CHAIRE EUROPEAN ELECTRICITY MARKETS/WORKING PAPER #55

SECURING INVESTMENT FOR ESSENTIAL GOODS

HOW TO DESIGN DEMAND FUNCTIONS IN RESERVATION MARKETS?

Leopold Monjoie, Fabien Roques

Dauphine | PSL

CHAIRE EUROPEAN ELECTRICITY MARKETS

Chaire de recherche soutenue par

Rte EDF epexspot TOTAL direct energie

SECURING INVESTMENT FOR ESSENTIAL GOODS HOW TO DESIGN DEMAND FUNCTIONS IN RESERVATION MARKETS?

Leopold Monjoie[1,*], Fabien Roques[2]

January 2022

Abstract

This paper studies the provision of an essential good with time-varying uncertain stochastic demand and capacity-constrained producers such as electricity. Due to price regulation, public good externalities, and market power, investments are typically under-procured by private agents. To restore efficient investment level, we analyze the design of reservation markets where producers can sell their capacity availability before the demand is known. While their direct effect on investment decisions is well known, we focus on indirect effects generated by their implementation, namely how the capacity price is allocated on the demand side and how the realized demand is accounted for in the market design. We develop a novel approach to studying the reservation market's interdependencies and the subsequent production and retail markets for the essential good. We provide a sequential analytical model of the three markets and describe how different market design regimes can indirectly affect the equilibria in the production and retail markets in terms of prices, investment level, and welfare. In particular, we demonstrate that the ability of the reservation market to restore the social optimum, or at least to reach a second-best optimum, crucially depends on the different design regimes of the reservation market, on the assumptions of policy interventions, and the various market inefficiencies. The model results and the associated policy implication are discussed first using a general framework and then in reference to electricity markets where capacity reservation is often used to ensure adequate investment to ensure the security of supplies.

Keywords: Market Design, Investment decisions, Imperfect competition, Regulation.

Acknowledgments: The authors are indebted to Françoise Forges, Sebastian Schwenen, Nicolas Astier, Vincent Rious and Jan-Horst Keppler for the interesting discussions. The suggestions from participants of the IAEE, the Microsoft Innovation Workshop 2021, and CEEM seminars are also much appreciated.

Disclaimer: This paper has benefited from the support of the Chaire European Electricity Markets (CEEM) of the Université Paris-Dauphine under the aegis of the Foundation Paris-Dauphine, supported by RTE, EDF, EPEX Spot, and Total Direct Energie. The views and opinions expressed in this paper are those of the authors and do not necessarily reflect those of the partners of the CEEM.

[1] Chaire European Electricity Markets, Paris-Dauphine University - PSL, France
[*] Corresponding author: leopold.monjoie@dauphine.psl.eu
[2] Chaire European Electricity Markets, and Associate Professor with Paris-Dauphine University - PSL and the Florence School of Regulation.

I. INTRODUCTION

For some essential goods with demand varying over time, wholesale markets' private incentives are not sufficient to ensure that producers make enough investments to meet peak demand in advance of the time when the demand materializes[1]. In such industries, due to the critical importance of these goods, policymakers tend to intervene and implement price caps or other types of regulation that distort the price signal and undermine investment incentives.[2] Moreover, these goods can be characterized as public goods, and they may exhibit externalities, for instance, a cold wave with peak electricity demand or a pandemic with peak demand for medicine or medical equipment.[3] In such circumstances, the absence of adequacy between the capacity and the peak demand, combined with the difficulty of implementing efficient rationing, leads to high costs for society.

One solution to restore the optimal level of investment lies in implementing a mandatory reservation market in which producers commit to having capacity available to meet the expected peak demand collectively, as prescribed by the regulator.[4] On the supply side, each participating producer makes a price-quantity offer for a capacity. If a producer sells capacity in this reservation market, he receives a capacity price and commits to being available to produce over future periods.

While the supply emerges naturally on those markets, the demand requires a regulatory intervention. Indeed, the public-good nature of investment during high-demand periods implies that consumers are unwilling to buy capacities in reservation markets.[5] Hence, the regulator must define the demand function administratively, so the market clears and provides producers' capacity prices. This paper establishes a framework describing the economic impacts of different demand designs for reservation markets and their policy implications. We focus on two interrelated questions that relate to (i) the cost allocation regime between capacity buyers and final consumers; and (ii) the degree to which the final consumers realized demand is accounted in the market allo-

[1]The intial framework of this work is electricity markets but our analysis fits into the more general analysis of industries in which a form of competition follows long-run investments such as communication network [1], or radio spectrum [24].

[2]Policy interventions such as price caps and non-economic distortions made by a public entity lead to a *Missing Money* issue that prevents sufficient revenue from being collected to cover costs [14].

[3]This inefficiency is associated with the existence of a *Missing Market* issue under which producers consider their revenue insufficient to invest optimally [20]. This can be caused by hedging markets being incomplete [7], or because of externalities associated with the public-good nature of investment and consumption choices [12], innovation spillovers, and climate change.

[4]Current implementations of such mechanisms have been the prerogative of the electricity sector under the name capacity remuneration mechanism, see for instance [8] for a technical description of potential implementations.

[5]The absence of a 'spontaneous' demand function has similar roots as the *Missing Money* issue previously discussed. Transaction cost and asymmetric information prevent adequate transactions up to the optimal level; see for instance [15] for a discussion in electricity markets. The insurance of having enough capacity has a private value (how much each consumer is willing to pay to avoid inadequacy) but also a social value, as an increase in investment reduces the probability of systemic costs [9]. Furthermore, it is sometimes technologically, socially, and economically impossible to know the willingness to pay for this insurance.

cation design.[6] In this paper, we first describe the channel through which each possible regime impacts the equilibrium on the demand side and ultimately how it changes investments decisions. Then, we demonstrate how those indirect effects constrain the regulator in choosing the first best economic level. In other words, our model allows analyzing the endogeneity between the welfare-maximizing market outcome that regulators aim to restore through their intervention and the design of the reservation market implemented.

The direct effect of an additional stream of remuneration on investment decisions is well understood. The current literature covers a significant range of issues : (i) the outcomes in investment decisions with and without reservations markets, (ii) the effect of market power on the capacity price determination, (iii) the relation with risk and business cycle (iv) the discrimination between different investment technologies.[7] However, to our knowledge, there has been no formal analysis of different reservation markets' demand designs, the incentive properties of these alternative approaches, and their ability to restore the socially optimal level of investment beyond the direct effect of the increase of the marginal investment value due to the additional capacity price.[8] [21] is the first paper to represent retailers' strategies in the reservation market. She develops a theoretical model to analyze the preferences regarding information precision for uncertain future demand. Contrary to our approach, she models heterogeneous price taker producers and homogeneous buyers competing for *à la Cournot* under uncertainty on their level of capacity obligation. In this paper, we take a step back from the direct price-setting approach and develop a model that sheds light on the complex interactions between the reservation market design and the incentives of producers and retailers. Our model provides some new and non-intuitive insights on their incentive properties and their ability to restore the socially optimal (welfare-maximizing) level of investment

An annex contribution of our model to the literature lies in constructing the supply function in the reservation market. To ensure the endogeneity of the investment decisions and the emergence of equilibrium in the three markets in our model, we derive an endogenous supply function in the reservation market. Namely, following the main theoretical view for reservation markets[9], we assume that producers offer their marginal opportunity cost of providing additional capacity. This opportunity cost equals the marginal loss of revenue incurred by the investment level beyond the profit-maximizing equilibrium. Our modeling proposition is central as any indirect effects generated by the reservation market can affect the expected revenue made by the

[6]We also use interchangeability term ex-post temporality, as reservations market are set ex-ante before demand is known; as well as the term capacity demand allocation. At the same time, while this is out of the scope of this paper, the dynamic nature of the provision of an essential good is central. It includes the decision to invest, which can span from many years to a few months, and the decision to consume the good. For instance, in electricity and vaccines, the good is almost immediately consumed. On the other hand, medical equipment, intensive care units, strategic energy reserves, or human capital are more durable goods.

[7]See [4] for a detailed literature review on the theory and implementations issues of reservation markets.

[8]On the other hand, the importance of the demand function design in the reservation market is well known. See for instance, [11] and [5] [9] [3]. However, those papers still only consider the effect of the reservation market directly on the supply side, while our paper underlines the indirect effect of this instrument on retailers and consumers, which in turn impacts producers.

[9]See for instance [6].

producers and can indirectly be captured during the formation of the supply function in the reservation market.

The first market design regime studied is the canonical reservation market. We build on the previous literature, and the design found in [18] and [12] which relies on the assumption that the reservation market does not have any effect beyond increasing the investment level. Following the approach of our paper, this canonical regime is similar to the cost passthrough using a lump-sum tax. In this case, and even when considering the endogenous supply function in the reservation market, this market design always restores the first-best optimum given the system inefficiencies. Namely, providing that the demand function is intersecting the supply function at the optimal level,[10] the equilibrium price in the reservation market and the suboptimal price in the subsequent markets is always equal to the price in an optimal system without price cap.[11]

We then investigate the case in which the reservation price impacts the consumers at the margin. In this case, the regime is similar to allocating the capacity price as a Pigouvian unitary tax. We show that the existence of the reservation market indirectly modifies the provision of the essential good on the wholesale market by redistributing the different states of the world between off-peak and on-peak periods[12], and by lowering consumers surplus. Therefore, we demonstrate that the first-best outcomes under this regime are always lower than under the canonical regime.

In the first extension, we analyze a second inefficiency. When the price cap is reached, the investment availability becomes a public good as the demand becomes inelastic. Due to the impossibility of efficiently rationing consumers, they incur a significant welfare loss.[13] This additional assumption regarding the inefficiency of wholesale-only market design has significant implications for the design of the reservation demand function. Indeed under this new assumption, we find that the indirect effect created by allocating the capacity price on a unitary basis is now ambiguous for social welfare. Under specific parameters value, the reservation market can bring more social welfare at the first-best level than the initial allocation regime or the optimal wholesale-only market.

As a third step, we extend our analysis to implementing a regime where the regulator allocates the cost based on actual retailers' market shares. It allows us to introduce the ex-post temporality in the current analysis, where the design of reservation markets considers the realized demand and analyzes the effect of retailers' market power in the model. We first show how this design affects at the margin the retailers who play *'à la Cournot'* on the retail market, and then we integrate the new equilibrium into

[10]In this paper and unlike [11], we do not analyze the risk of having regulatory errors.

[11]We also demonstrate that this result also holds for other types of inefficiencies. With the price cap, the equilibrium reservation price equals the expected lost revenue. In the case of capacity as a public good, the price is equal to the difference between the private value of the investment and its social value.

[12]That is when the capacity does not bind and binds.

[13]Using the same initial model [12] showed that additional capacity payment is necessary when the system includes the public-good nature of the investments. Indeed, the inadequacy between capacity and consumption generates negative externalities. Hence, to fully internalize the effect of the capacity inadequacy, it is necessary to generate an adder on the wholesale price. We also use in this paper the same representation of the public-good nature of the investment. This effect of a price cap is also closely related to the concept of reliability externality described by [23].

our model with investment decisions and the reservation market. We find that this allocation creates an intermediary outcome between the unitary tax and the lump-sum tax while having significant redistributional properties. Finally, we can also study the effect of the retail market structure on the equilibrium outcomes of the model.[14] Depending on the assumptions with respect to the model parameters and the inefficiencies assumptions, we find that lowering the number of retailers can provide additional social welfare.

Finally, we analyze the case of a reservation market entirely based on the realized demand level. To do so, retailers are obliged to cover their quantity sold on the retail market by buying directly on the reservation market, given a penalty system. We focus on how retailers' individual strategies can form an aggregated demand function in the reservation market, and we analyze the optimal capacity bought by retailers in the reservation market. We find that such an approach for the demand function can provide the optimal level of investment under specific conditions. The market equilibrium under this regime relies on the marginal value a capacity brings to retailers' profit which also depends on the market structure in the retail market, the consumers' demand function, and the penalty system.

We finally discuss the policy relevance of our findings in reference to electricity markets where capacity reservation is often used to ensure adequate investment to ensure the security of supplies. We conclude by reviewing future potential extensions of our work.

We provide in Section 2 a reminder of the benchmark model that describes investment decisions in generation capacity. In the same section, we implement the reservation market and build the theoretical supply function. Then, we model the market designs for the demand function in Section 3 for the different cost passthrough regimes and Section 4 for the different ex-post temporality regimes. The closed-form solution and the numerical application are presented in Section 5. To conclude, we discuss possible extensions of the model in Section 6.

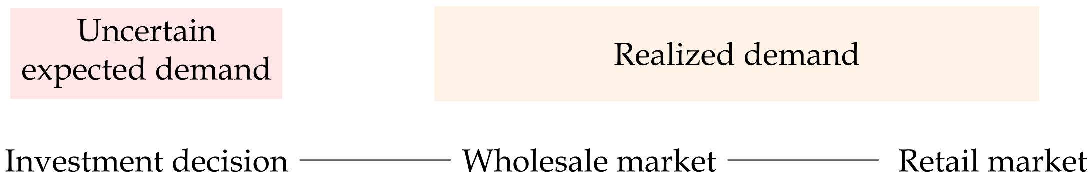
II. BENCHMARK MODEL

2.1. Assumptions

We consider an initial economic system with three agents: producers, retailers, and final consumers. Producers invest in capacities to produce a homogeneous good. They sell the goods on a wholesale upstream market to retailers. Then, retailers resell it on a downstream retail market to consumers.

Model stages. The model has three stages. First, producers choose the level of investment. Second, the wholesale market clears. Third, the retail market clears. We assume the final consumers' demand is uncertain for all agents when making investment decisions. On the other hand, the demand is known when the producers and retailers sell the goods. Those two stages can be interpreted as a repetition of multiple states of the world over a given period (for example, one year), drawn from the distribution [17].

[14]We do not consider market power on the supply side in our paper, as it is well documented in the literature, see for instance [25] and [17] for its effect on investment decision with a price cap, see [18] for its effect with a reservation market.



Producers. We assume perfect competition on the supply side. Producers use a single technology to produce the good. It is characterized by a variable unitary cost c and a fixed unity investment cost r. We normalized the capacity level, so one unit of capacity allows to produce one unit of the good. The level of capacity installed is k.

Retailers. Retailers compete *à la Cournout* to resell the goods to final consumers but do not behave as an oligopsony in the wholesale market. The imperfect competition is modeled using a finite number of retailers n. We model the retail as perfectly competitive for some parts of the paper to keep the analysis tractable (i.e., $n \to \infty$). The use of a finite number is always explicitly indicated. We assume that retailers incur no cost when reselling from the wholesale market to the retail market apart from the wholesale price [15].

Demand. On the retail market, final consumers, are characterized by the following assumptions:

- They have the same individual uncertain demand with an aggregate demand $D(p,t)$, t being the state of the world. The demand uncertainty is characterized by a distribution function $f(t)$ and a cumulative distribution function $F(t)$. The inverse demand function is $p(q,t)$, with q the quantity sold on the retail market [16], such as $D(p(q,t),t) = q$. For convenience, we assume that $p^s(q,t)$ is the price on the wholesale market, and $p(q,t)$ is the price on the retail market.
- The demand function have the following properties [17]: $\forall t \in [0, +\infty)$ (i) $p_t(q,t) > 0$ (*states of the world are ordered*), (ii) $p_q(q,t) < 0$ (*decreasing price with respect to* q) (iii) $p_q(q,t) + qp_{qq} < 0$ (*decreasing marginal revenue with respect to* q) (iv) $p_t(q,t) + qp_{qt}(q,t) > 0$ (*increasing marginal revenue with respect to* t) and (v) $\lim_{q \to +\infty} p(q,t) < c$ (*prices can be below the marginal cost for some* t).

To ensure producers invest in capacities we need additional conditions: $p(0,t) > c + r \quad \forall t$ and $\lim_{q \to 0} p(q,t) < c$.

2.2. Optimal investment decision

We now describe the three stages in reverse order. We define the equilibrium in each stage and find the final optimal level of investment by backward induction.

Third stage - Retail market. We assume that symmetric retailers can act strategically in the retail market, and they take the wholesale price as given. The retailer's profit made on the retail market is: $\pi_i^r(t) = q_i(p(q,t) - p^s(q,t))$. The first-order condition gives

[15] Therefore, perfect competition implies that prices are strictly equal in the wholesale and the retail market

[16] We assume the quantity sold on the retail market is strictly the same quantity asked on the wholesale market as storage is not available.

