Insuring against Disruptions of Energy Supply

The Question of Generation Adequacy in Liberalised Electricity Markets

L.J. de Vries

and

R.A. Hakvoort

THE QUESTION OF GENERATION ADEQUACY IN LIBERALISED ELECTRICITY MARKETS

INDES WORKING PAPER NO. 5/MARCH 2004

L.J. DE VRIES AND R.A. HAKVOORT[*]

Abstract

T his paper presents an overview of the reasons why unregulated markets for the production of electricity cannot be expected to invest sufficiently in generation capacity on a continuous basis. Although it can be shown that periodic price spikes should provide generation companies with sufficient investment incentives in theory, there are a number of probable causes of market failure. A likely result is the development of investment cycles that may affect the adequacy of capacity. The experience in California shows the great social costs associated with an episode of scarce generation capacity. Another disadvantage is that generation companies can manipulate price spikes. This would result in large transfers of income from consumers to producers and reduce the operational reliability of electricity supply during these price spikes. We end this paper by outlining several methods that have been proposed to stabilise the market, which provide better incentives to generation companies and consumers alike.[**]

[*] L.J. De Vries and R.A. Hakvoort are at the Delft University of Technology, Jaffalaan 5, 2628 BX Delft, The Netherlands +31(0)15.278.1137 (l.j.devries@tbm.tudelft.nlis).

[**] This paper was presented at the 26th annual IAEE Conference, Prague, 4-7 June 2003.

THE QUESTION OF GENERATION ADEQUACY IN LIBERALISED ELECTRICITY MARKETS

INDES WORKING PAPER NO. 5/MARCH 2004

L.J. DE VRIES AND R.A. HAKVOORT

1   Introduction

This paper addresses the question whether liberalised electricity markets tend to invest in generation capacity in a sufficient and timely manner, so that the chance of electricity shortages and the resulting service interruptions remain near the social optimum. There is no scientific consensus whether liberalised electricity markets can be expected to produce adequate capacity levels on a continuous basis. Although there are some cases, most notably California, in which a liberalised market appears to produce insufficient investment in generation capacity, practical experience is too limited and the available cases are far too convoluted by factors such as flawed market design or regulatory restrictions to provide convincing empirical evidence. Nevertheless, the social cost of capacity shortages is so high that a thorough analysis of the issue is available.

Caramanis et al. show that under ideal conditions, electricity spot markets provided efficient outcomes in both the short and the long term (Caramanis, 1982; Caramanis et al., 1982). This theory still stands; the question is whether it applies in practice or whether real market conditions deviate too much from the ideal situation. The belief that unregulated markets in electricity generation can produce an optimal outcome in the long term is widely shared (see Shuttleworth, 1997 and Hirst & Hadley, 1999). Generally, this school of thought asserts that underinvestment would be caused by obstacles to the proper functioning of the market mechanism, such as price restrictions or construction permits. The proper course of action, in this view, is to improve the investment climate by eliminating all extraneous sources of risk, such as regulatory risk and other obstacles to investment.

The lack of scientific agreement about the issue is reflected in the differences among existing electricity markets. Spain and a number of South American systems try to stimulate investment in generation capacity by providing payments for generation capacity (in addition to their revenues from the sale of electricity) (Vázquez et al., 2002). Three systems on the US East Coast – the Pennsylvania New Jersey and Maryland (PJM) system, the New York Power Pool and the New England Power Pool – use a system of capacity requirements to ensure there always is a certain reserve margin (PJM Interconnection, L.L.C., 2001; for an introduction, see Besser et al., 2002). Most European systems, on the other hand, have no specific provisions to ensure the adequacy of capacity. Instead, they rely on the electricity market to provide the incentive for investment. They can be characterised as *energy-only markets*, as the (expected) price of electric energy is the only driver for capacity investment.

We present arguments that cast doubt on the ability of unregulated markets to provide for the security of supply. A number of factors make it plausible that real markets do not follow the theory of spot pricing.

2   The narrow investment optimum

Electricity markets have a different dynamic from other markets because of three characteristics:

- Electricity is a strongly time-limited product. It cannot be stored, other than in pumped-hydro facilities, in a commercially viable way. Nevertheless, the electricity supply system can only function in a stable manner if supply and demand are continuously balanced.

- The supply of electricity is only partially characterised by a gradually increasing marginal-cost function. When all available generation units are producing electricity, no marginal increase is possible in the short term. As a result, the marginal-cost curve ends with a perfectly price-inelastic section.

- The demand for electricity also is highly inelastic. This inelasticity may be caused by the fact that there is no readily available alternative for most applications of electricity. Another factor that is at least as important is that few consumers receive the required price information in time to adjust their behavior. Moreover, electricity consumption is usually measured over long periods, so consumers have no incentive to shift consumption from peak hours to off-peak hours. As a result, few consumers adjust their electricity consumption to the current price of electricity, so that the observed price-elasticity of consumers is extremely low. There are multiple experiments aimed at increasing consumer price-elasticity, but in most electricity systems their impact still is small (Nilssen & Walther, 2001; Roberts & Formby, 2001; Sæle & Grønli, 2001).

The combination of these three characteristics is the reason that most mechanisms that aid the clearing of other markets, such as a delay in the delivery of the good, consumers switching to other goods or higher prices leading to a reduction in demand, are not available in current electricity markets. This situation has significant consequences: first, wholesale electricity prices are highly volatile; second, there is a chance of service interruptions.

