Market Design for Generation Adequacy: Healing Causes rather than Symptoms[1]

Fabien A. Roques[2]

International Energy Agency, Economic Analysis Division[3]

12 December 2007

Abstract

This paper argues that electricity market reform – particularly the need for complementary mechanisms to remunerate capacity – need to be analysed in the light of the local regulatory and institutional environment. If there is a lack of investment, the priority should be to identify the roots of the problem. The lack of demand side response, short-term reliability management procedures and uncompetitive ancillary services procurement often undermine market reflective scarcity pricing and distort long-term investment incentives. The introduction of a capacity mechanism should come as an optional supplement to wholesale and ancillary markets improvements. Priority reforms should focus on encouraging demand side responsiveness and reducing scarcity price distortions introduced by balancing and congestion management through better dialog between network engineers and market operators.

Keywords: electricity market, generation adequacy, market design, capacity mechanism

JEL-classification: D24, D43, D92, L94

---

[1] The author would like to thank Dominique Finon, Virginie Pignon and Ulrik Stridbaek for their helpful comments, as well as Alex Henney, Karsten Neuhoff, and David Newbery for their insights on an inspirational joint note on "Generation adequacy and investment incentives in liberalised markets" written in July 2005.

[2] Address: 9, Rue de la Fédération, 75739 Paris Cedex 15, France, Email: fabien.roques@iea.org, fabien.roques@cantab.net, Phone : +33 (0)140576641, Fax: +33 (0)140576659.

[3] The views in this article are those of the author alone. This paper is based on work that was done while the author was with the Electricity Policy Research Group, Judge Business School, University of Cambridge.

1 INTRODUCTION

While the focus of academic research and regulatory scrutiny concentrated mainly on short-term market efficiency and competitiveness in the early years of liberalisation, there is now much attention being paid to assessing the long-term dynamic performance of the liberalised electricity industry. In some U.S. states and European countries, there is also considerable political concern whether generating capacity will continue to be adequate to ensure security of electricity supply, and whether investment signals are timely and properly reflect the social profitability of investment. Industries such as the electricity industry that are capital intensive and produce homogenous products or services frequently exhibit investment booms and busts. Under the old-style vertically integrated regulated (or state-owned) franchise monopoly structure, investment booms and busts were not eliminated, but prices were regulated and were typically very stable.

Electricity liberalisation and restructuring have dramatically changed the industry structure. Electricity has an unusual set of physical and economic attributes that significantly complicate the task of successfully replacing hierarchies (vertical and horizontal integration) with decentralised market mechanisms. There is a consensus that the solution to the *short-run supply security* problem requires some form of *centralised* management (see e.g. Hunt, 2002, Stoft, 2002). In practice, this is commonly dealt with by designating the responsibility for system operation to the system operator (SO). Contrary to short term security of supply, there is as yet no consensus on market and institutional designs to ensure *long-term generation adequacy*. Should it be left to decentralised market forces or should there be mandatory resource adequacy standards? How should the role of the SO be defined? Is there any 'best-practice' market design that can ensure generation adequacy in the long run at least cost while minimising regulatory interference with the market? In particular, is a separate mechanism to remunerate capacity necessary to maintain generation adequacy?

This paper aims to complement the existing literature by taking a somewhat more institutional and empirical perspective. The paper sheds light on how market design, market structure and the supporting institutions affect investment incentives by contrasting different experiences. We show how a range of supporting market institutions (various kinds of contracts, reserve, balancing and ancillary services markets) influences the outcomes. The paper starts with an overview of the theory of energy and capacity pricing in electricity markets, highlighting the interdependencies in the sequence of markets involved in the provision of reliable electricity supplies in real time. It then details the different factors that distort prices at times of scarcity along this sequence of markets, emphasising the impact of operational constraints and of the SO reliability procedures on longer-term investment incentives. The paper then questions the usual conceptual separation of short-term reliability issues from longer-term generation adequacy. The second section puts the different market designs into context and argues that local organisational and institutional issues have a critical impact on investment incentives. It describes the main issues with existing capacity mechanisms, particularly some of their practical unintended side-effects. Besides, such capacity mechanisms have generally evolved from a specific pre-liberalisation institutional environment, casting doubts on their applicability and efficiency in a different context. The third section then details the priorities for market design reform to ensure generation adequacy, emphasising the need to revise SO procedures at times of scarcity and to integrate better ancillary services markets, balancing markets and contract markets.

2 MARKET DESIGN AND GENERATION ADEQUACY

The debate over security of electricity supplies is often oversimplified by isolating conceptually short-term issues from longer-term issues, which can hide some critical interactions between the two time horizons. According to standard definitions, s*upply adequacy* consists in ensuring optimal capacity investment in the medium- to long-term, and is often separated from the short-term issue of s*upply security* which consists in balancing consumption and generation on a continuous basis within existing capacity limits. Supply adequacy requires adequate investment in both transmission and distribution infrastructure and generation plant (de Vries, 2004). Although the focus of this paper is on generation adequacy, the interaction with the regulatory arrangements in place for transmission and distribution planning is a critical issue (Joskow, 2007). The System Operator (SO) is in general responsible for balancing supply and demand in operational timescales, which start after ‘gate closure’. In contrast, generation adequacy is a collective responsibility involving all industry actors. The paper will stick to these commonly accepted conceptual distinctions between short-term and longer-term aspects of security of supply, but will argue that many of the short-term reliability management procedures and short-term market price signals actually have a critical impact on longer-term generation adequacy.

The *generation adequacy* problem can be further separated in three different dimensions, namely ensuring: (i) an optimal *level* of overall generation capacity at the equilibrium consistent with socially optimal system reliability design criteria; (ii) an optimal *timing* of investment minimising the magnitude of investment cycles and the impact of transitory adjustment periods on security of supply; (iii) and an optimal *mix* of different generation technologies, both in terms of load profile (mix of baseload and peaking units) and in terms of fuel mix. Cost-efficient operation requires a mix of technologies with different variable to fixed cost ratios. An optimal capacity mix balances the gains from reducing variable operating costs by having more base-load units available against the higher fixed costs of such units. Another aspect of the optimal mix of generation technologies is related to the diversity of the fuels used for generation and the magnitude of the different risks associated with their supply for fuel importers countries. This paper, as most of the literature, concentrates on the first two issues, i.e. the level and timing of investment in liberalised markets.[4]

2.1 Market design and investment incentives

The theory of spot pricing claims that electricity spot markets can provide efficient outcomes both in the short and in the long run (Caramanis, 1982). The conditions for this to hold are strong – the usual General Equilibrium assumptions of a complete set of spot and forward markets or perfect foresight, price-taking behaviour by producers and consumers, risk neutrality (or adequate risk-sharing contracts), and convex production possibilities. In such a world, the spot prices should result in an efficient dispatch and allocation of available resources and, together with forward prices and rational expectations, should also signal the need for additional generating capacity. The optimal capacity stock should be such that scarcity payments to the marginal generators when demand exceeds supply exactly cover the capacity cost of these generators. A shortage of capacity will increase scarcity rents, producing profits in excess of what is needed to cover the amortised capacity cost. Such profits will attract generation expansion. On the other hand, excess generation capacity will eliminate scarcity rents driving prices to marginal cost. When this occurs, generators on the margin will not be able to

[4] For a discussion of the issues of the optimal mix of generation technologies in relation to market design and investment incentives, see e.g. Roques (2006) and Roques et al. (2007).

cover their investment cost. Unless such generators receive extra revenues through some form of capacity payments, this will result in early retirement or mothballing of plants which will reduce capacity and drive prices back to their long-run equilibrium level.