[17] For most of the functions $f(x,y)$, $f_x(x,y) = \frac{\partial f}{\partial x}(x,y)$, $f_{xx}(x,y) = \frac{\partial^2 f}{\partial x^2}(x,y)$,$f_{xy}(x,y) = \frac{\partial^2 f}{\partial x \partial y}(x,y)$

the equality between the marginal revenue and the marginal cost. Thus, the inverse demand function of retailers on the wholesale market is a downward rotation at the intercept of the final consumer demand function :

$$p^s(q,t) = p(q,t) + \frac{q}{n} p_q(q,t)$$

When the retail market is perfectly competitive, we have straightforwardly : $p^s(q,t) = p(q,t)$

Second stage - Wholesale market. Producers know the final consumer demand at this stage, so the retailers' inverse demand function is certain. The price is determined by the investment level k chosen during the first stage. We assume perfectly competitive producers, so when k is not binding, the price is equal to the marginal cost c (*off-peak periods*). When k is binding, the price has to rise above marginal to ensure that supply equals demand (*on-peak periods*). We denote $t_0(k)$ the first state of the world when capacity is binding, that is, when the price at the capacity level is equal to the marginal cost: $p^s(k, t_0(k)) = c$. We also define $q_0(t)$ as the quantity bought by final consumers when the retail price is equal to the marginal cost, such as $p^s(q_0(t), t) = c$. During off-peak periods, when $t_0(k) \geq t$, the price on the wholesale market is the marginal cost c and the price on the retail market is equal to $p(q_0(t))$. During peak periods, when $t > t_0(k)$, the demand function determines the price with $p(k,t) + \frac{k}{n} p_q(k,t)$ the price on the wholesale market, and $p(k,t)$ the price on the retail market.

First stage - Investment decisions. At this stage, final consumer demand is unknown, so is the wholesale and retail price. We find the optimal first-best[18] investment level k^* by maximizing the social welfare given in the following equation:

$$W(k) = \int_0^{t_0(k)} \int_0^{q_0(t)} (p(q,t) - c) dq \; f(t) dt + \int_{t_0(k)}^{+\infty} \int_0^{k} (p(q,t) - c) dq \; f(t) dt - rk \tag{1}$$

2.3. Market equilibrium

Essential goods are characterized by inefficiencies that prevent the market investment from reaching the first-best economically efficient. Two main reasons why private investors do not provide sufficient capacities: (1) the revenue collected on the market is insufficient to cover their production and investment costs, (2) prices do not consider the positive externalities implied by their availability during high demand periods. For the first rationale, we derived the inefficiency that typically characterized essential goods such as electricity: the suboptimality of the wholesale price modeled via a price cap[19]. In some extensions of the following sections, we present two other rationales: the public-good nature of capacity during peak-demand level and a concentrated retail market represented via retailers' market power.

[18]We use the term first-best, socially optimal and welfare-maximizing interchangeably.

[19]This modeling approach can represent both an explicit price cap and an implicit one. In the latter case, political interventions due to the essential nature of the good can artificially alter the price. For instance, when the power system operator needs to carry out technical interventions to avoid system failures.

The market outcome in terms of investments is found by estimating the expected rent $\phi(k)$. This rent is the net marginal revenue made on the wholesale market when the capacity is binding, which is the difference between the wholesale price and the marginal production cost. The following equation gives the expected unitary rent:

$$\phi(k) = \int_{t_0(k)}^{+\infty} (p^s(k,t) - c)\ f(t)dt \tag{2}$$

The market investment level $\bar{k}$ of investment under imperfect competition framework is found by solving: $\phi(k) = r$. Under perfect competition on the retail market we have $k^* = \bar{k}$.

We implement a price cap denoted p^w. In order to create inefficiencies, the price cap must be binding for some states of the world, so it needs to be below the highest price during the highest demand period; $p^w < \lim_{t\to\infty} p^s(0,t)$. However, to allow for investment, we also need that the price cap to be above the total unitary cost: $p^w > r + c$ [17]. Following the previous analysis, we introduce a second threshold $t_0^w(k)$. It is the first state of the world when the price cap is binding, that is, when the price at the capacity level is equal to the price cap: $p^s(k, t_0^w(k)) = p^w$. We also define $q_0^w(t)$ the quantity bought by retailers (or consumers under perfect competition) when the price is equal to the marginal cost, such as $p^s(q_0^w(t), t) = p^w$. The price cap does not change the social welfare function equal to $W(k)$ as it only redistributes surpluses between consumers, producers, and retailers. We find the investment level quantity by estimating the expected rent. The following equation defines this rent. It is shared between the states of the world when prices are above the marginal cost and below the price cap and when prices are above the price cap. The conditions on p^w relatively to the marginal cost c ensure that $t_0^w(k) \geq t_0(k)$.

$$\phi^w(k) = \int_{t_0(k)}^{t_0^w(k)} (p^s(k,t) - c)\ f(t)dt + \int_{t_0^w(k)}^{+\infty} (p^w - c)\ f(t)dt \tag{3}$$

We find the level of capacity installed in the system given the price cap k^w by solving: $\phi^w(k) = r$. The following reservation shows that a price cap in the wholesale market lowers the investment level and increases inefficiency.

Lemma 1. *A binding price cap leads to a lower installed capacity compared with the optimal investment level given by the social welfare maximization:* $k^w \leq k^* \quad \forall p^w \in [c{+}r, \lim_{n\to\infty} p(0,t)[$. *The optimal capacity payment* $z^w(k)$ *is equal to the expected difference between what should have been the wholesale price and the price cap when it is binding :*

$$z^w(k) = \int_{t_0^w(k)}^{+\infty} (p^s(k,t) - p^w) f(t)dt \tag{4}$$

Proof. See Appendix □

2.4. The supply curve on reservation markets

We set in place a reservation market to encourage producers to increase their investment. On the demand side, we assume in this subsection that the demand function regime in the reservation market is unspecified. We denote its inverse demand function $p^c(k)$ with k the level of capacity offered in the reservation market, the demand function is D^c such as $D^c(p^c) = k$. $p^c(k)$ should be decreasing in k and defined as $p^c(.)$ is twice derivable.

[18] defines the equilibrium conditions for the reservation market and the supply side: first, there are no short sells, meaning that producers cannot sell more capacity than they own. Second, it is optimal for producers to offer all their capacities if the first condition holds. Finally, decision timing does not matter given our current setting: results still hold if the reservation market is set before or after the investment decision as long as it is before final consumers' demand is known.

We build the supply function based on the assumption that producers offer their marginal profit loss associated with the reservation market's participation. It is the common approach in the literature as it represents the cost of investing beyond the optimal capacity level. However, to our best knowledge, this is the first time a supply function in a reservation market is directly modeled using the benchmark framework. As we assume perfect competition in the wholesale market compared to [18], and [25], there is no marginal effect of capacity choices on the rent.[20] The full profit with a reservation market for a producer is: $\pi_i^s(k) = \phi(k)k_i - rk_i + p^c(k)k_i$. Under perfect competition, the first-order condition gives: $\phi(k) - r + p^c(k) = 0$. Therefore, the reservation market's supply function equals the marginal cost associated with the deviation from the market investment level k^0, which would have been made without the reservation market.

Proposition 1. *We denotes the supply function $X(k)$ and the inverse supply function $X^{-1}(p^c)$ such as $X^{-1}(X(k)) = k$. Following a marginalist approach, the supply function on the reservation market is defined as follow :*

$$X(k) = \begin{cases} 0 \quad if \quad k \leq k^0 \\ r - \phi(k) \quad k > k^0 \end{cases} \tag{5}$$

Proof. See Appendix □

Below k^0, the marginal cost is positive, and the supply is null. Indeed, as the wholesale market's profit function is concave, any marginal revenues on the left side of the optimum are above the marginal cost of r. The marginal revenue is below the marginal cost on the right side of the optimal investment level. Therefore any deviation to the right creates a positive opportunity cost.[21]

[20]Indeed, under perfect competition, the rent appears only when total capacity is constraining. Under imperfect competition on the supply side, the rent also exists due to market power and can appear before the total capacity is binding.

[21]Our approach to the supply function in the reservation market is similar to the theory of supply function equilibria where bidders offer a function such as each point on this function maximizes their profit/utility [10]. In our paper, the supply function in the reservation market is built such that each producer is indifferent between providing their investment market equilibrium or any investment on the curve in return for the corresponding capacity price.

III. CAPACITY COST ALLOCATION DESIGN

3.1. Exogenous ex-ante requirements

We start our analysis of the demand function specification by assuming a capacity demand in which a single entity determines the whole demand of capacity in the reservation market[22]. To do so, she needs to forecast the future expected demand of final consumers first, and then she builds the demand function in the reservation market. Finally, she transfers the purchasing cost to the retailers using an exogenous ratio or directly to consumers via a lump-sum tax. This assumption corresponds to the traditional approach used in the literature on the reservation market. We call this market design the exogenous ex-ante regime because (i) the allocation of capacity costs does not depend, for instance, on retailers' realized strategy but rather on exogenous factors such as their past market share (ii) the design does not depend on realized demand for the final good. In other words, this regime only describes the reservation markets' direct effect via the incentive to invest by the capacity price. There is no effect on the final demand because this remuneration is simply a surplus transfer from consumers to producers. This approach's result is that the capacity price equals the optimal payment, allowing to restore an optimal level of capacity when the vertical demand function for capacity is calibrated to k^* is and whatever the type of inefficiency is considered [23]. This result is described in the following Proposition and implies that the cost of a reservation market is strictly equal to the transfer necessary to restore the optimal level of capacity.

Proposition 2. *Assuming that*

1. *Producers do offer the marginal opportunity cost on a reservation market (see eq. 5), and*
2. *The demand function is designed such that the clearing quantity is equal to the optimal level of investment, and*
3. *The underinvestment is caused by the price cap*

Then the clearing price is always equal to the optimal payment needed to restore efficiency.

Proof. See Appendix □

This result highlights the discussion between implementing a price or a quantity instrument to resolve the market inefficiencies or constraints [12, 22]. We show in this Proposition that the outcome of the reservation market is strictly equivalent to a capacity price set by the regulator. Under this regime, the exogenous ex-ante approach is optimal because it gives the right investment level given the inefficiencies.

Our model can also provide some comparative statics on the capacity price given the specification of section 6. For instance, when only the price cap is considered in the model, the results are intuitive and in line with previous works, with the capacity price being always positively impacted by an increase of the demand intercept or by the product costs (variable and fixed), while the price cap has a negative effect.

[22]We do not make any assumptions on the identity or on the role of the capacity buyer in this section as it is outside the scope of this paper.

[23]We demonstrate in the Appendix that the Proposition can be expanded to the inefficiencies created by retailers' market power and by the public-good nature of capacities.

3.2. Endogenous ex-ante requirements

We now introduce an indirect effect of the reservation market demand function. It implies that the capacity prices marginally impact the final consumer demand via capacity allocation. To compare with the previous setting without this indirect effect, the first previous case can be understood as an increase of the fixed part in a two-part tariff (or a lump sum tax), while the second case in this subsection can be understood as an increase of the variable part (or a unitary tax). The central idea is to enhance our understanding of an optimal investment level given a market design by highlighting this effect. In this case, the first-best solution is endogenous to the market design implemented to reach this first-best. We use a similar approach of the impact of a tax on a partial equilibrium model to illustrate this endogenous effect on investment decisions. We demonstrate the existence of the indirect effect by repeating the steps of the previous model and using backward induction.

Fourth stage - Retail market. **Fourth stage - Retail market**. Let $p^c(k)$ be the capacity price adder for final consumers, identical to a unitary consumption tax. The final consumers demand function shifts to the left with its new value equal to: $p^s(q,t) - p^c(k)$. k is still the quantity bought on the reservation market by the entity at a price $p^c(k)$. We denote $t_1(k)$ and $q_1(k)$ the new thresholds for respectively the states of the world between on-peak/off-peak periods such that $p^s(k, t_1(k)) - p^c(k) = c$, and the quantity such as prices are equal to the marginal cost such that $p^s(q_1(t), t) - p^c(k) = c$.

Third stage - Wholesale market Third stage - Wholesale market While the demand is always lower or equal to the initial demand function, the impact on the expected social welfare is not trivial. The following Proposition summarizes the main insight and states that the new welfare function is always lower or equal to the exogenous case.

Proposition 3. *Allocating the capacity price as a unitary tax only affects the share between on-peak and off-peak periods and the surplus's size during off-peak periods. Namely, only the occurrence of the two periods $t_0(k)$ and the intersection between the demand function and the marginal cost $q_0(t)$ change, the welfare function becomes:*

$$W_1(k) = \int_0^{t_1(k)} \int_0^{q_1(t)} (p(q,t) - c) dq \ f(t) dt + \int_{t_1(k)}^{+\infty} \int_0^{k} (p(q,t) - c) dq \ f(t) dt - rk$$

Proof. See Appendix □

We can rewrite the equation by showing the initial welfare function without endogeneity: $W(k) - W_1(k) = \Delta W_1(k)$. With

$$\Delta W_1(k) = \int_0^{t_0(k)} \int_{q_1(t)}^{q_0(t)} (p(q,t) - c) dq \ f(t) dt + \int_{t_0(k)}^{t_1(k)} \int_{q_1(t)}^{k} (p(q,t) - c) dq \ f(t) dt > 0$$

The first part of $\Delta W_1(k)$ represents the loss when it is off-peak periods for both cases (indeed we have $t_0(k) \leq t_1(k)$ as lower demand always means a higher chance of being off-peak): the consumers fully support the loss as producers receive the marginal cost. The second part represents the loss when the capacity level is such that it is an off-peak period with the endogenous case and an on-peak for the other case. Therefore, the loss is shared between consumers and producers, the former sustaining a higher price and receiving a lower margin. Note that there is no loss when both cases are in peak periods, as the quantity on the market is strictly equal to the capacity installed. This last remark is particularly interesting because recovering the capacity cost allocation only during peak periods does not generate a deadweight loss. Considering the price cap does not change the previous overall Proposition, as the price cap is binding only when both periods are on-peak. That is when no loss is generated.

The following reservation concludes on the new optimal investment level given this endogenous regime. It has a strong implication as we state that this regime also modifies the objective for the regulator in terms of final investment level. Moreover we find that the endogenous regime is always worse than the exogenous regime regarding social welfare.

Lemma 2. *The new first-best solution in terms of investment level under the endogenous regime defined as $k_1^* = \max_k W_1(k)$ is always lower or equal to the first-best solution under the exogenous level k^*. In terms of welfare analysis, the social welfare at the optimal investment level $W_1(k_1^*)$ is always lower or equal to the social welfare at the optimal investment level under the exogenous regime $W(k^*)$.*

Proof. See Appendix □

The result stems from the analysis of the derivative of $\Delta W_1(k)$ with respect to the level of investment k, which is always positive. We now provide some comparative statistics on the difference between the two welfare function $\Delta W_1(k)$ for the model specification. By construction, the comparative statistic for the difference between the welfare has the same effect on the new first-best solution in terms of investment level. We demonstrate that the price cap and the demand intercept always have a negative effect on $\Delta W_1(k)$. An increase of the price cap reduces the need for the capacity payment through the negative value of the derivative of $q_1(t)$, with respect to the price cap, hence the endogenous effect off the cost allocation. At first sight, a higher demand through the demand intercept has an ambiguous effect on the difference. An increase in its value increases the need for capacity, which also increases the payment.

On the other hand, it has a decreasing effect that materializes through the derivative of $q_1(t)$. We find that the second effect always dominates the first, hence the net decreasing effect of the demand intercept on the difference. Finally, an increase in the producer's cost (fixed and variable) always increases the delta, making the investment less profitable.

We continue the analysis of the endogenous regime by defining the main equilibrium variables of the economic system.

Second stage - Investment decisions Producers make their investment decision based on the expected net revenue, which is composed of the expected rent and the capacity revenue. The net revenue is similar to the exogenous case, except for the new state of the world thresholds and the wholesale price.

$$\phi_1^w(k) = \int_{t_1(k)}^{t_1^w(k)} (p^s(k,t) - p^c(k) - c)\ f(t)dt + \int_{t_1^w(k)}^{+\infty} (p^w - c)\ f(t)dt + p^c(k) \tag{6}$$

First stage - Reservation market When a producer participates in the reservation market, he bids its marginal opportunity cost without the capacity revenue equal to $r - \phi_1^w(k)$. Therefore, following the previous stage, the equilibrium is defined with the following equality $X(k) = r - \phi_1^w(k)$. The following Proposition states how the remaining equilibria are found and underline the endogenous nature of this regime where the choice of capacity both on the supply and demand side of a reservation market has indirect effects. In words, the endogeneity of the regime also changes the bidding behavior in the reservation market compared to the exogenous case.

Proposition 4. *When the capacity price enters the final consumers demand as a marginal cost, solving the following equation allows to find the supply function $X_1(k)$ in the reservations market:*

$$X_1(k) : p^c(k) = r - \left(\int_{t_1(k)}^{t_1^w(k)} (p^s(k,t) - p^c(k) - c)\ f(t)dt + \int_{t_1^w(k)}^{+\infty} (p^w - c)\ f(t)dt \right) \tag{7}$$

Moreover, the supply function is always higher under the endogenous regime than under the exogenous regime.

Proof. See Appendix ☐

Regarding the reservation market, as the demand is lower under this regime, the opportunity cost associated with providing another capacity is higher. By extension, the supply function on the reservation market is also higher. Therefore, this regime has an ambiguous effect on the reservation market equilibrium: capacity prices can be higher or lower than exogenous capacity prices, even though the quantity is always lower. This implication can be summarized by defining the optimal payment to restore the optimal level with an endogenous price. Recall that with only a binding price cap, the optimal payment is the expected difference between what should have been the wholesale price and the price cap when the price cap is binding. The following Lemma defines the new optimal payment, and we compare it with the previous with the exogenous regime.

