Energy Policy 105 (2017) 562–570

Contents lists available at ScienceDirect

Energy Policy

journal homepage: www.elsevier.com/locate/enpol

Rationales for capacity remuneration mechanisms: Security of supply externalities and asymmetric investment incentives

Jan Horst Keppler

Université Paris-Dauphine, Place du Maréchal de Lattre de Tassigny, 75775 Paris Cedex 16, France

ARTICLE INFO

Keywords:
Electricity markets
Capacity remuneration mechanisms
VOLL pricing
Investment
Security of supply externalities
Asymmetric investment incentives

ABSTRACT

Economics so far provides little conceptual guidance on capacity remuneration mechanisms (CRM) in deregulated electricity markets. Ubiquitous in real-world electricity markets, CRMs are introduced country by country in an ad hoc manner, lacking the theoretical legitimacy and the conceptual coherence enabling comparability and coordination. They are eyed with suspicion by a profession wedded to a theoretical benchmark model that argues that competitive energy-only markets with VOLL pricing provide adequate levels of capacity. While the benchmark model is a consistent starting point for discussions about electricity market design, it ignores the two market failures that make CRMs the practically appropriate and theoretically justified policy response to capacity issues. First, energy-only markets fail to internalize security-of-supply externalities as involuntary curbs on demand under scarcity pricing generate social costs beyond the private non-consumption of electricity. Second, when demand is inelastic and the potential capacity additions are discretely sized, investors face asymmetric incentives and will underinvest at the margin rather than overinvest. After presenting the key features of the theoretical benchmark model, this paper conceptualizes security of supply externalities and asymmetric investment incentives and concludes with some consideration regarding design of CRMs.

1. Introduction: the market failures at the heart of the capacity issues

This paper aims at providing coherent theoretical rationales for the introduction of capacity remuneration mechanisms (CRMs) in deregulated electricity markets. In other words, at current levels of demand elasticity there exist clearly identifiable market failures in the great majority of energy-only electricity markets that, if unaddressed, will lead to socially sub-optimal levels of capacity. In these cases, attaining optimal capacity requires regulatory intervention providing added incentives for capacity provision either through price or quantity instruments or a combination of the two, as, for instance, in capacity markets operating under a system-wide cap set by the regulator.

The standard argument that marginal cost pricing in combination with scarcity pricing at VOLL will provide optimal levels of capacity remains a conceptually coherent benchmark. However, the theoretical benchmark model neglects the two market failures that will be analysed below: security-of-supply externalities and asymmetric investment incentives in markets for non-storable goods with discretely sized equipment. Real-world CRMs have thus been obliged to develop with scant help from the theoretical literature. This has created a wide divergence of views at a time when the introduction of large amounts of variable renewables lends new urgency to the capacity issue, in particular in European electricity markets. The present paper thus aims at filling the gap between theory and practice in the area of optimal capacity provision.

There exist a number of previous attempts to come to terms with this contradiction between theory and practice. They broadly fall into three categories. Authors such as Oren (2003), De Vries and Hakvoort (2004) or Salies et al. (2007) identify potential shortcomings in energy-only markets and define security of electricity supply as a public good. However, these shortcomings relate either to indelicate behaviour by market participants (“under-reporting of true preferences”) or transaction costs in energy-only markets. Neither argument upholds closer scrutiny as a general case for CRMs.

The second strand of literature is organised around the notion of incomplete markets, see, for instance, Vázquez et al. (2001, 2002), Battle et al. (2006) or López-Peña et al. (2009). These authors correctly

---

E-mail address: jan-horst.keppler@dauphine.fr.

http://dx.doi.org/10.1016/j.enpol.2016.10.008
Received 9 June 2016; Received in revised form 24 September 2016; Accepted 8 October 2016

identifying the capacity issue as due to a “missing market”. The capacity issue does indeed arise due to the fact that no market exists to internalise security of supply externalities. However, they spend little time on the identification of the market failure as such, but concentrate on the design of the additional market that will internalise the failure.[1] Thus also the second strand fails to deliver a justification for the need of the decisive regulatory intervention that precedes the creation and operation of forward capacity markets.

A third strand of literature is constituted by economists and electricity market experts with backgrounds in industrial and institutional economics, concentrating on the failure of electricity markets to fully remunerate the provision of optimal amounts of capacity, the “missing money” problem. “The fundamental source of the missing money problem is the failure of spot energy and operating reserve markets to perform in practice the way they are supposed to perform in theory,” writes Joskow (2008, 166). A particular issue in this context is constituted by the price caps imposed by regulators during scarcity hours, often seen at the origin of less than sufficient capacity remuneration (Cramton and Stoft, 2006, 8; Joskow, 2008, 164; Finon and Pignon, 2008, 145; Finon and Roques, 2013, 112). However, Joskow also points out that empirically price caps are rarely the binding constraint (Joskow, 2008, p. 166) and cites several additional market imperfections related to institutional or informational shortcomings.

The present paper has benefitted from all three approaches cited, in particular the first and the third. In contrast to the first it provides a more complete notion of what constitutes an externality and hence a public good. In contrast to the third, it makes a general case for CRMs, in particular due to security of supply externalities, that is independent of the specific institutional failures of a given market or regulator that could be righted by an expert with superior knowledge.

Before discussing the fundamental market failures providing a pervasive rationale for CRMs, the present article sets out to reaffirm the conceptual validity of the theoretical benchmark model for energy-only markets with scarcity pricing at VOLL under the assumptions of perfect information and in the absence of externalities, market power and transaction costs. It will then conceptualise the two recurring features of real-world electricity markets, security of supply externalities and asymmetric investment incentives that challenge the conceptual benchmark model, which is ultimately too narrow a representation of electricity markets.

The fact that the case for capacity mechanisms ultimately depend on an externality argument also explains some of the difficulty that energy economics has had in organising a more systematic debate on the origins of the capacity issue. For methodological reasons, theoretical economics will always tend to disregard difficult to codify goods such as the security of electricity supply. However, the empirical pressure for CRMs requires addressing the capacity issue also on a conceptual level in a more definitive manner.

While security of supply externalities and asymmetric investment incentives exist, in principle, in all electricity markets, the magnitude of their impact depends crucially on structural factors such as the elasticity of demand and the size of generation equipment. There might thus exist electricity markets, where the tendency towards socially suboptimal levels of capacity might be too weak as to warrant dedicated regulatory intervention. In a system where storage is ample, the unit size of dispatchable generators is small and the elasticity of demand is high, it is unlikely that a CRM will be required.

The validity of the theoretical benchmark model for energy-only electricity markets thus depends on the presence, degree and precise form of the two market failures mentioned. Furthermore, as discussed in Keppler (2010) externalities always inscribe themselves in a dynamic of progressive internalisation. This is particularly true for capacity issues. In fact, CRMs have a tendency to progressively bring about a form of structural change – more storage, smaller unit sizes and more flexible demand – that will eventually reduce the tendency of energy-only markets to supply sub-optimal levels of capacity. Thus, while required in the majority of electricity systems today, CRMs have a tendency to render themselves obsolete over time.

