

Capacity choice, technology mix and market power¹

Guy MEUNIER¹

Ecole Polytechnique-GIS Larsen, route de Saclay 91128 France

Abstract

This paper investigates strategic capacity choices in electricity markets comprised of heterogeneous firms. Long term strategic investments are analyzed assuming that the wholesale market is competitive. There are two technologies available to produce electricity; both are efficient and used at a first best optimum. When not all firms can invest in both technologies, there can be over investment in either of these technologies. It is shown that if the number of firms that can invest in a particular technology is limited, the development of competition solely using the other technology can decrease welfare.

Key words: capacity choice, market power, heterogeneity

JEL: D43, D24, L13, Q41

1. Introduction

One of the motivations behind market reforms in the electricity industry was to encourage timely efficient investment in the optimal mix of technologies thereby avoiding the overcapacity situations observed in the former regulated monopoly regimes. Yet, there are concerns about the ability of the new liberalized market regime to induce sufficient investment in building capacity, in the optimal technology mix without penalizing some technologies. Ensuring enough generation capacity to meet future electricity demand by the optimal technology mix progressively became a contentious issue in the design of reforms (IEA, 2003).

¹Email adress: guy.meunier@polytechnique.edu

The literature on the subject highlights three major reasons to explain the potential suboptimality of investment in the generation mix: (1) “missing money”, (2) risk management, and (3) market power. The present paper focuses on the third reason: the strategic choice of generating capacity in a situation with presence of heterogeneous firms. The interrelationship of long-term investment choices with short-term market power is not considered, as it is assumed that the short-term market is perfectly regulated. In fact, the wholesale price is set at the variable cost of the marginal technology plus the value of the capacity constraint. It allows me to focus on the long term incentives for generators to distort the generation mix, and to consider the influence of the number of firms with access to a particular technology. This last point allows for a better understanding of the consequences of the development of competition through a unique technology.

The three explanations of underinvestment are briefly reviewed before discussing the issue of firms heterogeneity. The theoretical literature on capacity choice in electricity markets and the contributions of the present paper are then presented.

Sub-optimality of generation investment in electricity markets

A feature of the electricity industry is that many technologies are used to produce the same good. The variation of the load, the need to balance production and consumption permanently and the lack of storage facilities account for this feature. For any load duration curve, the optimal technology mix minimizes the cost of producing this load. A technology is described by the ratio of its variable cost to capacity cost³. A high capacity cost and a low variable one characterized a baseload technology that is used for production over long duration; whereas, a low capacity cost and a high operating one characterize peaking units that are thus efficient in producing for a few hours per year.

³This is a first approximation because others characteristics play important roles such as the ramping rate and start up costs.

Concerns on investment that were initially focused on peaking units have now extended to all technologies. Particularly, the ability of wholesale electricity markets to promote sufficient investment in capital-intensive technologies, such as nuclear, which is a typical baseload technology, is currently an issue under debate. A perfectly competitive wholesale electricity market should theoretically induce efficient investment, namely an optimal aggregate capacity and an optimal technology mix. Scarcity rents during peak periods are needed to ensure the profitability not only of the peaking units but also of all other technologies. Moreover the short-term system security of supply is a public good provided by the system operator using operating reserves (Joskow and Tirole, 2007). This implies that additional revenues are needed for the provision of this public good.

Three main arguments are found in the literature to explain the potential lack of investment in electricity markets. The first is the so called “missing money” argument, and refers to revenue deficit during peak hours (Cramton and Stoft, 2006). This deficit is attributed to the price caps decided by regulators for limiting skyrocketing prices during peak hours and to ill-designed regulatory procedures related to the use of operating reserves. Indeed, the technical rules used by system operators to guarantee system reliability by using operating reserves tend to erase revenues from the energy and reserves markets during these peak hours (Joskow, 2006). Even if the price cap were suppressed, a revenue deficit would remain.

The second reason is the exceptional volatility of electricity prices and market incompleteness. The volatility of electricity prices (ranging from 20€ to 5-10 000€/MW/h) is explained by the variations occurring in demand and supply and their short-run price inelasticity. Theoretically, volatility does not deter investment in due time and appropriate technologies, provided that future and forward markets exist to allow investors to manage these risks. However, these markets are poorly developed and this explains the difficulties for hedging in-

vestments in generation capacity⁴; these difficulties might be reinforced by the imperfections of capital markets.

The third reason relates to market power and is the subject of the present paper. Given the high concentrations observed in most electricity markets, particularly in Europe there are concerns that firms strategically limit their investment to keep prices high (D.G.Comp, 2007). Every generator can potentially benefit from sub-optimal investment in generation capacity in the system. This will in turn provoke profitable periods of price spikes. Indeed, because of the inelastic nature of demand and supply, even a slight shortfall of available capacity can provoke a dramatic increase in electricity prices in periods of tight supply.

Moreover, besides the incentive to underinvest in aggregate capacity, firms might also be tempted to profitably distort the technological mix. This distortion might be amplified by the heterogeneity of firms relative to the access to technologies. The present paper deals with this issue by analyzing the sub-optimality of investment related to the strategic choice of capacity, and its relationship with the number of firms with access to each technology.

Heterogeneity of firms

Implicit in the assumption of exercise of market power is the assumption of entry barriers. To assume that firms can exercise long-term market power one has to assume that entry to generation is restricted. This restriction might vary according to technologies, as it is natural to assume that firms are heterogeneous, that is some are ‘generalist’ and can invest in all types of power plants, whereas others are specialized in a particular technology. The concentrations observed in electricity markets suggest some entry barriers; furthermore, European markets reveal heterogeneity of the portfolios of firms’ technologies (D.G.Comp, 2007).

⁴For instance, Roques et al. (2006) argue that risk aversion can favor investment in combined cycle gas turbines (CCGT) over nuclear plants because of the correlation between prices of electricity and gas. Finon and Roques (2008) further analyze how risk issues can explain a perceived lack of investment in nuclear plants

This heterogeneity could be explained by the scarcity of inputs that some firms have historically acquired. Moreover, institutional settings might artificially maintain this scarcity.

A good example is the current situation of nuclear technology and gas generation technology (the combined cycle gas turbine or CCGT), the latter being a standardized technology that is perceived as the main vehicle of competition (Green and Newbery, 1992; Newbery, 1998), whereas only few firms seem able to invest in nuclear plants.

The previous form of organization of electricity sectors in industrialized countries (regulated vertically integrated monopolies) was justified not only by networks development but also by the increasing returns to scale in generation particularly for coal and nuclear technologies (Stoffaës, 1994). Generation was a part of the natural monopoly encompassing both generation and transmission. The development of CCGT technology was one of the technological conditions that allowed market liberalization because it took generation out of the natural monopoly of the wholesale supply⁵. The post-reform wave of investments in CCGTs in the UK and in the USA during the 1990s and early 2000s is a confirmation of this view. At the same time, the development of nuclear technology stopped in most industrialized countries (Nuttall, 2005). Moreover, several European countries (Italy, Netherlands, Germany, Sweden, Spain) unequivocally decided to impose a moratorium on new nuclear investment or even a phase-out policy to discontinue nuclear power generation. However, the rise of fossil fuels prices, concerns about long term security of energy supply, and the regulation of carbon dioxide emissions have renewed interest for nuclear generation in all

⁵Green and Newbery (1992) write: “ *The second defense of the lack of regulation is that entry into the industry will be open to any plausible supplier and that new technology using high-efficiency combined cycle gas turbines (CCGT) makes entry at modest scales (300-600 megawatts [MW]) simple and quick: construction times are short, and the technology is readily available from a number of suppliers and is competitive with existing larger thermal stations.*”

countries even those phasing it out⁶.

However, scarcity of human capital⁷ and sitting restriction might limit investment in new nuclear production plants. Consequently, firms that historically acquired specific know-how and who own a number of nuclear sites have an advantage to further develop nuclear capacity.

Other technologies than nuclear are liable to face similar constraints such as hydro and coal power production—the former because of the regulation of access to sites, and the latter because of the development of the technology of carbon capture and storage that will require large capital resource, skills, industrial planning and sites.

This difference in the nature of production technologies between the standard and non-standard marks out the electricity industry. This paper aims to clarify the consequences on competition of coexistence of different technologies when few firms are able to build and operate any of the technologies. The theoretical literature related to market power and capacity choice is reviewed next.