With some modifications, the theory of spot pricing still holds, even if demand is assumed to be completely inelastic. The main consequence of demand price-elasticity being insufficient is that there is a risk that the market does not clear – obviously revealed in the case that there is not physically enough generation capacity available to meet demand. At this point, the market does not reach price equilibrium and some load will need to be shed. The cost of installing so much capacity that the chances of load-shedding could be reduced to zero would thus exceed the social costs of the load-shedding, which are to be avoided.

Nevertheless, if the market does not clear, it may be necessary to institute a price cap to protect consumers against overcharging (e.g., Ford, 1999; Hobbs et al., 2001c; Stoft, 2002). If consumers are not involved in real-time price setting, they may otherwise find themselves paying more for electricity than their value of lost load. This price cap needs to be determined carefully, as it affects the attractiveness of investment in generation capacity. The price cap needs to equal the average value of lost load (VOLL), because at this price consumers should, on average, be indifferent about whether they receive electricity or not. Stoft (2002) shows that in a perfectly competitive market, this results in an optimal level of investment in generation capacity, with an optimal duration of power interruptions. Therefore, the theory of spot pricing still is valid, even if demand is fully inelastic. Price caps can be problematic, however, because it is difficult to determine the optimal level, as the value of lost load is difficult to measure (Willis & Garrod, 1997; Ajodhia et al., 2002).

Although theoretically sound, the reliance upon periodical price spikes to signal the need for peaking capacity has some significant weaknesses. To begin with, there is the risk that the price cap is set at the wrong level, resulting in over or underinvestment. There are, however, issues that are more fundamental. The first is that investment in peak generation units is quite risky, so that small distortions of the investment signal may have large consequences. The second is the argument that there is a positive externality associated with investment in peaking units, because the security of supply is a public good (owing to the network character of electricity supply). The third factor is the inevitable development of market power during periods of supply scarcity. These issues are addressed in the next sections.

3 Market failure

Now we discuss a number of factors that may disturb the narrow investment optimum. The following types of market failure can be discerned (based, in part, upon Hobbs et al., 2001b):

- 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 behavior by investors.

Price restrictions

The fact that a price cap may be needed to protect consumers against overcharging in times of scarcity represents a significant risk, because the optimal level of the price cap is difficult to determine. Although the theory is clear that the price cap needs to be equal to the value of lost load, there are many methods of measuring the value of lost load with widely varying outcomes (see for instance Willis & Garrod, 1997 and Ajodhia et al., 2002). The cost of erring is high. A price cap that is not equal to the value of lost load likely results in a sub-optimal level of investment in generation capacity.

Imperfect information

Producers lack the information needed for socially optimal investment decisions (Hobbs et al., 2001b; Stoft, 2000). This increases the investment risk and therefore reduces the willingness to invest. In order to calculate the probability that peak units will operate and to calculate the expected return on investment, generating companies need to know both the stochastic distribution of the demand function (so they know the distribution of the frequency, duration and height of price spikes) and the expected development of total available capacity (Hobbs et al., 2001a). The exact characteristics of the demand function are difficult to estimate, especially in newly liberalised markets for which no long-term sequences of empirical data are available. Moreover, the basic characteristics of demand change over time (for instance because of the introduction of new technologies), which reduces the validity of demand functions based upon historical data.

Regulatory uncertainty

Regulatory uncertainty increases investment risk and therefore adversely influences the willingness to invest. Regulatory uncertainty can be considered as a negative externality associated with changes in public policy. Especially in newly liberalised markets such as most electricity markets, regulatory uncertainty can be a significant factor. Consider, for example, a few of the policy changes that are currently underway in Europe:

- On 25 November 2002, the European Council on Transport, Telecommunications and Energy reached political agreement on amendments to the electricity and gas Directives and a regulation for electricity (EC, 2002).

- The European gas market is in the middle of a liberalisation process. Most notably the development of the gas transport-tariff system, including charges for flexibility and imbalance penalties, is highly uncertain. This has a considerable impact on a business plan involving today's state-of-the-art, gas-fuelled generators.

- Additionally, there is uncertainty about future European Union environmental rules, such as cooling water regulations or the specifics of the proposed CO_2 emissions trading scheme (EC, 2001).

A second source of regulatory uncertainty, with an equally significant impact upon the willingness to invest, exists with respect to the question of whether a period of high prices will give cause to the government or the regulator to implement a maximum price or, if a maximum already exists, to lower it. Volatile prices are not only a risk for investors, but also for regulators because of the public protests they give rise to. Most electricity systems start liberalisation with ample capacity. In fact, the desire to reduce excessive reserve margins was a motivation for liberalisation. If, after the initial excess capacity has disappeared, a period develops in which prices are many times higher than their historical levels, consumers may consider this a failure of liberalisation and demand intervention. Such events occurred in San Diego at the beginning of the crisis in California, when even a brief period of high consumer prices proved politically unacceptable (Liedtke, 2000). The political risk of being held responsible for high electricity prices, whether these are economically efficient or not, translates into a risk for investors of political intervention. Hence price volatility itself brings about regulatory risk, at least until sufficient experience has been gained with liberalised markets that investors know whether they should expect political or regulatory intervention or not (Oren, 2000; Newbery, 2001).