Of course, these ideal conditions never occur, but the same could be said of all other markets, where (apart from legitimate competition concerns) we are normally prepared to accept the outcome as workably competitive and feasibly efficient, and not requiring specific policy interventions. The practical question is not whether electricity markets will deliver perfect outcomes, but whether the specific characteristics of electricity introduces systematic biases in market behaviour that require more complex market designs and regulatory requirements. In practice, electricity supply and demand presents a number of special features which imply that with current technology (i.e. without individual advanced metering equipment), generation adequacy has public good attributes (Stoft, 2002, Finon et al., 2007, Joskow, 2007). These special attributes include the variability of demand, the non-storability of electricity, network constraints requiring the physical balance of supply and demand on each point of the network at all times, the inability to control power flows to a large part of consumers, limited use of real time pricing, and non-price rationing in some instances such as blackouts (see e.g. Stoft, 2002 and Joskow, 2007). A large body of literature details various market failures which could deter investment and have a negative impact on generation adequacy, particularly peaking units (see e.g. de Vries 2004, Stoft, 2002, Joskow, 2007, Arriaga et al., 2002). The most debated issue in the U.S. context is what is referred to as "the missing money" by Cramton and Stoft (2006): several power markets in the U.S. exhibit a significant gap between net revenues produced by electricity markets and the capital cost of investing in new peaking capacity measured over several years (Joskow, 2007).

These concerns revived the debate on whether an 'energy only' market can ensure generation adequacy, or whether an additional capacity mechanism is necessary. A useful categorisation is to distinguish mechanisms under which the regulator sets a price for capacity and lets the market determine the amount of capacity available from mechanisms under which the regulator sets the amount of capacity that has to be available and lets the market determine its price. These are known, respectively, as capacity payments and capacity requirements. The concept of capacity payment is rooted in the theory of peak load pricing, whose application to electricity supply was pioneered by Boiteux (1949, 1951) in the context of a regulated monopoly. It assumes that generation of electricity requires two factors of production, capacity and energy, where the amount of energy that can be produced in any given time period is constrained by the available capacity. Vardi et al. (1977) develop the theoretical approach for charging offpeak consumers with capacity costs by relating it to the reliability design criterion employed in planning for the capacity expansion of the power system, e.g. the loss of load probability (LOLP). Under optimal capacity planning the marginal cost of incremental capacity equals the marginal cost of unserved load, which can be approximated by the marginal Value of Lost Load (VOLL) times the probability or fraction of time that load must be curtailed due to insufficient capacity (LOLP).[5] The theory of peak load pricing has served as a theoretical reference in the design of most existing capacity mechanisms designs. Energy-only electricity markets have been

[5] Charging each hour relative to its contribution to the incurrence of these capacity costs, as given by its contribution to the LOLP, or equivalently, by its curtailment probability, yields:

$$C_i = \frac{K}{LOLP}.h_i \text{ , where}$$

C_i is the marginal capacity costs chargeable to hour *i*, h_i is the curtailment probability associated with hour *I*, and *K* the annualised charge on capital.

adopted in the original (defunct) California design, in ERCOT, in Nordpool, the Australian Victoria pool (although with an ex post Value of Lost Load, or VOLL) and in the British NETA design. In contrast, electricity markets in the US East Coast, in Latin America, in Spain, Ireland, and in the former British Pool market, include some form of capacity mechanism.

2.2 Energy and capacity: an artificial divide?

The supply of power on wholesale markets at a precise location and time and with a predetermined level of quality requires the sequential provision of a variety of complementary sub-products, which are often themselves procured through market mechanisms. The sequence starts with futures markets and a "day ahead" forward market, and involves nearer to real time markets or mechanisms for the management of congestion, operating reserves, and balancing. Glachant and Saguan (2007) point out that this sequence is highly modular, with different markets designs being characterised by differences both in the organisation of the different modules and within the modules themselves. Simplified textbook economic models are of little help to assess generators' revenues in electricity markets, as it is also important to capture the impact of the remuneration of reserves (e.g. regulation and spinning reserve capacity), and other ancillary services (e.g. transmission congestion revenues and balancing mechanism payments). It can be misleading to suppose that competitive prices are always set at avoidable cost, or that the energy market is the only market to consider when evaluating generators' revenues and investment incentives.

In practice, the distinction between energy-only markets and markets with a capacity mechanism needs also to be nuanced insofar as the procurement of short- or long-term operating reserve has an impact on plant revenues, investment incentives, and generation adequacy. The remuneration of operating reserves affect the supply curve for energy and produce a secondary source of income for generators (de Vries, 2004). The so-called "strategic reserve" and "long-term reserve contracting" mechanisms rely on such principles to maintain generation adequacy (Finon et al., 2007). In Europe, Sweden and Norway are using strategic procurement of long-term reserve by the system operator as a way to encourage investment in generation capacity. Under this mechanism the system operator enters into long-term contracts with generators in exchange for some agreed upon remuneration for new power plants. In some cases such as The Netherlands or New Zealand where systems are mature, contracts can be related to the operation of old units. The effectiveness of this mechanism depends on the implementation parameters, as well as the characteristics of the system such as the share of hydropower and the interconnexion capacity.[6] The short-term operating reserves market can also be used strategically by the SO to maintain old plants connected to the system (Stoft, 2002). The British SO used forward purchases of short-term operating reserve in 2003 to incentivise some previously mothballed plants to reconnect to the system (Roques et al., 2005).

2.3 Distortions of scarcity prices and generation adequacy

The usual distinction between the short-term issue of supply security and the longer-term issue of generation adequacy needs to be nuanced insofar as many of the reliability standards and procedures used by the system operator in the short-term have important implications on prices at time of scarcity and hence power generators' revenues, particularly for peaking units which earn most of their

[6] These include the triggers for dispatch of strategic reserves, the way strategic reserves are compensated, and the price paid for energy produced by strategic reserves of capacity (de Vries, 2004, Finon and Pignon, 2007).

revenues at time of high prices. The artificial divide between short-term reliability management under the responsibility of the SO and long term market based investment decisions overlooks the critical impact of scarcity price signals feeding from real time reserve, balancing, and congestion management markets through to contract markets. Many of the operating standards and practices inherited from vertically integrated utilities are widely accepted as necessary or desirable in the liberalized industry, but much of the economic analysis of the behaviour and performance of wholesale and retail markets has either ignored the impact of non-market mechanisms used in the short-term to maintain system reliability or failed to consider them in a comprehensive fashion (Joskow and Tirole, 2006 and 2007). Joskow (2007) details the different underlying issues characterising U.S. markets, including wholesale market imperfections, regulatory constraints on prices (price caps associated with market power mitigation procedures), and the non-market procedures used by System Operators in their use of operating reserves at times of scarcity.