Choice of capacity

The development of wholesale electricity markets created a renewed interest in the literature on the choice of capacity and the need to cope with demand fluctuation or uncertainty. Gabszewicz and Poddar (1997) analyze the choice of capacity by two competing producers in a linear model. They establish that firms invest more with uncertainty than without. In their model, one technology is available and short-term competition is a quantity game *à la* Cournot with capacity constraints. Their results are generalized by Zoetl (2008, chap1) who

⁶For instance, nuclear energy is recognized by both USA (EPACT, 2005) and UK (DBEER, 2008) governments as an essential element of their future energy policy and Sweden ended its nuclear phase-out policy at the beginning of 2009.

⁷According to Nuttall (2005): “*Of all the factors that shape the future of a nuclear renaissance in western Europe and North America only one has the status of potential ‘show stopper’.* This factor is not the safety of nuclear power, its environmental consequences or its economics: it is the supply of engineers, scientists and skilled trades needed to keep the industry operating and properly regulated...”

also considers the alternative assumption of a regulated, or perfectly competitive, spot market.

Several authors have addressed the issue of the technology mix and in most papers all firms have access to all technologies. In a major contribution, von der Fehr and Harbord (1997) analyze investments by symmetric producers for different price mechanisms or regulatory regimes. With non discriminatory auctions, they state that there is no symmetric equilibrium⁸ but only consider the duopoly case. If the spot market is efficient—which is the case considered here—they establish that firms will underinvest in aggregate capacity and profitably distort the technology mix toward peaking units. A firm has an incentive to restrict baseload capacity only to limit the period of marginality of this technology. An rise in the number of firms increases the aggregate quantity of capacity and improves the technology mix⁹.

These results contrast with those obtained with short-term Cournot competition and several technologies. Because of the strategic effect of lower variable cost¹⁰ there is a strategic incentive to invest in baseload capacity. Murphy and Smeers (2005) consider heterogeneous firms, namely a baseload producer and a peak producer. They emphasize the strategic effect of investment in a closed-loop equilibrium. Because of this strategic effect the baseload player invests more in a closed loop equilibrium than in an open loop one¹¹. Zoetl (2008,

⁸This is similar to the finding of Reynolds and Wilson (2000) on capacity choice under conditions of demand uncertainty and price competition. Fabra et al. (2008) further analyzed it to compare different auctions mechanisms.

⁹In a more recent paper, Arellano and Serra (2007) arrived at a similar conclusion; they took into account the incentive for firms to distort the technology mix when the aggregate quantity of capacity is fixed and the price of electricity is regulated. They extend the analysis by considering free entry of generators.

¹⁰The strategic effect refers to the decrease in rivals' production subsequent to the fall in a firm's marginal cost.

¹¹In the open loop equilibrium strategic effects are ignored: capacities and (conditional) production quantities are simultaneously chosen, as opposed to a closed loop equilibrium where capacity are chosen before production quantities.

chap3) considers symmetric firms that have access to a continuous technology set. The continuity property of the technology set allows tractability of the model and explains the symmetry of firms at equilibrium. He states that because of the strategic incentive firms might overinvest in baseload units.

To my knowledge, only Murphy and Smeers (2005) consider asymmetric firms, and no previous analysis performs comparative statics on the number of investing firms. This could be due to the analytical difficulties arising from dynamic games with discrete technology sets and strategic interactions at each stages. In conducting such an analysis here, short-term market power is ignored and so is the strategic effect mentioned above. The price is assumed set at the variable cost of the marginal technology when demand is not rationed and at the value of lost load (VoLL) when rationing occurs.

Empirically, observed prices are not as high as those predicted by theoretical models of imperfect competition (supply function, discrete auctions, Cournot) and are closer to marginal cost than to Cournot prices (Wolfram, 1999). That regulatory authorities closely scrutinized wholesale electricity markets or that firms not strategic in their short-term whereas their investment choices are typical strategic decisions could explain these results. Another explanation refers to the vertical relations between electricity producers and retailers (Bushnell et al., 2008). These relations are not considered here but further research should.

The simplification of the short-term competition together with the linearity of the load curve end up in a very simple and tractable model. My view of firms heterogeneity is that not all of them have access to all technologies — some are specialized and some are generalists. It is first shown that even if both technologies are efficient, generalist firms do not necessarily invest in both technologies. If firms specialized in a particular technology are sufficiently numerous they overinvest and deter generalist firms from investing in this technology. In such a case each generalist firm behaves as a specialist in the other technology.

In this case, both the aggregate quantity of capacity and the technology mix are suboptimal. The distortion of the technology mix is related to the number of firms that could invest in either technology. Many situations are possible:

overinvestment could arise at equilibrium in either baseload capacity or peaking power plants. The welfare consequences of a change in the number of firms that have access to a technology are investigated. It is established that even if both technologies are efficient an increase in a certain type of specialized firm can reduce welfare. Although, an additional firm is active increasing capacity could further distort the technology mix and this distortion could offset the gains arising from the increase of aggregate capacity.

The rest of the paper is organized as follow: I first introduce the model and then consider the first best optimum in the section that follows. The investment game is solved for the general case (Section 3) before analyzing the influence of the number of firms of each type (Section 4).

2. The model

2.1. Framework

I consider a simple electricity system without network constraints. Electricity demand (x) is assumed to be uniformly distributed in $[0, X]$. It is a rough representation of a load duration curve with a year's duration normalized at 1. If a quantity $y < x$ is served, the consumers' gross surplus is vy , where v is the value of lost load (VoLL). In Section 5 I discuss the underlying assumptions on consumers' utility function and the manner in which a more realistic load duration curve would influence the results.

There are two technologies to produce electricity and I label them $t = b, p$ respectively. Each technology t is characterized by a variable cost c_t (per kwh) and a capacity cost I_t (per kW per year). Although technology b is less costly to produce per unit of electricity all through the year than technology p , it is more costly for production over a short period:

$$c_b + I_b < c_p + I_p,$$

$$I_p < I_b.$$

Even if the sum of the variable and capacity costs of technology p is higher than that of technology b , it is efficient in production for a short duration in

the year. The difference of capacity costs is denoted by $\Delta = I_b - I_p$ and the difference of variable costs $\delta = c_p - c_b$. Both are positive and $\Delta < \delta$. The ratio $r = \Delta/\delta$ is the duration such that technology b (resp. p) is more efficient for production over longer (resp. shorter) period than r . These features are illustrated in figure 1.

Technologies b and p are, respectively, called baseload and peak throughout the paper, but the framework could be used to consider baseload- and midload technologies such as nuclear and CCGT.

For $t = b, p$, the ratio $r_t = I_t/(v - c_t)$ is the minimal duration of production with technology t such that the aggregate cost is lower than the consumer surplus. Value of each is assumed less than one i.e. $v > c_p + I_p$ and it is assumed that:

$$r_p < r;$$

this assumption ensures that technology p is used at equilibrium. The left hand side is the minimal duration of production with technology p such that costs are lower than the consumer surplus. To produce during a shorter period of time with this technology a unit of electricity consumed would imply a welfare loss. The right hand side is the maximal duration of production for which technology p is more efficient than technology b . The former being smaller than the latter it is optimal to use both technologies. It should be noticed that this assumption is equivalent to $r_b < r$ and to $r_p < r_b$, these two inequalities can be interpreted similarly.

Insert figure 1

Figure 1: Cost and production duration

There are n firms that produce electricity indexed $i = 1..n$. Individual quantities of capacity of firm i of each technology are denoted by k_b^i and k_p^i , and its aggregate quantity of capacity is $k^i = k_b^i + k_p^i$. Thus, aggregate quantities of capacities of all firms for each technology are $k_t = \sum_i k_t^i$ for $t = b, p$ and $k = k_b + k_p$.

As not all firms have access to both technologies, the set of firms is divided into three subsets. There are g ‘generalist’ firms that have access to both technologies and $n - g$ specialized firms; s_b baseload firms have access only to technology b , and s_p peaking firms that have only access to technology p . So the number of firms is $n = g + s_b + s_p$. Firms are ordered as follows: firm $i = 1, \dots, s_b$ are baseload firms, firms $i = s_b + 1, \dots, s_b + s_p$ are peaking firms and finally firms $i = n - g, \dots, n$ are generalists.

