TSE working paper series

12-341

Research Group: Industrial Organization

October 2012

“Is Mandating "Smart Meters" smart?”

Thomas-Olivier Léautier

Toulouse School of Economics (TSE) - Manufacture des Tabacs
Aile Jean-Jacques Laffont - 21, allée de Brienne - 31015 TOULOUSE Cedex 6
Tél : +33 (0) 5 61 12 85 89 - Fax : +33 (0) 5 61 12 86 37 - **www.tse-fr.eu** - contact@tse-fr.eu

Toulouse School of Economics

Is mandating "smart meters" smart?

Thomas-Olivier Léautier

Toulouse School of Economics (IAE, IDEI, University of Toulouse)

21 allée de Brienne

F31 000 Toulouse

thomas.leautier@iae-toulouse.fr and +33 6 33 34 60 68

October 8, 2012

Abstract

The advent of "smart meters" will make possible Real Time Pricing of electricity: customers will face and react to wholesale spot prices, thus consumption of electric power will be aligned with its opportunity cost. This article determines the marginal value of a fraction of demand (or a consumer) switching to Real Time Pricing. First, it derives this marginal value for a simple yet realistic specification of demand. Second, using data from the French power market, it estimates that, for the vast majority of residential customers whose peak demand is lower than 6 kVA, the net surplus from switching to Real Time Pricing is lower than 1 €/year for low demand elasticity, 4 €/year for high demand elasticity. This finding casts a doubt on the economic value of rolling out smart meters to all residential customers, for both policy makers and power suppliers.

Keywords: electric power markets, demand response, smart grid

JEL Classification: L11, L94, D61

1 Introduction

"Smart meters", which allow electric power users to respond to wholesale spot prices, are expected to transform the electric power industry. Consumers will reduce their consumption during peak hours, thus reducing installed capacity requirement and emissions of CO_2 and other pollutants. The potential value of these demand management benefits is significant. For example, Faruqui et al. (2009), estimate the annual potential value for all of Europe of reduced capacity cost at € 4.8 billions, and the value from reduced electricity consumption at € 600 millions. Similarly, the Department of Energy and Climate Change ($DECC$) estimates the Present Value for Britain of energy savings benefits at £ 4 400 millions, carbon savings at £ 1 100 millions, and peak load shifting at £ 800 millions. As a result, full deployment of "smart meters" is underway in many European countries and US states.

The policy discussion of smart meters appears to be framed as a one-or-zero problem: should we install meters for all users or for none? This is surprising. As all economic problems, it should be cast as an optimal share of deployment problem: which consumer groups should be equipped with smart meters? To answer that question, one should compare the marginal value of equipping a class of consumers against its marginal cost. A key ingredient in the analysis is the marginal value of Real Time Pricing (RTP), i.e., the marginal surplus generated by one customer becoming price responsive. This is precisely what this article estimates.

This article builds on and complements a rich literature. Reiss and White (2005) and Allcott (2011) estimate individual price elasticity of customers, and use these elasticities to estimate welfare effects. Reiss and White (2005), using data from California households, focus on the non-linearity of the pricing schedule and estimate demand for eight different types of electric appliances, e.g., electric space heating, room air conditioning, etc. They then estimate the welfare impact of a rate structure change proposed in California. Allcott (2011) estimates the demand function from consumers opting for Real Time Pricing in a pilot program in Chicago. He then estimates the annualized short term consumer surplus increase from RTP, assuming wholesale prices and producers profits are constant, at $10 per household.

Holland and Mansur (2006), Borenstein (2005), Borenstein and Holland (2005), and Allcott (2012) use existing estimates of price elasticity to estimate the welfare impact of RTP. Holland and Mansur (2006) estimate the short-term welfare impact of exposing 33%, 67%, and 100% of demand to RTP

in the Pennsylvania-New Jersey-Maryland market (PJM), a large power market in the North East of the United States. The estimated gross welfare gain if 100% of load is exposed to RTP is 0.24% of the total energy bill. Borenstein (2005) estimates the long-term welfare impact, including adjustment to the generation mix, of exposing 33%, 66%, and 99% of demand to RTP in the California market. For example if 33% of demand faces RTP, the estimated gross welfare gain ranges from 1.2% of the total energy bill for low elasticity to 7% for high elasticity. Allcott (2012) estimates the long-term welfare impact of moving 20% of demand to RTP in PJM, taking into account the impact of demand elasticity on producers' market power. He finds a gross welfare increase (excluding infrastructure cost) of 38.90 $ per kW of average demand equipped with smart meters.

This article follows a different approach, that proposes a closed form expression for the *marginal value of* RTP, then estimates it using the load duration curve of the French power system[1] and previous estimates of demand elasticity. Its contribution 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 marginal increase in net surplus of a customer switching to RTP. Using the load duration curve of the French power system and previous estimates of demand elasticity, this value is estimated at 1 to 4 €/customer per year for a small residential customer, whose peak demand is lower than 6 kVA. 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 article is structured as follows. The model used in this article is the one developed by Borenstein and Holland (2005) and Joskow and Tirole (2007), building on the earlier work by Boiteux

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

(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 RTP. 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. Technical proofs are gathered in the Appendix.

2 The model

2.1 Model structure

2.1.1 Uncertainty

Uncertainty is an essential feature of power markets. In this work, only demand uncertainty is explicitly modeled, since including production uncertainty does not modify the economic insights, although, as discussed in Section 6, it raises the value of RTP. 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.

2.1.2 Demand, supply, and rationing

Demand

Assumption 1 *Customers have the same load profile: in state t, all have the same underlying demand $D(p;t)$ up to a scaling factor. $D(p,t)$ is non increasing in price, and states of the world are ordered such that $D(p,t)$ is increasing in the state of the world:*

$$\frac{\partial D}{\partial p}(p;t) \leq 0 \text{ and } \frac{\partial D}{\partial t}(p;t) > 0.$$

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 $S(p;t) = \int_0^{D(p;t)} P(q;t)\,dq$.

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

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

Supply Different generation technologies are available, indexed by n. 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, both expressed in $€/MWh$. Generation technologies are ordered by increasing marginal cost: $c_n > c_m \ \forall \ n \geq m$. There is a trade-off between investment and marginal costs: if a technology produces at higher variable cost, it then requires lower investment cost, i.e., $r_n < r_m \ \forall \ n \geq m$.

Not all available technologies are included in the optimal investment plan. To simplify the exposition, I propose later necessary and sufficient conditions for technologies 1 to N to be used at the optimum.

Rationing and Value of Lost Load In some states of the world, it may be optimal to curtail constant price customers. 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: (i) $\mathcal{D}(p,0;t) = 0$, (ii) $\frac{\partial \mathcal{D}}{\partial \gamma} > 0$ for $\gamma \in [0,1]$, and (iii) $S(p;t) \equiv \mathcal{S}(p,1;t)$ 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 2 *1. The SO has the technical ability to curtail "constant price" customers while not curtailing "price reactive" customers.*

2. Rationing does not increase the net surplus: $\forall p > 0, \forall t \geq 0, \forall \gamma > 0$

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

3. 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:*

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

The first part of 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 met when "smart meters" are being rolled out, which is precisely the situation considered.

Parts 2 and 3 of Assumption 2 hold for example if rationing is anticipated, which yields $\mathcal{S}(p,\gamma) = \gamma S(p)$, and proportional, which yields $\mathcal{D}(p,\gamma) = \gamma D(p)$. Thus, (i)

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

for $\gamma \leq 1$; and (ii) $v(p,\gamma;t) = \frac{S(p)}{D(p)} > p$.

Parts 2 and 3 of Assumption 2 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, production, and consumption

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

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

The net surplus from consumption is then:

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

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 retail price, $k_n \geq 0$ the installed capacity of technology n. $W^h(\alpha)$ is the net expected surplus per hour. The yearly surplus is $W(\alpha) = 8760 \times W^h(\alpha)$ since there are 8 760 hours per year.

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 wholesale spot power 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 lower than the real time price.

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

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

As in Joskow and Tirole (2007), the optimal retail price is the weighted average wholesale price, where the weights are the marginal "rationed demand". The optimal 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:

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

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, price equals the 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:

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

Hence, for $n \in \{1, ..., , N\}$,

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

where γ^* is the optimal serving ratio.

Denote $\widehat{t}(K,c)$ the first state of the world such that $\rho\left(K, \widehat{t}(K,c)\right) \geq c$. By construction, we have:

$$\begin{cases} \rho(K_n; \bar{t}_n) = c_n \Leftrightarrow \bar{t}_n = \widehat{t}(K_n, c_n) \\ \rho(K_n; t_{n+1}) = c_{n+1} \Leftrightarrow t_{n+1} = \widehat{t}(K_n, c_{n+1}) \end{cases}.$$

The structure of these critical states of the world and prices is illustrated on Figure 1.

Define

$$\Psi(K,c) = \int_{\widehat{t}(K,c)}^{+\infty} (p(t) - c) f(t)\, dt,$$

the marginal social value of capacity K for marginal cost c. Since $p(t)$ 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:

$$\Psi(K_n, c_n) = \int_{\widehat{t}(K_n,c_n)}^{+\infty} (p(t) - c_n) f(t)\, dt = r_n. \tag{1}$$

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

$$\Psi\left(K_N, c_N\right) = r_N \tag{2}$$

The total installed capacity is solely determined by the long-run marginal cost of the last technology and the demand function.