However, before things might eventually get better, they are currently getting worse. In European electricity markets at least, the capacity issue is magnified by the decrease in average prices following the introduction of large amounts of variable renewable capacity. This has led to the decommissioning or mothballing of gas plants slated to work during periods of high demand. The shift of the load curve towards the right due to the influx of variable renewables (VRE) thus exacerbates the vulnerability of peak-load plants required to recover their capital costs during a small number of hours. Peakers have always been exposed to the stochastic nature of electricity demand. This effect is now doubled by the further increase in price volatility due to the equally stochastic nature of VRE production.

Price caps in energy-only markets contribute further to reducing privately provided capacity below socially optimal levels. While some administrative guidance is required for situations in which an inelastic short-term demand exceeds an inelastic short-term supply, current price caps are often unrealistically low. Even in the absence of the market failures discussed below, current price caps in the European market, for instance, are inconsistent with stated security of supply objectives. This can be easily demonstrated by looking at the French electricity market. The latter has the virtue of stating an explicit security of supply target, which is set at a level of three scarcity hours per year. The corresponding capacity will, at the margin, be provided by gas- or oil-fired combustion turbines. With prices on the French-German day-ahead market EPEX Spot capped at € 3 000 per MWh, VOLL pricing during three hours per year with a French peak demand of circa 100 GW will generate in an average year € 900 million. This however is not nearly enough to recuperate the missing money necessary to finance a peaking unit. The annualised capital cost of € 50 000 per MW for a combustion turbine would require revenues of € 5 billion during scarcity hours. Recuperating this amount would require at least 16 scarcity hours per year, far above the level deemed socially optimal.[2]

While these empirical considerations add urgency to the capacity issue, they do not affect the conceptual arguments of this article. Its structure is as follows. Section 2 will present the theoretical benchmark model according to which competitive energy-only markets in the absence of market or regulatory failures provide privately and socially optimal levels of capacity. It will also discuss its limits, in particular concerning the rarely fully spelled out details of scarcity pricing at VOLL. Section 3 will identify market failures in energy-only markets on the demand side in the form of security-of-supply externalities and the fundamental inability of consumers to properly hedge against system-wide security of supply risks. Section 4 will identify market failures on the supply side on the basis of the fact that discontinuities in electricity price formation will asymmetrically induce producers to underinvest rather than to overinvest in capacity, an effect that is exacerbated by uncertainty about the precise level of demand, risk aversion and the

---

[1] This approach comes under different headings such as “reliability options”, “price risk-hedging contracts”, “call options” or “forward reliability auctions”. They share the underlying idea that either market participants (decentralised approach) or the TSO (centralised approach) ensure by means of a call option the availability of the total maximum amount of electricity they will need at a fixed price (see Battle et al. (2006)). Since the overall cap is fixed by the regulator or the TSO, the security of supply externality is effectively internalised. The “market” element of this approach is limited to the efficient, least-cost provision of energy and capacity, not the system-wide capacity limit.

[2] In reality, the situation is even less satisfactory. Reaching the cap of € 3 000 per MWh during four hours in 2009 produced a political uproar and a serious questioning of the adequacy of liberalised electricity markets across the political spectrum.

discrete size of generation investments. Section 5 will conclude with some remarks on the necessarily dynamic nature of capacity mechanisms as instruments for internalising widely observable market failures of energy-only markets.

2. The theoretical benchmark model for optimal capacity provision in energy-only markets

Before spelling out the rationale for dedicated capacity remuneration mechanisms (CRMs) based on specifically identifiable market failures that prevent full cost recovery at socially optimal levels of security of supply, it is useful to recap briefly the "benchmark model" (Léautier, 2013) of optimal capacity provision and full cost recovery in energy-only electricity markets (see Boiteux (1949, 1960), Stoft (2002) and Joskow (2007), for expositions). The benchmark model applies both to a monopoly provider of electricity aiming at the maximisation of social welfare as well as to liberalised and competitive electricity markets with free price formation. On a level of first principles, assuming full information and no transaction costs under static optimisation, the two models are structurally identical. In practice, of course, differences pertain to dynamic incentives for efficiency gains and innovation, different levels of certainty for long-term industrial planning on the other and the number of hours over which the capacity outlays of the marginal technology are recuperated.[3]

The benchmark model is a good starting point for illustrating in contrast market failures in real-world electricity markets. However, identifying a structural tendency towards underinvestment on the basis of clearly identified market failures does not imply a logical fault in the benchmark model. It is incomplete rather than wrong and remains valid in the absence of the market failures identified in the sections below.

Under some aspects, electricity is an ideal good for competitive markets. Since electricity cannot be differentiated beyond very basic, easily observable and enforceable criteria (frequency, voltage, stability), it allows the functioning of a market that outside of the world of finance constitutes the rare example of a market without transaction costs or product differentiation.[4] Unsurprisingly, strict marginal cost pricing is the norm in competitive electricity markets outside the hours of extreme peak demand. During these hours of scarcity and rolling supply interruptions ("brown-outs") prices will be equal to the value of lost load (VOLL), which corresponds to the marginal utility of electricity.

In theory, decentralised decision-making in competitive electricity markets will then provide a level of capacity such that VOLL prices during hours of peak demand will top up revenues during ordinary operating hours to the extent that all costs of production, including fixed costs, are fully covered. VOLL prices thus cover the "missing money", which would otherwise affect all operators, as the marginal technology never earns more than its variable cost during normal operating hours.[5] As short-term demand is inelastic, in particular during extreme peak hours, the level of VOLL-prices needs to be fixed by regulatory fiat. In a market with free exit and entry, the available capacity will adjust such that the ratio between the fixed costs of the marginal technology and the level of the VOLL will correspond to the number of scarcity hours, i.e., the hours during which demand will not be fully covered due to *involuntary* demand response.[6]

When discussing capacity remuneration mechanisms, it is crucially important to understand the difference between involuntary and voluntary demand response. The difference in the economic costs of involuntary and voluntary demand reduction consists precisely of the security of supply externalities that will be discussed in Section 3. Typically, the expected average number of hours per year at which scarcity pricing prevails is measured in the single or low double digits, while prices that at this point correspond to the marginal utility of electricity are measured in the thousands of dollars or Euros. The higher the VOLL set by the regulator or the market operator, the smaller the average number of hours per year during which it will reached and vice versa. In a world with full information, no externalities and infinitely finely scalable investment, the sequence of balancing, intraday, day-ahead and forward markets will thus generate the socially optimal level of capacity.

The following equation summarises the principle of full cost recovery in energy-only markets as a combination of short-term marginal cost (variable cost) and long-term capacity cost (VOLL) pricing:

$$[FC_i+VC_i*h_i]*CAP_i=\left[\sum_m (VC_m-VC_i)*h_m) + (VOLL-VC_i)*h_{VOLL}\right]*CAP_i\ \forall\ m\,\mathrm{with}\,VC_m\geq VC_i.$$

where,

FC_i indicates the annualised investment costs of technology i.
VC_i indicates the variable costs per unit of output of technology i.
h_i indicates the hours of operation per year of technology i.
CAP_i indicates the installed capacity of technology i.
VC_m indicates the variable costs of the marginal technology that sets the price.
h_m indicates the hours of operation per year of technology m.
$VOLL$ indicates the value of lost load, and
h_{VOLL} indicates the number of VOLL hours per year.