More research is needed to assess the adverse impact on revenues and investment incentives of short term management procedures used by the system operator, as well as the lack of integration of reserve, congestion and balancing mechanisms with contract prices. It is now well understood that the source of the "missing problem" in the U.S. is that spot market prices do not rise high enough during scarcity hours to produce adequate revenues to cover the capital costs of investment in an efficient level and mix of generating capacity (Joskow, 2007, Cramton and Stoft, 2006). But Joskow (2007) points out that price caps might not be the main issue, as prices do not often even reach the price cap level at time of scarcity, and suggests that the non-market procedures and engineering requirements used by SOs constitute the critical issue. Chao et al. (2007) detail a number of engineering requirements for local reliability and procedures which lead to distortions of prices and revenues of generators, such as 'Out of Market' dispatch procedures used by SOs, penalties for excessive uninstructed deviations from schedules set day-ahead imposed in balancing mechanisms, and in the case of centralized markets, unit commitments selected by the SO distorting the market merit order.[7]

The interesting question when considering generation adequacy in electricity markets with a different institutional environment and market organisation – particularly in Europe – is which of the problems identified in the U.S. East Coast markets are likely to be shared by other electricity markets. While price caps in relation to market power mitigation procedures capping prices during scarcity periods might be less of a concern, the problems associated with engineering reliability rules and the procedures used by SOs during scarcity conditions are likely to be similar. There are also additional specific issues with European electricity markets which are likely to distort scarcity price signals. In particular, most European markets are based on bilateral trading and have balancing mechanisms which are not well integrated with contract markets and where SOs often impose penalties for excessive deviations from schedules set before gate closure.[8] Many European markets rely on 'balancing mechanisms' rather than actual 'balancing markets', which introduce penalties to discourage imbalances, and where imbalance prices are often not set based on marginal bids but rather

[7] Besides, price mechanisms adopted by SOs to alleviate some problems can themselves create price distortions and perverse incentives: for instance, transmission congestion charges can differ greatly from actual costs of redispatch, while the co-optimization of energy raises the price of energy even though reserve capacity is not scheduled to produce energy.

[8] An important difference is that balancing markets are managed by non-for-profit Independent System Operators in the U.S., while European markets are administered by Transmission System Operators.

using an average price formula (Table 1).[9] The main argument to justify such a mechanism is that it is necessary to maintain the security of the system, as paying balancing power at its real time market value would provide an incentive to market participants to intentionally create imbalances in their forward trading schedules. But such balancing mechanism has the negative side effects of reducing the liquidity and to mute scarcity price signals at time of scarcity. Glachant and Saguan (2007) find that the penalties introduced in the balancing mechanisms distort forward prices, increase TSO's revenues, increase market inefficiencies and adversely impact small and disintegrated agents. Moreover, the magnitude of these consequences increases as the gate closures moves further away from real time.

Table 1 – Electricity balancing mechanisms in Europe in 2005. Source: EC (2005)

	Market or fixed prices	Gate closure	Average TSO sell price	Average TSO buy price	Average spread
Austria	market	day ahead	51	24	27
Belgium	hybrid	"ex-post"	56	12	44
Denmark	market	½ hour	36	27	9
Finland	market	½ hour	32	27	5
France	market	6 during day	50	45	5
Germany	market	3 during day	70	2	68
Greece	fixed	day ahead	44	44	0
Ireland	hybrid	day ahead	69	60	9
Italy	market	ay ahead	102	23	79
Luxembourg	fixed	-	-	-	
Netherlands	market	1 hour	69	28	41
Portugal	fixed	2 during day	58	23	35
Spain	market	¼-3¼ hrs	0		
Sweden	market	1 hour	32	28	4
UK	market	½ hour	16		
Norway	market	1 hour	29	29	0
Estonia	-	day ahead	-	-	-
Latvia	-	2 hours	-	-	-
Lithuania	-	2 hours	-	-	-
Poland	market	day ahead	37	24	13
Czech Republic	market	1½ hours	51	0	51
Slovakia	-	day ahead	-	-	-
Hungary	market	day ahead	40	0	40
Slovenia	market	day ahead	-	-	-

[9] There are two types of imbalance price mechanisms: Dual imbalance pricing, where a different price is applied to positive imbalance volumes and negative imbalance volumes, and Single imbalance pricing where a single imbalance price is used for all imbalance volumes. There are also two methods of determining imbalance prices: average price and marginal price of energy balancing actions.

3 GENERATION ADEQUACY AND INSTITUTIONAL CONTEXT: LESSONS FROM CONTRASTED EXPERIENCES

Beyond market design, a range of other organizational and institutional factors are important drivers of investment incentives and generation adequacy. The debate on the optimal market design with regard to generation adequacy is therefore only insightful insofar as it is put into a broader perspective taking into account the local institutional and organisational environment.

3.1 Adapting market design to local constraints

The allocation of investment risks between the different stakeholders in the electricity industry has been radically transformed by liberalisation. Peaking units can be considered as the most capital intensive technologies and are therefore proportionally more affected by investment risks and their allocation among the different stakeholders. The question of the retail procurement framework is central to the allocation of market price and quantity risks between investors, intermediaries and consumers. The volatility of electricity wholesale prices, the lack of sufficiently liquid forward markets and of long term contracts are argued to render the financing of new capital intensive power plants difficult (de Vries, 2004 Neuhoff and de Vries, 2004). There are concerns that risk aversion, short-termism, and 'herd behaviour' of investors might lead to underinvestment and boom and bust construction cycles (Ford, 1999). Regulatory uncertainty can also be detrimental to investment, particularly peaking units (de Vries, 2004, Brunekreeft and McDaniel, 2005). Opportunistic behaviour by governments, regulators, or system operators could undermine spot market prices and producers' revenues. The perceived threat of regulatory intervention to curb scarcity rents alone may be sufficient to inhibit capital intensive investments and raise the capital cost for investment in generation capacity, leading to higher average prices but no greater security (Brunekreeft and McDaniel, 2005).

It is therefore important to recognise that there are institutional differences between countries that may preclude some desirable solutions. In jurisdictions with a fragile regulatory and/or commercial environment, private investment may be undersupplied without adequate contractual support and sovereign risk guarantees. Many developing countries may fall in this category, where long-term PPAs are the norm for private investors, and where state-owned enterprises may be the default option where the cost of capital is made too high by the excessive perceived risk of contract abrogation. In countries with a more stable regulatory and institutional environment, if there is a perception that the current market does not provide for the appropriate amount of generation capacity, then various approaches are available to resolve the issue. One approach could be to keep consumer franchises in those countries that have not yet introduced retail competition, or even reinstate the franchise in countries in which retail competition has been slow to progress and not lived up to its expectations. Franchise suppliers could sign long-term contracts (or run tender auctions) on behalf of their customers (although this might require that they had some obligation to do so). Long-term contracts reduce the price risk from scarcity pricing in short-term markets and provide the financial assurance to support investment in plant. The cost associated with the elimination of retail competition would have to be weighted against the benefit of improving the demand for longer-term contracting.

Various capacity payment mechanisms promise to offer a less interventionist alternative. The current commitment to retail competition, embedded for example in current EU legislation, reinforces the attraction of capacity mechanisms. They allow some of the difficulties created by retail competition to be addressed without eliminating competition. Again, important institutional differences between countries may make some solutions more adapted to some jurisdictions. If generation and transmission cannot be legally separated, or if there are many separately owned transmission systems, then it may not be acceptable to have a single Transmission System Operator (TSO) manage dispatch and balancing, and instead an Independent System Operator (ISO) may be preferred. It is difficult to provide significant financial incentives to an ISO, while a TSO has adequate assets to bear the risk associated with such incentives, and makes the task of the efficient organisation of balancing, ancillary services and reserve procurement more straightforward. It follows that different countries may need to adopt different solutions, although it also seems reasonable that as countries become better interconnected and address issues of market power, that there should be some convergence to a similar set of good practices.