Lemma 3. *The optimal payment to restore efficiency when a price cap is binding is defined as follow:*

$$z_1(k) = \int_0^{t_1(k)} \frac{\partial q_1(t)}{\partial k} (p(q_1,t) - c) f(t)dt + \int_{t_1(k)}^{t_1^w(k)} r - \phi_1^w(k) f(t)dt + \int_{t_1^w(k)}^{+\infty} (p(k,t) - p^w) f(t)dt \tag{8}$$

Proof. See Appendix ☐

Compared to the initial payment $z^w(k)$, only the third part of the optimal payment is directly related to the expected difference between the optimal wholesale price and the price cap. In this regime, the magnitude of the loss is impacted via $t_1(k)$, which means fewer periods during which the price cap is binding. The first part represents the loss associated with the threshold shift for on-peak/off-peak periods. It is negative as $\frac{\partial q_1(t)}{\partial k} \leq 0$. When $q_1(t)$ decreases due to the capacity price, the rent decreases when the capacity starts binding at $t_1(k)$. The second term is positive, and it is directly related to the loss associated with the decrease of the demand during on-peak, which also decreases the rent for any state of the world between $t_1(k)$ and $t_1^w(k)$.

3.3. Extension - endogenous ex-ante regime with inefficient rationing

We now introduce the public-good nature of capacity during peak demand. First, we use this rationale to revise the comparison between the endogenous and exogenous ex-ante design for reservation markets. Then, it will also be used to analyze the capacity demand allocation regime in the next section.

At the price cap level, when it is binding, the price-elastic demand becomes inelastic[24]. Therefore, we face the same rationing problem as in the literature with limited production capacities and inelastic consumers (see for instance [14]). The absence of efficient discrimination between consumers with heterogeneous willingness to pay implies that investment availability is a public good when the price cap is binding. Therefore, it is underprovided by producers when they make their investment decisions. There is various way to describe the cost of involuntary rationing in the literature.[25] [14] shows that it depends if the rationing is anticipated or not. [17] finds that the effect of involuntary rationing can be different if it has an impact on the expected demand level. From a modeling perspective, [12] uses a general function $J(.)$ to represent this negative externality. The function depends on the delta between the quantity bought at a price equal to the price cap and the investment level. We note this cost $M(k)$, namely :

$$M(k) = \int_{t_0^w(k)}^{+\infty} J(\Delta_0 k) f(t) dt \tag{9}$$

with $\Delta_0 k$ a function of the difference between the installed capacity k and the quantity bought by retailers at the price cap q_0^w.[26] For instance, we can model rationing using a ratio $(1-h)$, which represents the share of consumers selected indifferently that is forced to stop consuming [16]. When rationing occurs, an optimal ratio h should be endogenously chosen such as we have $(1-h(t))q_0^w(t) = k$. In this case, consumers sustain an additional loss proportional to their initial surplus with efficient rationing, namely :

$$M(k) = \int_{t_0^w(k)}^{+\infty} (1-h(t)) \int_0^k (p(q,t) - p^w) dq f(t) dt \tag{10}$$

[24]The introduction of retailers into the model does not change the intuition. At a price p^w, the *Cournot* competition between the retailers pushes them to ask a quantity equal to $q_0^w(t)$.

[25]Note this is an additional cost compared to the loss of the surplus that we described previously with the price cap.

[26]Regarding the sign of the cost and its derivatives: $M_k(.) \leq 0 < M_{kk}(.)$. Finally, note also that $\frac{\partial \Delta_0 k}{\partial k} \leq 0$ as an increase of capacity lower the difference for a given value of q_0^w.

Following [12], we use the general notation $M(k) = \int_{t_0^w(k)}^{+\infty} J(\Delta_0 k) f(t)dt$. The following equation describes the new social welfare function :

$$W^{bo}(k) \quad = \quad W(k) - \int_{t_0^w(k)}^{+\infty} J(\Delta_0 k) f(t) dt \tag{11}$$

We denote k^{bo} the optimal level of investment when we maximize social welfare. Contrary to the imperfect competition in the retail market and the price cap, the social cost of rationing directly affects the social welfare function. The expected rent collected on the wholesale market by producers remains unchanged when we include inefficient rationing, which only affects consumers' welfare.[27]The following reservation shows that when we cannot efficiently ration final consumers when the price cap is binding it implies a higher inefficiency.

Lemma 4. *When the price cap induces involuntary rationing, the inefficiency is greater than with voluntary rationing. In other words, the delta between the optimal level of investment and the market outcome is greater with the first than with the latter:* $k^* - k^w < k^{bo} - k^w$. *The optimal capacity payment* $z^{bo}(k)$ *depends on the representation of the involuntary rationing social cost. Under our assumption, it is equal to the marginal value of an additional capacity for the system, which decreases the cost of involuntary rationing:*

$$z^{bo}(k) = -M_k(k) = -\int_{t_0^w(k)}^{+\infty} J_k(\Delta_0 k) + \frac{J(\Delta_0 k)}{\partial \Delta_0 k} \frac{\partial \Delta_0 k}{\partial k} \tag{12}$$

Proof. See Appendix □

Note that the two parts of the optimal payment can be opposite. For instance, given the representation in equation 10, the first part of equation 12 is positive and is relative to the hypothesis behind the cost representation: as the rationing cost is proportional to the consumer surplus, a higher surplus is generated by the investment indirectly leads to a higher cost of inefficient rationing. The second part of the equation is negative and stands for the initial reduction in rationing: a higher investment level reduces the need to implement inefficient rationing. The rest of the paper assumes that the second effect is always higher than the first one to keep $M_k(k) < 0$.

What happens when we include the rationing cost in the previous model? Given the rationing cost of the form $M(k)$. The new welfare function becomes:

$$W_1^{bo}(k) = W_1(k) - \int_{t_1^w(k)}^{\infty} J(\Delta_1 k) f(t) dt$$

With $\Delta_1 k$, the new function represents the difference between the quantity consumed at the price cap $q_1^w(k)$and the investment level. By construction, the main results for the rationing hold under the endogenous regime, especially in terms of the first-best solution. It implies a higher investment level and lower welfare than the case without

[27]Some authors do include those costs in the producer profit, using a fixed reputational cost [19] or a market shutdown during which producers also lose profit [9].

inefficient rationing. Note that the quantity at the price cap $q_0(k)$ does not depend on the level of investment under the exogenous regime. However, when the capacity price is allocated under the endogenous regime, the quantity $q_1(k)$ is indirectly affected by the investment level.

We focus our analysis on comparing the first-best solution under the exogenous regime and the first-best under the endogenous regime. The following Proposition shows that with inefficient rationing, the effect of an endogenous regime is ambiguous on the social welfare, which depends on the size of the negative effect previously described of the capacity price and the gains in terms of avoided rationing cost.

Proposition 5. *The delta in welfare with respect to the delta with exogenous price is:*

$$\Delta W_1^{bo}(k) = \Delta W_1(k) - \int_{t_0^w(k)}^{t_1^w(k)} J(\Delta_0 k) f(t)dt - \int_{t_1^w(k)}^{+\infty} (J(\Delta_0 k) - J(\Delta_1 k))f(t)dt \qquad (13)$$

With $W^{bo}(k) - W_1^{bo}(k) = \Delta W_1^{bo}(k)$. The endogenous reservation market provide a higher social welfare compared to the exogenous case if and only if :

$$\int_0^{t_0(k)} \int_{q_1(t)}^{q_0(t)} (p(q,t) - c)dq \ f(t)dt + \int_{t_0(k)}^{t_1(k)} \int_{q_1(t)}^{k} (p(q,t) - c)dq \ f(t)dt < \int_{t_0^w(k)}^{t_1^w(k)} J(\Delta_0 k) + \int_{t_1^w(k)}^{+\infty} (J(\Delta_0 k) - J(\Delta_1 k))f(t)dt$$

Proof. See Appendix □

The new surplus when the optimal level of investment is reached is higher than with an exogenous capacity price when the endogenization effect on the rationing cost (i.e., the two negative parts of $\Delta W_1^{bo}(k)$, as we always have $J(\Delta_1 k) \leq J(\Delta_0 k)$) is higher than on the initial welfare (i.e., $\Delta W_1(k)$). Recall that $\Delta W_1(k)$ represents the loss associated with the effect of the capacity price on consumer demand. The first negative part in $\Delta W_1^{bo}(k)$ stands for the lower occurrence of on-peak periods due to the lower demand, which reduces the rationing cost. The second negative part represents the loss avoided because a lower consumer demand implies a lower consumer surplus, hence a lower cost than the exogenous regime.

While there is an ambiguity regarding the new value of the welfare function, we find that the new optimal investment level is always lower than without inefficient rationing. That is we always have $\frac{\partial \Delta W_1^{bo}(k)}{\partial k} \geq 0$. Regarding the social welfare at the optimal investment level, it depends on the ambiguous effects described in Proposition 5.

Regarding some comparative statistics, the effect of the model variables can be ambiguous when we consider inefficient rationing. We start with the optimal payment to the producer defined in equation 12 which is equal to the absolute value of the marginal surplus loss associated with the inefficient rationing. Recall that it is composed of two distinct parts (i) a positive value for the direct effect of the rationing on consumer welfare and (ii) a negative value for the indirect effect because the loss is based on consumer welfare. We find that the price cap and the demand intercept always have opposite signs between parts (i) and (ii). For instance, an increase in the price cap continuously decreases the positive value associated with the rationing, but because it also decreases consumer welfare when the price cap binds, it decreases the negative value, hence the ambiguous effect.

IV. CAPACITY DEMAND ALLOCATION DESIGN

4.1. Retailers realized market share allocation

4.1.1. Market equilibrium

The previous approach is based on ex-ante requirements, meaning that the quantity allocated to the retailers (or directly to the consumers) is independent of the demand's current realization. This section analyzes a specific implementation of the reservation market demand where the capacity allocation depends on the retailers' realized quantity sold to the final consumers. The main difference with previous ex-ante requirements lies in the retailer profit function, where the capacity cost allocation act as an additional marginal cost. We study how this new cost adder for retailers modifies the previous results in light of different degrees of competition in the retail market. Under competition *à la Cournot*, we find that having different numbers of retailers have a direct effect on the cost allocation sustained by final consumers. Therefore, the degree of competition determines the sign and the magnitude of the different outcomes described in the ex-ante regime.

The first implication of ex-post requirements concerns the last stage when the retail market clears. We rewrite the retailers' profit function by including an endogenous ratio in the retailer profit function, as shown in the following equation. Contrary to the previous section, we do not need to assume any tariff hypothesis for the capacity cost allocation as it directly affects retailers' profit at the margin. We focus our analysis on symmetric equilibrium.

$$\pi_i^r(q_i, k) = q_i(p(q) - p^s) - p^c(k)k\frac{q_i}{q_i + q_{-i}}$$

We find the equilibrium using the first-order condition. The main results are stated in the following Proposition:

Lemma 5. *When the retail equilibrium exists it is always unique and stable. The condition for the existence of an equilibrium is given by the following condition:*

$$-kp^c(k)\left(\frac{n-2}{n}\right)\frac{1}{q^2} \leq p_q(q) + \frac{q}{n}p_{qq}(q) - p^s$$

Proof. See Appendix □

Using the first-order conditions and the symmetry between the retailers, the Cournot equilibrium in the retail markets allows to define the endogenous retailer demand function in the wholesale market:

$$\tilde{p_n}(q) = p(q) + \frac{q}{n}p_q(q) - p^c(k)k\frac{1}{q}\frac{n-1}{n}$$

The equilibrium in this market is similar to the case with the ex-ante requirement. Therefore, we can define the periodic threshold between on-peak/off-peak/binding price cap periods. We respectively denote them $t_n(k)$ and $t_n^w(k)$, with also $q_n(t)$ and $q_n^w(t)$ the quantity at which the wholesale price $\tilde{p_n}$ is equal to respectively the marginal cost and the price cap. The wholesale market equilibrium does not necessarily exist as the demand function is not well defined on wholesale quantity and prices. To see this, recall the existing condition for the retail equilibrium. With a high value of q, some retail equilibrium does not exist. Therefore the retail threshold also applies to the wholesale market.

Given the investment and the variable revenue from the reservation market defined as follow, we find the marginal opportunity cost that defines the supply function in the reservation market and hence the final capacity price:

$$p^c(k) = r - \left(\int_{t_n(k)}^{t_n^w(k)} (p(k,t) - c - p^c(k)\frac{n-1}{n})\ f(t)dt + \int_{t_n^w(k)}^{+\infty} (p^w - c)\ f(t)dt \right)$$

Note the difference with the ex-ante requirement in the first part of the integrals, where the capacity cost adders are dependent on n. The following Proposition summarizes the main effect of an ex-post allocation:

Proposition 6. *When allocating the reservation market cost is based on retailers' realized market share, it generates a lower depreciating effect on the demand. The ex-post regime provides an intermediate indirect effect between an ex-ante regime with exogenous capacity price and an ex-ante regime with endogenous capacity.*

Proof. See Appendix □

To illustrate this Proposition, the capacity cost adder when $n = 2$ is equal to half of the cost adder of equation 7, and it is increasing with n. When $n \to +\infty$, the capacity cost is entirely allocated to the consumer, mimicking the ex-ante exogenous equilibrium. This Proposition states that increasing competition in the retail market increases the burden of consumers' capacity prices. Hence, the negative effect observes in the regime with endogenous capacity prices is now shared between retailers and consumers.

4.1.2. Extension - Reservation market and retail market structure

The previous results of our paper on cost allocation are independent of the market structure in the retail market. However, in this extension and using our analytical framework, we show that a change of market structure can have ambiguous effects both in the reservation market outcome and with respect to the determination of the optimal capacity level. This aspect of essential goods has been relatively less studied than the supply side. Therefore, in this extension, we describe the effect of different market structures on the optimal and market investment level and with and without inefficient rationing.

The market outcome under imperfect competition is the same as previously described: $\bar{k}$. The following reservation states that market power in the retail market lowers the investment level beyond market power's direct effect. The market investment level is different from the optimal investment level even when we maximize the welfare function given the market power in the retail market.[28]

Lemma 6. *Imperfect competition in the retail market leads to a lower installed capacity compared with the optimal investment level given by the social welfare maximization:* $\bar{k} \leq k^* \quad \forall n \in [2, +\infty)$*. The optimal capacity payment* $\bar{z}$ *is equal to the expected markup of retailers in the retail market:*

$$z_n(k) = \int_{t_0(k)}^{+\infty} \frac{-k}{n} p_q(k,t) f(t) dt \tag{14}$$

Proof. See Appendix □

We turn now to the analysis of the effect of the market structure on the model. First, suppose there is no inefficient rationing. In that case, an increase in the number of retailers has two effects of opposite sign : (i) an increase of the social welfare, which is the common effect of higher competition in a canonical model *à la Cournot* (ii) a decrease in the social welfare due to the lowering of the consumption associated with a higher capacity cost allocated to the consumers. Both effects are defined in the equation 15 with the derivative of the welfare function with respect to the number of retailers. Under our framework and without the ex-post regime, the demand is only affected by the degree of competition during off-peak periods when $t \leq t(k)$,[29] and the derivative is always positive. However, the capacity cost allocation also has a depreciating effect on the demand level at the marginal cost (recall equation 8), hence the ambiguous sign of $\frac{\partial q(k)}{\partial n}$.

$$\frac{\partial W(k)}{\partial n} = \int_0^{t(k)} \frac{\partial q(k)}{\partial n} (p(q(k),t) - c) f(t) dt \tag{15}$$

A similar but more complex effect arises when considering inefficient rationing. Paradoxically, lower competition in the retail market makes the rationing occurrences less likely, which lowers the surplus loss. Note that as the surplus loss is only associated with on-peak periods, the lowering of competition in the retail market does not affect the size of the rationing cost but only its occurrence. We provide in the following equation the marginal effect of an increase of the number of retailers on the social welfare when the rationing cost is based on a ratio as described in equation 10, as in the previous equation, the sign of $\frac{\partial q^w(k)}{\partial n}$ is also ambiguous:

[28]This result has important regulatory implications. Indeed, we state that the welfare-maximizing investment level given the imperfect competition in the retail market, is different from the welfare-maximizing investment level in a perfectly competitive market. Therefore, the will to necessarily reach a competitive investment level could cause significant harm, potentially greater than the welfare loss generated by the inefficient market equilibrium.

[29]We drop the subscript for simplicity.

$$\frac{\partial W(k)}{\partial n} = \int_0^{t(k)} \frac{\partial q(k)}{\partial n}(p(q,t)-c) - \int_{t^w(k)}^{+\infty} (\frac{\partial q^w(k)}{\partial n}\frac{k}{q^w(k)^2}\int_0^k (p(q,t)-p^w)dq)f(t)dt \quad (16)$$

It is sufficient that the signs of the two derivatives be different so that the effect of imperfect competition on social welfare is clear. For instance if the increase of n increases the surplus during the off-peak periods ($\frac{\partial q(k)}{\partial n} > 0$) but decreases the rationing occurrence due to the capacity cost adder ($\frac{\partial q^w(k)}{\partial n} < 0$) then the social welfare increases with n. On the other hand, if n has the same effect on both quantity thresholds $q(k)$ and $q^w(k)$, then the role of imperfect competition on the outcome is ambiguous. For instance, if the increase of competition increases the welfare during off-peak periods ($\frac{\partial q(k)}{\partial n} > 0$) but also increases the demand at the price cap despite the capacity cost adder ($\frac{\partial q^w(k)}{\partial n} > 0$), then the effect is unknown and depend on the relative size of the welfare gain during off-peak periods and the loss occurring during rationing.