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

$$\begin{aligned} r_n &= \int_{\widehat{t}(K_n,c_n)}^{\widehat{t}(K_n,c_{n+1})} \left(p\left(t\right) - c_n\right) f\left(t\right) dt + \int_{\widehat{t}(K_n,c_{n+1})}^{\widehat{t}(K_{n+1},c_{n+1})} \left(c_{n+1} - c_n\right) f\left(t\right) dt \\ &+ \int_{\widehat{t}(K_{n+1},c_{n+1})}^{+\infty} \left(p\left(t\right) - c_{n+1} + c_{n+1} - c_n\right) f\left(t\right) dt \end{aligned}$$

$\Leftrightarrow$

$$r_n = \int_{\widehat{t}(K_n,c_n)}^{\widehat{t}(K_n,c_{n+1})} \left(\rho\left(K_n; t\right) - c_n\right) f\left(t\right) dt + \int_{\widehat{t}(K_n,c_{n+1})}^{+\infty} \left(c_{n+1} - c_n\right) f\left(t\right) dt + \Psi\left(K_{n+1}, c_{n+1}\right)$$

$\Leftrightarrow$

$$\int_{\widehat{t}(K_n,c_n)}^{\widehat{t}(K_n,c_{n+1})} \left(\rho\left(K_n; t\right) - c_n\right) f\left(t\right) dt + \int_{\widehat{t}(K_n,c_{n+1})}^{+\infty} \left(c_{n+1} - c_n\right) f\left(t\right) dt = r_n - r_{n+1}. \tag{3}$$

A marginal substitution of technology $(n+1)$ by technology n increases net surplus by $\left(\rho\left(K_n; t\right) - c_n\right)$ for $t \in \left[\widehat{t}\left(K_n, c_n\right), \widehat{t}\left(K_n, c_{n+1}\right)\right]$ and by $\left(c_{n+1} - c_n\right)$ for $t \geq \widehat{t}\left(K_n, c_{n+1}\right)$. It also increases investment cost by $\left(r_n - r_{n+1}\right)$. The optimal capacity K_n exactly balances marginal benefit and marginal cost.

I have sofar assumed existence of the optimal cumulative capacities $0 < K_1 < ... < K_N$. I present below a set of necessary and sufficient conditions. First,

$$\rho\left(0, 0\right) > c_N$$

guarantees that $\widehat{t}\left(0, c_n\right) = 0$ for all n, hence simplifies the exposition. Second, $\Psi\left(0, c_N\right) > r_N$ is necessary and sufficient for the existence of $K_N > 0$ solution of equation (2). Since $\widehat{t}\left(0, c_N\right) = 0$,

$$\Psi\left(0, c_N\right) = \int_0^{+\infty} \left[\rho\left(0, t\right) - c_N\right] f\left(t\right) dt = \mathbb{E}\left[\rho\left(0, t\right)\right] - c_N.$$

Thus, $\Psi(0, c_N) > r_N$ is equivalent to

$$\mathbb{E}[\rho(0,t)] > c_N + r_N.$$

Third, for $n = 1, ..., N-1$, $\Psi(0, c_n) > r_n$ is necessary and sufficient for existence of $K_n > 0$ solution of equation (3). Since

$$\Psi(0, c_n) = \int_0^{+\infty} [c_{n+1} - c_n] f(t)\, dt + r_{n+1} = c_{n+1} - c_n + r_{n+1},$$

$\Psi(0, c_n) > r_n$ is equivalent to

$$c_{n+1} + r_{n+1} > c_n + r_n.$$

Finally, for $n = 1, ..., N-1$, $K_n < K_{n+1}$ is necessary and sufficient for technology $(n+1)$ to be included in the investment plan. Since Ψ is decreasing in its first argument, this is equivalent to

$$\Psi(K_{n+1}, c_n) < r_n.$$

These conditions are assumed to hold in Sections 3 and 4. They are verified by the numerical example proposed in Section 5.

3 Increasing the proportion of price reactive customers

This section examines the long-term impact of a marginal increase in α: as in Borenstein (2005) and Allcott (2012), all values, in particular, generation capacity and mix and retail prices, are optimal. One could challenge this choice as being unrealistic, since installed generation mix and retail prices are rarely optimal. However, I believe this is the appropriate analysis, as it isolates the impact of switching to RTP.

3.1 Impact on welfare

Result 1 *Increasing the proportion of price-reactive customers always increases the net surplus from consumption.*

P roof. *The envelope theorem yields:*

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

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

since rationing does not generate value; and

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

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

$$\frac{dW^h}{d\alpha} \geq \mathbb{E}\left[\left(S\left(p\left(t\right);t\right) - p\left(t\right)D\left(p\left(t\right);t\right)\right) - \left(S\left(p^R;t\right) - p\left(t\right)D\left(p^R;t\right)\right)\right] \geq 0.$$

■

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). Increasing their share thus always increases the net surplus. 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) assume that the retail price is such that the retail profit is equal to zero. In this work, following Joskow and Tirole (2007), we consider 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 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. In other words, welfare is measured with triangles, and not rectangles. This is illustrated on Figure 2.

3.2 Impact on average price

Result 2 *Increasing the share of price reactive customers has no impact on the expected price.*

P roof. *By construction, we have:* $t_1 = 0$*. Equation (1) for* $n = 1$ *then yields:*

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

$\Leftrightarrow$

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

Then,

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

hence

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

■

Since the first technology (the baseload technology) produces in all states of the world, the zero-profit condition implies that the expected spot price is simply the long run marginal cost of the baseload technology, 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 time weighted average power price.

3.3 A specific case: linear demand, no rationing at the optimum

Suppose demand is linear and is given by:

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

Suppose also that no rationing occurs at the optimum. As will be shown in Section 5, this assumption holds as soon as a small fraction of customers face real time prices, i.e., α larger than a small threshold.

Result 3 *If demand is linear and rationing does not occur at the optimum, the marginal net surplus is proportional to the spot price volatility.*

P roof. *Since demand is linear* $D(p,t) = \frac{a(t)-p}{b}$*, then*

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

The net surplus is

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

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

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

Thus,

$$\begin{aligned} \mathcal{S}\left(p^R, 1; t\right) - p\mathcal{D}\left(p^R, 1; t\right) &= S\left(p^R; t\right) - pD\left(p^R; t\right) \\ &= \left(a(t) - \frac{bD\left(p^R; t\right)}{2} - (a(t) - bD(p;t))\right) D\left(p^R; t\right) \\ &= \frac{b}{2}\left(2D(p;t) - D\left(p^R; t\right)\right) D\left(p^R; t\right). \end{aligned}$$

Then:

$$\begin{aligned} W^{h\prime}(\alpha) &= \frac{b}{2}\mathbb{E}\left[D^2(p;t) - \left(2D(p;t) - D\left(p^R;t\right)\right) D\left(p^R;t\right)\right] \\ &= \frac{b}{2}\mathbb{E}\left[\left(D(p(t);t) - D\left(p^R;t\right)\right)^2\right] \\ &= \frac{1}{2b}\mathbb{E}\left[\left(p^R - p(t)\right)^2\right] = \frac{1}{2b} Var(p(t)). \end{aligned}$$

■

If demand is linear, the marginal surplus is a exactly a triangle. The surface of the triangle can be expressed as the square of one of its side times the slope of the opposite angle. This then produces the result.

$Var\left(p\left(t\right)\right)$ does not depend on the average price, therefore transmission and distribution rates, which are constant across time, do not matter. This can also be understood by observing that these constant costs are included in the constant term a.

4 A closed form solution

A closed form solution is available if we assume that (i) demand is linear, with $a\left(t\right)=a_0-a_1e^{-\lambda_2 t}$ and $f\left(t\right)=\lambda_1 e^{-\lambda_1 t}$, and (ii) rationing is anticipated and proportional. 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

Define

$$y\left(K,c,\alpha\right)=a_1e^{-\lambda_2\widehat{t}(K,c)}=a_0-bK-\left(\alpha c+\left(1-\alpha\right)p^R\right).$$

Result 4 *Suppose rationing does not occur at the optimum. If*

$$a_1\geq\alpha\left(1+\lambda\right)r$$

where $\lambda=\frac{\lambda_1}{\lambda_2}$, *the optimal total capacity* K_N *is the solution of:*

$$y\left(K_N,c_N,\alpha\right)^{1+\lambda}=\alpha a_1^{\lambda}\left(1+\lambda\right)r_N. \tag{4}$$

For $n=1,...,N-1,$ *if*

$$a_1\left(1-\left(1-\frac{\alpha\left(c_{n+1}-c_n\right)}{a_1}\right)^{1+\lambda}\right)\geq\alpha\left(1+\lambda\right)\left(r_{n+1}-r_n\right),$$

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

$$y\left(K_n, c_n, \alpha\right)^{1+\lambda} - y\left(K_n, c_{n+1}, \alpha\right)^{1+\lambda} = \alpha a_1^{\lambda}\left(1+\lambda\right)\left(r_n - r_{n+1}\right). \quad (5)$$

P roof. *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 sufficient conditions ensure that $y\left(K_n, c_n, \alpha\right) \leq 1$. They are verified in the numerical example presented in Section 5. 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), while equations (5) determine the optimal capacity up to technology n, that depend only on the marginal and investment costs of technologies n and $(n+1)$.

4.2 No rationing conditions

Suppose rationing, if it occurs, is anticipated and proportional. Then:

Result 5 *No rationing occurs at the optimum if and only if*

$$\alpha \geq \alpha_{\min} = \frac{a_1\left[\left(1+\lambda\right)r_N\right]^{\frac{1}{\lambda}}}{\left[\frac{a_0-\left(2c_N-p^R\right)}{2}\right]^{1+\lambda}}. \quad (6)$$

P roof. *No rationing occurs at the optimum as long as price is lower than the $VoLL$. If rationing is anticipated and proportional,*

$$v\left(p^R;t\right) = \frac{S\left(p^R;t\right)}{D\left(p^R;t\right)} = \frac{a\left(t\right)+p^R}{2}.$$

Rationing can occur only when all capacity is constrained, i.e., $t \geq \widehat{t}\left(K_N, c_N\right)$. Algebra presented in the appendix shows that

$$v\left(p^R;t\right) \geq p\left(t\right) \;\; \forall t \geq \widehat{t}\left(K_N, c_N\right) \Leftrightarrow \text{ condition } (6)$$

■

As expected, rationing no longer occurs as soon as a sufficient share of demand is price reactive.