Regulatory restrictions to investment

Obstacles to obtaining the necessary permits may be another cause of underinvestment. Although the social benefits of a proper licensing process are not disputed here, it should be taken into account that they may incur negative side effects. First, the permitting process can be lengthy, thereby increasing the response time of generation investment to an increase in demand. Especially in a situation of incomplete information about the future development of supply and demand, this may contribute to investment risk. A second effect of increasing the lead-time for the construction of new plants is that it may contribute to investment cycles. This subject is further discussed below. A third effect of permits is that they may impose additional requirements on generators, leading to operational constraints to the response to market signals. An example is that cooling-water regulations may restrict operation during periods of hot weather.

Risk aversion

The theoretical approach by Caramanis et al. assumes that generating companies behave in a risk-neutral manner with respect to investment. This is not necessarily the case, especially when many risks themselves are not well understood. Given the many unquantifiable risks in a liberalised electricity market, it is not unlikely that investors in generation capacity choose a risk-averse strategy with respect to generation investment (Vázquez et al., 2002). If all investors do so, none of them lose market share, so the penalty is limited to a loss of sales during periods of supply shortage. Yet this loss of volume is small, when compared to the overall production of electricity and is likely to be more than compensated by the high prices that develop during a period of supply shortage. Therefore, a collective strategy of risk-averse investment behavior is beneficial to the generation companies, as long as this does not attract newcomers to the market. Such a risk-averse investment strategy would lead to less installed capacity than would be socially optimal.

4 Reserve capacity characterised as a public good

Another argument is that there is a positive externality associated with the provision of generation capacity and as a result, a competitive market will under-provide it. The existence of a chance of service interruptions is key to the existence of market failure in generation capacity, according to Jaffe & Felder (1996), because service interruptions are random. The more that generation capacity is available, the higher the reliability of the supply of electricity is. Therefore, they argue, the presence of generation capacity in excess of the capacity that is contracted by market parties ('reserve capacity') provides an additional benefit to all consumers of electricity in the form of higher reliability of service.

This benefit to all users of the system is a positive external effect of the provision of capacity, as the owner of the generation capacity cannot charge consumers for increasing the reliability of service. As the added reliability is non-excludable and non-rival (the reliability of service to all consumers increases), the reserve capacity can be characterised as a public good. The same is true for the withdrawal of firm load: when a consumer reduces his/her load, system demand goes down and the chance of a shortage decreases. Withdrawal of load and the provision of additional capacity have the same positive external effect: they both increase system reliability. Because part of the socially optimal amount of generation capacity is a public good, liberalised electricity markets will tend towards an equilibrium volume of installed generation capacity that is lower than the social optimum. This analysis is corroborated by Pérez-Arriaga & Meseguer (1997), who consider generators to deliver three distinct products: energy, operating reserves and capacity reserves. This implies that when generators are not paid for their capacity reserves, they provide an external benefit.

In market equilibrium, this positive externality would be reflected by consumers without 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. (This is the maximum percentage of load that was disconnected during the crisis in California in 2001.) This means that each individual consumer can expect to be without electricity for 2% of the time that the shortage lasts. 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 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.

Counterargument

This effect in principle could be offset by the phenomenon described in section 2, wherein prices may rise above consumers' willingness to pay during periods of scarcity, depending on the level of the price cap. Therefore, if the price cap is set high enough, the average price over time should still be high enough to recover investment in peaking units. Nevertheless, Jaffe & Felder are right that during periods of ample generation capacity, the electricity price in an energy-only market does not include a premium for the availability of unused capacity; it merely reflects the marginal cost of production of the marginal unit. The combination of the fact that ordinary prices do not reflect the full social value of electricity, while scarcity prices may be too high, may lead towards a tendency of cyclical investment behavior.

5   Investment cycles

A year before the California crisis started, Ford (1999) published a paper in which he used a computer model to show that investment in electricity generation facilities is inherently unstable in a system with rules such as those in California. His explanation is that investment is not aimed at dampening business cycles, which it would do if the right amount of new capacity became available at the right time, but at making a profit. Because investors tend to wait until they are reasonably certain that they can make a profit, and because they tend to overreact (in part because they do not know their competitors' plans), Ford considers the interaction between the price signal that a power exchange provides and investment to be inherently unstable.

Ford's argument is essentially that a combination of risk-aversion and an insufficiently long time-horizon leads to a delay of investment. Because of the low elasticity of supply and demand, the price signal will not indicate scarcity until the capacity margin is so slim that the chance of service interruptions has become unacceptably large. The long lead-time for new investment means that, once a shortage has developed, this shortage becomes worse before it is alleviated with new generation capacity. Ford's argument is reinforced by the argument from the previous section, that generation capacity is undervalued during periods of abundant supply and overvalued during periods of scarcity.

Visudhiphan et al. (2001) contend that investment cycles are not inevitable, as long as investors are able to anticipate market developments. Nevertheless, as we saw above, sufficient information about future supply and demand is lacking. In their simulation, Visudhiphan et al. also find that backward-looking investment, that is, investment based upon recent experience in the market, will lead to investment cycles. Stoft (2002) arrives at the same conclusion. He notes that the distribution of price spikes may be such that investors would need to have a time horizon of several decades to determine the real average revenues from price spikes. If they use a shorter time-horizon, they are bound to overestimate or underestimate their expected revenues.