Each generalist firm chooses quantities of capacity of each technology, whereas a peaking (resp. baseload) firm only chooses a quantity of technology p (resp. b).

Once capacities are fixed, short-term interactions are assumed to be ‘perfectly’ regulated, thus there is no modelling of short-term market power. The price is set at the marginal cost of the last unit called when all demand is satisfied, and at v in case of rationing. Firms produce with a technology when the price is above its operating cost. Rationing arises when the demand of electricity is higher than the aggregate capacity available. Therefore, when demand is less than k_b the wholesale price is c_b , and only firms with baseload capacities produce. When demand is greater than k_b and smaller than k the price is c_p , baseload capacities are fully utilized and the quantity $x - k_b$ that remains is produced by firms with peak capacities. For higher demand the price is v and there are only k units of electricity consumed, a part $x - k$ of the demand is not satisfied. Price and quantities are represented on Figure 2.

Insert figure 2

Figure 2: Supply curve and spot prices

Firms earn short-term positive revenue from capacities of a technology type only when the aggregate quantity of capacity of this type is fully utilized. The

profit of a firm $i = 1, \dots, n$ is:

$$\begin{aligned}\pi^i &= \frac{1}{X} \int_{k_b}^{k_b+k_p} \delta k_b^i dx \\ &+ \frac{1}{X} \int_k^X [(v - c_p)k_p^i + (v - c_b)k_b^i] dx \\ &- I_b k_b^i - I_p k_p^i.\end{aligned}\tag{1}$$

The net revenue of a firm has two terms: the first one is the net revenue from baseload capacity when the price is set at the variable cost of the peak technology, and the second one is the profit obtained from both technologies when rationing arises. Alternatively, one can rewrite the profit of a firm as a function of aggregate capacity and baseload capacity:

$$\begin{aligned}\pi^i &= \frac{1}{X} \int_{k_b}^X \delta k_b^i dx \\ &+ \frac{1}{X} \int_k^X (v - c_p)k^i dx \\ &- \Delta k_b^i - I_p k^i.\end{aligned}\tag{2}$$

This emphasizes the relation between the technology mix and the aggregate quantity of capacity of a generalist firm. In that case, The cost of a unit of capacity of technology b is net of the cost of the unit of capacity p it replaces, and similarly its short-term revenue is the difference of the variable costs in demand states where baseload capacities are fully utilized.

The sum of firms' profits only depends on the aggregate quantities of the capacities and not on its distribution among firms:

$$\begin{aligned}\Pi(k_b, k) &= \sum_{i=1..n} \pi^i \\ &= \frac{1}{X} \left[\int_{k_b}^X \delta k_b dx + \int_k^X (v - c_p)k dx \right] - I_p k - \Delta k_b\end{aligned}\tag{3}$$

2.2. Welfare optimum

Welfare is the sum of consumers' net surplus and firms' profit. Net consumers' surplus is denoted by $CS(k_b, k)$; it is the integral of the differences

between gross surplus and expenses:

$$CS(k_b, k) = \frac{1}{X} \left[\int_0^{k_b} (v - c_b) x dx + \int_{k_b}^k (v - c_p) x dx \right]. \quad (4)$$

This surplus is the sum of surpluses of final users and (possibly of) retailers; its splitting between the two classes of agents depends on the retail price, which is not made explicit here. Welfare is:

$$W(k_b, k) = CS(k_b, k) + \Pi(k_b, k). \quad (5)$$

Welfare could be maximized with respect to pairs of quantities (k_b, k_p) or (k_b, k) . The second method is used here because it fits in better with the notion of technology mix. Welfare (5) is concave and the first best quantities k^* and k_b^* solve following first order conditions:

$$\frac{\partial W}{\partial k} = \frac{1}{X} \int_k^X v - c_p dx - I_p = 0 \quad (6)$$

$$\frac{\partial W}{\partial k_b} = \frac{1}{X} \int_{k_b}^X \delta dx - \Delta = 0 \quad (7)$$

The first relation can be rephrased in the jargon of electricity systems. It is the relation between the VoLL v and the loss of load probability $(X - k)/X$ and the variable- and capacity costs of the peak technology. The second equation is related to the optimal technology mix that minimizes production costs. This technology mix is such that the time of use of each unit of capacity of technology p is less than r . From these equations, it is simple to obtain the expression of the optimal technology mix:

$$k^* = X(1 - r_p) \text{ and } k_b^* = X(1 - r), k_p^* = X(r - r_p). \quad (8)$$

Moreover, the assumption $r > r_p$ ensures that both technologies are used at the optimum. Furthermore, it is optimal to ration consumers during a fraction r_p of the year. This fraction is determined solely by the cost of the marginal technology and the value of lost load. The choice of first best quantities of capacities is depicted in Figure (3).

Insert figure 3

Figure 3: Load curve and optimal investment

If firms were price-takers in the long run, the optimal investment would be decentralized by the short-term prices system and long run profits would be null. In this sense the short-term is said to be perfectly regulated. With price-taking behavior, there is no need for further regulations to promote investment such as capacity payments or capacity markets. Such instruments might be justified, even in the context of perfect competition, by the public good nature of operating reserves which is related to the risk of network collapse¹², or by the ‘missing money’ brought about by the regulation of the energy market (Cramton and Stoft, 2006). For instance, if the price during rationing is below the VoLL v there is a revenue deficit that ensues that entails a suboptimal investment even with perfect competition. However, neither of these justifications is relevant here as there is no risk of network collapse and the price is set at the VoLL during rationing, but the investment is suboptimal because of the long-run strategic behavior of firms.

I consider several industry configurations where firms have access to both technologies or only to one. In the electricity industry some technologies are “standardized” and available to all firms whereas others are not. The concern regarding suboptimal investment in peaking units—technology p in the model—could be explained by market power and could be addressed within the present framework. In spite of the public good characteristics of operating reserves, few firms invest in peak capacity and hence might limit their investments in order to prolong the duration of high price periods. Such situations correspond to $s_p = 0$; $s_b > 0$ and $g > 0$ in this framework, that is all firms can invest in

¹²This risk is null in the present framework because no uncertainties remain on demand or on capacity availability when the system operator decides to ration consumers. The distinction between planned rationing and network collapse is fundamental; see Joskow and Tirole (2007) for a discussion of this distinction and a formal analysis of operating reserves.

baseload capacity but only a subset can invest in peaking units.

Nevertheless, concerns over underinvestment could be extended to all types of technologies and in particular to baseload technologies. Nuclear technology is a typical baseload technology and few firms are able to master it, and the situation is reversed; there is a potential lack of investment in baseload technology. An analysis of both situations followed.

First, the next section describes the equilibrium of the capacity game in the general case. Second, Section 4 describes the welfare consequences of an increase in the number of specialized firms of a certain type.

3. Oligopolistic equilibrium

Firms simultaneously choose their quantities of capacity in order to maximize their profit. While specialized firms invest only in one type of capacity, generalists invest in both. From expression (1), the first order conditions of both types of specialized firms are:

$$\text{for } i = 1, \dots, s_b : \quad \frac{\delta}{X} k_p + \frac{v - c_b}{X} (X - k - k_b^i) - I_b = 0, \quad (9)$$

$$\text{for } i = s_b + 1, \dots, s_b + s_p : \quad \frac{v - c_p}{X} (X - k - k_p^i) - I_p = 0. \quad (10)$$

As is usual in a quantity game, firms' incentive is to limit their investment. Here, a lower aggregate capacity prolongs the time with price at v . Baseload firms earn an additional revenue when the price is fixed at c_p , which is proportional to the quantity of peak capacity installed because of the linearity of the load curve.

As for generalist firms, each of them chooses quantities of both types of capacity, as they might not invest in one type of capacity positivity constraints are introduced. Therefore, the objective of a generalist firm is:

$$\begin{aligned} & \max_{k_b^i, k_p^i} \pi^i(k_b^i, k_p^i, k_b, k_p) \\ & \text{subject to } 0 \leq k_b^i, 0 \leq k_p^i. \end{aligned}$$

The Lagrange multiplier of the positivity constraint of baseload (resp. peak) capacity is ν_i (resp. μ_i). First order conditions of a generalist firm are for $i = n - g + 1, \dots, n$:

$$\frac{\delta}{X}k_p + \frac{v - c_b}{X}(X - k - k_b^i) - \frac{v - c_p}{X}k_p^i - I_b + \nu^i = 0; \quad (11)$$

$$\frac{v - c_p}{X}(X - k - (k_b^i + k_p^i)) - I_p + \mu^i = 0. \quad (12)$$

Several relationships between the quantities chosen by generalists and specialized firms could be deduced from these first order conditions. If a generalist firm invests in both types of technology, it chooses an aggregate quantity of capacity similar to a peaking firm because the marginal capacity of a generalist firm being a peaking unit, its marginal revenue is similar to that of a peaking firm. However, a part of a generalist firm's capacity is baseload.