4.3 Marginal value of Real-Time Pricing

Result 6 *Suppose $\alpha \geq \alpha_{\min}$. The annual marginal value of Real-Time Pricing is:*

$$W'(\alpha) = \frac{8760}{b}\left(-\frac{r_1^2}{2} + \frac{1}{\alpha a_1^\lambda (1+\lambda)} \sum_{n=1}^{N} \left(\begin{array}{c} (c_{n+1}-c_n)\left(\alpha a_1^\lambda (1+\lambda) r_n - y(K_n, c_n, \alpha)^{1+\lambda}\right) \\ +\frac{y(K_n,c_n,\alpha)^{2+\lambda} - y(K_n,c_{n+1},\alpha)^{2+\lambda}}{\alpha(1+2\lambda)} \end{array} \right)\right). \tag{7}$$

P roof. *If $N = 1$,*

$$\begin{aligned} Var\left[p(t)\right] &= Var\left[p(t) - c_1\right] = \mathbb{E}\left[(p(t)-c_1)^2\right] - \left(\mathbb{E}\left[p(t)-c_1\right]\right)^2 \\ &= \int_{\widehat{t}(K_1,c_1)}^{+\infty} (p(t)-c_1)^2 f(t)\, dt - r_1^2. \end{aligned}$$

Then, integrating by parts twice,

$$\begin{aligned} \int_{\widehat{t}(K_1,c_1)}^{+\infty} (\rho(t)-c_1)^2 f(t)\, dt &= 2\int_{\widehat{t}(K_1,c_1)}^{+\infty} (\rho(K_1;t)-c_1) \frac{\partial \rho}{\partial t}(1-F(t))\, dt \\ &= \frac{2a_1\lambda_2}{\alpha} \int_{\widehat{t}(K_1,c_1)}^{+\infty} (\rho(K_1;t)-c_1)\, e^{-(\lambda_1+\lambda_2)t} dt \\ &= 2\frac{a_1^2}{\alpha^2} \frac{\left(e^{-\lambda_2 \widehat{t}(K_1,c_1)}\right)^{2+\lambda}}{(1+\lambda)(2+\lambda)} = 2\frac{y(K_1,c_1,\alpha)^{2+\lambda}}{\alpha^2 a_1^\lambda (1+\lambda)(2+\lambda)}. \end{aligned}$$

Thus:

$$W'(\alpha) = 8760 W^{h\prime}(\alpha) = \frac{8760}{2b}\left(-r_1^2 + 2\frac{y(K_1,c_1,\alpha)^{2+\lambda}}{\alpha^2 a_1^\lambda (1+\lambda)(2+\lambda)}\right)$$

which yields expression (7) for $N = 1$, since $\alpha a_1^\lambda (1+\lambda) r_N - y(K_N, c_N, \alpha)^{1+\lambda} = 0$ by equation (4), and $y(K_N, c_{N+1}, \alpha) = a_1 e^{-\lambda_2 \widehat{t}(K_N, c_{N+1})} = a_1 e^{-\lambda_2 t_{N+1}} = 0$.

For $N > 1$,

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

Observing that

$$
\begin{aligned}
\left(\rho\left(K_n;t\right)-p^R\right)^2 &= \left(c_n-p^R+\rho\left(K_n;t\right)-c_n\right)^2 \\
&= \left(c_n-p^R\right)^2+2\left(c_n-p^R\right)\left(\rho\left(K_n;t\right)-c_n\right)+\left(\rho\left(K_n;t\right)-c_n\right)^2
\end{aligned}
$$

yields

$$
\begin{aligned}
Var\left[p\left(t\right)\right] &= \sum_{n=1}^{N}\left(\begin{array}{c} \int_{t_n}^{\bar{t}_n}\left(c_n-p^R\right)^2 f\left(t\right)dt+\int_{\bar{t}_n}^{t_{n+1}}\left(c_n-p^R\right)^2 f\left(t\right)dt \\ +2\left(c_n-p^R\right)\int_{\bar{t}_n}^{t_{n+1}}\left(\rho\left(K_n;t\right)-c_n\right)f\left(t\right)dt+\int_{\bar{t}_n}^{t_{n+1}}\left(\rho\left(K_n;t\right)-c_n\right)^2 f\left(t\right)dt \end{array}\right) \\
&= \sum_{n=1}^{N}\left(\begin{array}{c} \left(c_n-p^R\right)^2\int_{t_n}^{t_{n+1}}f\left(t\right)dt+2\left(c_n-p^R\right)\int_{\bar{t}_n}^{t_{n+1}}\left(\rho\left(K_n;t\right)-c_n\right)f\left(t\right)dt \\ +\int_{\bar{t}_n}^{t_{n+1}}\left(\rho\left(K_n;t\right)-c_n\right)^2 f\left(t\right)dt \end{array}\right).
\end{aligned}
$$

Algebra presented in the Appendix shows that integrating the last term by parts twice, then summing all terms yields expression (7) for $N>1$. ■

5 Application to the French power market

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 that of Borenstein (2005), Borenstein and Holland (2005), and Holland and Mansur (2006), and is consistent with the reality of the French power market: in 2009, most customers paid a constant power price, denoted p_0. Observed demand fluctuations are due therefore to variations in the states of the world ($a\left(t\right)$ and $f\left(t\right)$). As the share of price reactive demand increases, 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:

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

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 $f(t) = \lambda_1 e^{-\lambda_1 t}$:

$$\begin{aligned} G(x) &= 1 - \exp\left[\frac{\lambda_1}{\lambda_2} \ln \frac{a_0 - (p_0+bx)}{a_1}\right] \\ &= 1 - \left[\frac{a_0-(p_0+bx)}{a_1}\right]^{\lambda} \end{aligned}$$

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[2] leads to $Q^{\infty} = 92.4\ GW$ and $\phi = \frac{31.5}{92.4} = 0.34$. The average price paid by customers is $p_0 = 100$ €$/MWh$. Then, Maximum Likelihood estimation yields $\lambda = \frac{\lambda_1}{\lambda_2} = 1.78$.

Actual and fitted demand are presented on Figure 3.

5.1.2 Estimation of b and all other parameters

Lijesen (2007) provides an up to date survey of the empirical literature on price elasticity of electricity, as well as his own estimate.

The elasticity of demand $\eta(\delta, t)$ for a given price δ in state t is

$$\eta(\delta, t) = -\frac{\delta}{a(t) - \delta},$$

thus

$$\mathbb{E}[\eta(\delta, t)] = -\mathbb{E}\left[\frac{\delta}{a(t) - \delta}\right] = -\delta \int_0^{+\infty} \frac{\lambda_1 e^{-\lambda_1 t}}{a_0 - a_1 e^{-\lambda_2 t} - \delta} dt.$$

Setting $x = e^{-\lambda_1 t}$ yields

$$\mathbb{E}[\eta(\delta, t)] = -\delta \int_0^1 \frac{dx}{a_0 - a_1 x^{\frac{1}{\lambda}} - \delta} = -\delta \int_0^1 \frac{dx}{p_0 + bQ^{\infty} - (bQ^{\infty}(1 - \phi)) x^{\frac{1}{\lambda}} - \delta}.$$

From Lijesen (2007), we select as a base case $\mathbb{E}[\eta(\delta, t)] = -0.01$ at price $\delta = 100$ €$/MWh$, which corresponds to the upper estimate from Patrick and Wolak (1997) using UK data, and the lower bound of Lijesen (2007) own estimate on Dutch data. We also run a robustness check with $\eta = -0.1$, which corresponds to Allcott's (2012) estimate for customers who self-selected into an RTP pilot program. A higher elasticity of demand renders RTP more attractive.

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

$$\begin{cases} bQ^{\infty} = 18\ 727 \text{ €}/MWh \\ a_0 = 18\ 827 \text{ €}/MWh \\ a_1 = 12\ 360 \text{ €}/MWh \end{cases}$$

[2] Source: posted on the French transmission asset owner (RTE) website.

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 is likely to increase the marginal value of RTP, since including these technologies would likely increase generation flexibility, hence reduce the value of demand flexibility.

The International Energy Agency (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 RTP therefore includes the value of emissions reduction. 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 Estimation of the marginal surplus $W^{'}(\alpha)$

With the parameters estimated, curtailment of constant price customers does not occur if $\alpha \geq \alpha_{\min} = 3.93\%$ for $\eta = -0.01$ and $\alpha_{\min} = 14\%$ for $\eta = -0.1$. Figure 4 presents the marginal surplus $W^{'}(\alpha)$ for the base elasticity $\eta = -0.01$ and Figure 5 for the high elasticity $\eta = -0.1$, measured in € millions per year, for $\alpha \geq \alpha_{\min}$.

In both cases, $W^{'}(\alpha)$ is decreasing with α. For $\eta = -0.01$, a 1% increase in the fraction of price reactive customers increases welfare by € 6.7 millions annually for $\alpha = \alpha_{\min}$, and only by € 1.1 millions annually for $\alpha = 1$. With the selected specification, there are decreasing returns to RTP. This result is consistent with the economic intuition: the first percent of demand flexibility is highly valuable, as it dramatically reduces the cost of balancing demand with supply. As a larger share of demand becomes price responsive, the marginal value of RTP decreases. Further work will determine sufficient conditions for this result to hold.

$W'(\alpha)$ is slightly higher when demand elasticity increases: for $\eta = -0.1$, a 1% increase in the fraction of price reactive customers increases annual surplus between € 9.4 millions for $\alpha = \alpha_{\min}$ and 4.7 millions for $\alpha = 1$. If the price elasticity of demand is higher, customers react more to prices. Increasing price responsiveness thus has a higher impact.

5.4 Change in production technologies mix

We compare $W'(\alpha)$ for three production technologies mix: (*i*) a single production technology, Combined Cycle Gas Turbine ($CCGT$), (*ii*) the previous mix (nuclear and gas turbine), and (*iii*) a richer mix, including nuclear, $CCGT$, and gas turbines. Costs for the $CCGT$ are $r = 8$ €$/MWh$ and $c = 49$ €$/MWh$, also selected from IEA (2010), hence are consistent with the costs of the other technologies. Figure 6 presents $W'(\alpha)$ for these three technology mixes, and the base elasticity.

Changing the mix has only limited change on the marginal value of price responsiveness. For example, $W'(0.5) = 185$ millions per year for technology mix (*i*), 156 millions per year for mix (*ii*), and 144 millions per year for mix (*iii*). Thus, I expect the findings to apply to other power markets, where the technology mix is different.

As observed on Figure 6, $W'(\alpha)$ decreases for any value of α when the number of technologies increases. This result has a nice intuitive explanation: as the number of technologies increases, supply flexibility increases, hence the marginal value of demand flexibility decreases. Unfortunately, preliminary analysis suggests this result may not always hold. Establishing conditions under which this result hold in general is left for further research.