6   Long-term contracts

The investment risk in peaking capacity could be greatly reduced by the use of long-term contracts. Moreover, long-term contracts would reveal the expected future demand for peaking capacity to generating companies, as the retail companies (who purchase power on

behalf of their customers) would reveal their peak supply and demand when negotiating the contracts. This would improve the availability of information to generating companies and therefore reduce their investment risk. Long-term contracts would remove much of the price volatility, which is a risk for generators and consumers alike. An important additional benefit of long-term contracts for consumers would be the removal of the incentive to withhold capacity during periods of scarcity. So if there are so many benefits to long-term contracts to all involved parties, why are peaking units not covered by long-term contracts in practice? The answer is two-fold:

- first, there is an opportunity for consumers to 'free ride' the system, and as a result, retail companies are discouraged from engaging in long-term contracts for peaking capacity; and

- second, long-term contracts tend to have too short a duration to dampen the business cycle.

Free-riding

The opportunity for free-riding is caused by the fact that, to a large extent, generators do not sell their electricity directly to the final consumers but to retail companies that act as intermediaries. This situation creates an opportunity for free-riding by consumers (Neuhoff & De Vries, 2002), as follows. If long-term contracts are to provide an adequate signal to generators to install sufficient generation capacity, they must pay the generators the average cost of peaking capacity. During periods in which the peaking capacity is not used (which is most of the time), the spot market price will be below the cost of these contracts. Competitive spot-market prices reflect the marginal cost of production, which does not include capital cost. Retail companies that hold long-term, peak-load contracts will therefore have higher costs, as they contribute to financing the capital cost of the peaking unit. Since rational consumers, who are free to choose their retail company, generally choose the cheapest one, they will therefore normally prefer companies without long-term contracts for peaking capacity. Hence, retail companies will be reluctant to purchase long-term peak-load contracts during periods of ample generation capacity.

Long-term contracts for peak load become attractive again to retail companies during periods of scarcity, but then the reverse effect takes place. The retail companies try to obtain average-cost based, peak-load contracts from generators, as these are cheaper than spot prices. Yet the generators would then be unwilling to engage in these contracts, as they can sell their peak load for much higher prices in the spot market. Thus the fact that supply companies intermediate between generators and end-users creates an opportunity for free-riding by consumers, which discourages timely investment in peaking units.

If retail companies were vertically integrated with production companies the dynamics would change, but the companies would still not receive a sufficient incentive to invest in peaking capacity. Now the reason is the public-good character of the security of supply, as mentioned above in section 4. Vertically integrated companies would still have an incentive to keep insufficient peaking capacity, because this would allow them to sell electricity at a lower price than competitors with a complete portfolio. Consumers who purchase electricity from a company with insufficient peaking capacity would be free-riding on consumers who purchased (more expensive) electricity from a company with a complete generation portfolio.

Not long enough

The second failure of long-term contracts is that they generally are not long enough (Ford, 1999). Long-term contracts would need to extend beyond the current phase of the business cycle to cover at least the next phase in order to dampen the business cycle. The physical inertia of the electricity sector and the close relationship between demand growth and the general economy cause the business cycle of the electricity sector to be long, probably in the order of one or two decades. Because neither investors nor consumers have this long a time-horizon, the contract length is generally shorter. The result is, however, that long-term contracts do not represent a true average price of electricity, and thus they tend to extend the current phase of the business cycle.

Even if generating companies were to receive the appropriate long-term demand signals, there is a final problem with long-term contracts: the slow learning curve of consumers. The long time it takes to develop new capacity and the long life cycle of generating plants provides a serious obstacle to reaching an efficient equilibrium (Vázquez et al., 2000). If investment signals depend upon consumers entering into long-term contracts to hedge their risk of supply interruptions, consumers need to have the opportunity to learn how to find contracts that are attractive to them. As they would mainly learn through trial and error, this would require repeated periods of shortage and high prices. Owing to the length of the business cycle, consumers have few opportunities to learn how to find attractive contracts that hedge their risk of supply interruptions. Moreover, it is likely that each period of shortages will result in changes of the market rules by the regulator, so that the learning curve would need to start over.

The result is that consumers will probably never learn to cover all of their future demand with long-term contracts, and thus electricity shortages will occur time and again and the market will never reach equilibrium. Even if end-consumers or their suppliers have a proper incentive to enter into long-term contracts for peaking capacity, it would take an unacceptably long time before they would know what their actual (long-term) needs are and how to negotiate these contracts.

Even if consumers were to consider a long enough period to average the swings of the business cycle, the risk to generation companies of such contracts would probably be too large. In the course of several decades, the fuel markets are likely to change, generation technology and environmental regulations may change and the uncertainty about the development of demand is large. So here is the so-called '*Catch-22*' of long-term contracts: a short time-horizon does not isolate the contracts sufficiently from the business cycle, while a long time-horizon carries too much risk.

We may conclude that investors lack the incentive, the time horizon and the education to engage sufficiently in long-term contracts. Moreover, consumers have an option to free-ride. They also lack the time horizon and the required sophistication. Therefore, we may conclude that while long-term contracts may cover a significant portion of generation in a mature market, we cannot expect them to cover peak load capacity. Especially during a period of excess capacity and low prices, a shortage of long-term contracts for peaking capacity appears to be likely. Consequently, it is to be expected that this period of excess capacity is followed by a period of power shortages.