Compared to baseload firms, generalist firms' incentive to invest in baseload capacity is different because they use baseload capacity to distort favorably the technology mix and not to limit the aggregate capacity. The marginal baseload capacity of a generalist firm has the additional negative effect of lowering its revenue from its peaking plants; hence, a generalist firm invests less in baseload capacity than a baseload firm. The incentive to distort the technology mix might be best examined with an alternative method of rewriting the first order condition:

$$\frac{\delta}{X}(X - k_b - k_b^i) - \Delta + \nu^i - \mu^i = 0. \quad (13)$$

This form emphasizes that the comparison with peaking units decides the choice of a baseload capacity, thus marginal opportunity revenue and cost are δ and Δ . Equation (13) says that a generalist firm limits its share of baseload technology in order to increase the marginality period of peaking units. This is the effect at stake in the analysis of Arellano and Serra (2007) who focus on the incentive for a generator to under-invest in baseload capacity when the aggregate capacity is fixed equal to the maximum load.

However, whether generalist firms invest in both types of capacities depends upon the industry configuration. Rewriting the first order conditions (11) and

(12) of a generalist firm:

$$\begin{aligned} k_b^i &= k_b^* - k_b + (\nu^i - \mu^i) / \delta \\ k_p^i &= k_p^* - k_p + \left[\frac{v - c_b}{v - c_p} \mu^i - \nu^i \right] / \delta. \end{aligned} \quad (14)$$

If one group of specialized firms “overinvests”, i.e. invests more than the corresponding first best quantity, generalist firms do not invest in the corresponding technology. As specialized firms have only access to one technology their investment incentives are higher than those of generalist firms are, and such situations could arise at equilibrium. If there are too many specialized firms of one type, these firms may over invest and deter generalist firms from investing in their technology. In such cases, the situation is similar to an completely specialized industry.

Proposition (1) establishes expressions of equilibrium quantities when generalist firms invest in both types of technology. Situations where generalist firms do not invest in a certain technology are precisely stated in Proposition (2). A baseload (resp. peaking) firm’s individual equilibrium quantity of capacity is k_b^S (resp. k_p^S), and a generalist firm’s equilibrium quantities of baseload and peaking capacities are k_b^G and k_p^G .

Proposition 1. *The capacity game has a unique Nash equilibrium. If at this equilibrium generalist firms invest in both types of technology, the equilibrium quantities are:*

$$\begin{aligned} k_b^S(s_b, s_p, g) &= \frac{X}{A} \left[(g+1)(1-r_b) + s_p \frac{\delta - \Delta}{v - c_b} \right], \\ k_p^S(s_b, s_p, g) &= \frac{X}{A} [(g+1)(1-r_p) + s_b(r_b - r_p)], \\ k_b^G(s_b, s_p, g) &= \frac{X}{A} [(g + s_p + 1)(1-r) - s_b(r - r_b)], \\ k_p^G(s_b, s_p, g) &= \frac{X}{A} [(g + s_b + 1)(r - r_p) - s_p(1-r)] \end{aligned}$$

where $A(s_b, s_p, g) = (g + s_b + 1)(g + s_p + 1) - s_b s_p (v - c_p) / (v - c_b)$.

Thresholds for the number of specialized firms that deter generalist firms from investing in either of the two available technologies could be derived from expressions in Proposition 1.

Proposition 2. *Equilibrium quantities k_t^S and k_t^G for $t = b, p$ are:*

- if $s_b \geq (s_p + g + 1)(1 - r)(r - r_b)^{-1}$, baseload firms ‘over invest’ and generalist firms invest only in peaking capacities and:

$$\begin{aligned} k_b^S &= k_b^S(s_b, s_p + g, 0) \geq k_b^*/s_b \text{ and } k_b^G = 0 \\ k_p^S &= k_p^G = k_p^S(s_b, s_p + g, 0); \end{aligned}$$

- if $s_p \geq (s_b + g + 1)(r - r_p)(1 - r)^{-1}$, peaking firms ‘over invest’ and generalist firms only invest in baseload capacities and:

$$\begin{aligned} k_b^S &= k_b^G = k_b^S(s_b + g, s_p, 0) \\ k_p^S &= k_p^S(s_b + g, s_p, 0) \geq k_p^*/s_p \text{ and } k_p^G = 0; \end{aligned}$$

- else, quantities are those expressed in Proposition 1.

Proofs of both propositions are given in Appendix A. The term ‘overinvest’ is employed in a particular sense here. In the situations described, firms invest more than the first best quantity of a technology but less than a second best defined with a fixed quantity of the other technology. For instance, if the quantity of peak capacity is fixed at the equilibrium quantity, the quantity of baseload capacity that maximizes welfare is always higher than $s_b k_b^S$.

The ‘overinvestment’ results because of the limited access to a technology. It occurs if there are too many specialized firms of one type or too few generalist firms. If the number of generalist firms rises, overinvestment and specialization of generalist firms are less likely. As both conditions cannot be satisfied simultaneously, generalist firms always invest at least in one technology. This ‘overinvestment’ is different that the one analyzed by Zoetl (2008, Chap3), which is related to the strategic effect of baseload investment on short-term production.

It is worth emphasizing that specialization of generalist firms in one technology arises when there is overinvestment in the other technology. Therefore, the numbers of firms that invest in a technology are fewer than those likely (s_t instead of $s_t + g$) when the capacity of this technology is in excess. The conditions of overinvestment and specialization could be rewritten with the share of specialized firms in the industry included. Specialization of generalist firms to baseload or peak technology, respectively, occurs if:

$$\frac{s_p}{n+1} \geq \frac{r-r_p}{1-r_p} \text{ or } \frac{s_b}{n+1} \geq \frac{1-r}{1-r_b}.$$

These inequalities enable us to analyse the effect of r on specialization. Technology p can be viewed either as a midload- or a peak technology; the former case corresponds to a low r closed to r_b whereas the latter a high r close to r_p . Therefore, it appears from these equations that generalist firms are more likely to specialize in technology p if this technology is a peak technology, i.e. if r is large. And conversely, specialization into baseload technology is more likely when technology p is a midload technology.

Only generalist firms

In the particular case where there are only generalist firms, they invest in both types of technologies and from (14) the equilibrium quantities obtained are:

$$k_b^G(0, 0, n) = \frac{1}{n+1} k_b^* \text{ and } k_p^G(0, 0, n) = \frac{1}{n+1} k_p^*.$$

These expressions are similar to those found by von der Fehr and Harbord (1997); the oligopolistic quantities are qualitatively similar to those obtain in a linear Cournot model. The aggregate capacity chosen by firms is $n/(n+1)$ times the optimal one, and the technology mix is distorted by a similar proportion.

The interesting situations are those where some specialized firms are active. I briefly discuss the two cases where there is only one kind of specialized firm before devoting the last section to the case of an entirely specialized oligopoly and the welfare consequences of an increase of the number of specialized firms.

This last part is actually general because even if there are generalist firms they might specialize.

Generalist- and baseload firms

Originally, there was a concern regarding a potential lack of investment in peaking units, which was related to the effect of aggregate capacity on the frequency of rationing. The lack of investment in peaking units could be related to market power and the limited number of firms that invest in peakers¹³. Hence, one might consider that while all firms could invest in baseload plants only a subset could invest in peakers: $s_p = 0$.