5.5 Estimation of the marginal surplus per site

The Commission de Régulation de l'Energie (CRE) provides the total number of sites and the total consumption (MWh) for four categories of customers[3]:

1. *Large non residential.* Around 36 000 sites have peak demand higher than 250 kW: large industrials, hospitals, shopping malls, large buildings. They represent 0.1% of the total number of sites, and 42% of total demand

[3] The segmentation is based on net power (kW) for large users, and apparent power (kVA) for small users. The former is slightly smaller than the latter. This does not alter the results.

2. *Medium non residential.* Around 360 000 sites have peak-demand between 36 and 250 kW, mostly small companies. They represent 1% of the total number of sites, and 15% of total demand

3. *Small non residential.* Around 4.6 million sites have peak demand smaller than 36 kVA: professional offices, small workshops. They represent 13% of the total number of sites, 10% of total demand

4. *Residential sites.* Around 30.7 million sites, peak demand lower than 36 kVA. They represent 86% of the total number of sites, and only 32% of total demand.

Upon further request[4], CRE provided me with the distribution of maximum peak demand for the 30.7 million residential users, presented on Figure 7. The overwhelming majority of residential users have peak demand of less than 6 kVA (18.1 millions, 59% of residential sites), or 9 kVA (6.5 *millions*, 21% of residential sites). Under Assumption 1, this peak demand can be translated into in yearly load in MWh.

Denote $\delta\alpha$ the incremental increase in α from a single site. All sites in each class are assumed to have the same size, thus $\delta\alpha$ is constant for each class, and given by:

$\alpha\,(\%)$	$(0, 42)$	$(42, 57)$	$(57, 67)$
$\delta\alpha\,(\%/user)$	1.18×10^{-5}	4.21×10^{-7}	2.16×10^{-8}

for non-residential customers,

α (%)	$(67, 67.1)$	$(67.1, 67.2)$	$(67.2, 67.3)$	$(67.3, 69.3)$	$(69.3, 70.1)$
$Max\ demand\ (kVA)$	36	30	24	18	15
$\delta\alpha$ ($\%/user$)	5.23×10^{-8}	4.36×10^{-8}	3.49×10^{-8}	2.61×10^{-8}	2.18×10^{-8}

[4] I am grateful to Jean-Yves Ollier and Christophe Leininger from CRE for their invaluable help in obtaining these data.

for large residential customers (maximum demand higher than 12 kVA), and

α (%)	$(70.1, 74.7)$	$(74.7, 83.3)$	$(83.3, 99.1)$	$(99.1, 100)$
$Max\ demand\ (kVA)$	12	9	6	3
$\delta\alpha$ $(\%/user)$	1.74×10^{-8}	1.31×10^{-8}	8.71×10^{-9}	4.36×10^{-9}

for small residential customers.

$\delta W(\alpha)$, the incremental increase in net surplus from one site is estimated as $\delta W(\alpha) = W'(\alpha)\,\delta\alpha$. Since $\delta\alpha$ is discontinuous at the boundaries between users classes, so is $\delta W(\alpha)$.

The table below presents $\delta W(\alpha)$ for non residential sites for $\eta = -0.01$:

α (%)	$\alpha\,\mathrm{min}$	42	57	67
δW (€$/user/year$)	7 930	2 013	61	3

As expected $\delta W(\alpha)$ decreases rapidly as sites become smaller.

Figure 8 presents the marginal surplus per site for residential sites, expressed in € per site per year, for $\eta = -0.01$ and Figure 9 for $\eta = -0.1$. For the 18 millions residential sites whose peak demand is lower than 6 kVA, the marginal net surplus from price responsiveness is less than € 1 per year for $\eta = -0.01$, less than 4.3 € per year for $\eta = -0.1$. For the 6.5 millions residential users whose peak demand is lower than 9 kVA, marginal net surplus is less than € 1.7 per year for $\eta = -0.01$, less than 6.7 €/year for $\eta = -0.1$.

These estimates are in line with Allcott (2012), who, using an elasticity of $\eta = -0.1$, finds an annual welfare gain of \$39 per average kW of demand becoming price responsive. Since Allcott (2012) considers a 20% shift to real time pricing, his value is comparable to the average of $\delta W\ for\ \alpha \leq 20\%$, $\overline{\delta W}(0.2) =$€ 10 263 per site per year. Since the size of these sites is higher than 250 kW, the welfare gain per kW is lower than € 41 per kW, consistent Allcott's \$ 39.

To put these estimates in perspective, they can be compared to the cost of enabling one customer to switch to real-time pricing. Estimates of the cost of a real time meter (including installation) vary widely. A median estimate is € 250, higher than in Italy (€ 70) and lower than in the UK. Assuming a cost of capital at 10%, the annualized cost of each meter is € 25 */meter/year*. Furthermore, the cost of acquiring one customer and to convince him/her to switch pricing structure estimated around € 50.

5.6 Estimation of the total surplus and total surplus by users class

Finally, equation (7) can be used to estimate the total surplus from switching up to a fraction α of customers to real time pricing, assuming $\alpha \geq \alpha_{\min}$:

$$W(\alpha) = \int_{\alpha_{\min}}^{\alpha} W'(x)\, dx.$$

Numerical integration shows that $W(1)$ =€ 183 *millions* per year for low elasticity $\eta = -0.01$, and $W(1)$ =€ 507 *millions* per year for $\eta = -0.1$.

This number is consistent with other estimates. For example, *DECC* estimates at £ 6.3 billions the present value of the energy management benefits associated with installing smart gas and electricity meters. Discounting our value in perpetuity at 6.8%, *DECC*'s effective discount rate, we find a present value ranging between € 2 690 and 7 455 millions. Adjusting for the fact that the size of the British power system is roughly 80% that of France and converting into pounds provides a range between £ 1.7 and 4.8 billions for electricity only, not inconsistent with *DECC*'s estimate.

The next table presents the share of the total surplus captured for different values of α for $\eta = -0.01$:

α (%)	42	57	67	83.3
$\frac{W(\alpha)}{W(1)}$ (%)	58	71	78	89

As the previous results have led us to expect, the lion's share of the benefits is obtained with very few customers. If the 36 000 large industrial users all become price responsive, 58% of the total surplus is captured, if all non-residential users become price responsive, 78% of the surplus is captured. Capturing the remaining 20% of surplus carries a disproportionate share of the costs.

5.7 Policy implications

This analysis main finding is that the value of price responsiveness lies overwhelmingly with large customers. I believe this result, mostly driven by the differential in size between large industrial and small residential users, will prove robust as other specifications are tested on data from other countries, even though the exact values will vary.

This result has three main policy implications. First, it is essential that large customers face real-time prices, as is the case in most restructured power markets. Ensuring this captures most of the benefits of demand response.

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

When properly accounting for all these, it may be the case that the benefits still exceed the costs. For example, even if we adopt our low estimate for energy management benefits (£ 1.7 billions for electricity alone), the benefits still exceed the costs in $DECC$'s analysis. However, it appears that a large share of the energy management benefits can be obtained with a much more limited smart meters rollout.

Third, the economic case for exposing all residential customers to RTP appears weak. Policy makers have always been hesitant to do so, for fear that customers find the exposure to volatility of spot prices unbearable. The analysis presented here suggest that the economic benefit is small, less than € 4 per customer per year. Thus a voluntary approach may be preferred.

Finally, the above analysis has a commercial implication. Given the small value of price responsiveness compared the cost of convincing clients to adopt it, developing a profitable residential energy management offer will prove challenging.

6 Conclusion

The advent of "smart meters" will make possible Real Time Pricing of electricity: customers will face and react to wholesale spot prices, thus consumption of electric power will be aligned with its opportunity cost. This article determines the marginal value of a fraction of demand (or a consumer) switching to Real Time Pricing. First, it derives this marginal value for a simple yet realistic specification of demand. Second, using data from the French power market, it estimates that, for the vast majority of residential customers whose peak demand is lower than 6 kVA, the net surplus from switching to Real Time Pricing is lower than 1 €/year for low demand elasticity, 4 €/year for high demand elasticity. This finding casts a doubt on the economic value of rolling out smart meters to all residential customers, for both policy makers and power suppliers.

This analysis will be expanded in at least four directions. First, the methodology will be applied to other power markets, that present different shapes of load duration curves, different distributions of consumer sizes, and different generation mix.

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.

Third, intermittent generation will be included. This will increase the volatility of supply, hence the value of demand flexibility. Quantifying that impact will be extremely important.

Finally, alternative specification of demand will be tested. For example, demand can be assumed to be log-linear, or constant elasticity, 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.

7 Acknowledgments

I am grateful for insightful comments to Claude Crampes, Schmuel Oren, Richard Green, seminar participants at the Toulouse School of Economics and the Paris School of Economics, and participants to the 2011 Conference on Energy Markets in Toulouse, and to the 2012 Cambridge Electricity Policy Research Group Spring Seminar. A grant from Electricité de France that supported this research is gratefully acknowledged. All errors and omissions are mine.

References

[1] H. Allcott. Rethinking real time electricity pricing. *Resource and Energy Economics*, 33(4):820–842, November 2011.

[2] H. Allcott. The smart grid, entry, and imperfect competition in electricity markets. http://www.nber.org/papers/w18071, May 2012. NBER Working Paper No. 18071.

[3] M. Boiteux. La tarification des demandes en pointe: Application de la théorie de la vente au cout marginal. *Revue Générale de l'Electricité*, pages 321–40, August 1949.

[4] S. Borenstein. The long-run efficiency of real time pricing. *The Energy Journal*, 26(3):93–116, April 2005.

[5] S. Borenstein and S. Holland. On the efficiency of competitive electricity markets with time-invariant retail prices. *RAND Journal of Economics*, 36:469–493, 2005.

[6] A. Faruqui, D. Harris, and R. Hledik. Unlocking the Ă53 billion savings from smart meters in the EU. The Brattle Group Discussion Paper, October 2009.

[7] S. Holland and E. Mansur. The short-run effects of time-varying prices in competitive electricity markets. *The Energy Journal*, 27(4):127–155, October 2006.

[8] IEA. *Projected Cost of Generating Electricity.* OECD/IEA, 2010.

[9] P. Joskow and J. Tirole. Reliability and competitive electricity markets. *RAND Journal of Economics*, 38(1):60–84, Spring 2007.

[10] M. Lijesen. The real-time price elasticity of electricity. *Energy Economics*, 29:249–258, 2007.

[11] R. Patrick and F. Wolak. Estimating the customer-level demand for electricity under real time market prices, August 1997.