7   Market power

When high price peaks occur, the incentive may be quite strong to withhold generation capacity from the market, as was demonstrated during the electricity crisis in California (Joskow & Kahn, 2002). When the capacity margin is slim, or when acute shortages already exist, the low price-elasticity of demand means that a small reduction in the supply of electricity may lead to steep price increases. In that case, even a small market share may provide enough market power to raise prices by keeping some generation capacity off the market. The temptation will be strong to withhold generation capacity, for instance by listing generating units as requiring unscheduled maintenance.

Stoft (2002) points out that if the price cap is absent or very high, for instance equal to the value of lost load, the increase in profits from withholding can be so high that it becomes attractive even for small generators, who would have to withhold the majority of their generation capacity. As a result, many generating companies, not just the large ones, have market power during a period of scarcity. The increases in profit that result from withholding are large, while it is difficult to take judicial steps against this behavior (because one would have to prove which generators were withholding illegally for each hour during which the withholding occurred, and that these generators were not legitimately out-of-service). The attractive incentives to withhold capacity when it is needed most is a fundamental weakness of electricity markets that rely on price spikes to signal the need for investment, even in the absence of other forms of market failure.

Aside from the chance of being caught for abuse of market power, the only disadvantage of this strategy, from the point of view of generating companies, is that withholding electricity during a period of scarcity may cause such a crisis that it prompts a complete overhaul of the market design, as it did in California. Therefore, an established oligopoly of large generators may choose a more stable, long-term strategy.[1] If generation companies are able to keep prices above the competitive level during normal market conditions, they may opt to overinvest in order to discourage new entrants. The presence of excess generation capacity would serve as a threat to new market participants, wherein the incumbents would be able to meet all demand, if necessary, at a price equal to the marginal cost of production. A second reason for an oligopoly to invest more than would appear to be economic could be that it would place an extra value upon reliability, because service interruptions would attract undesired (political) attention. Allowing an oligopoly to develop is hardly an attractive option, however, because it would undo many of the gains from liberalisation (Newbery, 2002).

8   Trade among electricity systems

An entirely different aspect of the issue is how to ensure generation capacity in the presence of significant volumes of trade among systems. In theory, trade among liberalised electricity systems should not change the basic market dynamics. If the involved systems are liberalised in similar ways, trade among them only represents a scale increase. The scale of the system does not affect the question of generation adequacy, as it is addressed in this paper. A benefit of a larger, interconnected system is, however, more stability, as the relative impact of individual generators and capacity additions becomes smaller.

---

[1] Many markets are dominated by a few large suppliers. For example, the French, Belgian, Portuguese, Italian, Greek, Danish and Irish markets are dominated by one or two generators, while only three or four producers serve two-thirds or more of the markets in Germany, Austria, Sweden, the Netherlands and Spain (see the article "EU energy markets face cohesion barriers" in *Power Europe* (4), 4 January 2002.

In practice, interconnected electricity systems often have quite different market rules; the rules for using interconnectors are different from the regular transmission-access rules within the systems. Therefore, the markets that function within the interconnected systems are not fully integrated, but incompletely linked. This has repercussions upon the generation adequacy in the different markets.

In the case of California, for instance, part of the problem was that investment in generation was not only lagging in California itself, but also in neighboring states. There, however, it did not lead to a shortage, but to a reduction of capacity reserves. When the weather suddenly caused a shortage, these states used their own generation resources to meet their own demand first, selling to California only what excess electricity was left. As a result, California, the importing state, bore the full brunt of a crisis, the roots of which were actually spread among a number of states.

In the European Union, a similar scenario is possible. Article 23 of the Directive allows member states 'in the event of a sudden crisis' to take unspecified 'safeguard measures' (Directive 96/92/EC).[2] This can be interpreted as giving member states the right to close down interconnectors temporarily in an emergency. Although there may be sound technical reasons for doing so, this means that in the case of a crisis, the EU internal market may fall apart into a number of disconnected national markets. With respect to generation adequacy, this is an important issue, as each country should thus have sufficient reserves of its own to guarantee adequate supply under adverse conditions.

9 Risk asymmetry

The experience of California shows that the social costs of deviating from the optimum can be high. The outages in California totaled 30 hours, spread over six days. The largest amount of electricity not served at any given time was 1000 MWh, although much of the time the outages affected a much smaller volume of load (Hawkins, 2001). To place these figures in perspective: during 0.3% of the year, a maximum of 2% of electricity demand was not served. Despite the seemingly small proportion of time in which outages actually occurred, the estimated social costs of the crisis are $45 billion (Weare, 2003).

Clearly, a relatively small deficiency of the system can have significant, negative social consequences. This observation is corroborated by efforts to estimate the value of lost load, which is the cost to customers of not being served with electricity. The value of lost load is usually estimated to be some two orders of magnitude higher than regular electricity prices, i.e. well in excess of €1000/MWh, although it is difficult to estimate accurately (Willis & Garrod, 1997). For example, in Australia the value of lost load was determined at 20,000 AUS$/MWh, which is about €11,000/MWh (Australian Competition and Consumer Commission, 2000).

From these considerations, the conclusion may be drawn that the provision of electricity is characterised by a strongly asymmetric loss-of-welfare curve. The loss-of-welfare owing to underinvestment by a certain amount is higher in order of magnitude than the loss-of-welfare owing to overinvestment by the same amount.[3] In this view, the likelihood of

---

[2] Directive 96/92/EC of the European Parliament and of the Council of 19 December 1996 concerning common rules for the internal market in electricity, Official Journal 1997 L 27, pp. 20-29.