The condition $VC_m \geq VC_i$ indicates that technology i can itself be the marginal technology and it holds that $\sum_m h_m = h_i$. In cases, where $VC_m < VC_i$ technology i does not operate.

The equation above synthesises the three central features of the standard theory of optimal pricing in electricity systems, whether they are governed by the prices resulting from decentralised profit-maximisation in competitive markets or by the tariffs set by a benevolent monopolist aiming at maximising the social surplus:

1. Short-term marginal cost pricing, which corresponds to the variable cost of the marginal technology, outside of extreme peak hours;
2. Long-term marginal cost pricing during extreme peak or VOLL hours;
3. Full cost recovery and the satisfaction of budget constraints both at the level of the individual firm and the system in the sense that annual revenues are equal to total annual costs. In other words, there is no "missing money".[7]

---

[3] In practice and in the context of the capacity issue, the optimizing monopolist presented in Boiteux (1960 and, 1949) has the advantage of internalizing security of supply externalities by resorting to peak-load prices that "spread out the peaks and fill in the hollows" (p. 176), which is precisely what one would demand from a well-performing capacity mechanism.

[4] Physical network losses are precisely measurable and codifiable and thus form a competitive sub-market. This allows feeding them back into the main market without any economic efficiency losses.

[5] See the discussion on the term "missing money" below. In order to avoid terminological confusion we will call "marginal technology", the power plant with the highest variable costs. This clarification excludes "demand response" as the marginal production technology. This is done for reasons of readability only. In principle, demand response, in particular if triggered by dedicated technical hardware, can be considered a marginal technology since it is equating supply and demand at the margin.

[6] Nobody has done a better job educating economists about these fundamental relationships than Paul Joskow in "Competitive Electricity Markets and Investment in New Generating Capacity" (2007).

[7] The term "missing money" is highly ambiguous. In principle, it should not be part of the vocabulary of the theory of electricity markets as markets with free entry and exit will always adjust in order to ensure full cost recovery. In other words, operators will retire capacity in order to generate scarcity hours up to the point that their full costs are covered. However, the term is widely used precisely to designate the revenue shortfall that *would exist if the system provided electricity at socially desirable levels at all times*

The result is that the market is privately and socially optimal due to short-term marginal cost pricing outside of scarcity (VOLL) hours and that due to long-term marginal cost pricing during scarcity hours all actors satisfy their budget constraint. How does this square with the fundamental principle of Walrasian microeconomics that short-term marginal costs pricing should prevail at all times, even if it would mean subsidising the fixed costs of capacity? The answer is that in markets for non-storable goods the short-run marginal cost at peak time *is* the long-run marginal cost, i.e., the costs of an additional unit of capacity.[8] The central microeconomic principle that only short-run marginal cost pricing guarantees optimality is thus preserved. At the theoretical level, scarcity pricing at VOLL is economically efficient *and* socially optimal.

This unique result is due to the double nature of extreme peak (VOLL) prices. Prices at VOLL hours correspond both to the short-term marginal cost of not consuming electricity, which is equal to the marginal cost of making an additional unit of electricity available through demand restraint *and* the capital costs of producing an additional physical unit of electricity. Such demand restraint can also be considered as a particular technology of electricity production with very high marginal costs and zero fixed costs. In either case, VOLL corresponds to the disutility, the marginal utility lost, of not using the marginal unit of electricity. The key economic property remains, i.e., the coincidence of short-term and long-term marginal costs. As expressed by Marcel Boiteux, one of the founders of the theory of peak-load pricing:

> “Under the theory of selling at marginal costs, prices must be equal to the *differential costs* [short run marginal costs] for *existing plants*. Plant is of optimum capacity when the differential cost and the development costs [long run marginal costs of additional capacity] are equal, that is to say, when differential cost pricing covers not only working expenses but also plant assessed at its development cost (Boiteux, 1960, 167).”

How can this unique coincidence of short-term and long-term marginal costs come about? All that is necessary is that producers are capable of adding or subtracting generating capacity to and from the market such that the product of VOLL and the number of resulting VOLL hours corresponds to the balance of their fixed costs. An interesting question is whether the market needs to be competitive in order to achieve cost recovery. Stoft maintains that full cost recovery in a liberalised electricity market does not depend on the market being competitive.

> “The discussion of fixed-cost recovery does not depend on any details of the cost-functions or even on the market being competitive. It depends only on the ability of generators to enter and leave the market (Stoft, 2002, 123).”

True enough, but in an uncompetitive market, generators would recuperate *more* than full costs by restricting capacity and increasing VOLL hours beyond the level necessary to recuperate fixed costs. Stoft is thus not entirely correct that full-cost recovery “fails if there are barriers to entry (*ibid.*)”. Full-cost recovery would still work but it would no longer arrive at socially optimal outcomes. Léautier is thus correct in making the competitiveness of electricity markets the primary condition for the absence of underinvestment, as long as other market imperfections are absent (Léautier, 2013, 10). Needless to say, as long as Boiteux’ optimising monopolist is working with an objective function aiming at the maximisation of social welfare, its capacity will also be optimal in the absence of other imperfections.[9]

In the absence of market failures, the benchmark model, in which privately and socially optimal levels of capacity coincide in energy-only markets, is thus alive and well at the theoretical level. Arguments for CRMs substituting for scarcity pricing must hence transcend the benchmark model. The next two sections will show that security of supply externalities and informational asymmetries push the privately optimal level of capacity below the socially optimal level. Such an enlargement of the theory allows for radically different policy conclusions, in particular the introduction of CRMs in a conceptually coherent manner.

3. Market failures on the demand-side: security-of-supply externalities

Diverging from the theoretical benchmark model for energy-only markets by way of CRMs implies economic efficiency losses and must thus to be justified on the basis of market failures or externalities. Sections 3 and 4 develop the case for two types of market failures that both imply that energy-only markets will provide socially sub-optimal levels of capacity in energy-only markets. The first case pertains to the demand-side, the second to the supply-side. On the demand-side, consumers would prefer higher than privately contracted capacity due to the existence of security-of-supply externalities. On the supply-side, investors will provide on average less than socially optimal levels of capacity even in the absence of demand-side externalities due to asymmetric investment incentives in markets for non-storable goods. This effect is exacerbated by risk aversion and the increase in volatility caused by intermittent renewables. Demand-side and supply-side market failures do not imply each other and are additive.

Regarding demand, the existence of security-of-supply externalities implies that consumers and political decision-makers would be willing to pay for higher levels of security of supply than is individually contracted in energy-only electricity markets. However, issues arise at an even more fundamental level as real-world scarcity pricing implies efficiency losses that are not taken account of in the theoretical model. For theoretical economists, scarcity pricing at VOLL represents the inevitable, necessary and desirable moment, when operators recoup the revenue shortfall that would otherwise figure as “missing money”. For consumers, network operators and policy-makers, scarcity pricing represents the dreaded moment when electricity prices go haywire, electricity supply is cut and faith in the working of electricity markets breaks down.