[12] P. Reiss and M. White. Household electricity demand, revisited. *Review of Economic Studies*, 72(3):853–883, July 2005.

A Derivations of the closed-form solutions

A.1 Optimal investment

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

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

$\Leftrightarrow$

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

Adopt the convention $t_{N+1} \to +\infty$. For $n \in \{1, ..., N\}$, define

$$I_n = \int_{\bar{t}_n}^{t_{n+1}} (\rho(K_n; t) - c_n) f(t)\, dt.$$

Integrating by parts,

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

Determination of $\{K_n\}_{1 \le n \le N}$ proceeds by backwards induction. Start with technology N, the last technology available. Equation (2) is:

$$I_N = r_N.$$

Since $y(K_N, c_{N+1}, \alpha) \to 0$,

$$I_N = \frac{y(K_N, c_N, \alpha)^{1+\lambda}}{\alpha(1+\lambda)a_1^{\lambda}} = r_N.$$

Thus, total capacity K_N solves:

$$y(K_N, c_N, \alpha)^{1+\lambda} = \alpha a_1^{\lambda}(1+\lambda) r_N$$

$\Leftrightarrow$

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

Suppose now K_{n+1} is determined. Then, K_n solves $\Psi(K_n, c_n) = r_n$ where

$$\begin{aligned}
\Psi(K_n, c_n) &= \int_{\widehat{t}(K_n,c_n)}^{\widehat{t}(K_n,c_{n+1})} [\rho(K_n;t) - c_n] f(t)\, dt + \int_{\widehat{t}(K_n,c_{n+1})}^{+\infty} [c_{n+1} - c_n] f(t)\, dt + \Psi(K_{n+1}, c_{n+1}) \\
&= I_n + (c_{n+1} - c_n) e^{-\lambda_1 t_{n+1}} + \Psi(K_{n+1}, c_{n+1}) \\
&= \frac{y(K_n, c_n, \alpha)^{1+\lambda} - y(K_n, c_{n+1}, \alpha)^{1+\lambda}}{\alpha a_1^{\lambda}(1+\lambda)} + r_{n+1}.
\end{aligned}$$

Thus, if

$$\frac{y(K_{n+1}, c_n, \alpha)^{1+\lambda} - y(K_{n+1}, c_{n+1}, \alpha)^{1+\lambda}}{\alpha a_1^{\lambda}(1+\lambda)} \leq r_{n+1} - r_n,$$

$K_n \leq K_{n+1}$ is uniquely defined by:

$$y(K_n, c_n, \alpha)^{1+\lambda} - y(K_n, c_{n+1}, \alpha)^{1+\lambda} = \alpha a_1^{\lambda}(1+\lambda)(r_n - r_{n+1}).$$

A.2 No rationing condition

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

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

No rationing occurs at the optimum if and only if, for all $t \geq 0$:

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

Given the structure of the price, this is equivalent to:

$$\frac{a(t)+p^R}{2} \geq \rho(K_N;t) \ \ \forall t \geq t_{n+1}$$

$\Leftrightarrow$

$$\frac{a(t)+p^R}{2} \geq \frac{a(t)-bK_N-(1-\alpha)p^R}{\alpha} \ \forall t \geq t_{n+1}$$

$\Leftrightarrow$

$$bK_N \geq (2-\alpha)\frac{a(t)-p^R}{2} \ \forall t \geq t_{n+1}$$

$\Leftrightarrow$

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

Substituting in the expression of K_N, this is equivalent to:

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

$\Leftrightarrow$

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

A.3 Marginal value of α for $N \geq 2$

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

$$Var[p(t)] = 2L_N = \sum_{n=1}^{N}\left\{\left(c_n - p^R\right)^2\int_{t_n}^{t_{n+1}} f(t)\,dt + 2\left(c_n - p^R\right)I_n + \int_{\bar{t}_n}^{t_{n+1}}\left(\rho(K_n;t)-c_n\right)^2 f(t)\,dt\right\}.$$

Integrating by parts twice,

$$
\begin{aligned}
H_n &= \int_{\bar{t}_n}^{t_{n+1}} \left(\rho\left(K_n;t\right)-c_n\right)^2 f\left(t\right)dt \\
&= \left[-\left(\rho\left(K_n;t\right)-c_n\right)^2\left(1-F\left(t\right)\right)\right]_{\bar{t}_n}^{t_{n+1}} + 2\lambda_2 \frac{a_1}{\alpha}\int_{\bar{t}_n}^{t_{n+1}} \left(\rho\left(K_n;t\right)-c_n\right)e^{-(\lambda_1+\lambda_2)t}dt \\
&= -\left(c_{n+1}-c_n\right)^2 e^{-\lambda_1 t_{n+1}} + \frac{2}{1+\lambda}\frac{a_1}{\alpha}\left(\begin{array}{c} \left[-\left(\rho\left(K_n;t\right)-c_n\right)e^{-(\lambda_1+\lambda_2)t}\right]_{\bar{t}_n}^{t_{n+1}} \\ +\frac{a_1\lambda_2}{\alpha}\int_{\bar{t}_n}^{t_{n+1}} e^{-(\lambda_1+2\lambda_2)t}dt \end{array}\right) \\
&= -\left(c_{n+1}-c_n\right)^2 e^{-\lambda_1 t_{n+1}} - 2\frac{a_1}{\alpha}\left(c_{n+1}-c_n\right)\frac{e^{-(\lambda_1+\lambda_2)t_{n+1}}}{1+\lambda} + 2\left(\frac{a_1}{\alpha}\right)^2 \frac{\left(e^{-(\lambda_1+2\lambda_2)\bar{t}_n} - e^{-(\lambda_1+2\lambda_2)t_{n+1}}\right)}{\left(1+\lambda\right)\left(1+2\lambda\right)} \\
&= -\left(c_{n+1}-c_n\right)^2 e^{-\lambda_1 t_{n+1}} - 2\left(c_{n+1}-c_n\right)\frac{y\left(K_n,c_{n+1},\alpha\right)^{1+\lambda}}{\alpha a_1^{\lambda}\left(1+\lambda\right)} + 2\frac{y\left(K_n,c_n,\alpha\right)^{2+\lambda} - y\left(K_n,c_{n+1},\alpha\right)^{2+\lambda}}{\alpha^2 a_1^{\lambda}\left(1+\lambda\right)\left(1+2\lambda\right)}.
\end{aligned}
$$

Then,

$$
\begin{aligned}
S_1 &= \sum_{n=1}^{N}\left(c_n-p^R\right)^2\int_{t_n}^{t_{n+1}} f\left(t\right)dt = \sum_{n=1}^{N}\left(c_n-p^R\right)^2\left(e^{-\lambda_1 t_n} - e^{-\lambda_1 t_{n+1}}\right) \\
&= \sum_{n=1}^{N} e^{-\lambda_1 t_{n+1}}\left[\left(c_{n+1}-p^R\right)^2 - \left(c_n-p^R\right)^2\right] + \left(c_1-p^R\right)^2 e^{-\lambda_1 t_1} - \left(c_{N+1}-p^R\right)^2 e^{-\lambda_1 t_{N+1}} \\
&= \sum_{n=1}^{N} e^{-\lambda_1 t_{n+1}}\left(c_{n+1}-c_n\right)\left(c_{n+1}+c_n-2p^R\right) + r_1^2.
\end{aligned}
$$

Hence,

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

Since $p^R = c_1 + r_1$,

$$\begin{aligned} S_2 &= \sum_{n=1}^{N-1} \left(c_n - p^R\right)\left(r_n - r_{n+1}\right) + \left(c_N - p^R\right) r_N = \sum_{n=1}^{N} \left(c_n - p^R\right) r_n - \sum_{n=1}^{N-1} \left(c_n - p^R\right) r_{n+1} \\ &= \sum_{n=1}^{N-1} r_{n+1}\left(c_{n+1} - c_n\right) - r_1^2. \end{aligned}$$

Thus,

$$\begin{aligned} L_N &= \sum_{n=1}^{N-1} \left(c_{n+1} - c_n\right)\left(r_{n+1} - \frac{y\left(K_n, c_{n+1}, \alpha\right)^{1+\lambda}}{\alpha a_1^{\lambda}\left(1+\lambda\right)}\right) + \sum_{n=1}^{N} \frac{y\left(K_n, c_n, \alpha\right)^{2+\lambda} - y\left(K_n, c_{n+1}, \alpha\right)^{2+\lambda}}{\alpha^2 a_1^{\lambda}\left(1+\lambda\right)\left(1+2\lambda\right)} - \frac{r_1^2}{2} \\ &= \sum_{n=1}^{N}\left(\left(c_{n+1} - c_n\right)\left(r_n - \frac{y\left(K_n, c_n, \alpha\right)^{1+\lambda}}{\alpha a_1^{\lambda}\left(1+\lambda\right)}\right) + \frac{y\left(K_n, c_n, \alpha\right)^{2+\lambda} - y\left(K_n, c_{n+1}, \alpha\right)^{2+\lambda}}{\alpha^2 a_1^{\lambda}\left(1+\lambda\right)\left(1+2\lambda\right)}\right) - \frac{r_1^2}{2} \\ &= -\frac{r_1^2}{2} + \frac{1}{\alpha a_1^{\lambda}\left(1+\lambda\right)} \sum_{n=1}^{N} \left(\begin{array}{c} \left(c_{n+1} - c_n\right)\left(\alpha a_1^{\lambda}\left(1+\lambda\right) r_n - y\left(K_n, c_n, \alpha\right)^{1+\lambda}\right) \\ + \frac{y\left(K_n, c_n, \alpha\right)^{2+\lambda} - y\left(K_n, c_{n+1}, \alpha\right)^{2+\lambda}}{\alpha\left(1+2\lambda\right)} \end{array}\right). \end{aligned}$$

Consider now the units. Equation (7) can be rewritten as $W'\left(\alpha\right) = 8760 Q^{\infty} \frac{L_N}{bQ^{\infty}}$, 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, $\left[W'\left(\alpha\right) = 8.76 Q^{\infty} \frac{L_N}{bQ^{\infty}}\right] = €$ millions$/year$.

