THE INSTABILITY OF COMPETITIVE ENERGY-ONLY ELECTRICITY MARKETS

L.J. de Vries
Delft University of Technology, Fac. TPM, Section Energy & Industry
Jaffalaan 5, 2628 BX Delft, 015-278 1137 / l.j.devries@tbm.tudelft.nl

Abstract

In theory, competitive energy-only electricity markets should provide an optimal level of investment in generating capacity. This optimum, however, is easily disturbed, as a result of which there is a risk of an investment cycle. A second significant weakness of energy-only markets is that they are extremely vulnerable to price manipulation during periods of scarcity As a result, prices will be higher and supply less reliable than in the competitive benchmark. Therefore an adjustment to the market structure is called for to ensure generation adequacy and to improve the operational incentives for generating companies.

1. INTRODUCTION

This paper addresses the question whether liberalized electricity markets tend to invest sufficiently in generating capacity, in time, so that the chance of electricity shortages and the resulting service interruptions remains near the social optimum. Power shortages in California, Norway, Sweden, Brazil, New Zealand and, most recently, Italy, cast an increasing doubt upon the ability of liberalized markets to meet the public demands for a stable electricity supply at affordable prices. The different cases vary widely, however, due to the specific technical and regulatory circumstances, so they do not suggest obvious conclusions about the general ability of power markets to maintain generation adequacy. Moreover, most of the examples are in systems which depend heavily upon hydropower, which leaves the question whether systems without hydropower also risk shortages unanswered. The social cost of electricity shortages is so high, however, that a thorough analysis of the issue is in place.

Under ideal conditions (most notably perfect foresight and perfect competition), electricity spot markets can provide efficient outcomes both in the short and in the long term (Caramanis, 1982; Caramanis et al., 1982). The belief that unregulated produce optimal long-term results in practice, too, is widely shared (cf. Shuttleworth, 1997; Hirst and Hadley, 1999). Generally, this school of thought asserts that a shortage of investment in generating capacity is only 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 extraneous sources of risk, such as regulatory risk, and other obstacles to investment. There is a lack of scientific agreement about the issue, however, which is reflected in the differences between existing electricity markets. Spain and a number of South American systems try to stimulate investment in generating capacity by providing capacity payments to generation (Vázquez et al., 2002). Three systems on the East Coast of the USA (PJM, the New York Power Pool and the New England Power Pool) use a system of capacity requirements to secure a certain reserve margin (PJM Interconnection, L.L.C., 2001; see for an introduction Besser et al., 2002). Most European systems, on the other hand, have no specific provisions to ensure adequacy of capacity. Instead, they rely on the electricity market to provide the incentive for investment. They can be characterized as *energy-only markets*, as the (expected) price of electric energy is the only driver for capacity investment.

This paper presents arguments that cast doubt on the ability of energy-only markets to provide security of supply. A number of factors make it plausible that real markets do not follow the theory of spot pricing. Rather, there appears a risk of investment cycles, the damage of which may be aggravated by strategic withholding of generating capacity during periods of scarce supplies.

2. RELIABILITY AS A PUBLIC GOOD

Among the first to suggest the possibility of market failure were Jaffe and Felder (1996), who argued that reliability is a public good. Due to the random nature of service interruptions, the reliability of electricity service is non-excludable and non-rival, also to the extent that it is a function of generating capacity.[1] Because the reliability is the same for everyone on their network, consumers are not able to contract individually for a higher level of reliability. Therefore they have no incentive to pay generating companies for generating capacity which improves reliability, but only for the electricity which they consume. This means, in the reasoning of Jaffe and Felder, that liberalized electricity markets will tend towards an equilibrium volume of installed generating capacity which is lower than the social optimum.[2]

The same result is obtained from an analysis of the perspective of the generating companies. The more generating capacity is available within a certain electricity network, the higher the reliability of the supply of electricity is. Therefore, Jaffe and Felder argue, the presence of generating capacity in excess of the capacity which is contracted by market parties provides an additional benefit to all consumers of electricity in the form of higher reliability of service. However, this ‘extra’ generating capacity does not earn revenue. Therefore the contribution towards system reliability goes unrewarded.

3. THEORY: HOW IT SHOULD WORK

Electricity markets have some specific characteristics, which change their dynamics fundamentally:

- Electricity cannot be stored, other than in pumped-hydro facilities, in a commercially viable way, but supply and demand need to be balanced continuously.
- The supply of electricity is only partly characterized by a gradually increasing marginal cost function. When all available generation units are producing electricity, no increase is possible in the short term. As a result, the marginal cost curve ends with a perfectly price-inelastic section.
- The (observed) demand for electricity also is highly inelastic.