In such a case, when generalist firms invest in both technologies the expressions are relatively simple because generalist firms crowd out the investment of baseload firms. From Equations (14), the following relation is satisfied by equilibrium quantities:

$$\begin{aligned} k_b^G &= \frac{1}{g+1} (k_b^* - s_b k_b^S) \\ k_p^G &= \frac{1}{g+1} k_p^*. \end{aligned}$$

Corollary 1. *If $s_b (r - r_b) \leq (g + 1) (1 - r)$ generalist firms invest in both types of capacity, and equilibrium quantities satisfy:*

$$\begin{aligned} s_b k_b^S + g k_b^G &= \frac{g}{g+1} k_b^* + \frac{s_b}{g+1} \frac{1 - r_b}{n+1} \\ g k_p^G &= \frac{g}{g+1} k_p^*. \end{aligned}$$

Generalist- and peaking firms

The previous situation could be reversed when only few firms have access to baseload technology. This could be the case where only few firms master nuclear technology while development of competition waits upon investment in CCGT plants, a technology open to many more firms.

¹³For instance, Joskow and Tirole (2007) made this assumption when they analyse underinvestment in peakers with two regulations: price caps and capacity payment.

Generalist firms specialize when there are numerous firms that invest in peak technology. Otherwise, expressions of aggregate- and baseload capacities are simple and comparable to standard linear Cournot quantities.

Corollary 2. *If $s_p(1-r) \leq (g+1)(r-r_p)$, aggregate equilibrium quantities are:*

$$gk_b^G = \frac{g}{g+1}k_b^* \text{ and } g(k_b^G + k_p^G) + s_pk_p^S = \frac{n}{n+1}k^*$$

In that case the distortion of investment is simply related to respective numbers of firms. Because the incentives to invest in aggregate capacity of a generalist firm are similar to the incentives of a peaking firm, expressions are simpler than in the former case.

4. Number of firms and welfare

I consider here a fully specialized industry where $g = 0$ and $n = s_p + s_b$ and examine the welfare consequences of an increased number of firms of one type, the other type being fixed.

When firms are specialized, the number of firms with access to a particular technology influences both aggregate capacity and the technology mix. A larger number of firms always increases the aggregate capacity but can amplify the distortion of the technology mix. These effects imply that the number of firms that can invest in either technology does not have a monotonic effect on welfare.

Let us consider first that the number of baseload firms, s_b , is fixed. An increase in the number of peaking firms increases the quantities of aggregate capacity and peak capacity but reduces the quantity of baseload capacity. Therefore, even if aggregate capacity tends toward the optimal there is, as s_p grows, a loss because of the distortion of the technology mix. More precisely, welfare is quasi-concave with respect to the number of peaking firms, and there is an optimal number of peaking firms.

Some calculations (cf appendix B) yield the following derivative of aggregate

quantities with respect to the number of firms:

$$\begin{aligned}\frac{\partial}{\partial s_p} (s_b k_b^S) &= -\frac{1}{A} s_b \frac{v - c_p}{v - c_b} k_p^S \\ \frac{\partial}{\partial s_p} (s_p k_p^S) &= \frac{1}{A} (s_b + 1) k_p^S\end{aligned}$$

Abstracting from the integer constraint, the effect of a change of s_p on welfare is obtained by derivation of (5):

$$\frac{dW}{ds_p} = \left[\frac{\partial W}{\partial k_b} + \frac{\partial W}{\partial k} \right] \frac{\partial}{\partial s_p} (s_b k_b^S) + \frac{\partial W}{\partial k} \frac{\partial}{\partial s_p} (s_p k_p^S),$$

then, replacing partial derivatives by expressions (6) and (7) and inserting first order conditions (9) and (10) gives:

$$\frac{dW}{ds_p} = (v - c_b) k_b^S \frac{\partial}{\partial s_p} (s_b k_b^S) + (v - c_p) k_p^S \frac{\partial}{\partial s_p} (s_p k_p^S). \quad (15)$$

A rise in the number of peaking firms has opposite effects on baseload- and peak capacities, whether the gain from the increase in peak capacity compensates the loss from the decrease in baseload capacity depends on the relative numbers of the two types of firms.

Proposition 3. *For any s_b , welfare is quasi-concave with respect to s_p , and attains its maximum at*

$$s_p^*(s_b) = \frac{v - c_b}{s_b (\delta - \Delta)} [1 - r_b + (s_b + 1)^2 (r_b - r_p)].$$

The proof is in Appendix C. The reduction in welfare is related to the increase in production cost due to the technology mix distortion. A question that naturally arises is the manner this loss is shared between consumers and firms. At first, it is unclear whether net consumers' surplus increases with an increase of peaking firms. Higher aggregate capacity raises consumers' surplus but lower baseload capacity increases electricity prices in some demand states and leads to a decrease of consumers' surplus.

Net consumer surplus is given by (4), and the effects of an increase of each type of technology are:

$$\begin{aligned}\frac{dCS}{dk_b} &= \frac{\partial CS}{\partial k_b} + \frac{\partial CS}{\partial k} = \frac{1}{X} [\delta k_b + (v - c_p)(k_b + k_p)] \\ \frac{dCS}{dk_p} &= \frac{CS}{\partial k} = \frac{1}{X} (v - c_p)(k_b + k_p).\end{aligned}$$

Therefore, an increase in the number of peaking firms modifies net consumers' surplus of:

$$\begin{aligned}\frac{dCS}{ds_p} &= (v - c_p) \frac{s_b}{AX} k_p^S \left[(s_b k_b^S + s_p k_p^S) \left(\frac{s_b + 1}{s_b} - \frac{v - c_p}{v - c_b} \right) - s_b k_b^S \frac{\delta}{v - c_b} \right] \\ &= (v - c_p) \frac{s_b}{AX} k_p^S \left[s_p k_p^S \left(\frac{s_b + 1}{s_b} - \frac{v - c_p}{v - c_b} \right) + k_b^S \right].\end{aligned}$$

Thus, consumers' surplus always increases when an additive peaking firm is active. The welfare loss is entirely supported by firms. However, consumers pay electricity at a higher price for some level of demand because of the lowering of baseload capacity but the overall increase of available capacity this loss.

If a similar analysis is carried out on the number of baseload firms, similar results are obtained. If one more firm gains access to baseload technology, the quantity of baseload capacity increases whereas the quantity of peak capacity reduces. Even if the aggregate capacity increases, welfare is not monotonic and an optimal number of baseload firms exist.

Proposition 4. *For any s_p , welfare is quasi-concave with respect to s_b , and attains its maximum at*

$$s_b^*(s_p) = \frac{1}{s_p} \frac{1}{(r_b - r_p)} \left[1 - r_p + (s_p + 1)^2 \frac{\delta - \Delta}{v - c_p} \right].$$

The proof of this proposition is given in appendix D. The effect on aggregate welfare of an increase of baseload- or peaking firms produces similar qualitative properties. Likewise, consumers always benefit from the entry of a baseload firm because it raises both the aggregate quantity of capacity and the quantity of baseload capacity leading to less rationing and lower prices.

Those results can be compared to the analyses of a usual Cournot oligopoly with heterogeneous firms¹⁴. Within this standard framework, the addition of an inefficient firm increases the aggregate production but lowers the production of efficient ones, the overall effect could be negative because production is reallocated from efficient firms to the inefficient. Here, all firms are efficient, in the sense that both technologies are utilized at the first best optimum; nevertheless, an additional firm of any type can reduce welfare by modifying the technology mix in the wrong direction. The welfare loss is not related to an inefficiency of new firms but to a disequilibrium between the two types.

The welfare loss resulting from an excess entry is entirely supported by firms. Consumers are always made better off by an increase of the number of firms of any type. If consumers' surplus and firms' profit were weighted differently in the social welfare function—and competition agencies tend to weigh consumer surpluses more than firm profits—the optimal number of firms would be increasing with respect to the weight of consumers' surplus, and for a sufficiently large weight any increase in the number of firms of any type would increase welfare¹⁵.

This analysis has several policy implications. Given that it is far from obvious that all firms can use all technologies, some regulation might be aimed at promoting investment in some technologies. First, for any industry configurations it is theoretically feasible to decentralize the optimum technology mix by subsidizing investment. Such subsidies are justified solely by market power in the framework developed here. Two different subsidies should be established: one per technology. If the entry into a technology is unhindered, it is sufficient to regulate the other. Evidently, limiting entry in the “standard” technology is a second best solution that should be avoided. The specific technology with a limited access should be regulated and two types of regulation can be visual-

¹⁴In a recent paper Corchón (2008) presented a complete analysis of welfare loss with Cournot competition.

¹⁵Such an approach ignores the possibility of monetary transfers. If these are feasible at no cost the welfare function (5) is the only relevant.

ized. First, this technology could be subsidized or directly regulated to correct the investment deficit. Second, entry barriers could be reduced and particularly institutional ones.