Figure 1: Characterization of the equilibrium

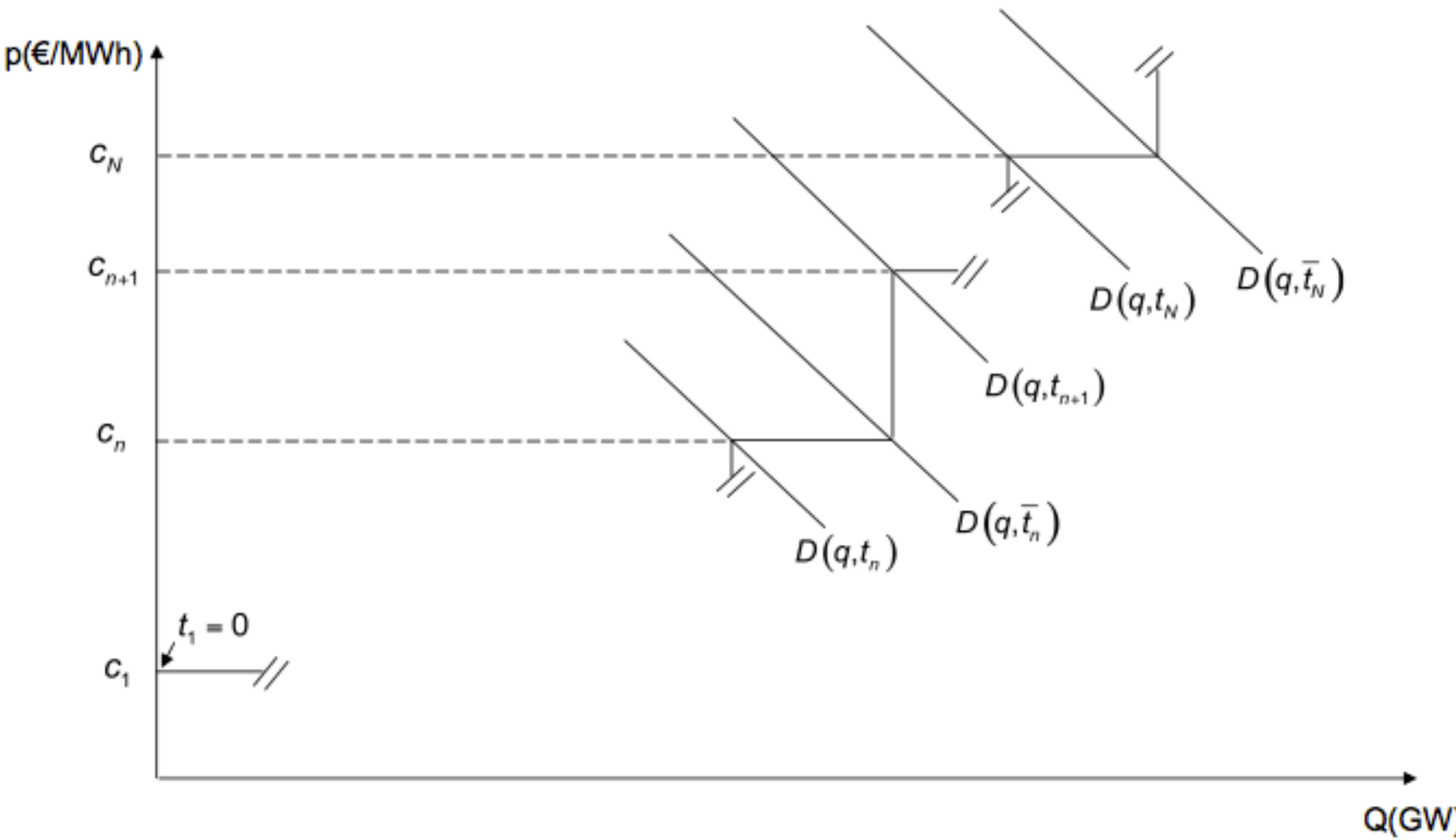


Figure 2: Welfare increase

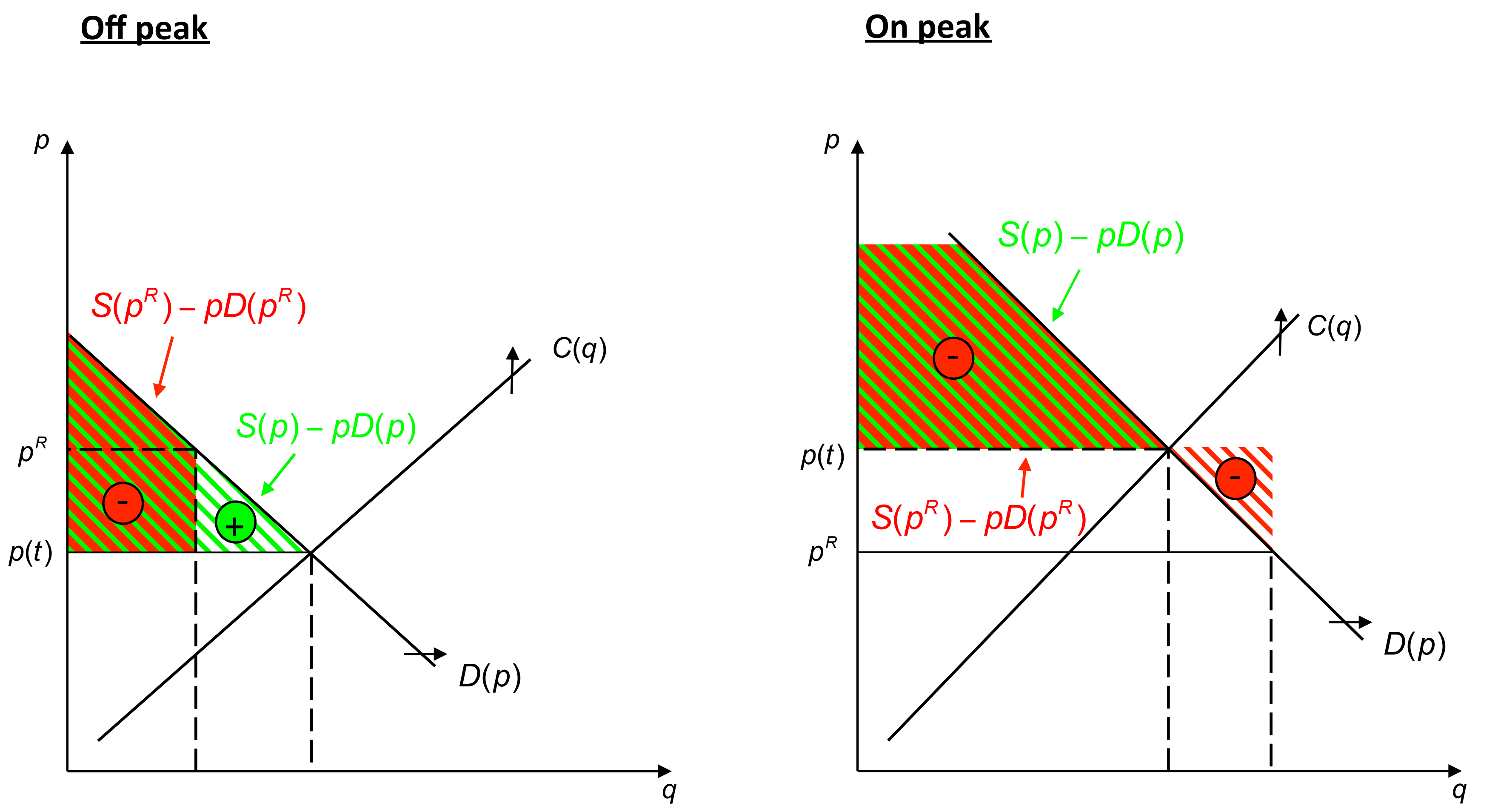


Figure 3: Actual and fitted demand

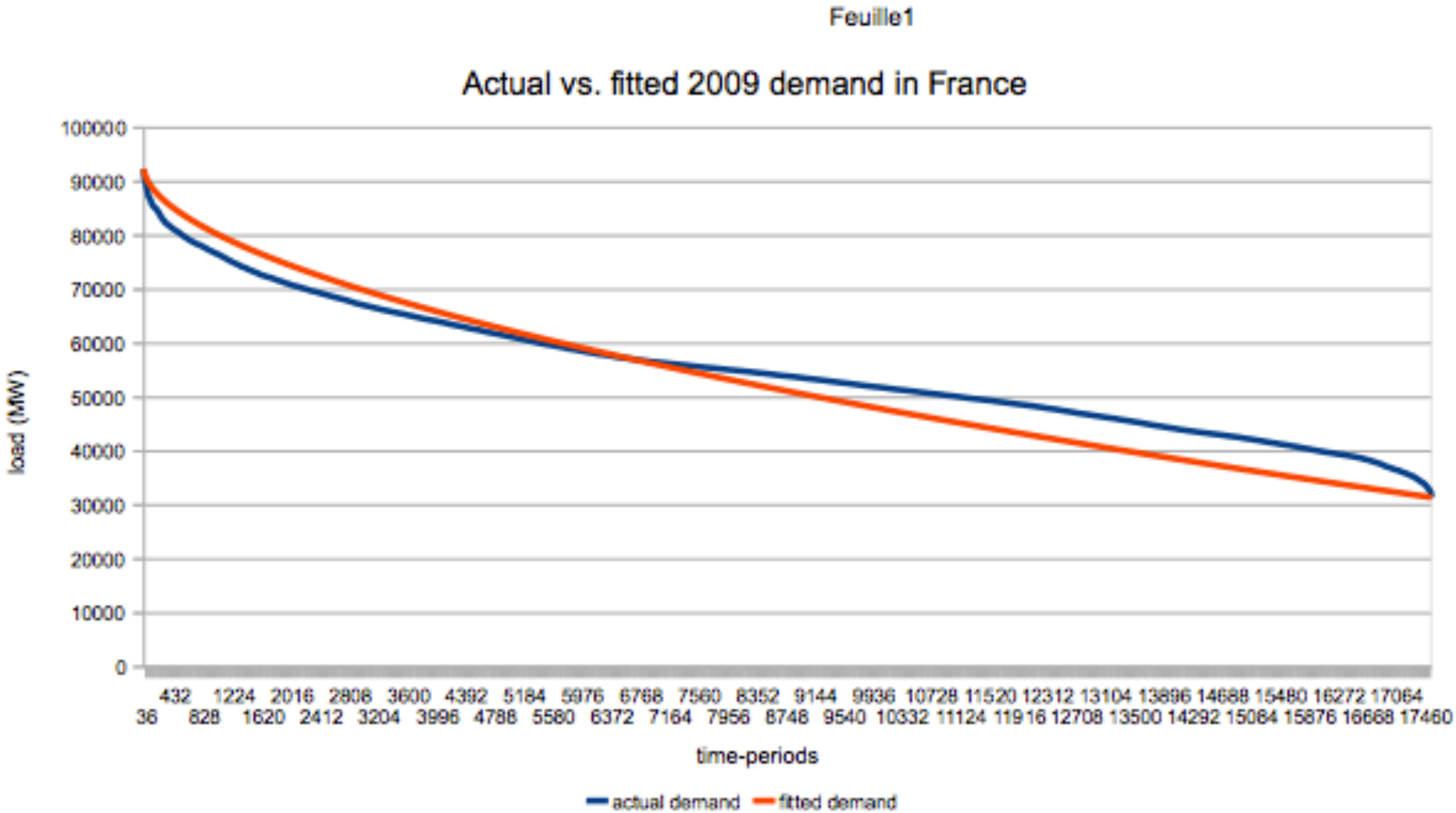


Figure 4: marginal surplus from price responsiveness, base elasticity