[3] For a similar view with respect to transmission capacity, see Hirst (2000).

underinvestment because of the factors described in the previous section is a serious risk, which it is worth considerable cost to avoid.

Two strategies to reduce the risk of underinvestment to society can be proposed. One is to deliberately over-invest in the electricity system. While the overinvestment would constitute a loss-of-welfare with respect to the social optimum, it can be considered as a social insurance against the much greater risk of underinvestment. The second strategy is to 'flatten' the investment optimum by changing the dynamics of the electricity system. If demand can be made more responsive to price, a shortage would result less quickly in random rationing and extreme prices. Instead, the least-valuable loads would first reduce their demand. This would reduce the social cost of a shortage from the average value of lost load to the value of lost load of the least valuable customers.

10 Capacity mechanisms

Despite the lack of consensus whether there is a problem, several solutions have been developed, some of which have been implemented.

Mothball reserve

One option that often is proposed is a so-called 'mothball reserve', a collection of mothballed, old plants that can be returned to service if necessary. The question is under what conditions this reserve is deployed. If the market is to perform its regular task and invest in generation capacity, it should be able to rely upon periodical price spikes to finance its investment in peaking units. This means that the reserve should only be deployed at a high price, namely a price equal to the value of lost load. This raises two issues. First, it may be politically unsustainable to allow prices to rise this high for any length of time if they can be lowered by deploying the mothball reserve. After all, the reserve will be something of a public facility. The second issue is that the incentive to withhold capacity by market parties will not be eliminated until the deployment price of the mothball reserve is reached. Nevertheless, if the reserve is to be deployed at any lower price, it will reduce the incentive to invest. This scenario would create a need for a larger reserve.

Capacity payments

A better prospect is offered by methods that convert the irregular revenues from price spikes to a constant revenue stream for generation companies. A primitive form consists of capacity payments, as tried in Spain and several South-American countries and, in a different form, in the former England and Wales Pool. A disadvantage of these payments is that their effect is uncertain: the payments do not necessarily lead to more investment. Instead of fixing the payment level and leaving the investment level to be decided by the generation market, it is more effective to do the reverse. The most promising capacity mechanisms provide a clear signal to the generation market regarding the demand for capacity, but leave it to the market to finance it.

Capacity requirements

The only system that has been tried in practice and that appears to work is the system of capacity requirements, such as the installed capacity (ICAP) system used by PJM on the US East Coast. (For an introduction, see for instance Besser et al., 2002.) In this system, the retail

companies are required to purchase a certain percentage of reserve capacity. The percentage is determined by the regulator. This reveals the demand for reserve capacity. The cost of providing the reserve capacity is passed along to the consumers by the retail companies who contract the capacity. The reserve capacity is tradable and may consist of an interruptible load.

Reliability contracts

A disadvantage of capacity requirements is that they do not provide an incentive to maximise the availability of reserve capacity. An improvement in this respect is provided by *reliability contracts*, a system of call options that the system operator purchases from the generation companies (Vázquez et al., 2002). When the options are called, the producers are required to pay the system operator the difference between the market price and the strike price. Operating power plants are a perfect hedge for the generators: their net income is equal to the strike price. Generation companies that have sold options that are not covered by available generation capacity when the options are called, lose on those options. This provides generation companies with an incentive to sell an option volume that is equal to the available volume of the generation capacity that they control. A second advantage is that the generation companies receive an incentive to maximise the availability of their generation units during periods of scarcity. Overall generation adequacy is determined by the system operator, by the volume of options that the operator purchases.

Capacity subscriptions

Where the previous two systems still contain an element of central coordination, a system of *capacity subscriptions* leaves all variables to the market (Doorman, 2000). In fact, this system may be considered more market-oriented than a traditional, unregulated electricity market, because it allows consumers to choose their level of generation adequacy. In this system, each customer needs to purchase an electronic fuse, which can limit his/her electricity use. The fuses are activated by the system operator during periods of scarcity. Customers can choose the size of their fuses. The fuses are sold by generation companies and need to be covered by available generation capacity. Thus, the market for fuses indicates the total demand for generation capacity and provides generation companies with fixed revenues to cover their investments. This system turns the security of supply into a private good: consumers can choose their own level of generation capacity that they want to have reliably available. The drawback of this system, compared with the previous options, is that it is more elaborate, as it requires the installation of an electronic fuse at each customer site.

A final option is not to end the consumer franchise (Newbery, 2002). With captive consumers, the free-rider problem is solved. Then the retail companies can cover their full projected demand with long-term contracts. But this option will not prove popular with those in favor of a fully liberalised market.

The discussed capacity mechanisms all have the limitation that they have limited effect when they are implemented in the presence of significant import volumes. To be both effective and economically efficient, an interconnected system would need to implement a capacity mechanism for its whole area.

11 Conclusions

In theory, it appears that there is sufficient incentive for generators in an energy-only market to invest in capacity. The recovery of investments, however, would depend upon a small chance of earning high returns during periodic episodes of power shortage. This delicate balance between investment and expected returns may easily be upset by a number of factors, some of which appear inevitable. Long-term contracts do not provide a solution, as there is an opportunity for consumers to free-ride the system and the required contract length is too great. Therefore, it appears likely that energy-only electricity markets will tend to lead to a shortfall of generation capacity over time, possibly resulting in investment cycles.