Even in the absence of security-of-supply externalities and asymmetric investment incentives, it is hardly obvious that VOLL pricing would work as indicated by theory. Scarcity pricing at VOLL may be a very imperfect way to provide appropriate investment signals by equating prices to willingness-to-pay in the very situation when demand is reaching capacity. As Joskow (2007) points out, the extreme demand and supply tensions necessary to induce load-shedding under VOLL are frequently characterised by disequilibria or even complete market breakdown that do not lend themselves to the discovery of the marginal cost of electricity provision at the level of the system:

(*footnote continued*)
without resorting to VOLL-pricing. *Using the term “missing money” thus unwittingly implies considering the level of capacity provided by competitive energy-only markets as socially suboptimal. There is money missing only if one supposes that levels of capacity should be higher than those provided by the market.*

[8] This is different from average cost pricing in industries producing storable goods. Here, firms must not take peak demand but total demand into account when choosing their optimal capacity. In such cases, only marginal cost pricing at variable costs of production is socially optimal. With increasing returns to scale, optimality thus would require subsidies for capital costs in order to ensure economic viability.

[9] An integrated monopolist aiming at welfare maximisation has an intrinsic advantage here over competitive electricity markets, even when both are based on the same underlying economic principles. Paradoxically, this advantage consists in the fact that contrary to a liberalised market, the monopolist is not obliged to pursue economic welfare optimisation in a narrow sense, i.e. to practice pure VOLL pricing. It also provides an opportunity to compensate for the stochastic nature of electricity demand which discourages private investment (see Section 4). Instead, it can integrate social preferences for less-than-VOLL prices but higher levels of security of supply in the form of less-than-VOLL but higher than marginal cost prices during a correspondingly longer number of hours. This explains why the two models are often seen as antipodes, even though they are structurally identical from a conceptual point of view.

> "There are a number of wholesale market imperfections… that appear collectively to suppress spot market prices… below efficient prices during the small number of "scarcity" hours in a typical year when wholesale market prices should be very high… Since the market also collapses in these situations, wholesale market prices are effectively zero and do not reflect consumer preferences to buy or generators' cost of supply (Joskow, 2007, 165)."

The rolling brownouts during scarcity hours thus correspond to market breakdown when trades are no longer made and pricing is absent. Under such circumstances, disconnected consumers no longer have the possibility to express their willingness-to-pay. Even in the absence of the two market failures presented below, scarcity pricing at VOLL remains thus a sub-optimal manner to allocate the scarce resource of electricity. Rolling brownouts will affect consumers with high willingness-to-pay, in precisely the same manner as consumers with low willingness-to-pay. Mutually profitable trades are impossible under such circumstances. The crucial point here is that disconnections due to brownouts during scarcity hours are *involuntary*.

Even without accounting for the technical risks involved in rolling brownouts, there is a clearly identifiable social disutility associated with scarcity pricing. The fact that consumers, network operators and politicians neither like nor accept scarcity pricing is not simply an irrational whim harboured by poorly informed non-experts but constitutes an intuitive grasp of the challenges connected with the transposition of theoretical scarcity pricing at VOLL into practice.

3.1. *Security of electricity supply as a public good: the discussion so far*

Energy-only markets provide less than socially optimal levels of capacity due to security-of-supply externalities. Such externalities are due to transaction costs and imperfect information, which prevent the creation of a working market for the good in question (see Coase (1988, 2008), Arrow (1970) and Keppler (1998)). Due to the complexity of the good "security of electricity supply", which depends on social preferences, political circumstances, the state of technology, behavioural structures and a slew of other factors it is impossible to price it out in energy-only markets. This is equivalent to stating that a market for security of supply is missing. As Arrow famously formulated, "the problem of externalities is … a special case of a more general phenomenon: the failure of markets to exist (Arrow, 1970, 76)." Due to the public nature of the good security of supply, which is precisely due to the externalities discussed below, it is inconceivable that such a market would spontaneously arise. Even where secure products such as long-term supply contracts exist, market participants will have no incentive to invest in them up to the optimal level since their private benefit of secure supplies will always be lower than the corresponding social benefit. The result is that market participants consider rightly that the security of electricity supply, the resulting number of acceptable scarcity hours and the corresponding level of capacity at current levels of technology, information and behaviour must be dealt with as a public good by a centralised authority, the regulator or the government.

The idea that security of electricity supply is in a yet to be defined sense a public good is not new. Oren (2003), Kiessling and Giberson (2004), Salies et al. (2007) and De Vries and Hakvoort (2004) have made this point in various forms. While all of these authors have a very good grasp of the practical issues, none of them formulates in a definitive manner the conceptual argument that due to security of supply externalities there exists a tendency for underinvestment in capacity in competitive energy-only markets. This leads to circular arguments such as "CRMs are needed when energy-only markets do not work properly." Oren's (2003) well-known paper on generation adequacy is a case in point:

> "When energy markets are not sufficiently developed to provide correct market signals for generation investment, setting capacity requirements with secondary markets that enable trading of capacity reserves is the preferred approach. It is more likely to produce correct market signals for investment than administratively set capacity payments which are likely to distort energy prices and result in over-investment (Oren, 2003, 21)."

Even qualitatively the notion of "over investment" requires a sharper definition. From a social point of view "over investment" beyond competitively supplied levels is precisely what one aims at in a CRM. Like a number of commentators, Oren also attempts to reduce the capacity adequacy issue to a question of indelicate behaviour by private operators:

> "An important concern that is often voiced in countries where there is no well developed institutional infrastructure that can enforce financial liability of corporation is that load serving entities or generators may assume more risk than they could handle reliably… We cannot ignore the reality that US bankruptcy laws provide a de facto hedge to load serving entities which may result in assumption of imprudent risk (Oren, 2003, 15)."

While US bankruptcy laws may or may not encourage excessive risk taking, this has nothing to do with the specific coordination failure related to the socially sub-optimal provision of capacity, which at low elasticities of demand persists even with perfectly hedged and risk averse operators. Conversely, with the right incentive structure even the most indelicate operator would provide adequate levels of capacity. There is an unfortunate tendency in discussions of electricity markets to moralise structural issue. This obscures, rather than clarifies the nature of socially adequate capacity provision. It is Coase's great merit to have irreversibly shifted the externality issue from an unwillingness to trade (a moral issue) to an inability to trade (a structural issue).

While the reasons for socially sub-optimal investment in capacity in liberalised electricity markets are more explicit in a paper by De Vries and Hakvoort, the authors also fall back on morally doubtful "free-riding" as the primary cause for this unsatisfying state of affairs. They first provide a useful list of "factors that may disturb the narrow investment optimum. The following types of market failure can be discerned (…):

- Price restrictions,
- Imperfect information e.g., regarding consumer willingness to pay or future supply and demand,
- Regulatory uncertainty,
- Regulatory restrictions to investment, and
- Risk-averse behaviour by investors (De Vries and Hakvoort, 2004, 4)."