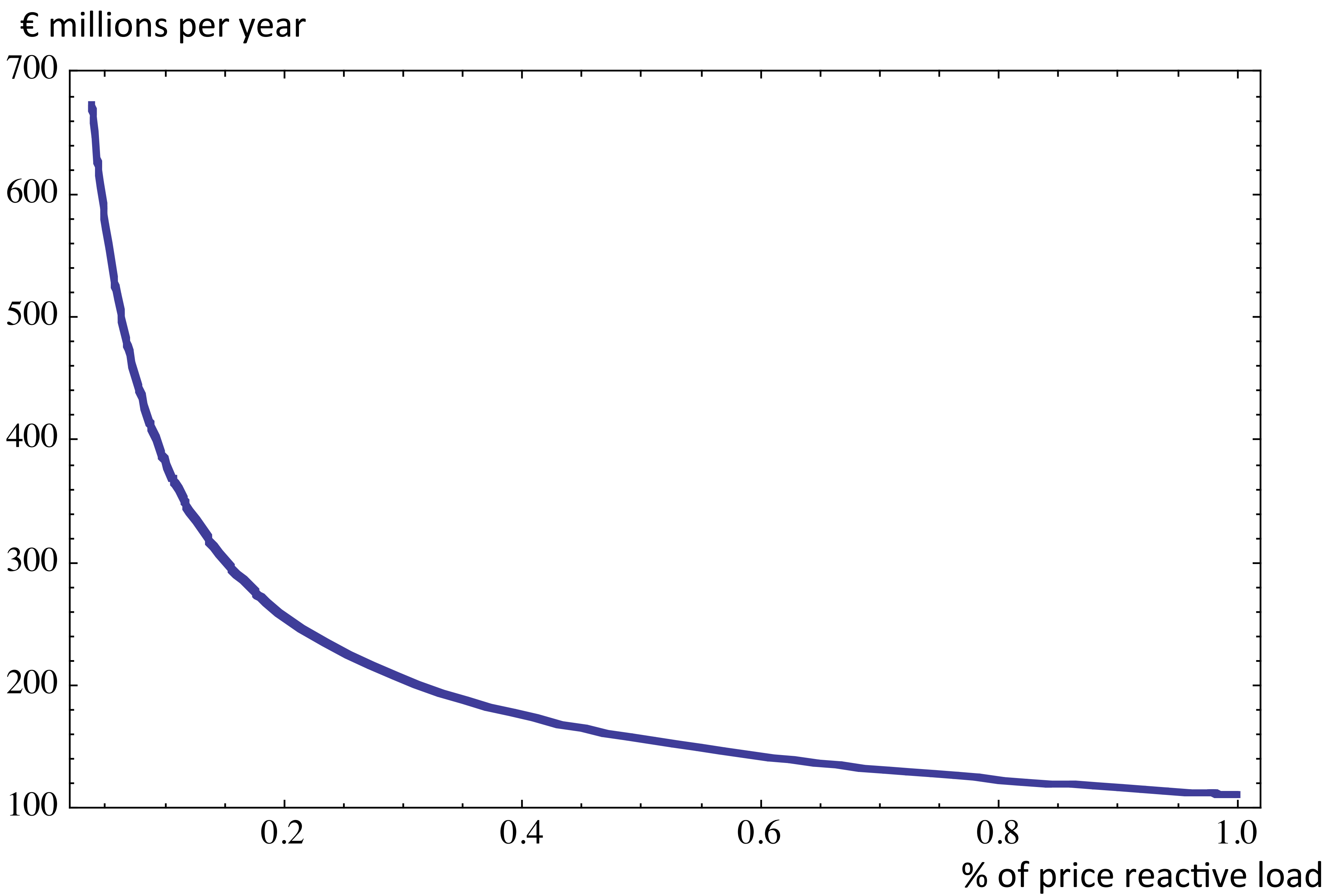


Figure 5: marginal surplus from price responsiveness, high elasticity

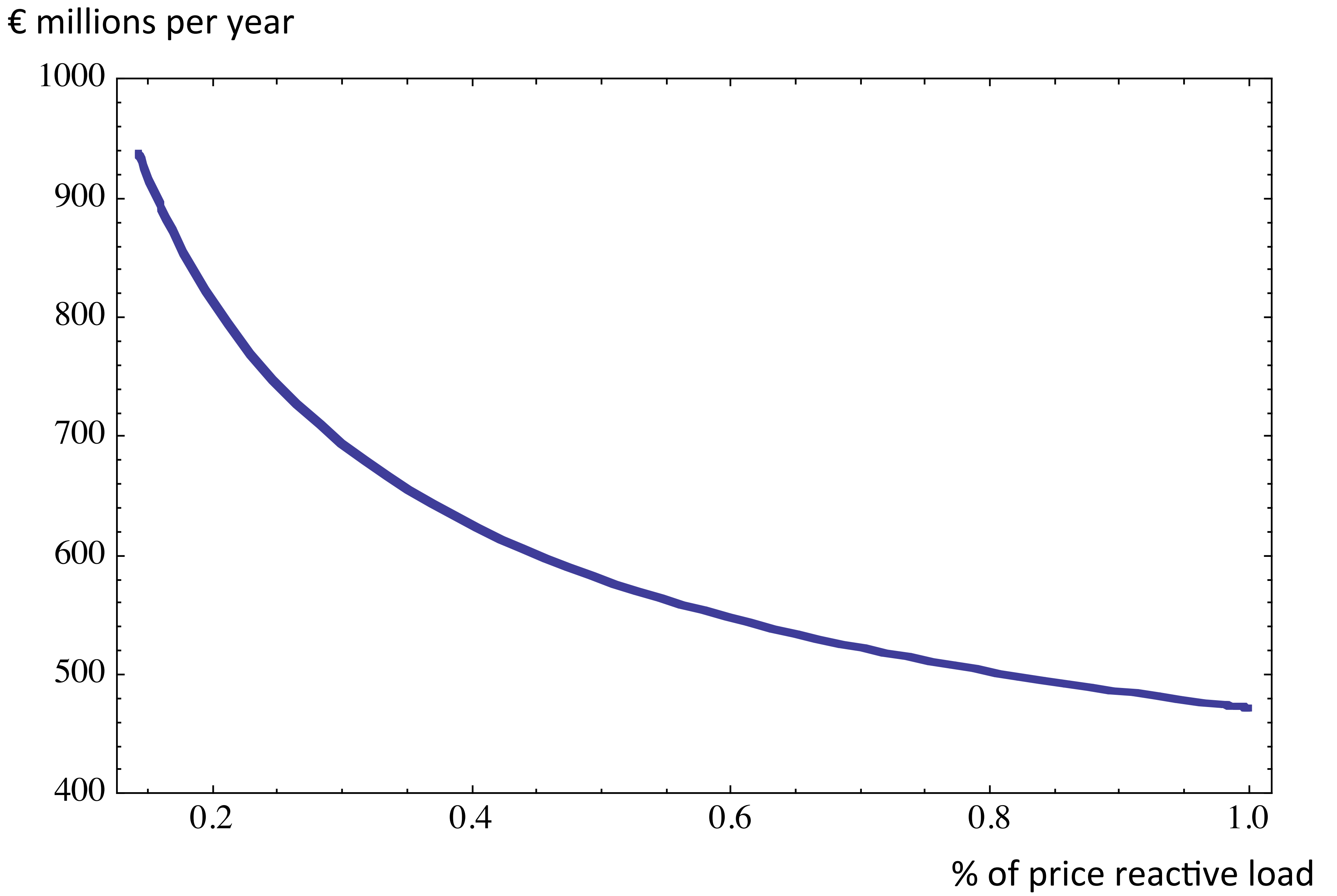


Figure 6: marginal surplus from price responsiveness, multiple technology mix, base elasticity