These three characteristics preclude most mechanisms which underpin the price mechanism in other markets, such as delaying the delivery of the good, consumers switching to other goods or higher prices leading to a reduction in demand. This has significant consequences: wholesale electricity prices are highly volatile, and secondly, there is a

1 Because power plants have a reliability of less than 100% and demand varies stochastically, there always is a chance that the available generation resources are insufficient to meet demand. This probability can be reduced by installing more generating capacity, but it cannot be eliminated. Therefore there is an optimum, from the perspective of society, between the cost of installing more generating capacity and the consequent reduction of the social costs of service interruptions.

2 This line of reasoning is only true when consumers are free to select their electricity provider. Providers of captive consumers can be given an incentive to maintain a certain level of reliability and can kept accountable; moreover, they are able to pass the cost of contracting reserve capacity along to their captive consumers.

chance that the market does not clear. In this case, when the supply and demand functions do not intersect, the only way to keep the system physically stable is to impose random service interruptions. Nevertheless, with some modifications, the theory of spot pricing still holds.

When the market does not clear, the market price is undetermined in theory, which means in practice that sellers have unlimited market power. Therefore a price cap is necessary to protect consumers against overcharging (e.g. Ford, 1999; Hobbs et al., 2001c; Stoft, 2002). The price cap needs to equal the average value of lost load (VOLL), because at this price consumers should, on average, be indifferent whether they receive electricity or not. Stoft (2002) shows that in a perfectly competitive market, this results in an optimal level of investment in generating capacity, with an optimal duration of power interruptions. During price spikes, when prices reach the level of the price cap, generating companies can recover their fixed costs.

By adjusting the level of the price cap, the government can influence the investment incentive for peaking capacity. This provides an opportunity to counter the tendency to under-invest which Jaffe and Felder observed. If the price cap is set high enough, the average price over time should still be high enough to recover investment in peaking units. Establishing the right price caps can be problematic, however, because the value of lost load is difficult to measure (Willis and Garrod, 1997; Ajodhia et al., 2002).

4. POTENTIAL CAUSES OF MARKET FAILURE

In the above theory, generating companies' revenues are insufficient to cover their fixed costs most of the time. Only during price spikes do they make sufficient profit so that on average they have a normal rate of return, if the generation volume is optimal. A number of factors may distort this investment optimum, however (based, in part, upon Hobbs et al., 2001b).

4.1 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 (e.g. Willis and Garrod, 1997; Ajodhia et al., 2002).

4.2 Imperfect information

In order to calculate the probability that peak units will operate and to calculate the expected return on investment, generating companies need to be able to estimate the distribution of the frequency, duration and height of price spikes) and the expected development of total available capacity (Hobbs et al., 2001a). In newly liberalized markets there are no long time series available to estimate these data. Moreover, the basic characteristics of demand change over time (for instance due to the introduction of new technologies) which reduces the validity of demand functions based upon historical data. Thus producers generally lack the information needed for socially optimal investment decisions (Hobbs et al., 2001b; Stoft, 2000).

4.3 Regulatory uncertainty

Regulatory uncertainty increases investment risk and therefore adversely impacts the willingness to invest. Regulatory uncertainty can be considered a negative externality associated with changes in public policy. A few examples are the impact of the new European Electricity Directive (Directive 2003/54/EC); the liberalization of gas markets, and future environmental rules such as the recently adopted CO_2 emissions trading scheme (EC, 2002).

A second source of regulatory uncertainty, with an equally significant impact upon the willingness to invest, exists with respect to the question whether a period with 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, as 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 risk to politicians 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 liberalized markets that investors know whether they should expect political or regulatory intervention or not (Oren, 2000; Newbery, 2001).

4.4 Regulatory restrictions to investment

Obstacles to obtaining the necessary permits may form another cause of underinvestment. While the social benefits of a proper licensing process are not disputed here, it should be taken into account that it may create negative side-effects. This may raise the barrier for new market entrants, which may reduce the incentive for incumbents to expand their volume of generating capacity, and the length of the licensing process may add to the already long lead time for new generating capacity.

4.5 Risk aversion

The theoretical approach by Caramanis et al. (1982) 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 non-quantifiable risks in a liberalized electricity market, it is not unlikely that investors in generating capacity choose a risk-averse strategy with respect to generation investment (Vázquez et al., 2002).