These are points well worth mentioning. However, De Vries and Hakvoort do not relate them to the distinction between private and public goods. They reduce the issue to individual consumers being unable to express their own private willingness-to-pay. In other words, they do not identify the externality, even though they introduce the term:

> "In a [socially optimal] market equilibrium, this positive externality would be reflected by consumers not revealing their true willingness to pay. If service interruptions are the consequence of, for instance, a 2% shortage of generation capacity, this means that service interruptions affect only about 2% of the customers at a time during a period of scarcity… The consumers who caused the shortage by under-contracting therefore do not suffer the full consequences; instead, they still can consume as much electricity as they want for 98% of the time. In a [private] market equilibrium, this means that those consumers who show a lower willingness to pay, benefit from those who show a higher willingness to pay and thereby attract more peak capacity. The public good character of reserve capacity therefore provides consumers with an incentive to understate their

willingness to pay (*ibid.*, 6–7).

This is not correct. First, security of supply externalities arise independently of the fact that consumers may deliberately underreport their true willingness to pay. They thus continue the argument introduced by Oren that less than socially optimal capacity is due to a minority of indelicate participants in the electricity market, although they shift the issue from the supply side to the demand side. Second, at stake is not the average demand of electricity but the demand for electricity at extreme peak times, which is equal to capacity. In other words, in question is not the average willingness-to-pay for electricity but the marginal willingness-to-pay for electricity at time of scarcity. If there are consumers that under-contract their true consumption they will suffer utility losses of their own. There are no externalities or public goods issues involved.

The public good issue was addressed head-on by Kiessling and Giberson (2004) in their presentation on “Is Network Reliability a Public Good?” Following Oren (2003), their contribution has the merit of highlighting the fact that there are several issues involved in network reliability such as adequate capacity provision, operational reliability and the provision of ancillary services. They also correctly point out that network reliability has both private and public good aspects.

However they subsequently set up the public good issue as a straw man to better take it down. Similarly to De Vries and Hakvoort they frame the problem in terms of heterogeneous preferences and free-riding (p. 10). The solution is then straightforward, better contracts and priority insurance (p. 19). This again misses the point. Without sufficiently elastic demand, the security of supply issue will persist even with perfectly honest, homogenous consumers as operators have no means of recuperating the full social willingness to pay for an additional unit of capacity in energy-only markets. In this manner no coherent argument for CRMs can be made.

3.2. *Security-of-supply externalities at the heart of the public good issue*

CRMs become necessary because the good security of electricity supply is too complex and transaction costs are too high to be traded bilaterally. This complexity is directly related to the short-term inelasticity of demand and the impossibility to store electricity at economically attractive costs. In general terms, a public goods or externality issue arises if

1. The non-consumption of electricity of consumer A affects the utility of consumer B and
2. The two are unable to move towards higher levels of capacity through appropriate side-payments.

The second condition is impossible to realise without that a third party codifies the good security of supply, for instance in terms of tradable capacity certificates, thus creating a capacity mechanism. Let us therefore concentrate on point 1. It is important to understand that such security-of-supply externalities arise only if the non-consumption is involuntary. With voluntary and possibly remunerated, demand restraint, the externalities will fade away. In other words, the underlying issue is again the inelasticity of demand linked to the fact that electricity in most markets cannot be stored in sufficient quantities at sufficiently low cost. If the demand side were elastic, it would work exactly like storage and the public goods issue would fade away.[10]

It is the presence of reciprocal externalities in electricity consumption that makes private contracting for the appropriate level of security of supply suboptimal. It also means that brown outs during scarcity hours have *higher* social costs than the product of private cost (equal to VOLL) times the number of disconnected customers. The aversion of customers and politicians towards scarcity pricing thus has a serious underlying rationale: due to network effects, the social costs of an interruption of electricity supplies are larger than the private costs. The network effects in question here do not relate to the physical networks of power transmission but to the economic and social networks of modern industrial societies. Electricity pervades every aspect of society. Preventing a fraction of consumers to participate in its social and economic networks will inevitably propagate and thus inflict damages, real and perceived, to larger sub-sections of the socio-economic system.

If an electricity customer, for instance, is a hospital or a restaurant, it is easy to see that the costs of an hour's outage will affect the well-being of many other people. The question is now whether the loss of utility of a hospital's patients or a restaurant's customers – a loss of utility that can stretch over far longer periods than the actual outage – is adequately taken into account in the decisions affecting electricity supply made by the manager of the hospital or the restaurant. If it is, there is no externality. If it is not, an externality exists.

A simple example may illustrate the point.[11] Imagine a visitor riding down the elevator in a multi-story office building after an afternoon meeting that stretched into the winter evening. Suddenly, the elevator stops and the lights go out due to a rolling brownout during evening peak hours. Even after electricity has come back after a few minutes, the stress is considerable and the evening is done for. Of course, the example can be expanded at will with a number of dramatic or hilarious ramifications. In the present context, there are two important points here:

1. Due to the inability for the electricity distributor to single out individual customers, this situation can arise *even if the building manager has correctly anticipated both his consumption and his capacity*. This is *not* an issue of free-riding or misrepresentation of true willingness-to-pay as implied by De Vries and Hakvoort or Kiessling *et al*. This is a classic externality issue where due to high transaction costs the building manager and its tenants are unable to transmit individual preferences for continuity of service. This holds a fortiori for the hapless visitor. The good in question (security of supply) is undersupplied.
2. If overall electricity demand was more elastic, the building manager and its tenants might have decided to partake in a demand-side management programme which organises *voluntary* (or contracted, which amounts to the same thing) load shedding at certain peak hours. A message sent several hours before would have reminded the building manager of his obligation to minimise electricity consumption and to shut down the elevators. In this case, a warning sign “Do not use elevators” would have fully internalised the potential externality.

The example illustrates that the security of supply externality is due to the *involuntary* character of the enforced load shedding. The difference between an involuntary disconnection with inelastic demand and a voluntary reduction or deferral of demand consists precisely of the positive externalities of electricity consumption. In reality, such security-of-supply externalities consist of a myriad of infinitely small impacts. Evening football matches, train and metro operations, public lighting and security as well as, ultimately, the investment climate and economic development depend on continuous, high quality electricity supply.[12]

Auto-generation or costly back-up systems are, of course, not a

---

[10] The inelasticity of the short-term electricity demand function is not only a result of technical and informational constraints but also of behavioural inertia at the level of individuals and households. All three are part and parcel of the “transaction costs” which impede the first best optimum to be realised without any externalities.

[11] I would like to thank Marc Bussieras, EDF, for providing this example. He is, of course, absolved from any responsibility for the usage made of it in this context.

[12] The straightforward models of economic theorists (see Léautier (2013), for example) treat electricity exclusively as a private good. They thus fail to make the difference between an expected voluntary and an unexpected involuntary reduction in demand.

solution. They would raise the overall cost of the electricity system above the cost of a centralised system with an appropriate level of capacity. Back-up just pushes costs towards individual customers and obliterates the efficiency gains from the mutualisation of resources that is the hallmark of electricity networks. Back-up is thus warranted only for those installations where the risk of massive externalities, think hospitals or data centres, is so high that it outweighs any concerns about the efficiency of electricity supply. The average consumer should not be forced to resort to costly back-up.