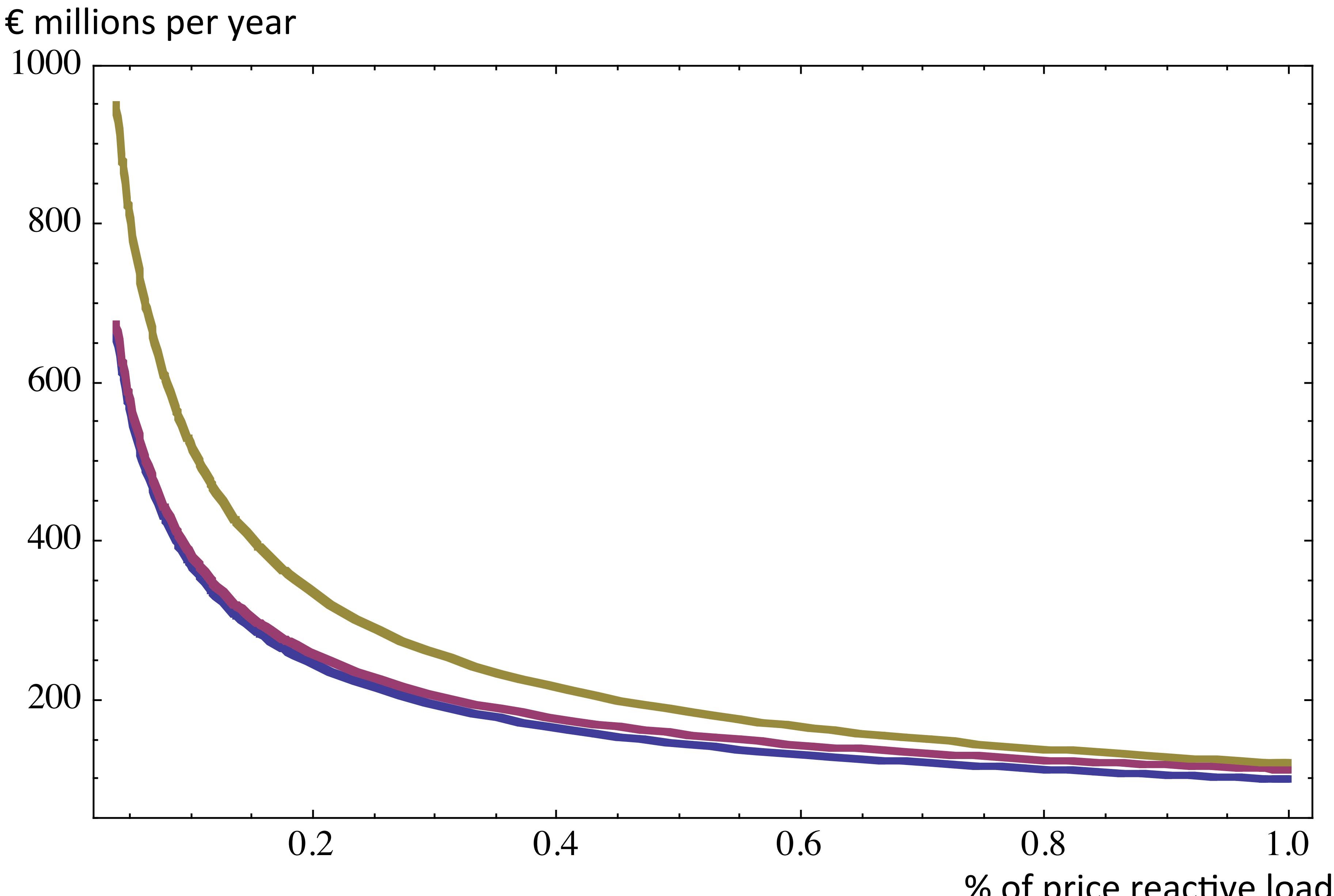
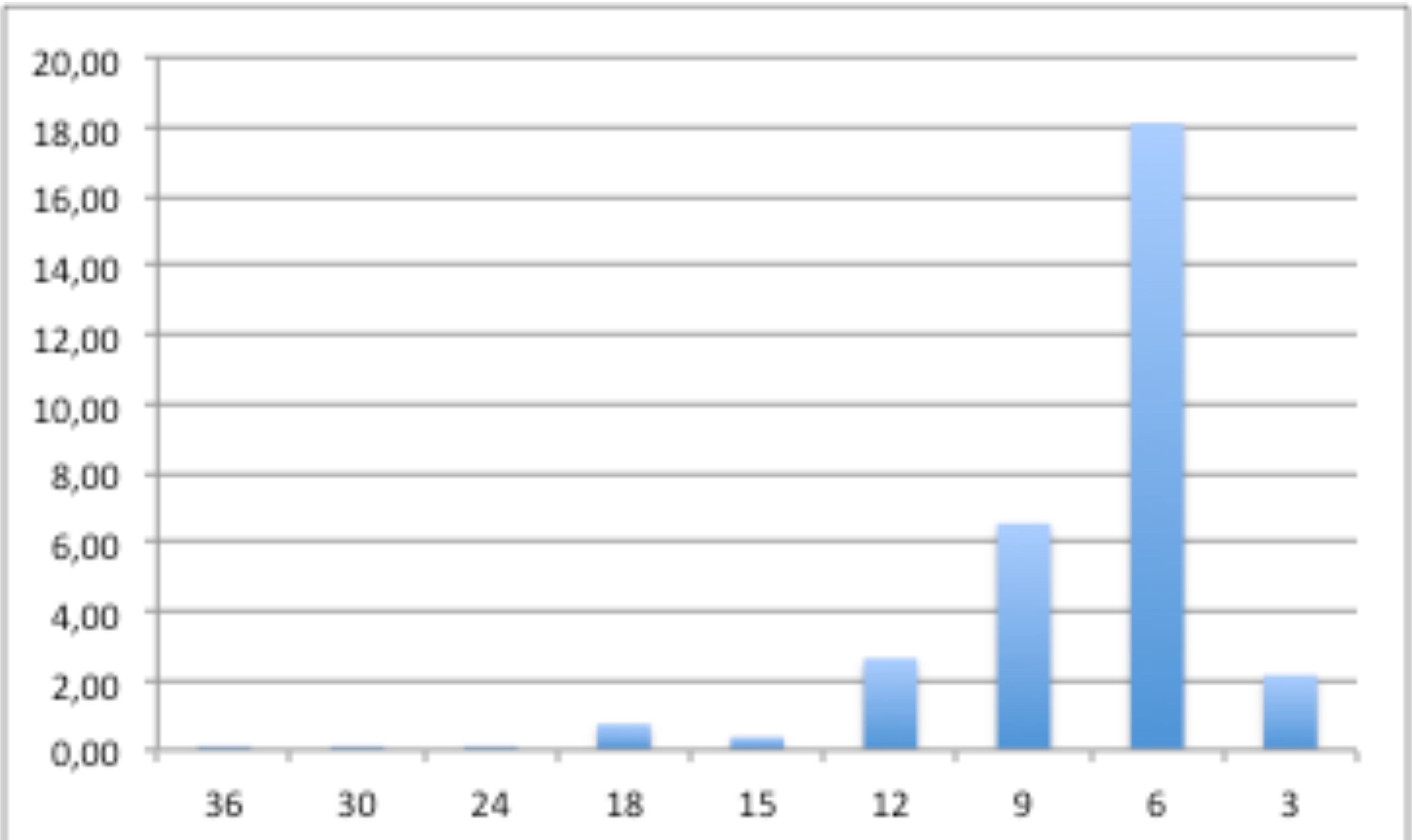


Figure 7: Distribution of residential consumers by size

millions of sites



Maximum peak demand

Figure 8: marginal surplus per residential consumer, base elasticity

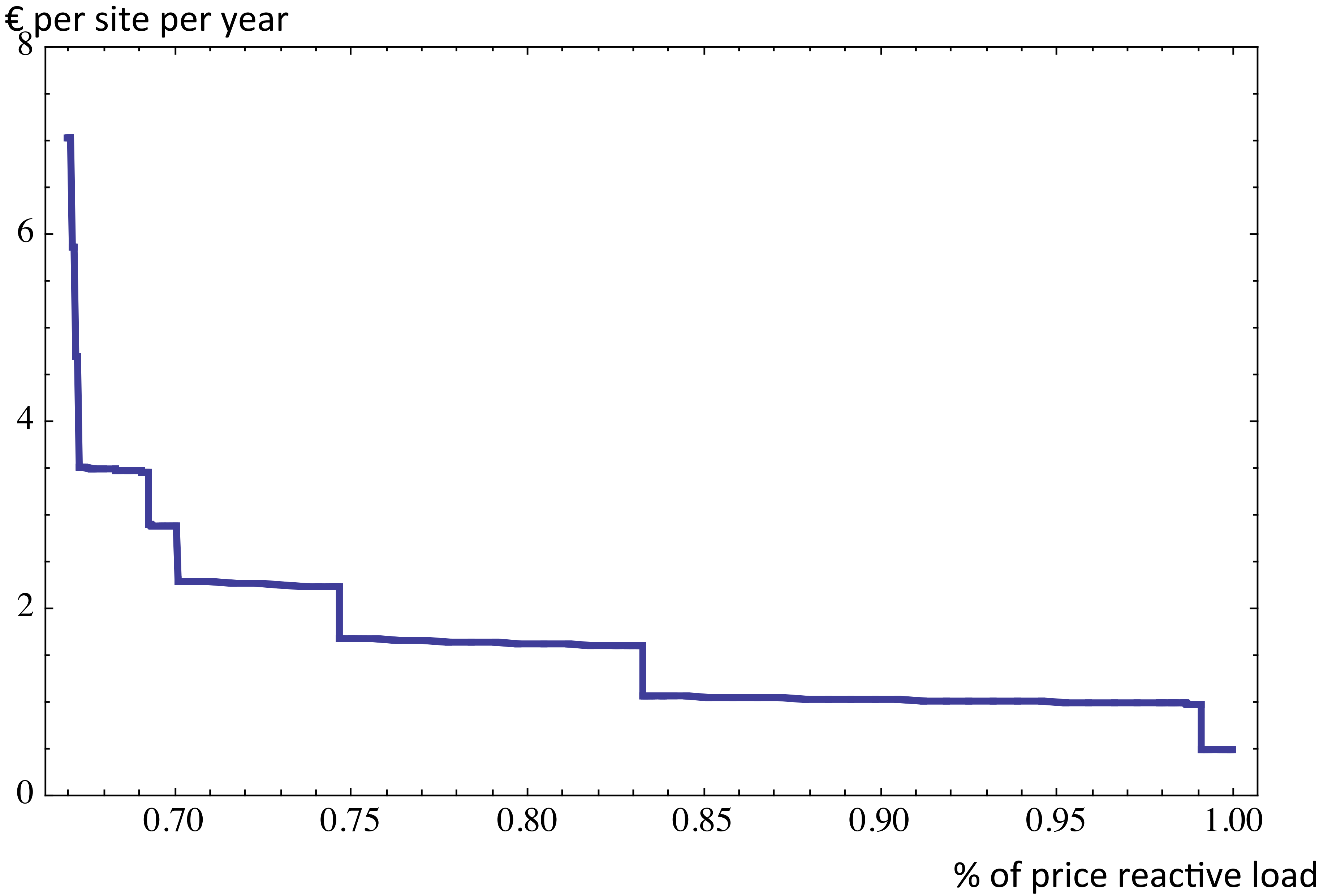


Figure 9: marginal surplus per residential consumer, high elasticity