In addition to these institutional issues, it is important to distinguish between systems that are largely hydro (as is the case to varying extents for some Latin America countries, Nordic countries in Europe, New Zealand and Austria), those that are relatively isolated or where imports are severely constrained (e.g. Britain, Ireland, Italy, Iberian peninsular in Europe), and those that are largely thermal or nuclear and well interconnected. In hydro or mixed system including reservoirs, generation adequacy is more complex than capacity adequacy as hydro energy reserves constitute a critical determinant of the ability of the system to cover the yearly energy demand. Energy availability can vary considerably over the years, which makes revenues from thermal peaking units more volatile and uncertain and raises the risk of investing in such units as compared to thermal systems. Other specific configurations can also magnify the risk of investing in peaking units, such as small systems with few very large generation units (e.g. the Finnish nuclear reactor representing more than 10% of the installed capacity in Finland) or systems with a large share of nuclear generation in case of a generic problem (e.g. France) (Finon and Pignon, 2007).

Another important issue to take into account is the extent of market power, which is the ability of generators to shift the wholesale price of electricity from a competitive equilibrium price, but also includes the prevalence of vertical integration of generation and transmission. The extent to which vertical integration is a problem will depend on the efficacy of transmission regulation and the ease of Third Party Access. Concentration in the supply of balancing and ancillary services can also distort prices and impede entry, reducing contestability. However, market power need not necessarily be detrimental to generation adequacy as it might be in the interest of the dominant generator(s) to maintain secure supplies to reduce the pressure on the regulator to change market design and structure. For instance, a monopoly or oligopoly might implicitly agree with a regulator to slightly over invest and maintain generation adequacy in return of a 'light-handed' regulation approach from the regulator. Finally, another dimension of market power to consider is its impact on investment timing, and the cyclicality of electricity prices and capacity reserve margins.

3.2 Putting market design in context: lessons from contrasted experiences

These institutional differences can now be used to contrast the different market design experiences with and without capacity mechanism. In Britain, National Grid as Transmission System Operator (TSO) has incentives to procure these balancing services at least cost, and is allowed (indeed, even encouraged) to contract ahead if this reduces the cost. Australia has an Independent System Operator (ISO), NEMMCO, a not-for-profit company managing the physical spot market and power system security over the entire national electricity market. To that extent, there is a less sharp distinction between these apparently energy-only markets and markets such as the Nordic countries where the system operator uses reserve capacity as a capacity mechanism. They are alike in relying on market signals feeding back from the balancing market to the spot, over-the-counter and contract markets as the scarcity value of existing capacity is revealed in the various markets. Besides, in Australia generators are paid a Value of Lost Load (VOLL) if the system cannot meet demand because of a shortage of generation capacity (but not, for example, because of a transmission failure).[10] This should produce the same *ex post* payments as the expectation of the *ex ante* capacity payments made under the English Electricity Pool until its termination in 2001.

The other markets that have or are introducing a capacity mechanism all have special features that might lead to distorted signals for timely investment. The Irish market is small and concentrated, generation is in large part state-owned, and the market is the course of evolutionary regulatory change as it slowly integrates with the North and possibly through interconnection with England and Wales. Norway and New Zealand are wrestling with the particular problem of hydropower yearly variability. The Spanish market was overlaid with Competition Transition Contracts that were designed both to recover stranded costs and mitigate the substantial market power of the two main incumbent generators (Crampes and Fabra, 2005, Batlle et al., 2007). Argentina had a system of audited cost-based bidding into the spot wholesale market that clearly needed a supplementary capacity payment. Not surprisingly, bidders could (and did) consider the combination of the energy and capacity bids when competing for space in the market, so the distinction between the two components was to some extent arbitrary.

The US is a special but important case of a prior long history of regulated franchise monopoly utilities under private ownership, still governed by the 1935 Federal Power Act. The Act imposes a duty on regulators to intervene if necessary to ensure that electricity prices are "just and reasonable". If a jurisdiction wishes to restructure its utility and de-regulate the wholesale market, it must provide FERC with evidence that the wholesale market is workably competitive, as competitive prices are, by definition, just and reasonable. If FERC is satisfied that the utility has been restructured to meet this condition, then it grants suppliers "market-based pricing authority" (Joskow, 2000). It is unclear whether markets deemed to pass this test initially would therefore be exempt from subsequent restraints (assuming no change in market structure through, e.g. mergers), or whether a market that, for a given level of spare capacity, would deliver effectively competitive prices, but with lower reserve margin would be susceptible to market manipulation, would force FERC to deem prices no longer "just and reasonable". The Californian events that have so coloured reactions to the ability of liberalised markets to deal with short term scarcity can only be understood in that context. If regulators are now predicted to intervene when prices rise, then investors and banks are likely to be unwilling to invest in states with liberalised markets if they are allowed to price at variable cost during times of adequate capacity but not to earn the rents needed to cover capital costs in times of scarcity. The peculiar US problem is that it remains a regulatory duty to ensure that market power is not unreasonably exercised while scarcity is adequately rewarded.

Given the particular difficulty in an interconnected electricity system such as the US of distinguishing between cases in which prices are high because of genuine scarcity or because of market

[10] In addition, this acts as a price cap on the spot price. There are disputes about the current value of A$10,000/MWh is a correct estimate of VOLL, which is reviewed each year by the Reliability Panel.

manipulation, it becomes attractive (and perhaps even necessary) to devise non-market mechanisms or obligations to reward scarcity. These may then be combined with varyingly aggressive market power mitigation procedures to deal with energy pricing. Jurisdictions (like the whole of the E.U.) that are not subject to such legal regulatory requirements start from a different position, and regulators and policy-makers should not be overly influenced by the special circumstances of the US. That does not necessarily imply that some additional mechanism to reward capacity availability is unwarranted, but it does mean that the legal and regulatory environments are relevant for any such design.

3.3 Side effects and pitfalls of existing capacity mechanisms

Institutional factors might lead policy makers to require stronger generation adequacy guarantees than energy-only markets, through the introduction of a capacity mechanism on a permanent or transitory basis. However, past experience with capacity payments and installed capacity markets revealed that their implementation might prove complex and have unintended incentive effects. Innovative reliability option approaches being implemented in the U.S. look promising but will need to be demonstrated.

3.3.1 Capacity payments in the UK, Spain and Latin America

In England and Wales, the capacity payment mechanism that was put in place after liberalisation was a direct heritage of the pre-liberalisation centralised despatch of the Central Electricity Generating Board. Capacity payments were made to each generator declared available to operate in each half hour, and were set equal to the loss of load probability (LOLP) multiplied by the excess of the value of lost load (VOLL) over the station's bid price (if not dispatched) or the System Marginal Price (SMP) if dispatched.[11] The Pool was successful in facilitating new investment – largely thanks to the franchise on which Regional Electricity Companies could write long-term Power Purchase Agreements – but suffered from the sustained market power of the incumbents (Newbery 1998, Sweeting 2001).

The capacity mechanism was criticised for being prone to manipulation through capacity withholding. The last years of the Pool witnessed the highest capacity payments, particularly 2000/01, while the capacity margin was as high as 25% (Figure 1). Green (2004) suggests that these abnormal payments were not the result of strategic capacity withdrawal, but rather were caused by the anomalous way in which the availability factor of new plants was calculated. The computer program used to calculate the LOLP overestimated the chance of a power failure and was easy to manipulate for generators, as probabilities of "disappearance" of each generating systematically underestimated the amount of capacity actually available and dispatched at peak times, and made it easy for generators to game (Roques et al., 2005).[12] The British Pool experience illustrates the difficulty of computing LOLP *ex ante* in a transparent (and not prone to manipulation) way.

---

[11] VOLL was set administratively at £2,000/MWh in 1990 and was then increased annually by the RPI – in 2000, it stood at £2,816/MWh (data from NGT).