In the absence of elastic demand in large swathes of the market, it will thus hold even in perfectly competitive energy-only markets with full information and producers as well as consumers expressing their true costs and preferences that the social willingness-to-pay for additional capacity is greater that the private willingness-to-pay for additional capacity. The social costs of supply disruptions thus exceed for the time being the value that can be captured in energy-only markets by the provider of the marginal unit of capacity. Hence, the number of VOLL hours in a liberalised energy-only market will be higher than the social optimum. We say, “for the time being,” as the inelasticity of demand or the availability of new technical solutions such as storage are not necessarily fixed. The progressive adoption of demand technologies that make the loss of utility due to *voluntary* reductions of electricity consumption amenable to compensation would indeed reduce the gap between privately and socially optimal levels of capacity. While the gap will never be zero, it may become negligible. Until then, capacity remuneration mechanisms (CRMs) are warranted.

4. Market failures on the supply-side: asymmetrical investment incentives in markets for non-storable goods

Other things equal, non-storable goods will always have more inelastic demand than storable goods. In addition to creating externality issues on the demand side, this inelasticity of demand provides also asymmetric incentives for investors in capacity on the supply side. This pushes actually provided levels of capacity further away from socially desirable levels of capacity. While the issue is also related to the inelasticity of supply, it is independent of the security of supply externalities spelled out above.

The reason is that electricity generation investments cannot be scaled to an arbitrarily fine degree. This means capacity investment will always either slightly over- or undershoot the privately optimal amount. However, due to the fact that during peak hours, when demand is at its most inelastic, every single producer has market power, the implications for profits are not symmetric for over- or underinvestment. Overinvestment creates small gains in added quantities sold and large penalties in terms of price declines, even for small amounts of excess capacity. Underinvestment creates small losses in terms of sales foregone but large gains in terms of more frequent scarcity pricing. Due to the extreme inelasticity of demand at peak time, the issue poses itself not only at the level of the industry, but at the level of the individual producer. Uncertainty and risk aversion reinforce this effect.

An example illustrates the point. Assume that in a given year extreme peak demand is expected to reach 101 GW for three hours and that the level of 100 GW is reached for 20 h. Abstracting from security of supply externalities, one may assume that three hours of VOLL are considered acceptable by the system operator as well as sufficient to recuperate the “missing money” for fixed investment costs. The socially optimal system size is thus 100 GW. Assume further that current capacity is 99 GW and that the minimum size of a generation investment is 2 GW.

In this constellation, every individual producer has to decide whether to invest in an additional 2 GW of capacity or not. If anybody invests, capacity will be at 101 GW with zero hours of scarcity. If nobody invests, capacity will be at 99 GW with 20 h of scarcity. Demand is assumed inelastic with respect to prices and prices are equal to variable cost at € 100 per MWh, if demand is below or equal to capacity, and that prices are equal to VOLL at € 3 000 per MWh, if demand exceeds capacity.

The issue of lumpiness or discreteness in power generation equipment is regularly mentioned in the literature on electricity market design, see, for instance, Batlle and Rodilla (2010, 2), Cepeda and Finon (2010, 10), De Vries (2006, 28), De Vries and Heijnen (2008, 217) or Stoft (2006, 101–103). Stoft, in particular, gives a thoughtful discussion of the issue but limits it to transmission investment under regulation. De Vries and Heijnen (2008) also spell out the asymmetry in profit expectations due to under- and overinvestment but without linking it to the shape of the demand curve. Usually, however, the notion of lumpiness is introduced in passing as a general source of market imperfections without linking it to the structure of demand. Lumpiness may also be considered a source of monopoly power. While it is true that in most markets indivisibilities in production give rise to monopolistic competition through the loop of product differentiation, the issue is different in power markets due to the fact that electricity is an undifferentiated good. During normal operations, lumpiness does thus not have adverse effects on consumer welfare. It is only through the interaction with inelastic demand at peak times that lumpiness, and the incentive for underinvestment that it creates, results in a distinct market failure. With elastic demand, there would be no policy-relevant market failure even in the presence of discretely-sized increments in investment, as underinvestment would be just as costly, in terms of profits foregone, as overinvestment. To our knowledge, this specific effect of lumpiness in electricity markets has not yet been developed elsewhere.

Underinvestment in capacity is not a question of barriers to entry. Competition will not change the issue, as long as the minimum investment size remains discrete. Even in markets with perfectly free entry, in the sense that incumbents and entrants face no difference in cost, a new competitor will not enter as long as demand remains inelastic. Just as the incumbents, they would not recuperate their investment outlay if the penalty for overinvestment were considerably larger than the penalty for underinvestment. Fig. 1 below illustrates the point.

This does not mean that strategic behaviour is wholly absent when it comes to capacity decisions in electricity markets. However, since wilful short-term capacity retention is illegal and closely monitored, operators have, as explained, every interest to keep their available capacity below levels that might contribute to exceeding peak demand. Strategic behaviour is thus confined to potential capacity, i.e., capacity that could be brought on-stream before a new competitor could enter if, for instance, peak demand increased. The widespread practice of “mothballing”, rather than dismantling, idle capacity includes, beyond the preservation of option value, precisely the sort of strategic overinvestment in capacity studied in papers such as Spence (1977) or Fudenberg and Tirole (1983). The strategic investment in potential rather than actual capacity referred to in the answer to the previous point can indeed result in a Cournot-Stackelberg equilibrium in capacity, combined with a Bertrand equilibrium in prices (due to the undifferentiated nature of electricity) for below-capacity demand and scarcity pricing thereafter. One of the function of CRMs is thus to provide the incentives to convert potential capacity into actual capacity, i.e., to return mothballed capacity to the market. The existence of such capacity implies that the clearing price in capacity auctions is usually

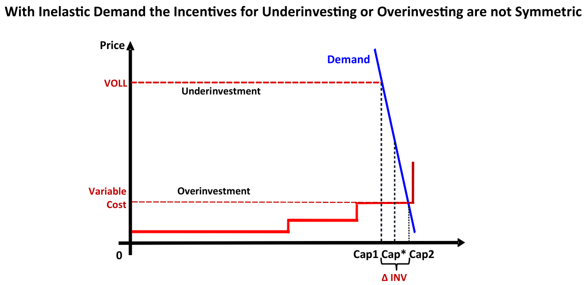

Fig. 1. With inelastic demand the incentives for underinvesting or overinvesting are not symmetric.

below the cost of genuinely new entry (CONE).

This said, if demand was perfectly known, even with inelastic demand the resulting underinvestment could be quite small. Uncertainty and risk aversion, however, significantly amplify the issue. Given the stochastic nature of electricity demand and the high costs of overcapacity, producers will strive to avoid overcapacity with the greatest possible probability. Capacity decisions will thus be made on the basis of an expected load that is less than the true mean of the probability distribution of peak electricity demand. This is observed also by De Vries and Heijnen in a general context:

> "Given the uncertainty about future market conditions, the socially optimal volume of generation capacity is higher than in the theoretical [private] optimum in the presence of perfect knowledge (De Vries and Heijnen, 2008, 226)."