A second disadvantage of relying upon periodical price spikes to signal the need for investment in generation capacity is that these price spikes can be manipulated by generation companies. This possibility dilutes their effectiveness as an investment signal. Worse, it may result in large transfers of income from consumers to producers and reduces the operational reliability of electricity supply during these price spikes. Several methods have been proposed to stabilise the market and provide better incentives to generation companies and consumers alike. The main effect of these methods is to make the demand for reserve capacity explicit, which reduces the investment risk for generation companies. For consumers, the benefits are better security of supply and lower price volatility.

Bibliography

Ajodhia, V., R.A. Hakvoort and M. Van Gemert (2002), "Electricity Outage Cost Valuation: A Survey", in *Proceedings of CEPSI 2002, Fukuoka (Japan), November 2002.*

Australian Competition and Consumer Commission (2000), *Determination, Applications for Authorisation; VoLL Capacity Mechanisms and Price Floor*, File No. C1999/865, (retrievable from www.accc.gov.au).

Besser, J.G., J.G Farr and S.F. Tierney (2002), "The Political Economy of Long-Term Generation Adequacy: Why an ICAP Mechanism is Needed as Part of Standard Market Design", *The Electricity Journal,* Vol. 15, No. 7, pp. 53-62.

Caramanis, M.C. (1982), "Investment decisions and long-term planning under electricity spot pricing", *IEEE Transactions on Power Apparatus and Systems,* Vol. PAS-101, No. 12, pp. 4640-48.

Caramanis, M.C., R.E. Bohn and F.C. Schweppe (1982), "Optimal Spot Pricing: Practice and Theory", *IEEE Transactions on Power Apparatus and Systems,* Vol. PAS-101 No. 9, pp. 3234-45.

Doorman, G. (2000), *Peaking Capacity in Restructured Power Systems*, Thesis (Ph.D.), Norwegian University of Science and Technology, Faculty of Electrical Engineering and Telecommunications, Department of Electrical Power Engineering, Trondheim.

European Commission (2001), Proposal for a Directive of the European Parliament and of the Council establishing a scheme for greenhouse gas emission allowance trading within the Community and amending Council Directive 96/61/EC, Brussels, 23 October.

European Commission (2002), Amended Proposal for a Directive of the European Parliament and of the Council amending Directives 96/92/EC and 98/30/EC concerning rules for the internal markets in electricity and natural gas, Brussels, 7 June.

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

Hawkins, D. (2001), *The California Report*, power point presentation for the California Independent System Operator (ISO).

Hirst, E. (2000), "Do We Need More Transmission Capacity?", *The Electricity Journal,* Vol. 13, No. 9, pp. 78-89.

Hirst, E. and S. Hadley (1999), "Generation Adequacy: Who Decides?", *The Electricity Journal,* Vol. 12, No. 8, pp. 11-21.

Hobbs, B.F., J. Iñón and M. Kahal (2001a), "Issues concerning ICAP and alternative approaches for power capacity markets", in *Proceedings of the Market Design 2001 Conference, Stockholm 7-8 June 2001,* pp. 7-18.

Hobbs, B.F., J. Iñón and M. Kahal (2001b), *A Review of Issues Concerning Electric Power Capacity Markets*, Project report submitted to the Maryland Power Plant Research Program, Maryland Department of Natural Resources.

Hobbs, B., J. Iñón and S.E. Stoft (2001c), "Installed Capacity Requirements and Price Caps: Oil on the Water, or Fuel on the Fire?", *The Electricity Journal,* Vol. 14, No. 6, pp. 23-34.

Jaffe, A.B. and F.A. Felder (1996), "Should Electricity Markets Have a Capacity Requirement? If So, How Should It Be Priced?", *The Electricity Journal,* Vol. 9, No. 10, pp. 52-60.

Joskow, P. and E. Kahn (2002), "A Quantitative Analysis of Pricing Behavior in California's Wholesale Market During Summer 2000", *The Energy Journal,* Vol. 23, No. 4, pp. 1-35 (retrievable from econ-www.mit.edu/faculty/pjoskow/).

Liedtke, M. (2000), "State Regulators Cut San Diego Power Rates 3-2 Vote: Utilities Commission Approves Cap on Costs in Effort to Ease Shock of Deregulation", Associated Press, 22 Aug.

Neuhoff, K. and L.J. De Vries (2002), *How Uncertainty Affects Investment in Electricity Generation*, mimeo, Cambridge University.

Newbery, D.M. (2001), *Regulating electricity to ensure efficient competition*, paper presented at the CEPR/ESRC Workshop on "The Political Economy of Regulation", London, 1 November.

Newbery, D.M. (2002), "Regulatory Challenges to European Electricity Liberalisation", paper presented at the conference "Regulatory Reform – Remaining Challenges for Policy Makers", Stockholm, 10 June.

Nilssen, G. and B. Walther (2001), "Market-based Power Reserves Acquirement, An Approach Implemented in the Norwegian Power System, with Participation from both Generators and Large Consumers", in *Proceedings of the Market Design 2001 Conference, Stockholm, 7-8 June 2001,* pp. 59-66.