5. OPPOSITE ASYMMETRIC LOSS OF WELFARE FUNCTIONS

The issue of investor risk aversion deserves closer scrutiny. A relatively small deficiency of generating capacity can create large social costs, as the experience in California demonstrated. This observation is corroborated by estimates of the value of lost load, which is usually estimated to be some two orders of magnitude higher than regular electricity prices. For example, SEO (2003) estimates the average value of lost load in the Netherlands to be 8600 €/MWh. The cost of excess generating capacity, on the other hand, is limited. The cost of generation typically contributes less than half to the end user price (depending on the share of taxes in the end user price) and the capital costs again are only a portion of total generation costs. Thus excess investment in generating capacity of 10% is likely to lead to an increase in consumer prices by only a few percent.

From these considerations, the conclusion may be drawn that the provision of electricity is characterized by a strongly asymmetric loss-of-welfare curve. The loss of welfare due to underinvestment by a certain amount is orders of magnitude higher than the loss of welfare due to over-investment by the same amount.[3] In this view, the likelihood of underinvestment due to the factors which were described in the previous section is a serious risk, which is worth considerable cost to avoid.

The interest of generating companies is opposite, however. From the perspective of a private firm, excess generating capacity constitutes a complete loss, whereas a shortage of generating capacity only constitutes a chance of a lower turnover. However, if all generating companies are equally risk-averse, and competitors also do not invest, none of them lose market share. Collec-

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

tive risk-averse behavior will lead to a lower than optimal volume of generating capacity, which results in more shortages, more frequent and longer price spikes and therefore higher generator revenues. Thus a lower level of investment may not have negative consequences for the generating companies at all, whereas erring on the side of too much investment certainly will.

Thus we may conclude that generating companies also face an asymmetric loss of welfare function with respect to the optimal volume of generating capacity, but one which is opposite to society's loss of welfare function. In the presence of some of the significant uncertainties mentioned in the previous section, generating companies would prefer to err on the side of less generating capacity, whereas the interest of society is to err on the side of excess generating capacity.

6. INVESTMENT CYCLES

Cautious behavior by generating companies, combined with any of the other potential causes of market failure which were described in Section 4, may lead generating companies to delay investment until the demand for new capacity is certain. Due to the long lead time for new generation facilities, however, and because investors tend to wait until they are reasonably certain that they can make a profit, the new generating capacity might not become available until after a prolonged period of scarce supplies and price spikes. These high prices could prompt an overreaction from investors, leading to an investment cycle. This mechanism was described for the Californian market (before the crisis started!) by Ford (1999). Ford's argument is essentially that a combination of risk-aversion and an insufficiently long time horizon leads to a delay of investment. Due to the low elasticity of supply and demand, the price signal does not indicate scarcity early enough. Ford's argument is reinforced by the Jaffe and Felder's (1996) argument that generating capacity is undervalued during periods of abundant supply: even if price spikes would compensate, the low electricity prices during regular circumstances discourage investment when investors cannot forecast the price spikes adequately. Stoft (2002) and Visudhiphan et al. (2001) also find that backward looking investment, that is, investment based upon recent experience in the market, will lead to investment cycles.

Another reason that investment cycles are likely to occur is that the electricity sector is characterized by steady (slow) growth, mature technology and high capital costs. Investment in new capacity will 'flatten' the supply curve, leading to lower margins. As a result, an 'optimal' level of investment will lead to market prices below the long-run marginal cost of new capacity, as a result of which a period with insufficient investment will follow (Boston Consulting Group, 2003).

7. LONG-TERM CONTRACTS

The standard economic solution to dampen investment cycles and to align investment with demand would be the use of long-term contracts. 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 demand when negotiating the contracts. However, they develop insufficiently because there is an opportunity for consumers to free-ride, and long-term contracts tend to have too short a duration to dampen the business cycle.

7.1 Free riding

The opportunity for free-riding is related to the externality which was signalled by Jaffe and Felder (1996). If long-term contracts are to provide an adequate signal to generators to install sufficient generating 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. Retail companies which hold long-term peak load contracts will therefore have higher costs than those who purchase their peak demand in the spot market, as long as there is no scarcity. Since rational consumers who are free to choose their retail company generally choose the cheapest one, they will 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 with ample generating capacity.

A similar effect exists with respect to consumers who contract directly with producers. Consumers have no incentive to engage in long-term contracts which enhance the security of supply, as they benefit all consumers connected to the same network the same, and their own benefit therefore is small. Consumers have an opportunity to free-ride on other consumers who engage in long-term contracts for peaking capacity.