Uncertainty coupled with risk aversion works as if the minimum discrete size of investment had increased. Coming back to the above example, an investor would invest if there was certainty about demand being 100 GW outside of extreme peak hours and if it was possible to invest in 1 GW increments of capacity. In principal, this would allow to equate demand at capacity at 100 GW. However, under uncertainty and risk-aversion, with an *expected* demand of 100 GW, the investor would no longer even countenance an investment of only 1 GW, because the risk of overshooting peak demand 50% of the time would simply be too costly. Needless to say, the social optimum would still be equal to expected demand at 100 GW.

Once it is accepted that uncertainty reinforces the tendency of producers to invest less than the social optimum, the case for CRMs becomes overwhelming, in particular, in the present context. Beyond permanent structural sources of uncertainty such as electricity demand or technological evolution, generators face enormous regulatory uncertainty in current electricity markets. This relates, in particular, to out-of-market support for variable renewables, but also to the rules governing carbon emissions trading or the complicated processes governing the establishment of new interconnections.

In conclusion, due to the very high inelasticity of demand at peak hours, uncertainty and risk aversion, there is practically no lower bound for the size of an incremental investment in generation capacity below which one could argue that it has no influence on the demand and supply balance. The discontinuity in the pay-off function when passing from under- to overinvestment will ensure that investors, whether incumbents or entrants, will always err on the side of caution. In industries, where investments that can be scaled to an infinitely fine manner, or in industries for storable goods with elastic demand, investors would have symmetric incentives to get as close as possible to expected capacity even under uncertainty. They would be indifferent between underinvesting and overinvesting. With inelastic demand and discretely sized investments or uncertainty, the pay-offs for over- or underinvesting are no longer symmetric and investors will always lean towards underinvestment.

5. Conclusions

The market failures identified in Section 3, security-of-supply externalities on the demand side, and in Section 4, asymmetric investment incentives on the supply side, will lead to a situation, in which actually provided levels of capacity are doubly inferior to socially optimal levels. Although both market failures are ultimately related to the fact that electricity cannot be easily stored at economically acceptable costs, they arise independently. In other words, short of inventing and deploying cheap and ubiquitous storage, resolving one would not resolve the other. The two market failures thus create independent and additive effects that cause privately supplied levels of capacity in liberalised electricity markets to fall short of socially desirable levels of capacity, Fig. 2 below.

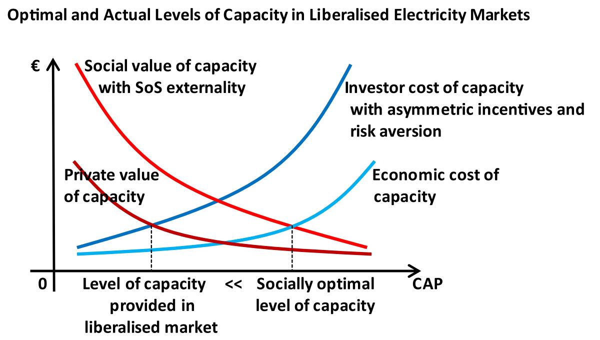

Fig. 2. Optimal and actual levels of capacity in liberalised electricity markets.

As soon as a significant capacity shortfall is identified in energy-only markets, CRMs become the appropriate tool to ensure the provision of socially optimal levels of investment. CRMs transform implicit social preferences for security of supply into explicit capacity objectives that translate through different channels, which depend on the form of mechanism chosen, into added remuneration for capacity providers above the level of remuneration generated in energy-only markets. In accordance with the theory of externalities (see Coase (1988), and Keppler (1994, 2010)), CRMs thus break down the complexity of the public good security of supply that otherwise is exposed to transaction costs that are too high for markets to handle.

Once CRMs are considered as instruments to reduce transaction costs, this implies a dynamic perspective. CRMs thus not only organise the optimal provision of the public good security of supply in a static sense, but by lowering transaction costs, they will also favour the introduction of new solutions that strengthen the market's ability to reduce the market failures of energy-only markets. This gives an aspect of endogenous obsolescence to CRMs, as they provide favourable framework conditions for the development of technologies and behaviour that render the demand curve more elastic. Especially, tradable capacity obligations or forward capacity markets may provide the sort of statistically treatable and commercially relevant price signal that transforms uncertainty into risk and thus advances the moment when peak-load pricing with *voluntary* demand response can wholly or partially substitute for involuntary demand response. Of course, there is no automatism here. Households may be permanently resisting participation in voluntary demand response programmes; industry may find itself exposed to rising marginal costs when trying to monetise demand response beyond the most obvious niches. Whether CRMs will be only temporarily needed, as a ratchet effect locks-in the structural change they engender, or will be required as a permanent complement to energy-only markets is ultimately an empirical question.

Whatever the final answer to the question, whether CRMs can generate sufficient demand flexibility to be no longer deemed necessary, they always set in motion structural changes in the electricity system that require periodic assessment and adjustment. Regular, transparent and pre-announced reviews are thus an indispensable feature of any well-conceived capacity mechanism. Precisely because CRMs are dynamic in nature, their fundamental set-up should be as simple and as robust as possible in order to allow taking all market participants along in a process of revision, whose rhythm, process and criteria are spelled out in advance.

Current evidence is stronger on the question whether existing short-term markets can provide the flexibility required to have supply and demand match at levels below VOLL. The frequent claim is, for instance, formulated in a recent *White Paper* on capacity mechanisms of the European Commission:

> "It has been argued that the downward pressure on day-ahead electricity prices in some markets leaves generators exposed to insufficient returns to cover their fixed costs… However, when intraday, balancing and ancillary services markets operate efficiently, such [mid-range and peaking] plants can participate in those markets, deriving additional revenue… Prices in those markets should be allowed to raise [sic] above short run marginal cost,

enabling generators to cover also part of their fixed costs (European Commission, 2013, 13)."

The attraction of the argument is obvious. Calling for better remuneration of flexibility allows on the one hand acknowledging the need for added remuneration for capacity providers, while on the other reaffirming that slightly tweaked energy-only markets, essentially shortening the time between trading, scheduling and dispatch, are best suited to do the job.

On a conceptual level, this call for short-term flexibility provision disregards that also demand response, storage and peaking units require fixed operating and capital costs. This means that even including these technologies into the mix some scarcity pricing would be necessary. In the absence of pervasive demand flexibility and real-time pricing for a significant share of customers, security-of-supply externalities would continue to exist. The issue of the effects of lumpiness would not be addressed at all.

On an empirical level, even cursory statistical evidence reveals that average prices in short-term markets for electricity are not significantly higher than in the day-ahead and future markets. In 2015, for instance, the average price on the French day-ahead market was € 38.48 per MWh. During the same period, the average price on the French balancing mechanism was € 37.91 per MWh. While the balancing mechanism plays a vital role in facilitating the physical equilibrium between demand and supply, it is unclear which role it can play in providing the surplus of remuneration to reach socially optimal levels of capacity.

At the current state of technology and behaviour, short-term markets are necessary but insufficient to provide socially optimal levels of capacity given security of supply externalities and asymmetrical investment incentives. Current policy choices such as the introduction of VREs financed out of the market and market arrangements such as price caps have further enlarged the gap between privately supplied and socially demanded levels of capacity. For the foreseeable future and, in particular in a European context, CRMs are a necessary complement to energy-only markets.