4.2. Retailers individual allocation

We provide in this section an analysis of a new regime for the demand function in reservation markets. Each retailer must purchase their capacities in the reservation market in this implementation. An entity only monitors the level of capacities and compares it to each retailer's consumption. We show that this market design can only bring additional welfare due to specific incentives when inefficient rationing exists. [30]

One of the critical features of this regime concerns the case when a retailer is in negative deviation, i.e., has sold more on the retail market than he has bought capacity in the reservation market. In this case, he suffers a penalty, which results in a payment from the retailer to the entity[31] by a unitary amount of S, with $S \geq 0$ being an administratively fixed value[32]. When every retailer has bought enough capacity, we are under the no penalty case, and no other mechanism is implemented [33]. The price in the reservation market is still noted as $p^c(k)$, and the individual quantity contracted by a retailer i is k_i. Under this regime, we process the following to describe the equilibrium: (i) We analyze the outcome with no uncertainty. We use a simple game theoretical framework to describe a game's equilibrium where agents must sequentially choose a fixed capacity first and then compete *à la Cournot* on a second game. (ii) Then, we extend our analysis to the initial framework developed in the model with investment and reservation decisions made before the demand is known.

[30] A comparison between the case with and without inefficient rationing for this regime would be relevant with additional hypotheses. For instance, the quality and quantity of information detained by private agents such as retailers in future states of the world can be seen as better or worse than the information detained by a single entity. While those specifications are outside the scope of this work, they should call for a deeper application of the model presented in this paper.

[31] Which acts like the government is the model.

[32] We use in this paper a linear form of the penalty system, but some implementation can encompass nonlinearities depending on the effect desired for the penalty system.

[33] Some remuneration mechanism can exist so to reward retailers who have provided additional capacity, but as we focus on symmetric equilibrium, they do not play a role in the outcome.

4.2.1. Market equilibrium without uncertainty

To provide the intuition for the general case with uncertainty, we start by describing the equilibrium in the case the demand is known when the retailer has to choose the level of capacity to be bought in the reservation market. We show that it is a dominant strategy to integrate the penalty value in the profit function as a marginal cost up to a point where it is optimal to stop buying capacity and sustain the penalty.

Given the assumptions and notation, the retailers' profit function during the last stage in the retail market is:

$$\pi_i^r(q_i, k_i) = q_i(p(q) - p^s) - p^c(k)k_i + \begin{cases} +0 & \text{if} \quad q_i \leq k_i \\ -S(k_i - q_i) & \text{if} \quad q_i > k_i \end{cases}$$

The equilibrium on the retail market is given by the first order condition:

$$p(q) + q_i p_q(q) - p^s = \begin{cases} 0 & \text{if} \quad \forall i \quad q_i \leq k_i \\ S & \text{if} \quad q_i > k_i \end{cases}$$

As the equation shows, the market design's penalty system implies different discontinuous retailers' reaction functions. The first case will be called the penalty case, while the second one the noremuneration case. It depends on the capacities bought in the reservation market by the retailer and his competitor and their strategies on the retail market. The penalty system always induces a lower reaction function, whatever is the sign of the retailer's deviation from its position in the reservation market. It is straightforward for the penalty system: a marginal increase in the retail market's quantity increases the marginal cost via the penalty. The capacities bought in the reservation market do not directly affect the reaction function's value, but it determines the form of the reaction function between the penalty/noremuneration cases. We summarize in the following Proposition the central insight of this game equilibrium.

Lemma 7. *The set of dominant strategies in the retail market is:*

$$\begin{cases} [q^p, q^r] & \textit{if} \quad p^c(k) \leq S \\ \{0,]q^p, q^r]\} & \textit{if} \quad p^c(k) > S \\ \forall q \in]q^p q^r[& q \textit{ is a solution of} \quad p(q) + q_i p_q(q) - p^s(q) - p^c(q) = 0 \end{cases}$$

With q^r the equilibrium quantity offered on the retail market when the retailers are in the noremuneration case. This value is given by the solution of $p(q) + q_i p_q(q) - p^s(q) = 0$. q^p is the equilibrium quantity offered on the retail market when the retailers are in the penalty. This value is given by the solution of $p(q) + q_i p_q(q) - p^s(q) - S = 0$.

Proof. See Appendix □

The optimal quantity on the retail and reservation markets depends on the difference between the penalty value S and the capacity price $p^c(k)$. We assumed a linear penalty system, so if the penalty is lower than the capacity price $S \leq p^c(k)$, then it is a dominant strategy (strict if $S < p^c(k)$) to buy no capacity and sustain a penalty on all the quantity sold on the retail market. Indeed, with strict inequality, the profit function is a decreasing non-concave function with respect to k_i. On the other hand, when the penalty is higher than the price, the profit function is an increasing non-concave function with respect to k_i. The dominant strategy is to buy the same amount of capacity as the quantity q^p, which corresponds to the retail market's corresponding equilibrium in the penalty case.

In this regime, the capacity price is passed on to the consumers, which increases the marginal cost equal to the capacity price. In turn, it reduces the demand function on the wholesale market. This result follows that the demand function in the reservation market strictly mimics an equilibrium quantity sold on the retail market when the marginal cost increases without a reservation market. When the cap S is not binding, those equilibria are given by the set $[q^p, q^r]$.

Under this no uncertainty assumption and relying on the set of dominant strategies, it is straightforward to extend the analysis to the optimal demand of capacity. Using the results of 7, the demand function on the reservation market is a linear decreasing function caped above at the penalty value while intersecting the null capacity price at exactly q^r which is the equilibrium without the reservation market. Any value between corresponds to the equilibrium given by the solution of $p(q)+q_ip_q(q)-p^s(q)-p^c(q)=0$. To say it differently, given a capacity price, retailers always buy the same amount of capacity that their equilibrium in the retail market unless the price is above the penalty value

On the supply side, following the marginalist approach, the supply function starts to be non-null and positive at the same value q^r. Therefore the equilibrium is always a null price with a level of investment strictly equal to the regime without any reservation market. Such counter-intuitive result stems from the fact that the reservation market is only a burden for retailers under no uncertainty and without any other model refinement, which does not encourage them to buy more capacities.

4.2.2. Social welfare under uncertain demand

In reality, retailers buy capacity before the demand is known. Hence, when building their demand function in the reservation market, they must consider a range of possible outcomes relative to the production and demand levels. They need to expect the occurrence of off-peak and on-peak periods and the rationing case when the price cap is binding. The last situation is critical as it determines the magnitude of the penalty payment bared by retailers. We start by providing the link between the previous analysis with no uncertainty and the general model. First, q^r is the value $q_0^w(t)$ as it is the *Cournot* outcome in the retail market[34]. Then, denote $q_d^w(t)$ the value equal to q_p, that is the threshold at which $p^s(q)-S$ is equal to the price cap p^w. In other terms, this is the *Cournot* equilibrium when the marginal cost for the retailers is equal to p^w+S. It is similar to assuming a decrease in demand when retailers pass the penalty cost onto

[34] In fact, every value between $q_0(t)$ and $q_0^w(t)$ are conceptually equal to q^r, but we focus on the threshold case between the periods where the price cap is binding and not binding.

the consumers. Finally, denote $t_d^w(k)$ the state of the world when the price cap starts binding under the case the demand of final consumers is equal to $p^s(q) - S$.

Following the no uncertainty case, the set $[q_d^w(k), q_0^w(k)]$ defines the set of dominant strategies in the retail market for a wholesale price equal to p^w and any capacity level between 0 and the value q_0^w. To see this, let us distinguish three cases depending on the value of the installed capacity.

- (Case 1) When $k > q_0^w$, the price cap is never binding, and the outcome is strictly identical as a regime without a reservation market.
- (Case 2) For a value of k between $q_d^w(k)$ and $q_0^w(k)$, we observe a paradoxical outcome; rationing should have occurred as soon as k is below $q_d^w(k)$ without a penalty. It implies that retailers sustain the penalty, which is then passed to consumers as a marginal cost, which lowers their demand. However, rationing is not happening, which contradicts the demand's decrease due to the penalty. Therefore, as in the no uncertainty case, retailers follow the level of investment. To do so, they increase the price of their consumers by a unitary amount of $T(k)$ so that at any state of the world between $t_0^w(k)$ and $t_d^w(k)$ the demand is equal to the capacity k,[35] that is we have $p^s(q) - T(k) = p^w$.
- (Case 3) Finally, k is below $q_d^w(k)$, it is now optimal for the retailers to keep their strategy at q_0^w as it is the *Cournot* equilibrium given the penalty value S.[36].

The distinction between the two different cases ((2) and (3)) is crucial as in the former case (2) there is no inefficient rationing as the quantity sold by the retailers is equal k, while in the latter case (3) inefficient rationing necessary occurs because retailers ask for $q_d^w(k)$ which is above k.

We can deduce the wholesale market outcomes using this set of outcomes on the retail market. The following Proposition describes how the new welfare function encompasses the previous implications regarding the retailer strategy when we assume that the initial capacity price is not passed onto the consumers via a unitary payment. [37]

Proposition 7. *The regime ambiguously impacts the welfare function, with a negative distributional effect due to the penalty system and a positive effect due to reducing rationing costs. The rest of the welfare function equals the no reservation market regime.*

$$W_d(k) = W^{bo}(k) \quad - \int_{t_d^w(k)}^{+\infty} S(q_d^w(t) + k) + \int_{t_0^w(k)}^{t_d^w(k)} J(\Delta_0 k) + \int_{t_d^w(k)}^{+\infty} J(\Delta_0 k) - J(\Delta_d k) f(t) dt \tag{17}$$

With $\Delta_d k$ the new difference between the installed capacity and the quantity bought by the retailers $q_d^w(k)$

Proof. See Appendix □

[35]We could also assume the reverse mechanism where retailers pay consumers $T(k)$ to reduce the demand in order to avoid the penalty.

[36]Similarly, it is identical to the case where the benefice to pays consumers to lower their demand is above the cost generated by the penalty.

[37]This additional feature of this regime would cause a similar effect as the outcome presented in the endogenous ex-ante regime.

The intuition behind the Proposition is as follows. For the second negative part of the welfare function, the penalty S entirely affects the retailers' profit margin. It implies that both the retailers and the consumers suffered a surplus loss due to the demand reduction. Added to the penalty cost borne by the retailer of $S(q_d^w(k) - k)$, it gives the net loss for the welfare function. The rationing cost reduction is directly linked to the shift from $q_0^w(k)$ to $q_d^w(k)$. Indeed, as stated before, any capacity $k \in]q_d^w(k) q_0^w(k)]$ implies equality between the quantity sold by the retailers and the capacity level k. Therefore, there is a net increase in welfare. Note that the last term of the welfare function is always positive as $q_d^w(k) \leq q_0^w(k)$. It is similar to the effect observed in the endogenous ex-ante capacity price, but it is limited to the rationing period in this regime. Therefore, this regime avoids the negative effect of lower demand when there is no rationing and optimal wholesale price.

4.2.3. Investment decision and reservation market equilibrium with uncertainty

Assuming that the capacity price is not included in the variable price in the retail market, we state that the supply function is not impacted by the market design. Indeed, the previous analysis shows that the demand function on the wholesale market is impacted only when the price cap starts binding under no reservation market regime (i.e., when $k \leq q_0^w(k)$). In case (2), the adjusted demand follows the capacity level, while in case (3), inefficient rationing exists, but it still implies that demand equals capacity. Therefore, the market design does not impact the occurrence and the magnitude of the rent. The supply function is given by the equation 5 and the undistorted $\phi(k)$. Under this configuration, we define the aggregated retailer profit function when the demand is uncertain in the following equation:

$$
\begin{aligned}
\pi^r(k,t) = \quad & \int_0^{t_0(k)} -\frac{q_0(t)^2}{n} p_q(q_0(t),t) f(t)dt + \int_{t_0(k)}^{t_0^w(k)} -\frac{k^2}{n} p_q(k,t) f(t)dt && \text{Case (1)} \\
& + \int_{t_0^w(k)}^{t_d^w(k)} k(p(k,t) - p^w - T(k,t)) f(t)dt && \text{Case (2)} \\
& + \int_{t_d^w(k)}^{+\infty} k(p(k,t) - p^w - S) f(t)dt - \int_{t_d^w(k)}^{+\infty} S(q_0^w - S) f(t)dt && \text{Case (3)} \\
& - p^c(k) k_i
\end{aligned}
$$

The expected profit function comprises three main parts related to different values of k given a demand level (or a different level of demand given a value of k). The two first terms are the same with and without a reservation market as the price cap is not binding. The retail price rises while the wholesale price is fixed and equal to the price cap for the second term. As explained in the previous analysis, between the two states of the world $t_0^w(k)$ and $t_d^w(k)$, the demand decreases due to the retailers' actions to avoid paying the penalty. It is materialized by the transfer $T(k,t)$ [38]. When it is not prof-

[38] Note that if we assume that retailers pay the consumers to reduce their consumption, only the sign changes.

itable to reduce the demand given the penalty (when $t_d^w(k)$ is reached), then the new demand is given by $q_d^w(k)$, the retailer profit is in the fourth term, inefficient rationing is implemented, and the retailers pay the penalty in the fifth term. The last term is the capacity cost due to the retailer's obligation to buy their capacities. Given this expected profit, we can define the marginal value of a capacity for the retailer, which serves as the retailers' willingness to pay for an additional capacity.

Proposition 8. *Under the market design, the retailers aggregated demand function in the reservation market is equal to the marginal value of an additional capacity for their profit function.*

$$
\begin{aligned}
D^c(k) = \quad & -\int_{t_0}^{t_0^w(k)} \left(\frac{2}{n}kp_q(k,t) + \frac{k^2}{n}p_{qq}(k,t)\right) f(t)dt \quad + \\
& -\int_{t_0^w(k)}^{t_d^w(k)} \left(\frac{2}{n}kp_q(k,t) + \frac{k^2}{n}p_{qq}(k,t)\right) f(t)dt \\
& + \int_{t_d^w(k)}^{+\infty} \left(p(k,t) - p^w + kp_q(k,t)\right) f(t)dt
\end{aligned}
$$

Proof. See Appendix □

The value of a capacity for a retailer depends on the effect a marginal variation brings to its profit function. When the level of capacity increases, we can distinguish three effects:

- (i) The decrease in the cost of the penalty during the case (3)
- (ii) An increase of the oligopolistic profit during on-peak periods when the quantity offered is equal to the capacity; and
- (iii) A change in the occurrence of off-peak/on-peak and price cap-binding periods.

The effects (i) and (ii) do not directly appear in the demand function as they are entirely offset. Indeed, for the penalty effect, while the increase of capacity lowers the marginal cost of penalty by S, the retailer gains at the same time a marginal profit equal to $p(k,t) + kp_q(k,t) - p^w - S$. Therefore, the marginal effect of the penalty is null, and the effect during the case (3) is limited to an increase of the marginal profit, as illustrated in the third term of the demand function. The effect (iii) of the occurrence of the different periods cancel each other out because when $q_0^w(k)$ is reached, the value of $T(k,t)$ is null. While, when $q_d^w(k)$ is reached the value of $T(k,t)$ is equal to the penalty. Finally, note that the first and second parts, which represent the net gain from an increase of capacity for retailers, are always positive as the marginal revenue is always decreasing following the initial assumptions regarding $p(.)$ with respect to q (i.e., $p_q(.) + qp_{qq}(.) < 0$).

The last part is ambiguous and depends on the value of k relatively to the monopoly outcome in this model. Indeed, note that $p(k,t) + kp_q(k,t) - p^w$ is the first-order condition of the retailer monopoly profit function with marginal cost equal to the price cap. Given its concavity, any value of k below the monopoly quantity implies a positive third part in the demand function, while a value of k above implies a negative third part. It is a sufficient condition that the *Cournot* outcome $q_0^w(k)$ bounds from below the monopoly outcome for the third part be always positive. In other terms, this condition holds whenever the penalty value is sufficiently high or with a low number of retailers. This last part of the demand shows that retailers are willing to pay for capacity if it allows them to reach (in expectation) the monopoly quantity when the price cap is binding. In other words, they bid the expected marginal revenue that makes them indifferent between being at the monopoly outcome or not buying more capacity, given that this third part is positive.

Finally, the general equilibrium in the system is found by solving $D^c(k) = p^c(k)$. As the supply function is entirely independent of the outcome, the general outcome analysis is identical to the reservation market demand function analysis. For instance, it is sufficient to state that an increase in the demand function due to the increase of the penalty also increases the investment level. On the other hand, it can also indirectly affect the optimal level of investment given by maximizing the welfare function. Therefore, the comparative statistic for this regime's welfare effect boils down to the same approach described in the previous ex-ante case.

V. CLOSED-FORM SOLUTION AND APPLICATION TO THE ELECTRICITY MARKET

5.1. Data for calibration to French capacity market

We now provide a closed-form solution of the model with a numerical illustration in this section. We focus on the electricity system and consider the reservation market as a capacity market, where producers can sell the availability of their power plant in return for a capacity price. Following our general model, this transaction forces them to have a given level of investment in the wholesale electricity market. This section will provide insight into the effect of each market design on an actual economic system and allow us to give general policy recommendations for the power sector.