Policies actually implemented to reduce access costs to some technologies could be interpreted on following these lines. Concerning nuclear power, in France it seems that the first type of regulation is favored, and implemented through central planning and direct control of the dominant and historic operator (Electricité de France), although, participation of other companies (Enel and Gdf-Suez) in new plant development is a means of spreading knowledge and promoting entry of competitors. In UK, the White Paper on nuclear (DBEER, 2008) stresses the UK government's reliance on markets and private investors while insisting on the need to clarify procedures for new nuclear investment. The procedures entailed are: licensing, decommissioning, waste management and insurance for nuclear incidents. Such a clarification should facilitate entry but could be also interpreted as indirect subsidies of nuclear. Similar policies are implemented in the US.

Finally, technological evolution can also reduce entry barriers by standardizing a technology. For instance, the development of small-scale standardized nuclear reactors could help develop competition and consequently investment. Such technological development is interrelated with liberalization; it should facilitate competition in generation, and is driven by the potential market created by liberalization. Indeed, the development of CCGT that justifies liberalization was also accelerated by first attempts to introduce competition in electricity generation in the US in the 80s (Winskel, 2002).

5. Discussion

This section is devoted to the discussion on the assumptions of the model: the constant surplus v and the load duration curve.

5.1. The demand function

It was assumed that v represents gross consumer surplus from electricity consumption. The most direct method of justifying this approach and the subsequent welfare analysis is to consider that demand x is derived from a time-varying utility function of the form $U(q, R, x) = \phi(q, x) + R$ where R is consumer revenue net of electricity expenses, q is the quantity of electricity consumed and x represents the demand state. The gross surplus from the consumption of q in a demand state x is

$$\phi(q, x) = \begin{cases} vq & \text{if } q < x \\ vx & \text{otherwise} \end{cases} .$$

The demand is sensitive to the retail price; if it is lower than v the demand varies and x represents both the demand state and the quantity demanded. It is not necessary to introduce the retail price because retail expenses are only a transfer from consumers to retailers that is welfare-neutral. If entry were free in the retail market, the retail price would be the average of the wholesale prices weighted by quantities consumed; this price would be lower than v and net consumer surplus similar to (4). However, with this justification, it would actually be appropriate to consider that consumers are sensitive to the wholesale price and price rationed.

Furthermore, a linear cost of rationing could be added to the framework. This cost could be related to the negative effects of unforeseen rationing. If consumers bear a cost αx when rationed in a state x their aggregate surplus would be:

$$\begin{aligned} S(k) &= \frac{1}{X} \int_0^k vxdx + \frac{1}{X} \int_k^X vk - \alpha(x - k) dx \\ &= \frac{1}{X} \int_0^k vxdx + \frac{1}{X} \int_k^X (v + \alpha)k - \alpha x dx. \end{aligned}$$

With such a specification the analysis would be similar if v is replaced by $v + \alpha$. These specifications amount to assuming a linear cost of rationing as Crew and Kleindorfer (1976) do.

Another suitable specification to consider is a reduced form of consumers' gross surplus and a third technology. The third technology is a peaking technology with variable cost $c_3 > c_p$ and capacity cost $I_3 < I_p$. The capacity of technology 3 is chosen by a competitive fringe—i.e. there is free entry in this technology. Consumers' gross surplus in a state x could be expressed in a reduced form $V(x, y)$ —possibly related to a fixed retail price—where x denotes the quantity demanded and y the quantity offered.¹⁶ $V(x, y)$ is maximized at $y = x$ and if $y < x$ consumers are rationed and $\partial V(x, y)/\partial y$ is the VoLL.¹⁷ Thus, welfare as a function of k_b , $k = k_b + k_p$ and $K = k_b + k_p + k_3$ is

$$\begin{aligned}
W = & \frac{1}{X} \left[\int_0^K V(x, x) + \int_K^X V(x, K) \right. \\
& - \int_0^{k_b} c_b x dx - \int_{k_b}^k c_p x dx - \int_k^K c_3 x dx - \int_K^X c_3 K dx \\
& \left. + \int_{k_b}^X \delta k_b dx + \int_k^X (c_3 - c_p) k dx \right] - \Delta k_b - (I_p - I_3)k - I_3 K
\end{aligned} \tag{16}$$

The cost parameters, c_3 and I_3 , and consumers' surplus solely determine the optimal total quantity of capacity K^* :

$$I_3 = \frac{1}{X} \int_{K^*}^X \frac{\partial V(x, K^*)}{\partial y} - c_3 dx \tag{17}$$

The optimal quantity of baseload capacity k_b^* and of baseload plus peak capacity

¹⁶The notation $V(x, y)$ is somewhat abusive, a more precise notation would be $V(q(x), x, y)$ where q is the quantity demanded and x the state of the world that determines—together with the retail price—the demand $q(x)$. The demand varies because some exogenous parameters do. However, x can be chosen so that $q(x) = x$.

¹⁷For instance, if $S(x)$ is gross consumer surplus, with perfect rationing the function V would be $V(x, y) = S(y)$ if $y < x$ and $S(x)$ otherwise; this rationing method has been used by Brown and Johnson (1969). With random rationing (Visscher, 1973), $V(x, y) = S(x)y/x$ if $y < x$. These two kinds of rationing methods are equivalent with the linear assumption— $S(x) = vx$ —used in the present paper. The interested reader should refer to Joskow and Tirole (2007) for a more detailed discussion on rationing costs.

k^* satisfy

$$\Delta = \frac{1}{X} \int_{k_b^*}^X \delta dx, \text{ and } I_p - I_3 = \int_{k^*}^X c_3 - c_p dx.$$

To decentralize this investment with perfect competition, the regulator should fix the price at the variable cost of the marginal technology when no rationing is required, that is, at c_b when $x < k_b$ at c_p when $k_b < x < k$ and at c_3 when $k < c_3 < K$, and at the VoLL ($\partial V/\partial y$) when rationing occurs, that is, when x is higher than the total capacity K . In the former case, short-term revenue of capacities of type 3 is null and in the latter case all capacities are used and received VoLL- c_t , $t = b, p, 3$. For any k chosen by strategic firms, k_3 is such that the total revenue of a capacity of type 3 is equal to its cost because of perfect competition. Therefore, equation (17) is satisfied and the competitive fringe invests in $k_3 = K^* - k$. For a strategic producer $i = 1..n$ that invests in capacity of type b and p , its expected revenue during rationing for a capacity of type b and p is I_3 because (17) is satisfied. Its long-term profit is

$$\begin{aligned} \pi^i &= \frac{1}{X} \int_{k_b}^k \delta k_b^i dx \\ &+ \frac{1}{X} \int_k^X [(c_3 - c_p)k^i + \delta k_b^i] dx \\ &- \Delta k_b^i - (I_p - I_3) k^i. \end{aligned} \quad (18)$$

This expression is similar to (2). Similarly, welfare can be written as a function of k_b and k by replacing K by K^* into (16). Thus, the derivatives of welfare with respect to k and k_b are similar to (6) and (7) if v , and I_t for $t = b, p$ are respectively replaced by c_3 and $I_t - I_3$. Therefore, with this specification, the analysis can be reproduced by replacing v by c_3 and capacity cost I_t by $I_t - I_3$ for $t = b, p$.

A third possibility,¹⁸ is to consider that the objective of the regulator is to satisfy the load with the three technologies $b, p, 3$ and without rationing. In this case, for any capacity $k = k_b + k_p$ the remaining quantity $X - k$ should be added with the third technology. To ensure that it is the case the regulator should set

¹⁸I thank an anonymous referee for suggesting this possibility.

the wholesale price at the variable cost of the marginal technology and set a subsidy I_3 on all capacities. This is the framework used by Arellano and Serra (2007). With such scheme, the profit of strategic firms is similar to (18) and the total cost of satisfying the load is:

$$C(k_b, k) = \frac{1}{X} \left[\int_0^{k_b} c_b x dx + \int_{k_b}^k c_p x dx + \int_k^X c_3 x dx - \int_{k_b}^X \delta k_b dx - \int_k^X (c_3 - c_p) k dx \right] + \Delta k_b + (I_p - I_3)k + I_3 X.$$

As previously, the analysis of sections 3 and 4 can be used by replacing v by c_3 and capacity cost I_t by $I_t - I_3$ for $t = b, p$.