Oren, S.S. (2000), "Capacity Payments and Supply Adequacy in Competitive Electricity Markets", in *Proceedings of the VII Symposium of Specialists in Electric Operational and Expansion Planning*, Curitiba (Brazil), 21-26 May 2000.

Pérez-Arriaga, I.J. and C. Meseguer (1997), "Wholesale marginal prices in competitive generation markets", *IEEE Transactions of Power Systems,* Vol. 12, No. 2, pp. 710-717.

PJM Interconnection, L.L.C. (2001), *Reliability Assurance Agreement Among Load-Serving Entities in the PJM Control Area*, First Revised Rate Schedule, Federal Energy Regulatory Commission (FERC), No. 270.

Roberts, L. and R. Formby (2001), "Market Participant Experiences with Demand Side Bidding and Future Direction", in *Proceedings of the Market Design 2001 Conference, Stockholm, 7-8 June 2001,* pp. 67-76.

Sæle, H. and H. Grønli (2001), "Small customers as active peak power providers in periods of capacity problems", in *Proceedings of the Market Design 2001 Conference, Stockholm 7-8 June 2000,* pp. 77-84.

Schweppe, F.C. (1978), "Power Systems '2000': Hierarchical Control Strategies", *IEEE Spectrum*, July, pp. 42-47.

Shuttleworth, G. (1997), "Getting Markets to Clear", Letter to the Editor in *The Electricity Journal,* Vol. 10, No. 3, p. 2.

Stoft, S. (2002), *Power System Economics: Designing Markets for Electricity*, Piscataway, N.J: Wiley-IEEE Press.

Willis, K.G. and G.D. Garrod (1997), "Electric Supply Reliability, Estimating the Value of Lost Load", *Energy Policy,* Vol. 25, No. 1, pp. 97-103.

Vázquez, C., M. Rivier and I.J. Pérez-Arriaga (2002), "A market approach to long-term security of supply", *IEEE Transactions on Power Systems,* Vol. 17, No. 2, pp. 349-357.

Visudhiphan, P., P. Skantze and M. Ilic (2001), "Dynamic Investment in Electricity Markets and Its Impact on System Reliability", in *Proceedings of the Market Design 2001 Conference, Stockholm, 7-8 June 2001,* pp. 91-110.

Weare, C. (2003), *The California Electricity Crisis: Causes and Policy Options*, Public Policy Institute of California, San Francisco (retrievable from *http://www.ppic.org/content/pubs/R_103CWR.pdf).*

INDES Working Papers Series

No. 1   *Market-based Options for Security of Energy Supply*,
          C. Egenhofer et. al., March 2004

No. 2   *Security of Supply for Natural Gas Markets: What is it and what is it not?*,
          G. Luciani, March 2004

No. 3   *Investments in Gas Pipelines and LNG Infrastructure: What impact on the
          security of supply?*, P. Cayrade, March 2004

No. 4   *Transport Energy Security – the Unseen Risk?*, D. Fisk, March 2004

No. 5   *The Question of Generation Adequacy in Liberalised Electricity Markets*,
          L.J. de Vries & R.J. Hakvoort, March 2004

No. 6   *Social Costs of Energy Disruptions*, V. Constantini & F. Gracceva, March
          2004

No. 7   *Oil Security, Short- and Long-Term Policies*, V. Constantini & F. Gracceva,
          March 2004

About INDES

This publication, part of the INDES Working Paper Series, originates from the project "Insuring against Disruptions of Energy Supply – Managing the Risks Cost-Effectively" (INDES). INDES has been a one-year joint research project under the initiative of the Centre for European Policy Studies (CEPS) together with the Energy research Centre of the Netherlands (ECN) and the Fondazione Eni Enrico Mattei (FEEM). The project was supported by the Fifth Framework Programme and funded by the European Commission Directorate-General for Energy and Transport.

The INDES project focuses on market-compatible, cost-effective security of supply responses by the European Union. Security of supply is understood as insurance against risks, in which responsibility is shared between the EU, member states, energy companies and customers. Thus security of supply is seen as an economic risk-management strategy. Critical to such an approach is first the minimisation of the insurance 'premium' to achieve the degree of security that is politically called for. Second, there is a need to identify the best systemic actor able to 'hedge' the risk. This can be governments, companies, consumers or in some cases, the market itself subsequent to careful design. Based on these premises, INDES research has emphasised two areas: i) costs of energy supply disruptions and ii) costs of potential policy responses. Towards this end, robust methodologies to assess costs and a sound empirical basis for cost data were used as the precondition for informed policy choices reflecting both effectiveness and cost-efficiency. Following this work, INDES research sought to identify the appropriate market-compatible instrument and the associated actors that would convey the process, be they governments, companies or consumers.

INDES has operated around three axes. The first was academic workshops that developed and refined the methodological framework and empirical base. The second was stakeholder workshops that presented and discussed findings with policy-makers and other stakeholders. The third axis has been the promotion of publications – both academic and policy-relevant – that aim at participating in the existing academic debate and influencing policy-makers. For more information on the project and the series of working papers, visit the INDES website at http://www.energymarkets.info/indes/index.html.

About CEPS

Founded in 1983, the Centre for European Policy Studies is an independent policy research institute dedicated to producing sound policy research leading to constructive solutions to the challenges facing Europe today. Funding is obtained from membership fees, contributions from official institutions (European Commission, other international and multilateral institutions, and national bodies), foundation grants, project research, conference fees and publication sales.