The main specification is the final consumers demand function, which is assumed linear, and where the uncertainty comes from the intercept of the linear function. We define the inverse demand function as follow :

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

Where $a(t)$ is the uncertain intercept such as $a(t) = a_0 - a_1 e^{-t}$. We assume that t follows an exponential distribution which is characteristic of electricity consumption : $f(t) = e^{-t}$.

Regarding the supply side of the system, we assume the marginal technology is a gas power plant (CCGT) with variable cost (c) equal to 50.24 € /MWh, which includes fuel and carbon costs. The fixed cost of the technology (r) is equal to 26.70€ /MWh, which includes investment and fixed operation costs. Those data come from the last report for the International Energy Agency and represents an average investment in this technology for the OECD countries [13]

For the demand side, the calibration and the model results are highly sensitive to the assumption regarding the value of the demand function. While there are many studies on the value of the actual demand function of electricity, we still lack a proper view on the future characteristics of the demand when all consumers are price reactive, as we assume in this paper. Therefore, we provide in this numerical illustration some scenario sensitivity. We use three different demand function scenarios : (i) Low elasticity such as the Value of Lost Load (VoLL)[39] is equal to 20000 € /MWh (ii) Medium elasticity such as the VoLL is equal to 3000€ /MWh (iii) High elasticity with a demand characterization from [18]. In the first two cases, we proceed as follows to find the coefficients a_0, a_1, and b. First, we find the maximum and minimum value of the hourly quantity consumed on the electricity market. Given the French data, we find an amplitude of 30 000 MW and 102 000 MW for 2012-2020. Next, we assume that a price cap exists in the wholesale market, and if its value is equal to the VoLL, then the market equilibrium should be equal to the first-best solution in terms of investments level. Given this condition and the magnitude of the demand, we can find the value of the coefficient a_0, a_1, and b. To give an order of magnitude, the demand elasticity with respect to a price of 100€ /MWh is equal to respectively: -0.0009, -0.04, and -0.53. The first scenario uses a value assumed in our current system with most consumers highly inelastic. The second scenario uses the actual price cap on the French wholesale market. We use this value as it is never reached and given that the elasticity using this value is close to what has been empirically found for a small share of price reactive consumers [2]. The last scenario allows assuming a hypothetical system's future state where consumers are fully priced reactive and highly elastic.

Finally, regarding the price cap value, we use it as a control variable to assess the effect of different market designs; the range used in this paper is from 200 € /MWh to 3000 € /MWh. It allows us to represent both the current explicit price cap existing in the wholesale market (at 3000 € /MWh) and other indirect inefficiencies that can have the same effect on the expected rent for the producers. For the inefficient rationing, we use the specification defined in equation 10.

The results of the numerical simulation are as follow. First, we provide the results with respect to the first-best and market equilibrium, and we recall the main effect of this canonical framework. Then, we discuss implementing a capacity market in the system with no indirect effect (exogenous ex-ante design). Next, we show how allocating the capacity price as a tax can modify the equilibrium (endogenous ex-ante design). Finally, we discuss the calibration of the retailers individual regime, and we compare its outcome with the ex-ante design. An analysis of the imperfect competition in the retail market is provided in the Appendix section, where we study its effect for the canonical model, the retailers realized market share regime, and the individual allocation one.

5.2. First-best and market equilibrium investment

Following the model specification, the welfare function is concave, ensuring a first-best solution exists. We start with the first-best analysis of the simulation without inefficient rationing, and we provide in table 1 some critical indicators for different scenarios. First, we express optimal investment levels in absolute terms and for the maximum demand level. Intuitively and in line with the theory, when comparing the different demand scenarios, we find that a higher elasticity leads to a lower optimal

[39] This concept, widely used in the electricity sector, relates to the case of inelastic demand. The VoLL is the maximum price at which inelastic final consumers are willing to buy electricity; above it, the demand is null. Therefore, in this numerical illustration, we expand the link between this concept and the calibration of the elastic demand.

investment, ranging from a delta of 0.19% of the maximal demand to 4%. However, counter-intuitively, at first sight, the equilibrium price in the wholesale and retail market is strictly the same between the different scenarios. Indeed in the expectation, it should always cover both the fixed and variable costs. The difference has to be found in the magnitude of the average offpeak and on-peak prices and the occurrence of those periods. When the elasticity increases, the average peak price significantly decreases, from 10025 €/MWh to 456€/MWh, with a respective maximum value equal to the VoLL in the two first scenarios and to 863 €/MWh in the last scenario. While the average peak prices are inversely correlated with the elasticity, we find that peaks periods occur more frequently under high elasticity. Therefore, investment is more binding with a high elasticity system but generates lower prices.

When we introduce a price cap in the wholesale market, the market equilibrium can be impacted. The effect depends on the value of the cap and the elasticity, which has the opposite effect. Namely, a higher elasticity given a price cap tends to lower the inefficiency, while a more constraining price cap given an elasticity level tends to increase it. The inefficiency seems to be stable for the different scenarios, with a delta in installed capacity ranging from 0.65 % (compared to 0.19%) to 11.9% (compared to the 4%) of the maximum demand. As expected, the social and consumer welfare is always lower than the absence of a price cap. In terms of price equilibrium, we still have the equality of the average price with the sum of the production costs. However, due to the price cap, the expected peak prices are lower, and due to the effect on the investment, peaks periods are occurring more frequently.

The existence of inefficient rationing in the system significantly affects the first-best solution. In the low elasticity case, the optimal level of investment needs to be significantly close to the maximum level of demand (0.03%). We observe the same effect for the high elasticity with an optimal value of 0.71% compared to the 4% without inefficient rationing. Under the market equilibrium, this inefficiency does not impact the market outcome as it only affects the consumer surplus. Following our assumption, the occurrence of inefficient rationing is the same as when the price cap binds on the wholesale market. Under the low elasticity scenario, rationing occurs only 0.89% of the time. This value increases when the elasticity is lower given a price cap, and of course, increases when a price cap is lower. For instance, we observe a significant increase between the two price caps under the high elasticity scenario.

From a welfare perspective, the cost associated with a price cap and the inefficient rationing crucially depends on the assumption on the elasticity and the price cap level. The cost can vary between almost 2 million € under the medium elasticity scenario with only the direct effect of the price cap, up to 150 million € when inefficient rationing is considered. The consumer surplus is lower when the elasticity is higher, hence a lower cost. However, we still find significant figures ranging from 600k€ to 9 million € without and with inefficient rationing.

Demand scenario		Low elasticity		Medium elasticity	High elasticity	
Price cap	***€/MWh***	***3000***	***1000***	***1000***	***500***	***200***
Investment level (FB)	*GW*	101.8	101.8	100.7	97.8	97.8
% of maximum demand		0.19%	0.19%	1.26%	4.07%	4.07%
Average price	*€/MWh*	76.94	76.94	76.94	76.94	76.94
Average peak price	*€/MWh*	10025.12	10025.12	1525.12	456.45	456.45
Maximum price	*€/MWh*	20000.00	20000.00	3000.00	862.66	862.66
% of offpeak	%	1.00	1.00	0.98	0.93	0.93
% of peak	%	0.00	0.00	0.02	0.07	0.07
Investment level (MM & IR)	*GW*	101.3	100.0	99.8	97.0	89.9
% of maximum demand (MM & IR)	%	0.65%	1.99%	2.17%	4.85%	11.90%
Average peak price (MM & IR)	*€MWh*	2936.89	997.85	910.79	394.49	195.08
% of offpeak (MM & IR)	%	0.99	0.97	0.97	0.92	0.82
% of peak (MM & IR)	%	0.01	0.03	0.03	0.08	0.18
change in soc welfare (MM)	*k€*	-56.4	-1949.9	-22.0	-4.1	-616.2
change in soc welfare (IR)	*k€*	-14918.8	-150440.5	-2633.8	-521.8	-9498.8

Table 1: Simulation outcomes for the first-best and market equilibrium (FB: first-best , MM : price cap inefficiency, IR : inefficient rationing

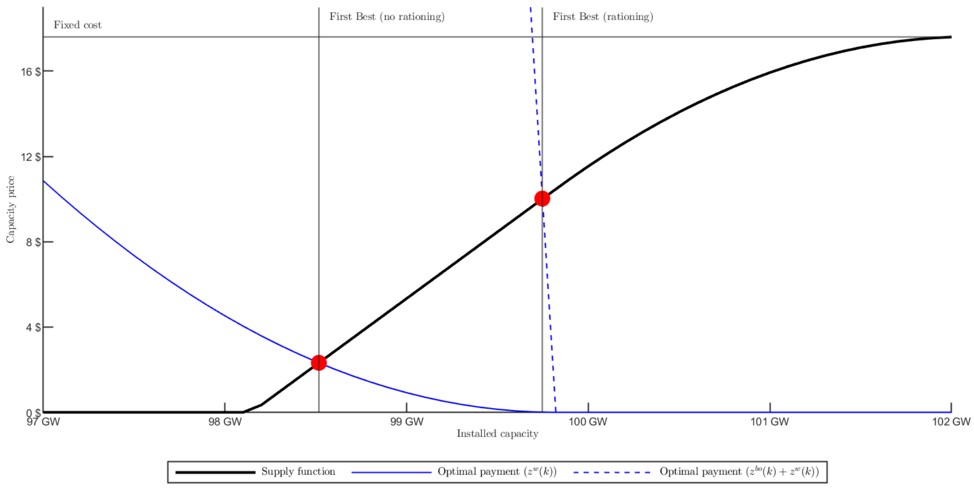


Figure 1: Supply function on a reservation market with the optimal payments ($p^w = 500$)

5.3. Price equilibrium with the canonical capacity market

One of the critical components of the framework is the supply function built on the reservation market. Using a marginal approach, we assume that if the producer were to bids their investment availability, the supply function should represent the marginal opportunity cost of providing a capacity at the margin as in equation 5. We show in figure 1 the supply function (thick black line) on a reservation market with perfect competition on the retail market and with a price cap at 500 $\$$. The supply function is null below a critical investment value equal to the market investment equilibrium. Note that the supply function converges toward the fixed cost as the additional rent collected on the wholesale market becomes null at a certain level of investment. On the same figure, we have also represented the two optimal payments necessary to restore the first-best solution, both when the price cap does not generate inefficient rationing and when it does (with perfect competition in the retail market). As proven in the Proposition 2, the intersection between the supply curve and the optimal payments coincides with the first-best solution in terms of investment (represented with the two vertical lines). Consequently, the capacity price at the equilibrium under the exogenous regime is located at the two red dots depending on the inefficiency.

Following this framework, we derive some comparative statistics of the capacity price with respect to the price cap level on the wholesale market with perfect competition in the retail market. Figure 2 shows the different value for the capacity price given the price cap when we both consider only the direct price cap effect and then add the inefficient rationing. Note the respective convexity and concavity of the capacity price. This specific curvature stems from the value of the first-best solution with respect to the price. Indeed, when only the price cap is considered, the first-best solution is constant with respect to the price cap. On the other hand, when inefficient rationing is considered, the first-best decreases with respect to the price cap.

We now compare the outcome of the simulation for the different scenarios, which are provided in table 2. As described in the analytical section, the canonical framework

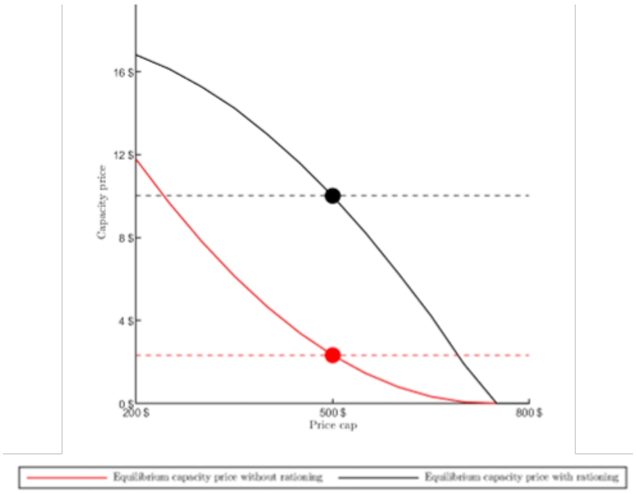


Figure 2: Capacity prices with respect to the price cap

does not generate any indirect effect on the equilibrium compared to the first solution regarding investment level and welfare, both for the consumer and the social surplus. However, the capacity market impacts prices in the wholesale market and, by extension, in the retail market. First, we observe a decrease in the average wholesale price. Therefore, the model aligns with the theory that a capacity market depresses electricity prices compared to the optimal equilibrium. By construction, the difference with the first-best equilibrium is strictly equal to the capacity price. Hence, we should observe an inverse relationship between capacity prices and wholesale prices from a policy perspective. However, this is not strictly the case for on-peak prices. While the price cap always implies a lower expected price than the first-best outcome, the average on-peak price is also lower than the market equilibrium with no capacity market. We observe, for instance, that in the low elasticity scenario, a difference of 154 € /MWh (a decrease of 6%) in the average peak price compared to a capacity price of 19.39 € /MW. In the high elasticity scenario, a similar effect is observed, with a decrease in the peak price of 8.88€ /MWh (5%) compared to a capacity price of 17.76 € /MW. Note that when we consider inefficient rationing in the simulation, we should expect the effect of a capacity market on the equilibrium prices to be stronger. Indeed, the new optimal investment level is always higher with rationing. Therefore the capacity price should also be higher.

Finally, we observe a conflicting result for the consumer under our framework. Indeed, when assuming an inefficiency from the price cap, the consumer surplus is strictly the same as in the optimal situation. Indeed, the capacity price simulates the missing transfer that should have happened under the optimal framework, which has a net null effect on the consumer. On the other hand, we find in all scenarios a negative effect for the consumer, compared to the social welfare[40]. Such results stem from the fact that the consumers directly remunerate producers for their availability to avoid inefficient rationing under the capacity market. To say it differently, this loss can be understood from a Pigouvian view as the necessary cost of internalization of the positive externality generated by producers, net of the benefits brought by such internalization. Under our simulation, it ranges from 2 565 € to 52 656.36 € .

[40]Consumers always gain from a capacity market compared to the market equilibrium.

Demand scenario		**Low elasticity**		**Medium elasticity**	**High elasticity**	
Price cap	***€/MWh***	***3000***	***1000***	***1000***	***500***	***200***
Consummer welfare (MM)	*M€*	247814.51	247814.51	5418.39	464.15	464.15
Consummer welfare (IR)	*M€*	247814.51	247814.51	5418.38	464.13	464.10
Rationing costs without capacity market	*k€*	14862.42	148490.66	2611.82	517.79	8882.59
Rationing costs with capacity market	*€*	655.18	461.72	150.98	1067.83	2222.23
Average price (MM)	*€/MWh*	57.55	52.72	64.67	71.62	59.18
Average price (IR)	*€/MWh*	51.00	50.35	53.19	59.26	51.55
Average peak price (MM)	*€/MWh*	2781.93	977.39	847.10	375.51	186.20
Average peak price (IR)	*€/MWh*	1715.68	657.57	540.19	286.06	138.61
Consummer gains vs inefficiency (MM)	*k€*	56.42	1949.85	21.96	4.06	616.18
Consummer gains vs inefficiency (IR)	*k€*	14915.62	150436.96	2621.57	503.00	9443.90
Consummer loss vs first-best (IR)	*€*	-2565.15	-3091.25	-12055.04	-17773.15	-52646.36
Capacity price (MM)	€ /MW	19.39	24.22	12.27	5.32	17.76
Capacity price (IR)	€ /MW	25.94	26.59	23.75	17.68	25.39
Capacity cost (MM)	k€	1973.85	2465.61	1236.25	520.58	1738.10
Capacity cost (IR)	k€	2645.36	2712.13	2412.79	1762.62	2571.79

Table 2: Simulation outcomes for the exogenous capacity market regime (FB: first-best , MM : price cap inefficiency, IR : inefficient rationing

5.4. Welfare effect of an endogenous ex-ante reservation market

We now study the effect of allocating the capacity price directly on the consumers via an increase in the unitary electricity price. We have demonstrated in Proposition 3 and Lemma 2 that the effect is negative when we only consider the price cap as a source of inefficiency. On the other hand, when rationing is considered, then an endogenous capacity price can bring benefits in terms of social welfare at the first-best investment level, as shown in Proposition 5. We use the methodology in the Proposition 4 to derive the equilibrium for both cases.

We find two notable results with our numerical simulation : (i) the results are in line with the analytical model, however (ii) the indirect effect does not bring significant changes compared to the exogenous case. We provide in table 3 the main results for the simulation. First, note that we do find a lower first-best investment level given this endogenous framework. However, the change is only significant with the high elasticity case with a maximum deviation of 0.14% from the initial optimal investment level, and only if we consider inefficient rationing.

We also find the opposite effect on the social welfare described in the analytical section from a welfare perspective. Without inefficient rationing, the loss from allocating the cost onto the consumers ranges from a minimum value of -1.8 € under the low elasticity case to -821.54 € for the high elasticity case. Moreover, we find that the cost associated with this market design is higher when the elasticity is higher and the price cap is lower. On the other hand, implementing this market design for the capacity market always brings more social welfare for all scenarios when inefficient rationing is considered. The gains range from 58.60 € under the medium elasticity case to 1741.02 € under the high elasticity scenario.