[12] The calculation used average availability, ignoring the fact that some plants are fully available at peak times but less available off peak. Besides, the disappearance ratio represented the probability of a plant not being available on any random day of the year, given that it was available in the previous week: it made no allowance for the various reasons why plant was not available. Lastly, demand reduction offered by large customers received capacity payments, but this demand reduction was anomalously not included within the LOLP calculation, thus increasing capacity payments.

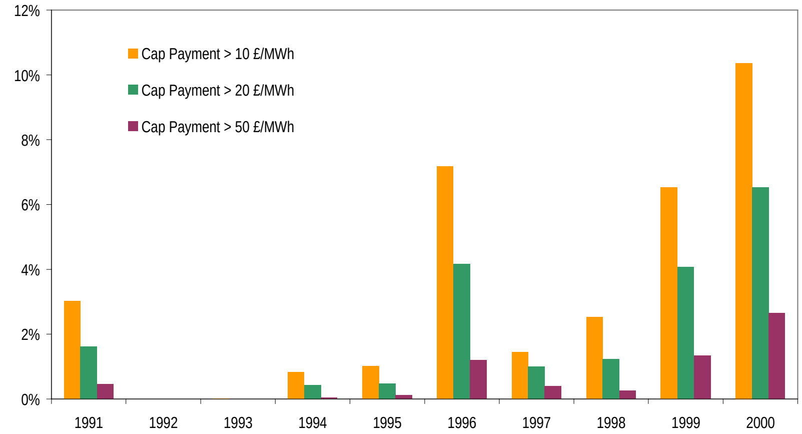


Figure 1: Percentage of time in the year when Capacity payments exceeded a certain threshold (data from Elexon)

Contrary to the British Pool capacity payment which is based on an *ex post* computation, the capacity payments which were first used in Chile in 1982 and later adopted in Argentina, Colombia, Peru and also in Spain, are based on a simplier *ex ante* computation. The method consists in awarding to each generating unit a daily payment (only when it is available) which is computed by multiplying the firm capacity of each generating unit times a per unit capacity payment that may be uniform or may vary with the season. Different approaches are used to determine the firm capacity of the generating units, which represents a measure of the contribution of each generating unit to the reliability of the power system (Batlle et al., 2007). Frequent conflicts have arisen because of the rules of definition of firm capacity of hydro plants and also of different technologies and vintages of thermal plants. In hydro dominated systems like Chile, price based capacity mechanisms provide less insurance than quantity based instruments directly targeting capacity, and the yearly fluctuations of hydro energy makes the definition of the capacity equivalent of hydro plants difficult (Ernesto et al., 2006).

Similarly, the Spanish capacity mechanism suffered from not providing generators with an incentive to be available and producing electricity at times of scarcity, and it was based on a contestable formula for the definition of firm capacity.[13] Moreover, the design of the capacity payment introduced perverse incentives, as a generator would loose the payment if it was declared unavailable, thus creating an incentive to rather bid high enough to be excluded from the dispatch. Finally, it is important to emphasize the impact of the stranded cost recovery contracts in distorting investment incentives in the Spanish market. Their complex definition and distorting impact on bidding behaviour undermined the role of spot market prices and capacity payments in conveying scarcity through price signals (Crampes and Creti, 2005). The stranded cost recovery contracts and the capacity payment rendered the Spanish market susceptible to market power leverage from incumbent generators and demonstrate how the overlay of different – apparently simple regulatory mechanisms – can create unintended bidding and investment strategies incentives.

[13] The procedure consisted in multiplying an average availability rate times by a capacity value that, schematically, is the installed capacity for thermal units and the energy produced in an average year for hydro plants (Batlle et al., 2007).

3.3.2 *The US experience with installed capacity markets (ICAP)*

In the U.S. East Coast markets, ICAP markets were implemented as a natural follow-up of the pre-liberalisation installed capacity obligations that were imposed on regulated utilities. These capacity markets have been very volatile and failed to provide appropriate additional remuneration to peaking units (Oren, 2005, Cramton and Stoft, 2006, Joskow, 2007). The fundamental issue with ICAP markets is that they are disconnected from energy markets, and their short time horizon which is inappropriate to provide adequate incentives for capacity investment. As a result, ICAP markets have proven to be very volatile, with prices either close to zero, when total available capacity exceeds the total demand for ICAP, or very high, when there is capacity storage. Cramton and Stoft (2005) detail other issues with ICAP markets, including: i) the impossibility for new entrants to participate in the auction for ICAP products, increasing the potential for incumbent generators to exercise market power; ii) poor incentives to be available during peak periods as capacity payments were computed using historical availability rather than actual performance; iii) and the failure to incorporate network congestion.

Recent reforms in East Coast ICAP markets attempt to integrate better the capacity markets with energy markets, particularly at times of scarcity. One priority for reform in New England has been to extend the duration of the ICAP obligation to provide better incentives for new capacity investment through the creation of forward capacity markets (Henney and Bidwell, 2007). The New York ISO has put in place artificial demand curves for ICAP that adjust the price smoothly when supply falls short or exceeds the target quantity, with the objective to lower the volatility of ICAP markets by articulating in a smoother way the energy and capacity products. In both cases, reformed capacity market designs rely on generating reserve adequacy criteria, explicitly as in the New England forward capacity markets, or implicitly through the use of an artificial reserve capacity demand curve based on an assessment of the distribution of the loss of load probability and the value of lost load in the New York ISO case. Changes have also been suggested to associate performance obligations and non-performance penalties with ICAP in order to reduce the scope for gaming. Finally, locational ICAP have been proposed to take into account the different value of capacity to the system at different geographical points (Cavicchi and Kolesnikov 2005).

There remains however questions about the compatibility and efficiency of such improved ICAP markets in decentralised systems, particularly with those with retail competition. The implementation of capacity obligations on suppliers in markets with retail competition raises a number of issues. Decentralised contracting, particularly for small suppliers, might entail high transaction costs that make them unattractive. Besides, such approach would also place additional financial burdens on suppliers since it would increase their credit obligations to become credible counterparties to long-term supply contracts (Joskow, 2007). This would risk slowing down an already slow diffusion process of retail competition. Another issue is the compatibility of ICAP markets with interconnected markets (de Vries, 2004). Creti and Fabra (2007) show that if neighbours set the ICAP at different levels then the country with the lower cap will subsidise reserves for the country with the higher cap. Any payments for availability in one market must therefore be associated with adequate penalties for non-delivery if called, to reduce the temptation to sell the output of that capacity into another market.

3.3.3 Innovative Reliability Options

The U.S. debate on the reform ICAP markets revived interest in alternative innovative capacity mechanism designs such as Reliability Options (RO) (Arriaga et al., 2002, Bidwell, 2005, Oren 2005, Cramton and Stoft, 2006). A RO is a call option that requires a plant to be generating or to be supplying reserves (*i.e.*, to be available to generate) when the system is stressed by setting a strike price that is higher than the most expensive unit on the system and, thus, is higher than the maximum price that would normally be seen in a competitive market in non-shortage conditions. The price for RO is set trough a centralised auction run periodically by the system operator to which both existing and potential new plants can participate. The SO sets the RO strike price and the non-performance penalty. If the spot price equals or exceeds the strike price and a plant is selling into the spot market at more than the strike price, the plant pays back the difference to the SO. ROs would be paid for by the SO, which would then pass the net costs (RO purchase price less penalties) through to retailers and, ultimately, to consumers.