5.2. Load curve

The load duration curve used is a very rough simplification of a real one. It was used to derive explicit formula of quantities and other results. A more realistic representation that does not significantly modify results is to consider that demand x is uniformly distributed across the set $[x^-, x^+]$ with $x^+ - x^- = X$. With this distribution of demand, the first best optimum is depicted on figure (4), baseload capacity is raised by x^- whereas peak capacity is unchanged compared to quantities (8).

Insert figure 4

Figure 4: Load curve and optimal investment

Firms' choices are also modified; they all invest in greater quantities with this load duration curve. For instance, the equilibrium individual capacity of a specialized baseload firm becomes:

$$k_b^S = \frac{X}{A} \left[(g+1) \left(\frac{x^+}{X} - r_b \right) + s_p \frac{\delta - \Delta}{v - c_b} \right]$$

Moreover, other quantities change similarly. This change is not as benign as it might seem at first glance. The main consequence of this change is that the profit of a baseload or a generalist firm is not differentiable at $k_b = x^-$; hence,

they might invest in precisely the minimal quantity x^- of baseload technology and no more. In such a case, in the short-term, the baseload technology is never marginal and the price of electricity is never fixed at c_b .

Such situations arise if there are few baseload- and generalist firms and $X = x^+ - x^-$ is small enough, *i.e.* the load duration curve is sufficiently flat. Else, the analysis is not modified and results still hold.

6. Conclusion

In this paper, I have analyzed investment by strategic firms in a simple electricity market that was assumed perfectly regulated (or competitive) in the short-term. This framework has allowed us to understand how firms have the incentive to both reduce the aggregate quantity of capacity and distort the technology mix. The way the technological mix is distorted is related to the industry configuration. If one technology is not accessible to all firms, there may be an overinvestment in the other technology.

While both technologies are efficient, a firm with access to both technologies does not invest in either of them if there are too many specialized firms. If all firms are specialized, an increase in the number of firms of one type could decrease welfare. Therefore, if the access to one technology is limited, investment in this technology should be regulated otherwise the development of competition via a unique technology is questionable and can decrease welfare.

Two major extensions are visualized: the first would be to endogenize specialization by adding an initial stage where firms decide which technology to master, the second would be to analyse capacity markets and capacity payments that are implemented to correct investment incentives in electricity markets.

References

Arellano, M., Serra, P. A model of market power in electricity industries subject to peak load pricing. *Energy Policy* 2007; 35 (10); 5130–5135.

- Brown, G. and Johnson, M. B. Public Utility Pricing and Output Under Risk. *American Economic Review* 1969; 59 (1); 119–128.
- Bushnell, J., Mansur, E., Saravia, C. Vertical Arrangements, Market Structure and Competition: An analysis of Restructured U.S. Electricity Markets. *American Economic Review* 2008; 98 (1); 237–266.
- Chao, H. Peak load pricing and capacity planning with demand and supply uncertainty. *The Bell Journal of Economics* 1983; 14(1); 179–190.
- Corchón, L. Welfare losses under Cournot competition. *International Journal of Industrial Organization* 2008; 26(5);1120–1131.
- Cramton, P., Stoft, S. The Convergence of Market Designs for Adequate Generating Capacity. A White Paper for the Electricity Oversight Board 2006.
- Crew, M. A., and Kleindorfer, P. R. Peak Load Pricing with a diverse Technology. *The Bell Journal of Economics* 1976; 7(1); 207–231.
- D.G.Competition Report on Energy Sector Inquiry. European Commission 2007.
- Department for Business, Enterprise & Regulatory Reform. Meeting the Energy Challenge. A White Paper on Nuclear Power. Crown Copyright 2008.
- Fabra, N., von der Fehr, N., de Frutos, M. Investment Incentives and Auction Design in Electricity Markets. CEPR Discussion Paper No. DP6626; 2006.
- Gabszewicz, J., Poddar, S. Demand Fluctuations and Capacity Utilization under Duopoly. *Economic Theory* 1997;10 (1); 131–146.
- Green, R.J. and Newbery, D.M. Competition in the British Electricity Spot Market. *The Journal of Political Economy* 1992; 100 (5); 929–953.
- IEA Power generation Investment in Electricity Markets. OECD/IEA, Paris; 2006.
- Joskow, P. L. Competitive Electricity Markets and Investment in New Generating Capacity. MIT, W.P. 06-009; 2006.

- Joskow, P. L., Tirole, J. Reliability and Competitive Electricity Markets. *RAND journal of economics* 2007;38 (1); 60–84.
- Murphy, F., Smeers, Y. Generation Capacity Expansion in Imperfectly Competitive Restructured Electricity Markets. *Operations research* 2005; 53 (4); 646.
- Newbery, D. M. Competition, Contracts, and Entry in the Electricity Spot Market. *Rand Journal of Economics* 1998; 29 (4); 726–749.
- Nuttal, W.J. Nuclear renaissance: technologies and policies for the future of nuclear power. Taylor & Francis; 2005.
- Reynolds, S., Wilson, B. Bertrand–Edgeworth Competition, Demand Uncertainty, and Asymmetric Outcomes. *Journal of Economic Theory* 2000; 92 (1); 122–141.
- Roques, F., Nuttal, W., Newbery, D., de Neufville, R., Connors, S. Nuclear power: A hedge against uncertain gas and carbon prices? *The Energy journal* 2006; 27 (4); 1–23.
- Finon, D. and Roques, F. Financing arrangements and industrial organisation for new nuclear build in electricity markets. Working Paper Larsen 13; 2008.
- Stoffaës, C. Entre monopole et concurrence: la régulation de l'énergie en perspective historique. Editions PAU; 1994.
- Congress of The United States of America; Energy Policy Act of 2005.
- Visscher, M., L. Welfare-Maximizing Price and Output with Stochastic Demand: Comment. *American Economic Review* 1973; 63(1); 224–229.
- von der Fehr, N.-H., Harbord, D. Capacity Investment and Competition in Decentralised Electricity Markets. Department of Economics University of Oslo, memorandum 27/97; 1997.

Wolfram, C. Measuring Duopoly Power in the British Electricity Spot Market. *American Economic Review* 1999; 89(4); 805–826.

Winkel, M. When Systems Are Overthrown: The 'Dash for Gas' in the British Electricity Supply Industry. *Social Studies of Science* 2002; 32(4); 563–598

Zoetl, K. Investment Incentives under Uncertainty. Ph.D. thesis, CORE, Louvain la Neuve; 2008.

Appendix

The following notation is introduced to facilitate exposition of calculations:

$$\gamma = \frac{v - c_p}{v - c_b}$$

A. Proof of propositions 1 and 2

Let assume that an equilibrium exists. It is clear from symmetry that all firms of a particular type invest in similar quantities at equilibrium. To establish the existence and unicity of equilibrium, I consider the three subcases where generalist firms invest in both types of technologies or specialize in peak- or baseload and show that these subcases cannot coexist.

Equilibrium quantities are: k_b^S, k_p^S and k_b^G, k_p^G which are, respectively, the individual capacity of a baseload firm, the individual capacity of a peaking firm, and the peak- and baseload capacities of a generalist firm.

1. Let assume that at equilibrium generalist firms invest in both technology types.

First, I simplify the problem in order to obtain a simple linear system. As the individual aggregate quantity of a generalist firm is equal to the individual quantity of a peaking firm: $k_b^G + k_p^G = k_p^S$, the problem is already limited to three quantities: k_b^S, k_p^S , and k_b^G .