Interestingly, when allocating the capacity price under this framework, the consumer always loses in both cases compared to the previous market design. This loss can range from 982€ under the low elasticity scenario to 63239€ under the high elasticity scenario. Three different effects can explain the change in the consumer surplus between the exogenous and endogenous market design : (i) the cost associated with the capacity market (ii) the avoid loss due to the inefficient rationing (iii) a direct welfare loss due to the increase in the electricity price. As expected and similar to the social welfare effect, allocating the capacity price onto the consumer via a unitary tax always generates the opposite effect (ii) and (iii). The first is always positive, and the second is always negative. With the simulation, we find that the gain in rationing costs is lower with higher elasticity, and the loss in terms of welfare is higher under higher elasticity. Therefore, the negative effect (iii) tends to be higher when the elasticity is also higher. On the other hand, we find the sign of the first effect to be ambiguous in our numerical illustration. Given our assumption and changing the market design from the exogenous to the endogenous case, we should expect a loss in terms of higher capacity market cost for consumers when the elasticity is low and a net gain when the elasticity is high. It is explained by the ambiguous effect of changing the market design on the capacity price. We find that under the low elasticity scenarios, capacity prices are higher, and lower under the high elasticity scenario. [41]

[41] Recall investment level is always lower under the endogenous market design.

Demand scenario		**Low elasticity**		**Medium elasticity**	**High elasticity**	
Price cap	***€/MWh***	***3000***	***1000***	***1000***	***500***	***200***
Change in social welfare (Max)	€	654.60	461.12	150.95	1067.83	2222.22
Change in social welfare (Min)	€	-217.80	-149.26	-28.82	-134.54	-294.35
Change in social welfare (Equal)	€	0.22	0.03	0.00	-0.04	0.05
Gains in avoided rationing costs	€	655.18	461.72	150.98	1067.83	2222.23
Penalty value (Max)	€	11006000	11108000	355000	27800	28120
Penalty value (Min)	€	19328000	19760000	1430600	149740	150000
Penalty value (Equal)	€	13154000	14054000	1008400	97020	94280
Penalty ratio (Max)	%	100.000%	100.000%	100.000%	100.000%	100.000%
Penalty ratio (Min)	%	99.998%	99.998%	99.998%	99.986%	99.980%
Penalty ratio (Equal)	%	99.997%	99.998%	99.997%	99.979%	99.970%

Table 3: Simulation outcomes for the endogenous capacity market regime (FB: first-best , MM : price cap inefficiency, IR : inefficient rationing

5.5. Retailers individual allocation

5.5.1. Calibrating the model

We conclude the analysis of our numerical application with the retailers individual regime. The critical concept to understand this regime is how retailers formulate the demand function on the reservation market and transmit incentives to consumers because of the penalty system. Given the initial assumptions on the investment technology, we can find the equilibria on the reservation market in terms of price and quantity for different penalty levels, price cap, and market structure. One of the main challenges this numerical illustration raises is that the regime never restores efficiency unless retailers do not entirely pass the penalty cost onto the final consumers. More specifically, we introduce in the general framework a ratio such as the final consumer demand function is only partially impacted when the price cap starts to bind and rationing occurs. [42] In other words, retailers need to sustain an additional cost with the penalty system to induce sufficient incentives to buy capacities.

Our first takeaway is that this penalty ratio needs to be very high for the regime to provide enough incentives to retailers to buy capacities on the reservation market. To say it differently, the marginal value of a capacity is relatively low for retailers whenever they can entirely pass the penalty cost onto consumers. In this case, as shown in the general framework and under perfect competition, the marginal value for a retailer relies on a monopoly quantity. In this numerical application, such quantity, when computed, is also very low, which explains the findings. However, with a high α, the penalty system induces a new value which stems from a tradeoff between buying more capacities and sustaining the penalty cost. It translates retailers' demand function to the right and allows an intersection with producers' supply function at a price above 0. We use the numerical illustration to compute the couple value between the penalty value and the penalty ratio that brings the same equilibrium as the first-best optimal value with inefficient rationing. Under the three demand scenarios, the couple translates into a convex function with a minimum value above 99.8%. To say it differently, when designing the penalty system under this regime, policymakers need to impose to retailers to sustain 99.8% of the penalty system when inefficient rationing occurs. Otherwise, retailers do not buy sufficient investment in the capacity market. To give some order of magnitude, if the policy market aims at ensuring the investment level equal to the first-best without inefficient rationing, which is always lower than the previous one, then the ratio is still significantly high with a minimum value of 99%. Given this model calibration, we now turn to the comparison with the ex-ante regimes.

5.5.2. Ex-ante vs ex-post regimes

This section discusses under which conditions an retailers individual market design can generate more or less efficiency than an ex-ante market design. Note that we do not have any possibility for regulatory errors in our model. Therefore, our results rely solely on the set of effects described in the theoretical framework.

[42]Initially, the final consumer demand when the penalty is allocated is $p(q,t) - S$, with a ratio $\alpha \in [0,1]$ it boils down to the same model except that the final demand is $p(q,t) - (1 - \alpha S)$. If $\alpha = 1$, then retailers fully sustain the cost of the penalty system.

This approach allows us to derive the following implication: the regulator can choose the correct combination set of penalty value, and penalty ratio such as the equilibrium in the capacity market given the supply function and the aggregate demand function of retailers is equal to the first-best investment level with inefficient rationing [43] However, we find that while it leads to the same equilibrium in terms of investment level, the indirect effect of each set can be significantly different. More precisely, we can differentiate between the set of penalty values and ratio, which gives higher social welfare at the retailers individual regime equilibrium than the first-best social welfare in the ex-ante regime equilibrium. We provide in table 4 the results of the simulation for the former design under the three demand scenarios. The *Max* case represent the highest possible welfare given the set of penalty value, the *Min* the minimum case, and *Equal* the penalty value such as the equilibrium under the retailers individual regime is the same as under the ex-ante regime

We first find that it is always possible to have a set of values that gives higher welfare under the retailers individual market design. For the three scenarios, those values vary significantly, which also illustrates the sensitivity of the retailers individual market design concerning the demand assumptions when choosing the penalty system. Namely, the penalty range from 11 million € /MWh for the low elasticity case to only 28 120.00€ /MWh under the high elasticity to provide higher social welfare than the ex-ante design. Those values are well above the actual penalty in France, for instance. Under perfect competition in the retailer market, those findings stem from the fact that when assessing the marginal value of an additional capacity, retailers only assess the benefits and the costs during rationing periods when the price cap binds (see Proposition 8). Therefore, their demand function results from an expected assumption on the rationing periods. Recall that this period should only be occurring at a maximum of 0.266% under a first-best optimum.

Regarding the origin of the gains in terms of social welfare, which vary between 150 € to 2222 €, we find that it comes mainly from the avoided rationing cost when the price cap binds. More precisely, our numerical results show that the retailers individual market design never generates a rationing period compared to the ex-ante market design, which only minimizes those costs. To say it differently, the indirect effect underlines in the theoretical framework for the retailers individual regime fulfills its role by reducing the occurrence of such periods. Indeed, recall that the penalty system incites retailers to reduce their consumption whenever the investment level is between the realized demand at the price cap and the realized demand adjusted by the penalty at the price cap, hence avoiding inefficient rationing.

Finally, we find that the gain from a retailers individual market design is higher under the high elasticity case. This result is particularly relevant with respect to a policy perspective as it demonstrates that such a regime seems to be more beneficial when final consumers have reached a sufficient level for their demand elasticity with respect to price.

[43] As in the theoretical framework, we do not discuss here a retailers individual market design without inefficient rationing.

Demand scenario		**Low elasticity**		**Medium elasticity**	**High elasticity**	
Price cap	***€/MWh***	***3000***	***1000***	***1000***	***500***	***200***
Change in social welfare (Max)	€	654.60	461.12	150.95	1067.83	2222.22
Change in social welfare (Min)	€	-217.80	-149.26	-28.82	-134.54	-294.35
Change in social welfare (Equal)	€	0.22	0.03	0.00	-0.04	0.05
Gains in avoided rationing costs	€	655.18	461.72	150.98	1067.83	2222.23
Penalty value (Max)	€	11006000	11108000	355000	27800	28120
Penalty value (Min)	€	19328000	19760000	1430600	149740	150000
Penalty value (Equal)	€	13154000	14054000	1008400	97020	94280
Penalty ratio (Max)	%	100.000%	100.000%	100.000%	100.000%	100.000%
Penalty ratio (Min)	%	99.998%	99.998%	99.998%	99.986%	99.980%
Penalty ratio (Equal)	%	99.997%	99.998%	99.997%	99.979%	99.970%

Table 4: Simulation outcomes for the retailers individual capacity market regime (FB: first-best , MM : price cap inefficiency, IR : inefficient rationing

VI. CONCLUSION AND DISCUSSION

This paper built a tractable framework to analyze multiple markets' interdependences for an essential good prone to underinvestment. We showed how the investment decisions are affected by those markets, their structure (such as the degree of competition), and, most importantly, their design. Our case study is the reservation markets that were implemented to encourage producers to invest by providing an additional remuneration. Most of the literature on reservation markets has focused on the supply side, where producers offer their availability on future transaction periods on the wholesale market. Therefore, the demand side has been overlooked, even though some system efficiency effects are well known. Current implementations show many options regarding the demand side's design on reservation markets, as consumers do not have proper incentives to buy capacities. Using our framework, we compare multiple market designs and their implications. The first set of regimes is based on differentiating the capacity cost allocation. The second set of regimes is represented by how the design can account for current demand realization. We underline the different parameters that can significantly affect the outcomes of a reservation market on investment decisions. The single entity's quantity can significantly affect prices and quantities on the three markets and redistribution welfare between agents for the ex-ante case. One of the advantages of this framework relies on the possible extensions that we can implement, besides providing a simple but complete vision. The rest of the section discusses two issues that could be addressed in future research using this framework.

First, we initially assumed that consumers were fully reactive to retail prices. Such assumptions do not describe the reality yet as illustrated in the electricity system, as most small final consumers such as households are still under fixed-price contracts. The study of final consumers' heterogeneity and its implications for investment decisions in the power system is an emerging trend. [16] and [18] provide a relevant model close to the one presented in this paper. They show the effects of having those two types of consumers with different investment decisions and a reservation market. However, the author does not compare demand design options for reservation markets or consider retailers. Therefore, implementing this new extension in our model could shed light on the issue associated with power systems' investment decisions. It could also have a significant impact on retailers individual market design options. Indeed, if we consider that some consumers cannot react to price, but retailers are still forced to cover their consumption, the demand function's formation in the reservation market is significantly impacted.

Finally, we assume that future final consumer demand is commonly shared between the different agents. A single entity, potentially regulated, and retailers could access a different quantity and quality of information. For instance, we can assume that the entity has only a global vision of the future demand, and hence she is prone to make a more significant error forecast than retailers. On the other hand, retailers have private access to more precise information on their client portfolio while sharing common information on the world's future global states. Therefore, introducing these private/common elements in our model could shed new light on the effect of reservation markets and their market design options. Finally, in some current implementations, the entity based its global forecast on retailers' information. Consequently, the comparison between the various regimes' cases could be analyzed using game theory and signaling.

References

[1] Daron Acemoglu, Kostas Bimpikis, and Asuman Ozdaglar. "Price and capacity competition". In: *Games and Economic Behavior* 66.1 (2009), pp. 1–26.

[2] Hunt Allcott. "Real-time pricing and electricity market design". In: *NBER Working paper* (2012).

[3] David P Brown. "Capacity payment mechanisms and investment incentives in restructured electricity markets". In: *Energy Economics* 74 (2018), pp. 131–142.

[4] Andreas Bublitz et al. "A survey on electricity market design: Insights from theory and real-world implementations of capacity remuneration mechanisms". In: *Energy Economics* 80 (2019), pp. 1059–1078.

[5] James Bushnell, Michaela Flagg, and Erin Mansur. "Capacity Markets at a Crossroads". In: *Energy Institute at Haas* (2017). . ISSN: 1556-3693. DOI: `10.1097/JFN.0b013e31827a1e9d`. URL: `http://ei.haas.berkeley.edu`.

[6] Anna Creti and Natalia Fabra. "Supply security and short-run capacity markets for electricity". In: *Energy Economics* 29.2 (2007), pp. 259–276. ISSN: 01409883. DOI: `10.1016/j.eneco.2006.04.007`.

[7] Gauthier De Maere d'Aertrycke, Andreas Ehrenmann, and Yves Smeers. "Investment with incomplete markets for risk: The need for long-term contracts". In: *Energy Policy* 105 (2017), pp. 571–583.

[8] Gerard Doorman et al. "Capacity mechanisms: needs, solutions and state of affairs". In: *CIGRÉ, Paris* (2016).

[9] Natalia Fabra. "A primer on capacity mechanisms". In: *Energy Economics* 75 (2018), pp. 323–335.

[10] Richard J. Green and David M. Newbery. "Competition in the British Electricity Spot Market". In: *Journal of Political Economy* 100.5 (1992), pp. 929–953. ISSN: 00223808, 1537534X. URL: `http://www.jstor.org/stable/2138629`.

[11] Benjamin F. Hobbs et al. "A dynamic analysis of a demand curve-based capacity market proposal: The PJM reliability pricing model". In: *IEEE Transactions on Power Systems* 22.1 (2007), pp. 3–14. ISSN: 08858950. DOI: `10.1109/TPWRS.2006.887954`.

[12] Pär Holmberg and Robert A Ritz. "Optimal Capacity Mechanisms for Competitive Electricity Markets". In: *The Energy Journal* 41.Special Issue (2020).

[13] IEA. *Projected Costs of Generating Electricity*. Tech. rep. IEA, 2020.

[14] Paul Joskow and Jean Tirole. "Reliability and competitive electricity markets". In: *The Rand Journal of Economics* 38.1 (2007), pp. 60–84.

[15] Jan Horst Keppler, Simon Quemin, and Marcelo Saguan. "Why the Sustainable Provision of Low-Carbon Electricity Needs Hybrid Markets: The Conceptual Basics". In: (2021).

[16] Thomas-Olivier Léautier. "Is Mandating". In: *The Energy Journal* 35.4 (2014).

[17] Thomas-Olivier Leautier. "On the long-term impact price caps: Investment, uncertainty, imperfect competition, and rationing". In: *International Journal of Industrial Organization* 61 (2018), pp. 53–95.

[18] Thomas-Olivier Léautier. "The visible hand: ensuring optimal investment in electric power generation". In: *The Energy Journal* 37.2 (2016).

[19] Gerard Llobet and Jorge Padilla. "Conventional Power Plants in Liberalized Electricity Markets with Renewable Entry". In: *The Energy Journal* 39.3 (2018).

[20] David Newbery. "Missing money and missing markets: Reliability, capacity auctions and interconnectors". In: *Energy Policy* 94 (2016), pp. 401–410.

[21] Charlotte Scouflaire. "Capacity remuneration mechanisms: analytical assessment of contemporary experiences and lessons for the future design of electricity markets". PhD thesis. PSL - Paris Dauphine University, 2019.

[22] Martin L Weitzman. "Prices vs. quantities". In: *The review of economic studies* 41.4 (1974), pp. 477–491.

[23] Frank A Wolak. *Long-Term Resource Adequacy in Wholesale Electricity Markets with Significant Intermittent Renewables*. Working Paper 29033. National Bureau of Economic Research, 2021. DOI: `10.3386/w29033`. URL: `http://www.nber.org/papers/w29033`.

[24] Haomin Yan. "Auctions with quantity externalities and endogenous supply". In: *International Journal of Industrial Organization* 71 (2020), p. 102638.

[25] Gregor Zöttl. "On optimal scarcity prices". In: *International Journal of Industrial Organization* 29.5 (2011), pp. 589–605.

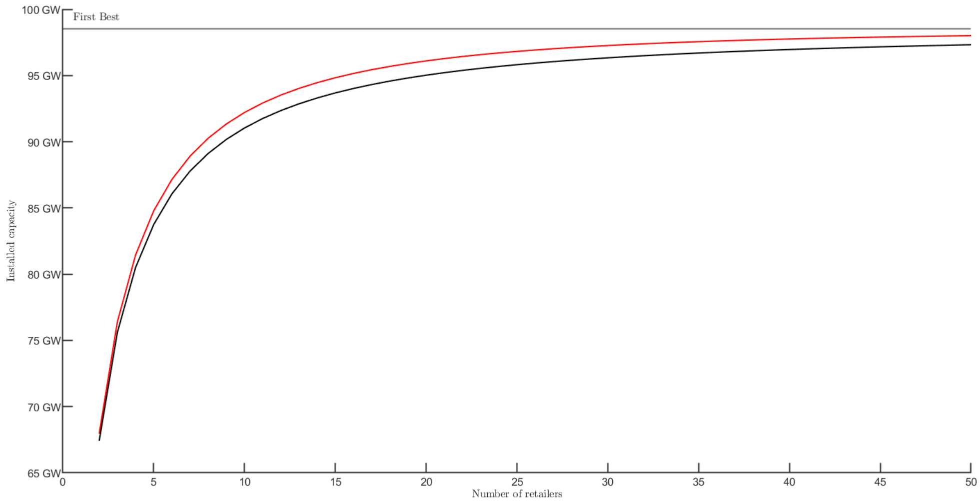


Figure A.1: First-best investment level under retail Cournot competition between an ex-ante and ex-post market design

APPENDICES

APPENDIX - SECURING INVESTMENT FOR ESSENTIAL GOODS. HOW TO DESIGN DEMAND FUNCTIONS IN RESERVATION MARKETS?