The critical advantage of ROs is that they introduce a link between energy and capacity by creating a specific product that connects short-term energy scarcity with longer-term capacity adequacy. The precise product that the SO buys is capacity when the energy price is higher than the strike price, i.e. at times of scarcity. The RO mechanism is therefore a hybrid price and quantity instrument which can be expected to be more efficient than capacity payments or ICAP markets. RO capacity mechanisms also have some additional side benefits, as they directly suppress part of price volatility and mitigate market power in spot markets (Cramton and Soft, 2006).

While interesting in theory, ROs need to be considered with caution when looking at electricity markets which do not have the institutional legacy of the US ICAP markets. ROs are a relatively complex mechanism which is likely to be more easily implemented in systems with an institutional legacy of capacity requirement imposed on suppliers. In systems with an energy only market institutional history, or with a price based capacity mechanism, implementing a RO mechanism might be too radical a change. Most regulators recognize the importance of a stable regulatory framework with incremental improvements, as radical market design changes may have unforeseen effects. The case of Spain is exemplary, where an improvement of the existing capacity payment scheme has been preferred to the radical changes associated with the implementation of a RO scheme (Batlle et al., 2007). While ROs look promising, they are yet untested, and their implementation should provide an insightful precedent. Two areas seem to require further research to assess the efficiency of the promising ROs schemes. First, the central issue is the design of the auction: the lack of bidders may exacerbate the potential for market power leverage in concentrated markets, and the specific terms of the auction (e.g. numbers of years in advance of plant delivery) might favour technologies with a short lead time. The interaction of RO schemes with contract markets and the resulting impact on investment incentives is another issue that would warrant further research. The impact on the equilibrium price in spot markets concentrates most discussions, but assessing the impact on the contract market is equally important, for the extent and type of contract coverage reflects market participants' desire for hedging, and influences the incentive to exercise market power.

4 IMPROVING GENERATION ADEQUACY BY TACKLING THE ROOTS OF THE PROBLEM

With these institutional differences and historical experiences in mind, we now return to the question of how to improve electricity market design to ensure generation adequacy. We argue that capacity mechanisms should come as an optional supplement to reforms of ancillary services, balancing, and spot markets to reduce scarcity price distortions.

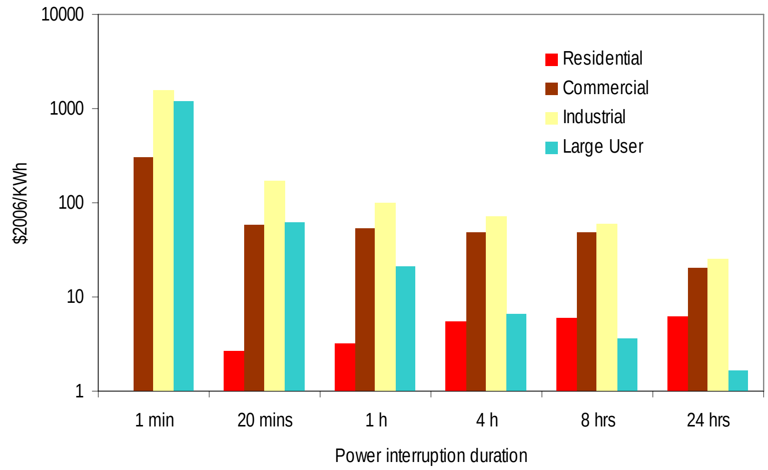
4.1 Setting cost efficient resource adequacy standards?

One of fundamental premises of the restructuring was that market forces will determine the socially desirable level of reliability and will facilitate customer choice between reliability and price. But unless demand is made more responsive to prices, and unless engineering standards and system operator procedures are revised to reduce price distortions at times of scarcity, electricity markets cannot be relied upon to select the socially optimal level of reliability. This does not necessarily imply that electricity markets without additional capacity mechanism will fail to maintain a sufficient level of generation adequacy, as other institutional and organisational factors also have a great impact on investment incentives. While in some electricity systems with specific characteristics (such as hydro-dominated system) or with a particular institutional legacy (such as the US East Coast markets), a capacity mechanism might be needed on a permanent basis to guarantee generation adequacy, in most thermal systems, capacity mechanism should be seen as a transitory and optional tool which could ultimately be phased out in the future if markets design improves and the dissemination of advanced metering equipment and evolved network control systems technology progresses.

The importance of sustaining policy confidence and commitment to the process of electricity market liberalisation seems to have been largely overlooked in the debate on generation adequacy to date. In many countries, the potentially damaging political impact of blackouts and the critical importance of electricity to modern societies make it unlikely that policy makers will trust markets to determine the optimal level of reliability in the near future. It follows that explicit resource adequacy standards may be needed to sustain policy makers' confidence in the liberalisation process. Following the series of blackouts in Europe and on the U.S. East Coast in 2003 and fading confidence in liberalised electricity markets, some countries have reconsidered the need for explicit mandatory resource adequacy criteria: the U.S. Energy Policy Act of 2005 introduced mandatory reliability criteria which are set by the regional reliability councils.

Setting a resource adequacy standard rather than allowing market forces and demand response to set these levels arguably forfeits part of the long term efficiency gains from restructuring (Chao et al., 2007). In countries which choose to impose mandatory resource adequacy criteria, it is therefore important to consider the cost effectiveness of such standards. In theory, resource adequacy criteria should be based on some kind of estimate of the value of lost load (VOLL), but measuring VOLL without market valuation data is a difficult exercise. Cramton and Stoft (2006, p. 40) suggest that conventional US planning reliability criteria based on keeping the probability of blackout very low (e.g. once every ten years) imply a VOLL of $267 000/MWh, well above the range estimated from customer surveys ($2 000/MWh to $50 000/MWh). Another issue is the choice of the socially optimal VOLL level among different categories of consumers. VOLL varies among consumer categories, with the duration of the supply interruption, and also depends as to whether the interruption is scheduled and consumers forewarned or not (Figure 5). In the absence of VOLL figures based on demand side bidding, choosing a second-best average VOLL is the most that is likely to be feasible, coupled with demand-side bidding and/or interruptible contracts for large consumers.

Figure 5: Value of Lost Load by consumer type and outage duration (Source: Kariuki and Allan, 1996)



4.2 Integrating ancillary services with contract markets and encouraging demand participation

Short term reliability criteria and procedures used by the SO have a large impact on the ability of market prices to convey scarcity signals. It is therefore essential to bridge the gap between economists focused on designing competitive market mechanisms and engineers focused on the physical attributes and engineering requirements they perceive as being needed for operating a reliable electric power system (Joskow and Tirole, 2006 and 2007, and Chao et al., 2007). Many of the engineering standards regarding the allowed variation of such power quality parameters as frequency, voltage, etc., have not been subjected to a benefit-cost analysis in the years since wholesale markets were restructured, and would benefit from a joint critical review by both engineers operating the system and economists running the markets. The implementation of good engineering practices depends indeed on close coordination with the design of markets for grid resources and the specification of procedural rules and contract forms, and it must be done within the many constraints imposed by tariff provisions. Increasingly, these problems are eased by technical developments such as more reliable communication and real-time monitoring, more complete state estimators, and frequent re-optimization of dispatch.

