N}\left\{\int_{t_n}^{\bar{t}_n}\left(c_n - p^R\right)^2 f(t)\,dt + \int_{\bar{t}_n}^{t_{n+1}}\left(\tilde{P}(K_n;t) - p^R\right)^2 f(t)\,dt\right\}\right\} \\
&= \sum_{n=1}^{N}\left\{\begin{array}{c} \int_{t_n}^{\bar{t}_n}\left(c_n - p^R\right)^2 f(t)\,dt + \int_{\bar{t}_n}^{t_{n+1}}\left(\tilde{P}(K_n;t) - c_n\right)^2 f(t)\,dt \\ +2\left(c_n - p^R\right)\int_{\bar{t}_n}^{t_{n+1}}\left(\tilde{P}(K_n;t) - c_n\right) f(t)\,dt + \int_{\bar{t}_n}^{t_{n+1}}\left(c_n - p^R\right)^2 f(t)\,dt \end{array}\right\} \\
&= \sum_{n=1}^{N}\left\{\left(c_n - p^R\right)^2\int_{t_n}^{t_{n+1}} f(t)\,dt + 2\left(c_n - p^R\right) I_n + H_n\right\}
\end{aligned}$$

where $H_n = \int_{\bar{t}_n}^{t_{n+1}}\left(\tilde{P}(K_n;t) - c_n\right)^2 f(t)\,dt$ is determined as:

$$\begin{aligned}
H_n &= \left(\frac{a_1}{\alpha}\right)^2\left\{\left[-\left(e^{-\lambda_2\bar{t}_n} - e^{-\lambda_2 t}\right)^2 e^{-\lambda_1 t}\right]_{\bar{t}_n}^{t_{n+1}} + 2\lambda_2\int_{\bar{t}_n}^{t_{n+1}}\left(e^{-\lambda_2\bar{t}_n} - e^{-\lambda_2 t}\right)e^{-(\lambda_1+\lambda_2)t}dt\right\} \\
&= \left(\frac{a_1}{\alpha}\right)^2\left\{-\left(e^{-\lambda_2\bar{t}_n} - e^{-\lambda_2 t_{n+1}}\right)^2 e^{-\lambda_1 t_{n+1}} + 2\lambda_2 J_n\right\} \\
&= -\left(c_{n+1} - c_n\right)^2 e^{-\lambda_1 t_{n+1}} + 2\left(\frac{a_1}{\alpha}\right)^2\lambda_2 J_n
\end{aligned}$$

where $J_n = \int_{\bar{t}_n}^{t_{n+1}} \left(e^{-\lambda_2 \bar{t}_n} - e^{-\lambda_2 t}\right) e^{-(\lambda_1+\lambda_2)t} dt$ is determined as:

$$\begin{aligned} J_n &= \left[-\left(e^{-\lambda_2 \bar{t}_n} - e^{-\lambda_2 t}\right) \frac{e^{-(\lambda_1+\lambda_2)t}}{\lambda_1+\lambda_2}\right]_{\bar{t}_n}^{t_{n+1}} + \frac{\lambda_2}{\lambda_1+\lambda_2} \int_{\bar{t}_n}^{t_{n+1}} e^{-(\lambda_1+2\lambda_2)t} dt \\ &= \frac{1}{\lambda_1+\lambda_2} \left\{-\frac{\alpha}{a_1}\left(c_{n+1}-c_n\right) e^{-(\lambda_1+\lambda_2)t_{n+1}} + \frac{\lambda_2}{\lambda_1+2\lambda_2}\left(e^{-(\lambda_1+2\lambda_2)\bar{t}_n} - e^{-(\lambda_1+2\lambda_2)t_{n+1}}\right)\right\} \end{aligned}$$

H_N and J_N are determined by the same expressions, with $t_{N+1} \to +\infty$. Then:

$$\begin{aligned} S_N &= \sum_{n=1}^{N} \left(c_n - p^R\right)^2 \int_{t_n}^{t_{n+1}} f(t)\, dt = \sum_{n=1}^{N} \left(c_n - p^R\right)^2 \left(e^{-\lambda_1 t_n} - e^{-\lambda_1 t_{n+1}}\right) \\ &= \sum_{n=1}^{N} e^{-\lambda_1 t_{n+1}} \left[\left(c_{n+1}-p^R\right)^2 - \left(c_n - p^R\right)^2\right] + \left(c_1 - p^R\right)^2 e^{-\lambda_1 t_1} - \left(c_{N+1}-p^R\right)^2 e^{-\lambda_1 t_{N+1}} \\ &= \sum_{n=1}^{N} e^{-\lambda_1 t_{n+1}} \left(c_{n+1}-c_n\right)\left(c_{n+1}+c_n-2p^R\right) + r_1^2 \end{aligned}$$

Hence, for $N > 1$:

$$\begin{aligned} 2L_N &= \sum_{n=1}^{N-1} \left\{ \begin{array}{c} e^{-\lambda_1 t_{n+1}} \left(c_{n+1}-c_n\right)\left(c_{n+1}+c_n-2p^R\right) - \left(c_{n+1}-c_n\right)^2 e^{-\lambda_1 t_{n+1}} + 2\left(\frac{a_1}{\alpha}\right)^2 \lambda_2 J_n \\ +2\left(c_n-p^R\right)\left(r_n - r_{n+1} - \left(c_{n+1}-c_n\right) e^{-\lambda_1 t_{n+1}}\right) \end{array} \right\} \\ &\quad +2\left[\sum_{n=1}^{N-1} \left(c_n - p^R\right)\left(r_n - r_{n+1}\right) + \left(c_N - p^R\right) r_N + \left(\frac{a_1}{\alpha}\right)^2 \lambda_2 J_N\right] + r_1^2 \\ &= 2\left\{\sum_{n=1}^{N-1} \left[\left(\frac{a_1}{\alpha}\right)^2 \lambda_2 J_n + r_{n+1}\left(c_{n+1}-c_n\right)\right] + \left(\frac{a_1}{\alpha}\right)^2 \lambda_2 J_N - \frac{r_1^2}{2}\right\} \end{aligned}$$

Solving for $N = 1$, we obtain:

$$L_N = \left\{ \begin{array}{ll} \sum_{n=1}^{N-1} \left[\left(\frac{a_1}{\alpha}\right)^2 \lambda_2 J_n + r_{n+1}\left(c_{n+1}-c_n\right)\right] + \left(\frac{a_1}{\alpha}\right)^{\frac{\lambda}{1+\lambda}} \frac{\left[(1+\lambda) r_N\right]^{\frac{2+\lambda}{1+\lambda}}}{(1+\lambda)(2+\lambda)} - \frac{r_1^2}{2} & N > 1 \\ r_1 \left[\left(\frac{a_1}{\alpha}\right)^{\frac{\lambda}{1+\lambda}} \frac{\left[(1+\lambda) r_1\right]^{\frac{1+\lambda}{1+\lambda}}}{2+\lambda} - \frac{r_1}{2}\right] & N = 1 \end{array} \right.$$