A. DISCUSSION ON THE RETAIL MARKET STRUCTURE

1.1. Retail market structure and ex-post requirements

Proposition 6 states that allocating the capacity price onto the retailers and their realized market share provide an intermediary indirect effect on the demand side. Namely, with a less concentrated retail market, retailers tend to pass more the capacity cost onto the consumers, mimicking the ex-ante regime with an endogenous price. In contrast, we state that when the number of retailers tends to converge towards the same outcome as the ex-ante regime, the difference with an ex-ante regime with exogenous capacity price at the same degree of retail competition might differ with respect to the market structure. This relation is illustrated in figure A.1 where we draw the evolution of the first-best investment level for different numbers of retailers. The red line stands for the exogenous case, while the black line considers the realized market share in the outcome. While the difference in terms of investment is relatively small for a very concentrated retail market, it tends to widen as the number of retailers increases. Therefore, the point of convergence is different, as shown in the figure.

One key element raised in this paper is the link between different market structures and the evolution of social welfare. In the analysis of the welfare function with respect to the market structure, we found that the latter can have an ambiguous effect. Indeed, recall that increasing the number of retailers increases welfare during off-peak periods and increases the capacity price allocated towards the consumers, hence decreasing welfare.

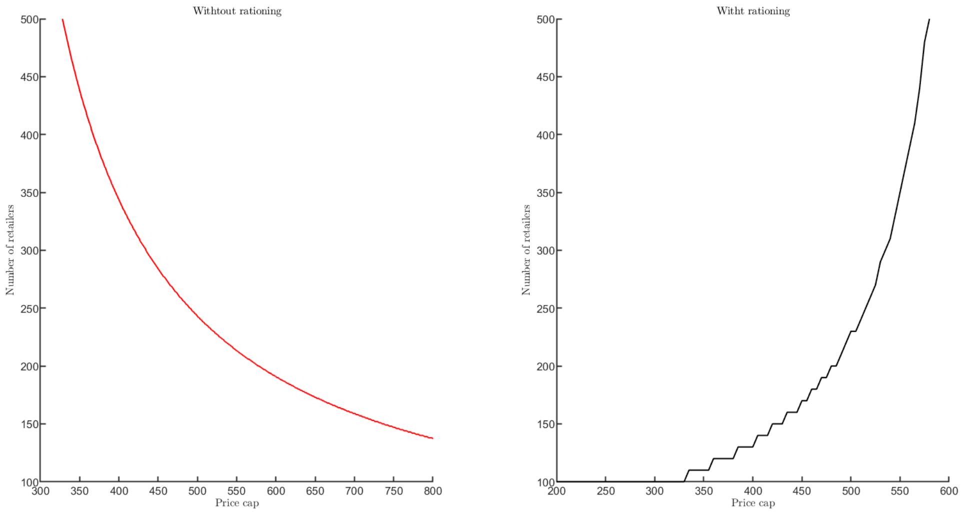


Figure A.2: Threshold price cap and the market structure couples for the sign of the welfare function derivative with respect to the number of retailers

A similar effect can be found when we take into account inefficient rationing. In figure A.2, we show the couple price cap - number of retailers for which the effect is null. Namely, below the line on the first sub-figure, an increase in the number of retailers always increases welfare, while above the line, it continuously decreases welfare. This threshold is decreasing with respect to the price cap. Note the inverse relationship when we consider inefficient rationing (with positive values below the line). In the first case, only one effect is studied, while the second to opposite effect are considered. As shown in the previous figures, the cost of inefficient rationing is high in social welfare, which explained the outweigh of the first term by the second in the equation.

B. PROOF

2.1. Proof of Lemma 1

Similar to the proof of Lemma 1. Given the price cap, we define the inframarginal rent:

$$\phi^w(k) = \int_{t_0(k)}^{t_0^w(k)} (p(k,t) - c) f(t) dt + \int_{t_0^w(k)}^{+\infty} (p^w - c) f(t) dt \tag{B.1}$$

The market investment level is equal to the optimal investment level only if the solution of the following equality has the same solution as with the first-order condition of the welfare function:

$$\phi^w(k) + z^w(k) = r \tag{B.2}$$

Which is the case if and only if :

$$z^w(k) = \phi^w(k) - \int_{t_0(k)}^{+\infty} \int_0^k (p(q,t) - c) dq \; f(t) dt \tag{B.3}$$

Therefore the optimal payment is equal to the expected lost revenue between the optimal price and the price cap:

$$z^w(k) = \int_{t_0^w(k)}^{+\infty} (p(k,t) - p^w) f(t) dt \tag{B.4}$$

2.2. Proof of Proposition 1

The proof is straightforward and relies on the first order condition of the producer profit function. Recall that without a reservation market the profit function:

$$\pi_i^s(k) = \phi(k)k_i - rk_i \tag{B.5}$$

As we are in a perfect competitive framework, $\frac{\partial \phi(k)}{\partial k_i} = 0$. Therefore the first-order condition implies that the maximum is reached when $\phi(k) = r$. It is the standard equilibrium condition for investment decisions. Given that the profit function is concave, any deviation from this point generates a loss. Consequently, the marginal cost of bringing an additional capacity when $\phi(k) < r$ is null as it is always profitable to do so for the producer. On the other hand, when $\phi(k) > r$, the producer is always incited to reduce the capacity level unless given a remuneration. By construction, at the margin and indifferent between an additional capacity against a remuneration, the producer should receive a payment equal to $r - \phi(k)$.

2.3. Proof of Proposition 2

Recall that when the supply function in the reservation market with a price cap in the wholesale market is positive, it is equal to

$$X(k) = r - \phi(k) = r - \int_{t_0(k)}^{t_0^w(k)} (p(k,t) - c) f(t) dt - \int_{t_0^w(k)}^{+\infty} (p^w - c) f(t) dt \tag{B.6}$$

We then find the intersection between the supply function and the optimal payment functions given by the following equations:

$$z^w(k) = \int_{t_0^w(k)}^{+\infty} p(k,t) - p^w f(t) dt \tag{B.7}$$

For instance, when $X(k) = z^w(k)$, then it gives:

$$r - \int_{t_0(k)}^{t_0^w(k)} (p(k,t) - c) f(t) dt - \int_{t^w}^{+\infty} (p^w - c) f t dt = \int_{t_0^w(k)}^{+\infty} (p(k,t) - p^w) f(t) dt \tag{B.8}$$

When rearranged, we have:

$$r = \int_{t^w(k)}^{+\infty} (p(k,t) - c f(t) dt \tag{B.9}$$

Which is the condition for the first-best to be reached. The exact process can be applied to the two other optimal payment functions. Therefore having a demand function at the optimal investment level given the supply function in the reservation market is strictly the same as providing producers with a payment whose value is given by the payment function at the optimal investment level.

2.4. Proof of Proposition 3

First, we state in this proof that allocating the reservation price on the consumer as a tax only affects the share between on-peak and off-peak periods and the surplus's size during off-peak periods.

We start by defining the consumer surplus without reservation market:

$$W^c = \int_0^{t_0(k)} \int_0^{q_0(t)} (p(q,t) - c)dqf(t)dt + \int_{t_0(k)}^{t_0^w(k)} \int_0^k (p(q,t) - p(k,t))dqf(t)dt + \int_{t_0^w(k)}^{+\infty} \int_0^k (p(q,t) - p^w)dqf(t)dt \tag{B.10}$$

Next, we define the new thresholds between offpeak and on-peak periods, when the price cap starts to bind, and the new quantity of the good asked at a price c: t_1, t_1^w and q_1.

Then we define both the expected welfare for consumers and producers. We show that they encompass at the same time the direct welfare loss and gain from the transfer due to the reservation market but also the indirect effect due to the reservation price allocation. The expected surplus for the consumer during off-peak periods is, therefore:

$$\begin{aligned} W^c_{offpeak} &= \int_0^{t_1(k)} (-p^c(k - q_1) + \int_0^{q_1(t)} (p(q,t) - c - p^c)dqf(t)dt \\ &= \int_0^{t_1(k)} -p^c k + \int_0^{q_1(t)} (p(q,t) - c)dq)f(t)dt \end{aligned} \tag{B.11}$$

With p^c the reservation price.

For the on-peak periods when the price cap is not binding, the consumer welfare is:

$$W^c_{peak} = \int_{t_1(k)}^{t_1^w(k)} -p^c k + \int_0^k (p(q,t) - p(k,t) - p^c)dqf(t)dt = \int_{t_1^w(k)}^{+\infty} \int_0^k (p(q,t) - p(k,t))dqf(t)dt \tag{B.12}$$

For the periods when the price cap is binding, the consumer welfare is:

$$W^c_{cap} = \int_{t_1^w}^{+\infty} -p^c k + \int_0^k (p(q,t) - p^w)dqf(t)dt \tag{B.13}$$

On the other hand, the producer welfare during offpeak periods is:

$$W^p_{offpeak} = \int_0^{t_1(k)} p^c k f(t)dt \tag{B.14}$$

For the on-peak periods when the price cap is not binding, the producer welfare is:

$$W^p_{peak} = \int_{t_1(k)}^{t_1^w} p^c k + k(p(k,t) - p^c - c)f(t)dt = \int_{t_1(k)}^{t_1^w} k(p(k,t) - c)f(t)dt \tag{B.15}$$

For the periods when the price cap is binding, the producer welfare is:

$$W^c_{cap} = \int_{t_1^w}^{+\infty} p^c k + k(p^w - c)f(t)dt \tag{B.16}$$

When we add the different expected welfare for consumers and producers, we have the $p^c k$ parts canceled and the price cap p^w. It gives the following expected social welfare:

$$W_1(k) = \int_0^{t_1(k)} \int_0^{q_1(t)} (p(q,t) - c)dq\ f(t)dt + \int_{t_1(k)}^{+\infty} \int_0^k (p(q,t) - c)dq\ f(t)dt - rk \tag{B.17}$$

The proof of $\Delta W_1(k)$ relies on the difference between $W_1(k)$ and $W(k)$. There is no ambiguity when calculating it with the integrals as we always have $q_1(r) < q_0(t)$, $t_0(k) < t_1(k)$ and $t_0^w(k) < t_1^w(k)$.

Recall that:

$$W_1(k) = \int_0^{t_1(k)} \int_0^{q_1(t)} (p(q,t) - c)dq\ f(t)dt + \int_{t_1(k)}^{+\infty} \int_0^k (p(q,t) - c)dq\ f(t)dt - rk \tag{B.18}$$

We rearrange the occurrence terms:

$$\begin{aligned} W_1(k) = &\int_0^{t_0(k)} \int_0^{q_1(t)} (p(q,t) - c)dq\ f(t)dt + \int_{t_0(k)}^{t_1(k)} \int_0^{q_1(t)} (p(q,t) - c)dq\ f(t)dt \\ &+ \int_{t_0(k)}^{+\infty} \int_0^k (p(q,t) - c)dq\ f(t)dt - \int_{t_0(k)}^{t_1(k)} \int_0^k (p(q,t) - c)dq\ f(t)dt - rk \end{aligned} \tag{B.19}$$

We next rearrange the first term associated with the quantity at the marginal cost and we add the the second term with the last one :

$$\begin{aligned} W_1(k) = &\int_0^{t_0(k)} \int_0^{q_0(t)} (p(q,t) - c)dq\ f(t)dt - \int_0^{t_0(k)} \int_{q_1(t)}^{q_0(t)} (p(q,t) - c)dq\ f(t)dt \\ &+ \int_{t_0(k)}^{+\infty} \int_0^k (p(q,t) - c)dq\ f(t)dt - \int_{t_0(k)}^{t_1(k)} \int_{q_1(t)}^k (p(q,t) - c)dq\ f(t)dt - rk \end{aligned} \tag{B.20}$$

Note that the first and third term with is $W(k)$. therefore

$$W_1(k) = W(k) - \int_0^{t_0(k)} \int_{q_1(t)}^{q_0(t)} (p(q,t) - c)dq\ f(t)dt - \int_{t_0(k)}^{t_1(k)} \int_{q_1(t)}^{k} (p(q,t) - c)dq\ f(t)dt \tag{B.21}$$

2.5. Proof of Lemma 2

The proof relies on the analysis of the derivative of $\frac{\partial \Delta W_1(k)}{\partial k}$. The derivative is equal to

$$\begin{aligned}\frac{\partial \Delta W_1(k)}{\partial k} &= \frac{\partial t_0(k)}{\partial k} \int_{q_1(t_0)}^{q_0(t_0)} (p(q,t_0) - c)dqf(t_0) + \int_0^{t_0(k)} -\frac{\partial q_1(k)}{\partial k}(p(q_1,t) - c)f(t)dt \\ &- \frac{\partial t_0(k)}{\partial k} \int_{q_1(t_0)}^{k} (p(q,t_0) - c)dqf(t_0) + \frac{\partial t_1(k)}{\partial k} \int_{q_1(t_1)}^{k} (p(q,t_1) - c)dqf(t_1) \\ &+ \int_{t_0(k)}^{t_1(k)} (-\frac{\partial q_1(k)}{\partial k}(p(q_1,t) - c) + (p(k,t) - c))f(t)dt\end{aligned} \tag{B.22}$$

Note that the three terms with the derivatives of t_0 and t_1 are null. Therefore the equation boils down to:

$$\frac{\partial \Delta W_1(k)}{\partial k} = \int_0^{t_0(k)} -\frac{\partial q_1(k)}{\partial k}(p(q_1,t) - c)f(t)dt + \int_{t_0(k)}^{t_1(k)} (-\frac{\partial q_1(k)}{\partial k}(p(q_1,t) - c) + (p(k,t) - c))f(t)dt \tag{B.23}$$

Finally, as k increases, the reservation price also increases as producers need a higher remuneration to be indifferent. Therefore, the indirect effect on the demand side is magnified. It implies that the sign of $\frac{\partial q_1(k)}{\partial k}$ is negative. In turn, the derivative $\frac{\partial \Delta W_1(k)}{\partial k}$ is fully positive. Hence, the new first-best solution in terms of investment level given under the endogenous regime is always lower or equal to the first-best solution under the exogenous level.

2.6. Proof of Proposition 4

The process is straightforward and relies on the inframarginal rent:

$$\phi_1^w(k) = \int_{t_1(k)}^{t_1^w(k)} (p_1(k,t) - p^c - k) - c)\ f(t)dt + \int_{t_1^w(k)}^{+\infty} (p^w - c)\ f(t)dt + p^c(k) \tag{B.24}$$

We determine first the system's main components in terms of an exogenous value representing the reservation price. We then express the inframarginal given this value, and we equalize the supply function $r - \phi_1^w(k)$ to this value. We solve the equation and isolate the reservation price.

Again, we express all the main components of the system with this new value. The rest of the equilibrium follows.