The first order condition of peaking firms (10) is written as:

$$s_b k_b^S + (s_p + g + 1) k_p^S = X(1 - r_p). \quad (19)$$

Furthermore, with the first order conditions (9) of baseload firms and (11) of the baseload capacity of generalist firms it appears that $k_b^S = k_b^G + \gamma k_p^G$; hence, by introducing the relation $k_p^G = k_p^S - k_b^G$, the capacity of baseload firms becomes: $k_b^S = (1 - \gamma) k_b^G + \gamma k_p^S$. This relationship could be used to establish a second relation between the capacity of a baseload firm and one of a peaking firm:

$$(s_b + g + 1) k_b^S + \gamma s_p k_p^S = X(1 - r_b). \quad (20)$$

Thus, the two quantities k_b^S, k_p^S satisfy equations (19) and (20):

$$\begin{bmatrix} s_b + g + 1 & \gamma s_p \\ s_b & s_p + g + 1 \end{bmatrix} \begin{bmatrix} k_b^S \\ k_p^S \end{bmatrix} = X \begin{bmatrix} 1 - r_b \\ 1 - r_p \end{bmatrix}$$

The determinant of the matrice is:

$$A(s_b, s_p, g) \stackrel{def}{=} (s_b + g + 1)(s_p + g + 1) - \gamma s_b s_p$$

It is strictly positive so there is a unique solution of the system. Finally, some calculations lead to:

$$k_b^S = X \frac{1}{A} \left[(g + 1)(1 - r_b) + s_p \frac{\delta - \Delta}{v - c_b} \right] \quad (21)$$

$$k_p^S = X \frac{1}{A} [(g + 1)(1 - r_p) + s_b(r_b - r_p)] \quad (22)$$

For generalist firms, the individual quantity of baseload capacity can be obtained with the relation $(1 - \gamma) k_b^G = k_b^S - \gamma k_p^S$ by noting that $(1 - r_b) - \gamma(1 - r_p) = (1 - \gamma)(1 - r)$. The peak capacity then is simply $k_p^S - k_b^G$:

$$k_b^G = X \frac{g}{A} [(g + s_p + 1)(1 - r) - s_b(r - r_b)] \quad (23)$$

$$k_p^G = X \frac{g}{A} [(g + s_b + 1)(r - r_p) - s_p(1 - r)] \quad (24)$$

Hence, if there is an equilibrium with generalist firms that invest in strictly positive quantities of both types of capacities the equilibrium quantities are defined by equations (21), (22) and for generalist firms by (23) and (24).

Furthermore, if $s_b (r - r_b) \leq (g + s_p + 1) (1 - r)$ and $s_p (1 - r) \leq (g + s_b + 1) (r - r_p)$, the quantities defined by these equations described equilibrium strategies, that is each firm's profit is concave and maximum at these quantities.

2. Let assume that generalist firms only invest in peak capacities at equilibrium.

Equilibrium quantities could be found from the above calculations by replacing s_p by $g + s_p$ and g by 0. Thus, if such an equilibrium exists, it is fully described by the quantities:

$$\begin{aligned} k_b^S &= X \frac{1}{A(s_b, s_p + g, 0)} \left[1 - r_b + (s_p + g) \frac{\delta - \Delta}{v - c_b} \right] \\ k_p^S &= X \frac{1}{A(s_b, s_p + g, 0)} [1 - r_p + s_b (r_b - r_p)] \\ k_b^G &= 0 \text{ and } k_p^G = k_p^S \end{aligned}$$

These quantities describe an equilibrium only if a generalist firm has an incentive not to invest in baseload capacity, this is the case if the aggregate baseload capacity is above the first best optimal quantity: $s_b k_b^S > k_b^* = X(1 - r)$, and this inequality is equivalent to:

$$s_b \geq (g + s_p + 1) (1 - r) / (r - r_b).$$

3. Let assume that generalist firms invest only in baseload capacities at equilibrium.

If at equilibrium generalist firms invest only in baseload capacity, equilibrium strategies are:

$$\begin{aligned} k_b^S &= X \frac{1}{A(s_b + g, s_p, 0)} \left[1 - r_b + s_p \frac{\delta - \Delta}{v - c_b} \right] \\ k_p^S &= X \frac{1}{A(s_b + g, s_p, 0)} [1 - r_p + (s_b + g) (r_b - r_p)] \\ k_b^G &= k_b^S \text{ and } k_p^G = 0. \end{aligned}$$

These quantities describe an equilibrium only if the aggregate quantity of peak capacity is greater than the first best quantity. This is so if and only if:

$$s_p \leq (g + s_b + 1) (r - r_p) / (1 - r).$$

Propositions 1 and 2 are directly obtained from these results.

B. Quantities of capacity derivatives

I establish expressions for quantities derivatives with respect to the number of firms s_b and s_p when there are only specialized firms. In this case, the first order conditions (19) and (20) are:

$$(s_b + 1)k_b^S + \gamma s_p k_p^S = X(1 - r_b), \text{ and } s_b k_b^S + (s_p + 1)k_p^S = X(1 - r_p) \quad (25)$$

- Derivatives of quantities with respect to s_p :

Derivation of (25) with respect to s_p gives:

$$\begin{aligned} (s_b + 1) \frac{\partial s_b k_b^S}{\partial s_p} + s_b \gamma \frac{\partial s_p k_p^S}{\partial s_p} &= 0 \\ s_p \frac{\partial s_b k_b^S}{\partial s_p} + (s_p + 1) \frac{\partial s_p k_p^S}{\partial s_p} &= k_p^S \end{aligned}$$

This leads to the following expressions of derivatives:

$$\frac{\partial s_b k_b^S}{\partial s_p} = -\frac{s_b}{A} \gamma k_p^S \text{ and } \frac{\partial s_p k_p^S}{\partial s_p} = \frac{s_b + 1}{A} k_p^S. \quad (26)$$

- Derivatives of quantities with respect to s_b :

From first order conditions (25):

$$\begin{aligned} (s_b + 1) \frac{\partial s_b k_b^S}{\partial s_b} + s_b \gamma \frac{\partial s_p k_p^S}{\partial s_b} &= k_b^S \\ s_p \frac{\partial s_b k_b^S}{\partial s_b} + (s_p + 1) \frac{\partial s_p k_p^S}{\partial s_b} &= 0 \end{aligned}$$

This give the expression of quantities evolution with respect to the number s_b :

$$\frac{\partial s_b k_b^S}{\partial s_b} = \frac{s_p + 1}{A} k_b^S \text{ and } \frac{\partial s_p k_p^S}{\partial s_b} = -\frac{s_p}{A} k_b^S \quad (27)$$

C. Proof of proposition 3

I relax the integer constraint and consider the effect of change of s_p on welfare. Replacing the expressions of capacities derivatives (26) into the expression (15) found in the main text and factorizing gives:

$$\frac{dW}{ds_p} = (v - c_p) k_p^S [(s_b + 1) k_p^S - k_b^S] / AX.$$

It is unclear whether welfare is concave or not (generally, it is not) but it is quasi-concave because its derivative is nil only once and strictly positive (resp. negative) for smaller (resp. greater) values of s_p . This can be demonstrated with the formula of equilibrium quantities in Proposition 1:

$$\begin{aligned} A [(s_b + 1) k_p^S - k_b^S] / X &= (s_b + 1) [1 - r_p + s_b (r_b - r_p)] \\ &\quad - s_b [1 - r_b + s_p (\delta - \Delta) / (v - c_b)]. \end{aligned}$$

As this expression is decreasing with respect to s_p , welfare is quasi concave. Furthermore, it is maximum at:

$$s_p^* = \frac{1}{s_b} \frac{v - c_b}{\delta - \Delta} \left[(1 - r_b) + (s_b + 1)^2 (r_b - r_p) \right].$$

D. Proof of proposition 4

This proof is close to the previous one. I relax the integer constraint and examine the effect of a change of s_b on welfare. The derivative of welfare with respect to s_b is

$$\frac{dW}{ds_b} = \frac{v - c_b}{X} k_b^S \frac{\partial s_b k_b^S}{\partial s_b} + \frac{v - c_p}{X} k_p^S \frac{\partial s_p k_p^S}{\partial s_b}.$$

Replacing the expressions (27) of capacities derivatives into above formula and factorizing gives:

$$\frac{dW}{dn} = (v - c_b) k_b^S [(s_p + 1) k_b^S - \gamma s_p k_p^S] / AX,$$

and with the formula of equilibrium quantities (Proposition 1):

$$\begin{aligned} [(s_p + 1) k_b^S - \gamma s_p k_p^S] / X &= 1 - r_b + s_p (s_p + 2) (\delta - \Delta) / (v - c_b) \\ &\quad - s_p s_b (r_b - r_p) \gamma. \end{aligned}$$

As this expression is decreasing with respect to s_b , welfare is quasi-concave and maximum at:

$$s_b^*(s_p) = \frac{1}{s_p} \frac{1}{(r_b - r_p)} \left[(1 - r_p) + (s_p + 1)^2 \frac{\delta - \Delta}{v - c_p} \right].$$