Another priority should be to treat symmetrically load and demand and to remove the barriers to demand participation in demand response and possibly ancillary services provision, where possible by integrating them better with the SO needs. The current lack of much demand-side bidding or evidence of price responsiveness is partly because for most activities electricity costs are small and the demand for insurance against price volatility low. While it may be too costly to require 100% of load to be so metered and hence confronted at the margin with real time prices (although ENEL is implementing such a scheme in Italy, as it brings many side-benefits for customer monitoring and billing), in the UK and many other EU markets, about half the load is on interval meters and could be offered a greater variety of sophisticated tariffs, including load management or interruptible tariffs. The way in which demand response is brought into the system at time of scarcity is particularly important in determining the impact on the price formation process. As pointed out by Joskow (2007), demand should be an active component of the price formation process and compete directly with resources on the supply side. ETSO (2005) prescribes that the development of demand responsiveness for as little as 5 % of the peak demand "would substantially contribute to the balance and price during the peak hours".

More generally, the debate on the "missing money" problem in U.S. East Coast markets highlighted the need to improve wholesale markets price formation process at times of scarcity. This requires in the U.S. context to raise price caps used as part of market power mitigation procedures, with perhaps more reliance on other approaches to mitigating market power (Joskow, 2007). There is also a need to revise procedures for procurement and use of operating reserve by the system operator during scarcity times in order not to suppress scarcity price signals. This would require a broader range of operating reserve products procured through ad hoc markets or tendering procedures by the system operator. Perhaps more importantly, a better integration of the procurement and dispatch of reserve and congestion management products with real time spot or balancing markets is necessary to feed scarcity price signals through to contract markets. Most European markets have balancing mechanisms are not well integrated with contract as they rely on "balancing mechanisms" rather than actual "balancing markets", with penalties to discourage imbalances. The initial rationale for such balancing mechanism design penalising imbalances was to encourage parties to self-balance so as to facilitate the system operator's task in maintaining security of supply in real time. Now that SOs have gained experience and confidence, regulators should consider changes in the design of the balancing mechanisms to improve liquidity and integration with spot markets, e.g. by reducing imbalance penalties. In the U.K. for instance, there is much debate as to whether the balancing mechanism should be modified to become more alike a balancing market, bringing better integration with contract markets (Littlechild, 2007). This would involve both moving toward marginal bid pricing of imbalances and removal of imbalance penalties.

4.3 Which role for the System Operator (SO)?

The reform of the balancing mechanism is interlinked with changes in the mandate of the system operator (SO), particularly with the question of whether the SO should have a more active role, through e.g. greater scope and discretion to buy operating reserves ahead of gate closure (Roques et al., 2005). Electricity requires a central SO to balance the system (or sub-system) in real time, and the question then resolves into how that SO function is designed, and whether anything else is needed. An alternative could be to extend the system operator's scope and discretion in balancing markets, possibly coupled with responsibilities extend to generation adequacy. In Finland, Norway, and Sweden for example, SOs have duties which extend to generation adequacy. The main advantage associated with giving the SO stronger incentives to balance the system at least cost and wider discretion could be to reduce the overall system balancing costs. Chao et al. (2007) suggest that the wider use of long-term options on short term reserve capacity could benefit security of supply by easing the TSO's dependence on bids offered in daily markets, enhancing its control over dispatch, and reducing the volatility of the prices paid. But this might also undermine the bilateral markets and the market participants (Littlechild, 2007).

Giving a broader responsibility to the SO however raises important issues of design, monitoring, and evaluation. ETSO (2005) recommends indeed that "to the extent that TSOs have these responsibilities, they must be clearly separated from the system responsibility and financed transparently to avoid distorting impacts on the market.” The difference between systems with a not-for-profit ISO and systems with a TSO that can be financially incentivised is determinant when considering an extension of the role of the SO for security of supply beyond operational time scales. Making the SO responsible for delivering generation adequacy efficiently requires at least three things: first that the SO be legally responsible for maintaining generation adequacy (i.e. a pre-determined reliability criteria), second that the SO be properly incentivised to do so at least cost, and third, that the SO is adequately credit-worthy to bear the risk of the contract position it may need to take. Britain provides an interesting test case. Since 2003, the costs incurred by the UK SO (NGT) when procuring ‘Supplemental Standing Reserve’ as it judges necessary to maintain security of supply for the next winter were not subject to incentive regulation. If NGC were to be made formally responsible for maintaining generation adequacy, then NGC’s price control should include such costs of procuring ‘uneconomical’ reserve ahead of real time (Roques et al., 2005). NGC would then have an incentive to find the most cost-effective way to maintain generation adequacy, be it to contract forward, or to propose any more specific capacity mechanism such as a Reliability Option approach.

Other less controversial changes in the role of the SO could be greatly beneficial. A SO’s choice of an operational protocol usually allows some latitude. The specific choice might be dictated by engineering necessity, but if different procedures are cost-effective then the impact on incentives in the markets can be the decisive criterion. SOs should within their responsibility enhance utilisation of market-oriented approaches for the procurement of the operational reserves and the operation of the balancing market (Chao et al., 2007). Conversely, the use of procedures that distort market prices (such as out-of-market despatch procedures) should be limited as much as possible. More generally, SOs should recognise that their role within operational timescales has a critical impact on scarcity price signals which inform long term investment decisions. These interactions between short term system management procedures and longer term investment incentives somewhat blurs the distinction between system security and generation adequacy. SOs’ actions have far reaching impacts on investment incentives, and SOs therefore need to actively engage with regulators to improve market design and short-term system management procedures.

Finally, SO have an important role to play to monitor investment plans and disseminate information among market players, both with regard to transmission and to generation assets. Under the regulated vertically integrated industry, utilities undertook integrated resource planning that coordinated investments in generation and transmission to meet predicted load growth. Coordination is now more difficult because transmission expansion must be based on predictions about the location and magnitude of new generation units; similarly, investors in generation must rely on predictions about investments in grid resources. The transmission system operator (TSO) can facilitate coordination by proposing a long-term plan for expansion of the grid, thus providing investors in transmission and generation with shared expectations about future developments. Such a plan can anticipate strategic generation investments in response to transmission expansion and recapture some of the lost gains from abandoning the integrated resource planning. Examples of such coordination role in Europe include the British TSO National Grid’s Seven-year Statement, which monitors and disseminates information about future transmission and distribution investments, while the French SO’s (RTE) is in charge of informing and implementing the multi-annual programming of investment by the French Government (through e.g. tendering procedures and if needed to maintain generation adequacy).

5 CONCLUSION

There is as yet no consensus on which market design provides the least distorting long-term investment incentives. Energy only market designs can in theory deliver appropriate investment incentives, but they rely on high prices during periods of capacity scarcity to remunerate peaking units. In practice, some of the system operator reliability engineering practices and standards inherited from pre-liberalisation times distort prices at times of scarcity, which are critical to remunerate peaking units. Besides, with current technology, market signals provide little information on consumers' marginal willingness to pay for supply reliability. But depending on the market structure and the institutional environment, such market failures may not necessarily prevent an energy only market from maintaining generation adequacy. Most liberalised markets remain fairly concentrated, and maintaining generation adequacy is likely to be of mutual interest for incumbent generators and regulators. If there is adequate capacity and market power, there is less urgency to devise new ways of rewarding possibly over-paid generators. If revenue appears inadequate one should investigate the cause of such shortfall, and whether other market design changes might not be more cost effective than the introduction of such supplementary capacity mechanism. The priority should be to ensure that reliability criteria are adequate and cost effective, that responsibilities are clearly allocated, that ancillary services and balancing markets work well and that demand response is encouraged and integrated into the system operator reliability procedures.