2.7. Proof of Lemma 3

We define the expected producer welfare in the following equation:

$$W_1^p(k) = \int_{t_1(k)}^{t_1^w(k)} \int_0^k (p(q,t) - c - p^c) dq f(t) dt + \int_{t_1^w(k)}^{+\infty} \int_0^k (p^w - c) dq f(t) dt - rk \quad \text{(B.25)}$$

The first order condition of the expected producer welfare function gives:

$$\begin{aligned} \frac{\partial W_1^p(k)}{\partial k} = &-\frac{\partial t_1(k)}{\partial k} \int_0^k (p(q,t_1) - c - p^c) dq f(t_1) dt + \\ \frac{\partial t_1^w(k)}{\partial k} \int_0^k (p(q,t_1^w) - c - p^c) dq f(t_1^w) &+ \int_{t_1(k)}^{t_1^w(k)} (p(k,t) - c - p^c) f(t) dt \\ &-\frac{\partial t_1^w(k)}{\partial k} \int_0^k (p^w - c) dq f(t_1^w) dt + \int_{t_1^w(k)}^{+\infty} (p^w - c) f(t) dt - r \end{aligned} \quad \text{(B.26)}$$

Note that $p(q,t_1) = c$ and $p(q,t_1^w) = p^w$ so the the first term is null, and the terms with the derivative of t_1^w cancel each other. It implies that:

$$\frac{\partial W_1^p(k)}{\partial k} = \int_{t_1(k)}^{t_1^w(k)} (p(k,t) - c - p^c) f(t) dt \int_{t_1^w(k)}^{+\infty} (p^w - c) f(t) dt - r \quad \text{(B.27)}$$

On the other hand, the expected social welfare is:

$$W_1(k) = \int_0^{t_1(k)} \int_0^{q_1(t)} (p(q,t) - c) dq \; f(t) dt + \int_{t_1(k)}^{+\infty} \int_0^k (p(q,t) - c) dq \; f(t) dt - rk$$

The first order condition of the expected welfare function gives:

$$\begin{aligned} \frac{\partial W_1(k)}{\partial k} = \frac{\partial t_1(k)}{\partial k} \int_0^{q_1(t_1)} (p(q,t_1) - c) dq \; f(t_1) + \int_0^{t_1(k)} \frac{\partial q_1(k)}{\partial k} (p(q_1,t) f(t) \\ -\frac{\partial t_1(k)}{\partial k} \int_0^k (p(q,t_1) - c) dq f(t_1) + \int_{t_1(k)}^{+\infty} (p(k,t) - c) f(t) dt - r \end{aligned} \quad \text{(B.28)}$$

Again, note that $q_1(t_1) = k$, therefore the first and the third term of the equation cancel each other. It implies that:

$$\frac{\partial W_1(k)}{\partial k} = \int_0^{t_1(k)} \frac{\partial q_1(k)}{\partial k}(p(q_1,t)f(t) + \int_{t_1(k)}^{+\infty}(p(k,t)-c)f(t)dt - r \tag{B.29}$$

Reaching the equality between the first-best and market equilibrium requires that the current derivative with the additional capacity remuneration is equal to the derivative of the social welfare, that is $\frac{\partial W_1(k)}{\partial k} = \frac{\partial W_1^p(k)}{\partial k} + z_1(k)$. Therefore we need to have the following equality

$$\begin{aligned}\int_0^{t_1(k)} \frac{\partial q_1(k)}{\partial k}(p(q_1,t)f(t) + \int_{t_1(k)}^{+\infty}(p(k,t)-c-p^c)f(t) = \\ \int_{t_1(k)}^{t_1^w(k)}(p(k,t)-c-p^c)f(t)dt \int_{t_1^w(k)}^{+\infty}(p^w-c)f(t)dt + z_1(k)\end{aligned} \tag{B.30}$$

Which gives the optimal payment $z_1(k)$

2.8. Proof of Lemma 4

In this case, the inframarginal rent is the same as the case with a price cap only. On the other hand, the new welfare function is:

$$W^{bo}(k) = W(k) - M(k) \tag{B.31}$$

The first order condition is similar to those in the previous proof:

$$\int_{t_0(k)}^{+\infty}\int_0^k (p(q,t)-c)dq\ f(t)dt - M_k(k) = r \tag{B.32}$$

The market investment level is equal to the optimal investment level only if the solution of the following equality has the same solution as with the first-order condition of the welfare function. Because the inefficient rationing is due to the price cap, the optimal payment includes de facto $z^w(k)$ to restore the optimal investment level. We derive here only the additional part relative to the rationing.

$$\phi(k) + z^{bo}(k) = r \tag{B.33}$$

Which is the case if and only if :

$$z^{bo}(k) = \phi(k) - (\int_{t_0(k)}^{+\infty}\int_0^k (p(q,t)-c)dq\ f(t)dt - M_k(k)) \tag{B.34}$$

Therefore the optimal payment is equal to the marginal surplus loss due to inefficient rationing.

$$z^{bo}(k) = M_k(k) \tag{B.35}$$

2.9. Proof of Proposition 5

The first part of the difference can be found in the proof of Proposition 3. Now recall that, in the case of an exogenous market design, or without reservation market, the system cost associated with inefficient rationing is equal to:

$$M_0(k) = \int_{t_0^w(k)}^{+\infty} J(\Delta_0 k) f(t) dt \tag{B.36}$$

In the case of an endogenous market design, it becomes:

$$M_1(k) = \int_{t_1^w(k)}^{+\infty} J(\Delta_1 k) f(t) dt \tag{B.37}$$

With $\Delta_1 k$, the new function representing the difference between the quantity consumed at the price cap $q_1^w(k)$and the investment level.

Therefore, it is straightforward to show that the gain or the loss can be expressed as follow:

$$M_0(k) - M_1(k) = \int_{t_0^w(k)}^{+\infty} J(\Delta_0 k) f(t) dt - \int_{t_1^w(k)}^{+\infty} J(\Delta_1 k) f(t) dt \tag{B.38}$$

Recall that $t_0^w < t_1^w$, therefore we rearrange the expression as follow:

$$M_0(k) - M_1(k) = - \int_{t_0^w(k)}^{t_1^w(k)} J(\Delta_0 k) - \int_{t_1^w(k)}^{+\infty} (J(\Delta_0 k) - J(\Delta_1 k)) f(t) dt \tag{B.39}$$

Finally, we know that $J(\Delta_0 k) > J(\Delta_1 k)$, as an exogenous market design, always implies higher system costs when rationing occurs given a value of k. Indeed, everything else being equal, the demand is lower under the endogenous case.

2.10. Proof of Lemma 5

The profit function of retailers is defined as follow:

$$\pi_i^r(q_i, k) = q_i(p(q) - p^s) - p^c(k) k \frac{q_i}{q_i + q_{-i}} \tag{B.40}$$

Given the retail market structure, the first-order condition of the profit function under a competition *à la Cournot* implies that

$$p(q) + q_i p_q(q) - p^s - p^c(k) k \frac{q_{-i}}{(q_i + q_{-i})^2} = 0 \tag{B.41}$$

And the second order condition:

$$2p_q(q) + q_i p_{qq}(q) - p^c(k)k\frac{q_{-i}}{(q_i + q_{-i})^3} = 0 \quad \text{(B.42)}$$

The cross derivative of the profit function with respect to the competitor q_{-i} is:

$$p_q(q) + q_i p_{qq}(q) - p^c(k)k\frac{q_{-i} - q_i}{(q_i + q_{-i})^3} = 0 \quad \text{(B.43)}$$

With n symmetric retailers and q the total quantity, the last equation becomes:

$$p_q(q) + \frac{q}{n}p_{qq}(q) + p^c(k)k\frac{(n-2)}{n}\frac{1}{q^2} = 0 \quad \text{(B.44)}$$

The condition for existence requires that the cross derivative be positive, which establishes the condition of the Lemma. The stability and the uniqueness of the equilibrium are given by the second-order condition, which is always negative.

2.11. Proof of Proposition 6

The proof is straightforward and relies on the depressing effect generated at the margin on retailers' profit function. In this case, this form of allocation is similar to an increase in the marginal cost of production passed onto consumers.

Similar to the endogenous market design, this additional marginal cost is sustained whatever is the realization of the demand level for final consumers. Therefore, it lowers the quantity bought in the offpeak periods and lowers the prices in the peak periods. However, the effect is not as significant as in the endogenous case, as, under our framework, retailers do take into their welfare part of the capacity cost allocation, and they do not fully transfer this new marginal cost onto the consumers.

2.12. Proof of Lemma 6

The welfare function is given by

$$W(k) = \int_0^{t_0(k)} \int_0^{q_0(t)} (p(q,t) - c)dq\ f(t)dt + \int_{t_0(k)}^{+\infty} \int_0^k (p(q,t) - c)dq\ f(t)dt - rk \quad \text{(B.45)}$$

The first order condition gives the optimal investment level:

$$\int_{t_0(k)}^{+\infty} \int_0^k (p(q,t) - c)dq\ f(t)dt = r \quad \text{(B.46)}$$

Regarding the market investment level, we define the inframarginal rent under Cournot competition:

$$\phi(k) = \int_{t_0(k)}^{+\infty} (p^s(k,t) - c)f(t)dt \quad \text{(B.47)}$$

The market investment level is equal to the optimal investment level only if the solution following equality has the same solution as with the first order condition of the welfare function:

$$\phi(k) + z_n(k) = r \tag{B.48}$$

Which is the case if and only if :

$$z_n(k) = \phi(k) - \int_{t_0(k)}^{+\infty} \int_0^k (p(q,t) - c)dq\ f(t)dt \tag{B.49}$$

Therefore the optimal payment is equal to the expected markup of retailers.

$$z_n(k) = \int_{t_0(k)}^{+\infty} \frac{-k}{n} p_q(k,t) f(t)dt \tag{B.50}$$

2.13. Proof of Lemma 7

We denote the value q^r, the equilibrium quantity offered on the retail market without reservation market. This value is given by the solution of $p(q)+q_i p_q(q)-p^s = 0$. We also denote the value q^p the equilibrium quantity offered on the retail market when the retailers sustain the penalty. This value is given by the solution of $p(q)+q_i p_q(q)-p^s-S = 0$. By construction, we necessary have $q^r > q^p$. Now that we found the equilibrium in the retail market, we can derive the reservation market's demand function. We characterize the symmetric equilibrium, and we left the asymmetric equilibrium for future extensions of this paper.

By construction, if retailers buy the same quantity in the reservation market, the retail market equilibrium is also symmetric. We start by defining which equilibrium exists in the retail market, given the capacity level. There are only three possible cases, each with a unique equilibrium on the retail market:

- i $k > q^r$. All retailers are in positive deviation. However, the equilibrium in the retail market is still q^r as it is the *Cournot* equilibrium.
- ii $q^r \geq k > q^p$. All retailers have bought the same quantity in the reservation market that they have sold on the retail market. The equilibrium in the retail market is k.
- iii $q^p \geq k$. All retailers have not bought enough capacity, and all retailers have to pay penalties. The equilibrium in the retail market is q^p.

We provide in figure B.1 an example of the three cases based on a different level of capacity. The capacity level k_i and k_{-i} are represented respectively by the horizontal and vertical line. In the first subplot, the level capacity is low, and it is an equilibrium to stay in the penalty system (case (iii)). Notice in particular that retailer profit function is continuous but not differentiable at $q_i = k_i$, which induces the case without the penalty at k_i and k_{-i}. In the second figure, retailers buy the same amount of capacity (case (ii)). In the third, equilibrium is the case (i).

Next, to find the equilibria set in the reservation market, we start by determining the dominant strategies for the three previous cases. The equilibrium is then the strategy

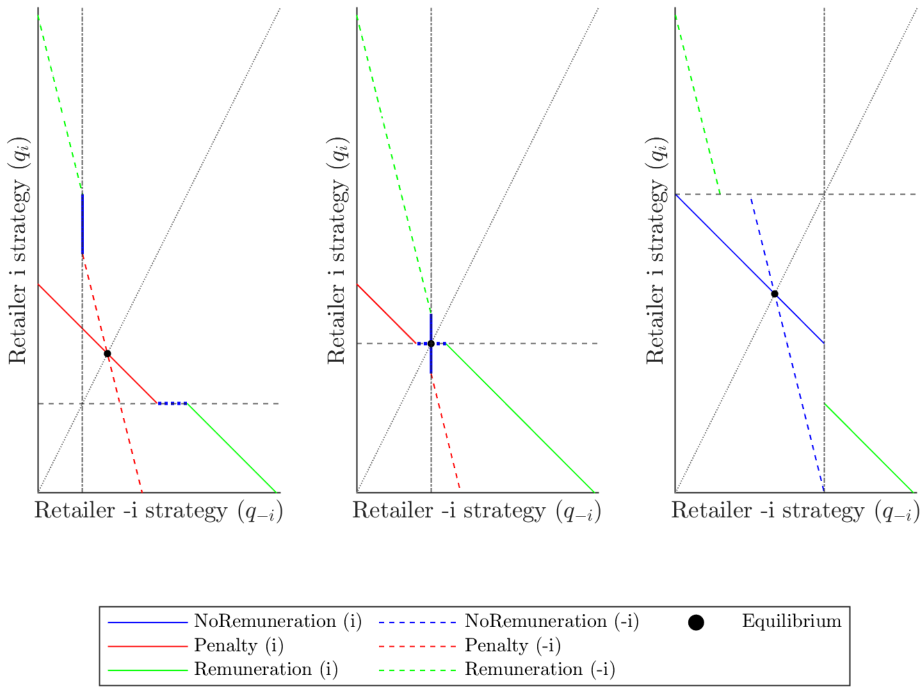


Figure B.1: Best reaction functions for two retailers in a linear case

that brings the highest profits. It is straightforward that in cases (i) and (ii), it is a dominant strategy (strict if $p^c > 0$) to buy the same amount of capacity as the quantity in the retail market. Indeed, the retail's optimal quantity is the same as without reservation market, and buying additional capacities does not bring any remuneration. Therefore, it is always optimal to buy what is strictly necessary [44].

The optimal quantity of capacity bought in case (iii) depends on the difference between the penalty value S and the reservation price $p^c(k)$. We previously assumed a linear penalty system, so if the penalty is lower than the reservation price $S \leq p^c(k)$, then it is a dominant strategy (strict if $S < p^c(k)$) to buy no capacity and sustain a penalty on all the quantity sold on the retail market. Indeed, with strict inequality, the profit function is a decreasing non-concave function with respect to k_i. On the other hand, when the penalty is higher than the price, the profit function is an increasing non-concave function with respect to k_i. The dominant strategy is to buy the same amount of capacity as the quantity q^p which corresponds to the retail market's corresponding equilibrium in the penalty case. Therefore, the set of dominant strategies in the retail market is:

$$\begin{cases} [q^p, q^r] & \text{if} \quad p^c(k) \leq S \\ \{0,]q^p, q^r]\} & \text{if} \quad p^c(k) > S \\ \forall q \in]q^p q^r[& q \text{ is a solution of} \quad p(q) + q_i p_q(q) - p^s(q) - p^c(q) = 0 \end{cases}$$

[44]Retailer profit function is a decreasing function with respect to k as we assume retailers do not behave strategically in the reservation market.

2.14. Proof of Proposition 7

The proof relies on the three cases describes in the corresponding section.

First, we showed that when the price cap is never binding (case 1), the market design does not indirectly affect the demand side.

Second, when the price cap starts to bind, but retailers prefer to lower their demand rather than sustaining the penalty (case (2), the indirect effect does not appear in the social welfare function. Indeed, under this case, the capacity is binding, so a decrease in the demand due to a price increase only redistributes the welfare between consumers and producers.

Finally, the only indirect effect of this design exists when inefficient rationing occurs, that is, when retailers always sell at a quantity equal to $q_d^w(k)$. This case implies that both the retailers and the consumers suffered a surplus loss due to the demand reduction. Added to the penalty cost borne by the retailer of $S(q_d^w(k) - k)$, it gives the net loss for the welfare function. The gains in terms of rationing cost avoided are similar to the demonstration of Proposition 5.

2.15. Proof of Lemma 8

The three indirect effects described in the market design also impact how retailers' profit functions can be expressed. Using the same logic as for the expected producer and consumer welfare, we have the following expression for retailers' expected welfare.

$$\begin{aligned}\pi^r(k,t) = \int_0^{t_0(k)} -\frac{q_0(t)^2}{n} p_q(q_0(t),t) f(t)dt + \int_{t_0(k)}^{t_0^w(k)} -\frac{k^2}{n} p_q(k,t) f(t)dt + \\ \int_{t_0^w(k)}^{t_d^w(k)} k(p(k,t) - p^w - T(k,t)) f(t)dt + \int_{t_d^w(k)}^{+\infty} k(p(k,t) - p^w - S) f(t)dt - \\ \int_{t_d^w(k)}^{+\infty} S(q_0^w - k) f(t)dt - p^c(k)k_i\end{aligned} \tag{B.51}$$

The demand function of retailers in the reservation market is found by the first order condition of their profit function:

$$
\begin{aligned}
\pi^r(k,t) = &-\frac{\partial t_0}{\partial k}\frac{q_0(t_0)^2}{n}p_q(q_0,t_0)f(t_0)+\frac{\partial t_0}{\partial k}\frac{k^2}{n}p_q(k,t_0)f(t_0)-\frac{\partial t_0^w}{\partial k}\frac{k^2}{n}p_q(k,t_0^w)f(t_0^w)\\
&+\int_{t_0(k)}^{t_0^w(k)}(\frac{2k}{n}p_q(k,t_0^w)-\frac{k^2}{n}p_{qq}(k,t_0^w))f(t)dt-\frac{\partial t_0^w}{\partial k}(k(p(k,t_0^w)-p^w-T(k,t_0^w))f(t_0^w))\\
&+\frac{\partial t_d^w}{\partial k}(k(p(k,t_d^w)-p^w-T(k,t_d^w))f(t_d^w))+\int_{t_0^w(k)}^{t_d^w}(p(k,t)-p^w-T(k,t)+kp_q(k,t)-\frac{\partial T(k,t)}{\partial k})f(t)dt\\
&-\frac{\partial t_d^w}{\partial k}(k(p(k,t_d^w)-p^w-S)f(t_d^w))+\int_{t_d^w}^{+\infty}(p(k,t_d^w)-p^w-S+kp_q(k,t))f(t_d^w)dt\\
&-\frac{\partial t_d^w}{\partial k}S(q_0^w(t_d^w)-k)f(t)-\int_{t_d^w}^{+\infty}-Sf(t)dt-p^c(k)
\end{aligned}
\tag{B.52}
$$

Note that: $q_0(t_0) = k$, so the two first terms cancel each other. The third term is null. The fifth term is also null as $p(k,t_0^w) = p^w$, and $T(k,t)$ is null at t_0^w as we are at the boundary of the case (2). Similarly, at the opposite boundary t_w^d we have $T(k,t_w^d) = S$, so the sixth and the eighth term cancel each other. Finally, the eleventh term with the expected value of the penalty enters and cancels the penalty value in the seventh term. Finally, the tenth term is null as $q_0^w(t_d^w) = k$.

The first-order condition allows isolating the reservation price, which is the demand function in the reservation market.