The precise form of the appropriate CRM will vary widely from country to country in function of its load curve and the existing mix of generation and demand-side technologies. The interplay of existing supply and demand will in fact define a number of hours during which a capacity shortfall may be arising. This in return determines an optimal technology (e.g., demand-side technologies in France, gas-fired power plants in Germany or nuclear baseload in the United Kingdom). Depending on the technology and its capital intensity, and hence the need to provide certainty to investors, the optimal CRM will be chosen, whether in form of a decentralised capacity obligation, a centralised auction for capacity payments or, straightforwardly, a guaranteed tariff. While different CRMs have different advantages and drawbacks, they all share the key quality that the system-wide level of capacity is set by a central decision-maker legitimised by the appropriate institutional processes. As long as the inelasticity of electricity demand generates security of supply externalities and asymmetric investment incentives, CRMs constitute the appropriate manner to bring private investment decisions in line with social optimality in the electricity system.

Acknowledgement

The author would like to thank Dominique Finon and Marc Bussieras for their helpful comments. Discussions at the "European Workshop on Capacity Mechanisms in EU Power Markets" in April 2013 at Université Paris-Dauphine and the 8th Session of the Research Seminar in Energy Economics in December 2013 at Paris-Sciences-Lettres (PSL) on capacity mechanisms around Thomas-Olivier Léautier also provided important building blocks towards this paper. Two anonymous referees provided detailed and insightful comments, which have considerably contributed to improve the final version of this paper. This paper has benefited from the support of the Chaire European Electricity Markets (CEEM) of the Paris-Dauphine Foundation, supported by RTE, EDF, EPEX Spot and the Caisse des Dépôts. The views and opinions expressed in this Working Paper are those of the authors and do not necessarily reflect those of the partners of the CEEM.

References

Arrow, K.J., 1970. The organization of economic activity: issues pertinent to the choice of market versus nonmarket allocation, In: Haveman, R., Margolis, J. (Eds.), Public Expenditures and Policy Analysis, Markham; Chicago

Batlle, C., Rodilla, P., 2010. A critical assessment of the different approaches aimed to secure electricity generation supply. Energy Policy 38, 7169–7179.

Batlle, C., Vázquez, C., Rivier, M., Pérez-Arriaga, I.J., 2006. A regulatory instrument to enhance security of supply in the Spanish wholesale market. IIT Work. Pap. IIT, 06–025A.

Boiteux, M., 1949. La tarification des demandes en pointe: application de la théorie de la vente au coût marginal. Rev. générale de l'électricité 58, 321–340.

Boiteux, M., 1960. Peak-load pricing. J. Bus. 33 (2), 157–179, (Translation of).

Cepeda, M., Finon, D., 2010. Generation capacity adequacy in interdependent electricity markets. Larsen Work., 29.

Coase, R.H., 1988. The Firm, the Market and the Law. University of Chicago Press, Chicago.

Cramton, P., Stoft, S., 2006. The convergence of market designs for adequate generating capacity with special attention to the CAISO's resource adequacy problem: a white paper for the Electricity Oversight Board. April 2006. (available at ⟨http://www.cramton.umd.edu/papers/electricity/⟩).

De Vries, L., Heijnen, P., 2008. The impact of electricity market design upon investment under uncertainty: the effectiveness of capacity mechanisms. Uti. Policy 16 (3), 2015–2227.

De Vries, L.J., 2006. Generation adequacy: helping the market to do its job. Uti. Policy 15 (1), 20–35.

De Vries, L.J., Hakvoort, R.A., 2004. The question of generation adequacy in liberalized electricity markets FEEM Nota di Lav 2004. vol. 120. (at ⟨http://www.feem.it/⟩).

European Commission, 2013. Generation Adequacy in the internal electricity market – guidance on public interventions. Commission Staff Working Document, Brussels., SWD(2013). (438 final at ⟨http://ec.europa.eu/energy/sites/ener/files/documents/com_2013_public_intervention_swd01_en.pdf⟩).

Finon, D., Pignon, V., 2008. Electricity and long-term capacity adequacy: the quest for regulatory mechanism compatible with electricity market. Uti. Policy 16 (3), 2–14.

Finon, D., Roques, F., 2013. European electricity market reforms: the 'visible hand' of public coordination. J. Econ. Energy Environ. Policy 2 (2), 106–124.

Fudenberg, D., Tirole, J., 1983. The fat cat effect, the puppy dog ploy and the lean and hungry look. Am. Econ. Rev. Pap. Proc. 74, 361–368.

Joskow, P.L., 2007. Competitive electricity markets and investment in new generating capacity. In: Helm, D. (Ed.), The New Energy Paradigm. Oxford University Press, Oxford, 76–121, (at ⟨http://economics.mit.edu/files/1190⟩).

Joskow, P.L., 2008. Capacity payments in imperfect electricity markets: need and design. Uti Policy 16 (3), 159–170.

Keppler, J.H., 1998. Externalities, fixed costs and information. Kyklos 52 (4), 547–563.

Keppler, J.H., 2010. Going with coase against coase: the dynamic approach to the internalization of external effects. In: Lasry, J.M., Fessler, D. (Eds.), The Economics and Finance of Sustainable Development. Economica, Paris, 118–138.

Kiessling, L., Giberson, M., 2004. Is network reliability a public good?. In: Proceedings of the Paper Submitted for Publication to CMU Conference on Electricity Transmission in Deregulated Markets. Carnegie-Mellon Electricity Industry Center, Pittsburgh. 15–16 December.

Léautier, T.-O., 2013. The visible hand: ensuring optimal investment in electric power generation. IDEI Work., 605 ⟨http://idei.fr/display.php?a=22628⟩.

López-Peña, Á., Centeno, E., Barquín, J., 2009. Long term issues to be addressed by regulators in liberalised electricity systems: generation adequacy and indicative planning. Justification, available mechanisms, and a simulation study on some concrete policies. European University Institute, Robert Schumann Centre for Advanced Studies, EUI Working Paper RSCAS 2009/67.

Oren, S., 2003. Ensuring Generation Adequacy in Competitive Electricity Markets. University of California Energy Institute, Berkley, CA, (Working Paper EPE-007, June 2003), www.ieor.berkeley.edu/~**oren**/pubs/**adequacy**.pdf.

Salies, E., Kiessling, L., Giberson, M., 2007. L'électricité est-elle un bien public? Revue de l'OFCE 101. pp. 399–420.

Spence, M., 1977. Entry, capacity, investment and oligopolistic pricing. Bell J. Econ. 8, 534–544.

Stoft, S., 2002. Power System Economics. IEEE Press, Piscataway (NJ).

Stoft, S., 2006. Problems of transmission investment in a deregulated power market. In: Lévêque, F. (Ed.), Competitive Electricity Markets and Sustainability. Edward Elgar, Cheltenham, 87–130.

Vazquez, C., Rivier, M., Pérez-Arriaga, I.J., 2001. If pay-as-bid auctions are not a solution for California, then why not a reliability market? Electric. J. 14 (4), 41–48.

Vazquez, C., Rivier, M., Pérez-Arriaga, I.J., 2002. A market approach to long-term security of supply. IEEE Trans. Power Syst. 17 (2), 349–357.