If policy makers' risk aversion requires stronger generation adequacy guarantees, alternative institutional and regulatory arrangements should also be examined. One simple option is to determine explicit resource adequacy criteria and to give a clear duty to the system operator to maintain this reliability level (and incentivising the SO to do so at least cost, which might be more feasible with a TSO than in a non-for-profit ISO context). Capacity mechanisms can also be considered on a permanent or transitory basis, although past experience with capacity payments and installed capacity markets revealed that their implementation might prove complex and have unintended incentive effects. Promising new approaches have been implemented in the US East Coats markets, including Reliability Options and forward capacity markets which improve existing designs by linking capacity to energy at times of scarcity and taking a more forward looking perspective. However, it is important to underline the institutional legacy of such jurisdictions, where capacity requirements on utilities were already in place prior to liberalisation. In many other markets – particularly in Europe or in Australia – liberalisation started with fewer distortionary regulatory legislation hanging over from a previous era of private regulated utilities. In many countries there is a working assumption that the market should be allowed to work, and market failures should be addressed at source, not by overlaying offsetting regulations.

6 REFERENCES

Arriaga, I., C. Vasquez, and M. Rivier (2002). A market approach to long term security of supply. *IEEE Transactions on power systems* 17(2): 349-357, May 2002.

Batlle, C., C. Vázquez, M. Rivier, I. J. Pérez-Arriaga (2007). Enhancing power supply adequacy in Spain: migrating from capacity payments to reliability options. Energy Policy 35 (9): 4545-4554, Sept. 2007.

Bazilian, M. and F. Roques (2008). *Analytic Methods for Energy Diversity and Security: A Tribute to Shimon Awerbuch*. In press, *Elsevier* series *Energy Policy and Economics*.

Bidwell, M. (2005). Reliability options. A market-oriented approach to long-term adequacy. *The Electricity Journal* 18(5): 11-25.

Boiteux, M. (1949). La tarification des demandes en pointe: Application de la théorie de la vente au coût marginal. *Revue générale de l'électricité.*

Boiteux, M. (1951). La tarification au coût marginal et les demandes aléatoires, Cahiers du Séminaire d'Econométrie, reprinted in English in 1960 as Peak load pricing, *The Journal of Business* 33(2): 157-179.

Brunekreeft, G. and T. McDaniel (2005). Policy Uncertainty and Supply Adequacy in Electric Power markets. *Oxford Review of Economic Policy* 21 (1): 111-127.

Caramanis, M.C. (1982) "Investment decisions and long-term planning under electricity spot pricing", *IEEE, Transactions on Power Apparatus and Systems* 101 (12): 4640-4648.

Cavicchi and A. Kolesnikov (2005). Ensuring the Future Construction of Electricity Generation Plants: The Challenge of Maintaining Reliability in New U.S. Wholesale Electricity Markets. *IAEE Newsletter*, First Quarter 2005.

Chao, H.-p, Oren, S., Papalexopoulos, A., Sobajic, D., and R. Wilson (2007). Interface between Engineering and Market Operations in Restructured Electricity Systems. *Proceedings of the IEEE*, Vol. 93(11), Nov. 2005: 1984 – 1997.

Crampes, C. and N. Fabra (2005). The Spanish Electricity Industry: Plus ça change… *The Energy Journal,* 26 (European Energy Liberalisation Special Issue): 127–153.

Cramton, Peter and Steven Stoft (2006), "The Convergence of Market Designs for Adequate Generating Capacity," white paper for the California Electricity Oversight Board, March 2006.

Creti, A. and N. Fabra (2007). Supply security and short-run capacity markets for electricity. *Energy Economics* 29(2): 259-276.

de Vries (2004), "Securing the public Interest in electricity generation markets: the myths of the invisible hand and the copper plate", PhD thesis, mimeo.

Ernesto, R., J. Huber, M. Rodrigo, V. Espinoza, and R. Palma-Behnke (2006). Hydrothermal Coordination and Capacity Payment Schemes in Chile - Current Discussion and Future Challenges. Working Paper, Mimeo.

European Commission (2005). Report on Progress in Creating the Internal Gas and Electricity Market. Technical Annex to the Report.

European Transmission System Operators, ETSO (2005). ETSO position paper on TSO role in maintaining balance between supply and demand. Available at www.etso-net.org

Finon, D., G. Meunier, and V. Pignon (2007). The Social Efficiency of Long-term Capacity Reserve Mechanism. Working Paper LARSEN.

Finon, D. and V. Pignon (2007). Electricity and Long-Term Capacity Adequacy. The Quest for Regulatory Mechanism Compatible with Electricity Market. Working Paper LARSEN 2006-2bis, available at www.gis-larsen.org.

Ford, A. (1999). Cycles in competitive electricity markets: a simulation study of the western United States. *Energy Policy* 27 (11): 637-658.

Glachant, J.-M., and M. Saguan (2007). An Institutional Frame to Compare Alternative Market Designs in EU Electricity Balancing. EPRG Working Paper No. 07/11. www.electricitypolicy.org

Green, R. (2004). Did English generators play Cournot? Capacity withholding in the Electricity Pool. CMI Working Paper 41, March 2004. www.econ.cam.ac.uk/electricity

Hunt, S. (2002), Making Competition Work in Electricity, Wiley, New York.

Joskow, P. (2000). Deregulation and Regulatory Reform in the U.S. Electric Power Sector," in *Deregulation of Network Industries: The Next Steps.* S. Peltzman and C. Winston, eds., Brookings Press.

Joskow, P. (2007). Competitive Electricity Markets and Investment in New Generating Capacity. *The New Energy Paradigm* (Dieter Helm, Editor), Oxford University Press.

Joskow, P. and J. Tirole (2006). Retail Electricity Competition. *Rand Journal of Economics* 37(4), 799-815, Winter 2006.

Joskow, P. and J. Tirole (2007). Reliability and Competitive Electricity Markets. *Rand Journal of Economics*, 38(1), 60-84, Spring 2007.

Kariuki, K. and R. Allan (1996). Evaluation of reliability worth and value of lost load. Generation, Transmission and Distribution, *IEEE Proceedings* 143 (2): 171 – 180.

Litlechild, S. (2007). Electricity Cash Out Arrangements. EPRG and OFGEM document. www.electricitypolicy.org

Neuhoff, K. and L. De Vries (2004). Insufficient Incentives for Investment in Electricity Generation, *Utilities Policy* 12 (4): 253-268.

Newbery, D., (1998). Pool Reform and Competition in Electricity, ch.5, pp. 117-166 in M. Beesley (ed.) *Regulating Utilities: Understanding the Issues*, London: Institute of Economic Affairs.

Oren, S. (2005). Generation Adequacy via Call Option Obligations: Safe Passage to the Promised Land. *Electricity Journal* 18(9): 28-42.

Roques, F. (2006). Investment Incentives and Security of Supply in Liberalised Electricity Markets. PhD thesis, University of Cambridge.

Roques, F., D. Newbery and W. Nuttall (2005). Investment incentives an electricity market design: the British experience. *The Review of Network Economics* 4(2) : 93-128.

Roques, F., D. Newbery and W. Nuttall (2007). Fuel mix diversification incentives in liberalized electricity markets: A Mean–Variance Portfolio theory approach, *Energy Economics* (2007), doi:10.1016/j.eneco.2007.11.008

Stoft, S. (2002), Power Systems Economics - Designing Markets for Electricity, IEEE Press, John Wiley & Sons.

Sweeting A. (2001). The effect of falling market concentration on prices, generator behaviour and productive efficiency in the England and Wales Electricity Market. MIT Working Paper, May 2001, Mimeo.

Vardi, J., J. Zahavi and B. Avi-Itzhak (1977). Variable load pricing in the face of loss of load probability. *Bell Journal of Economics* 8(Spring 1977): 270-288.