7.2 Not long enough

The second failure is that long-term contracts 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. However, neither investors nor consumers have this long a time-horizon, so that the contracts do not represent the long-term average cost of electricity. Rather, they tend to represent the current phase of the business cycle, instead of signaling the future need for generating capacity.

A second problem with long-term contracts is the slow learning curve of consumers (Vázquez et al., 2000). Consumers learn through trial and error, but the long business cycle means that they have few learning opportunities. Moreover, the costs of such a learning opportunity are high: only a period of scarcity indicates the need to secure sufficient generating capacity. Finally, it is likely that a period of shortages will result in changes of the market rules by the regulator, so that the learning curve would need to be started over.

We may conclude that investors lack the incentive and the time horizon to engage sufficiently in long-term contracts. Moreover, consumers have an option to free-ride and 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.

8. MARKET POWER

A significant vulnerability of electricity markets is that generating companies have both ample opportunity and strong incentives to manipulate price spikes, as was demonstrated during the electricity crisis in California (Joskow and 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 generating capacity off the market (Stoft, 2002). The temptation will be large to withhold generating capacity, for instance by listing generating units as requiring unscheduled maintenance. The strong incentives to withhold capacity when it is needed most is a fundamental weakness of electricity markets which rely on price spikes to signal the need for investment, even in the absence of other forms of market failure. It undermines reliability, creates undesired wealth transfers from consumers to generating companies and distorts the investment signal of price spikes.

9. CONCLUSIONS

Theory suggests that there is sufficient incentive for generators in an energy-only market to invest in capacity. However, the recovery of investment 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 and the required length is too great. Therefore it appears likely that energy-only electricity markets will tend to lead to a shortfall of generating capacity over time, possibly resulting in investment cycles.

A second disadvantage of relying upon periodical price spikes to signal the need for investment in generating capacity, is that these price spikes can be manipulated by generation companies. This 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.

For these reasons, the market structure should be adjusted to provide a more clear and stable investment signal and to remove the incentive to withhold generating capacity during periods of scarcity.[4]

REFERENCES

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

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 15 (7): 53-62.

Bijvoet, C., de Nooij, M. and Koopmans, C. 2003, *"Gansch het raderwerk staat stil." De kosten van stroomstoringen*, Amsterdam, Stichting Economisch Onderzoek (SEO), available at www.tennet.nl/images/14_5477.pdf.

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.

Caramanis, M.C., R.E. Bohn and F.C. Schweppe (1982) *Optimal Spot Pricing: Practice and Theory*, IEEE Transactions on Power Apparatus and Systems PAS-101 (9): 3234-3245.

Directive 2003/54/EC of the European Parliament and of the Council of 26 June 2003, concerning common rules for the internal market in electricity and repealing Directive 96/92/EC, Official Journal of the European Union 2003, L 176: 37-55.

De Vries, L.J. and Hakvoort, R.A. (2003) *Opties voor voorzieningszekerheid,* Economisch Statistische Berichten 87 (4396): 108-111.

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.

EC (Commission of the European Communities) 2002, *Amended proposal for a Directive of the European Parliament and of the Council establishing a scheme for greenhouse gas emission allowance trading COM(2002) 680 final within the Community and amending Council Directive 96/61/EC*, available at http://europa.eu.int/eur-lex/en/com/pdf/2002/com2002_0680en01.pdf.

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

[4] For an introduction (in Dutch), see De Vries and Hakvoort (2003). For a discussion of the legal aspects of mechanisms to secure generation adequacy, see Knops (paper presented at this symposium).

Jansen, Y., Brogneaux, C. and Whitehead, J. (2003) *Keeping The Lights On, Navigating Choices in European Power Generation*, Boston, The Boston Consulting Group.
Hirst, E. (2000) *Do We Need More Transmission Capacity?*, The Electricity Journal 13 (9): 78-89.
Hirst, E. and S. Hadley (1999) *Generation Adequacy: Who Decides?* The Electricity Journal 12 (8): 11-21.
Hobbs, B., 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 and 8 June 2001: 7 – 18.
Hobbs, B., J. Iñón and M. Kahal (2001c) *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 14 (6): 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 9 (10): 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 23 (4): 1-35, available at econ-www.mit.edu/faculty/pjoskow/.
Knops (2003) *Securing electricity supply: what is the potential of national measures in the European market?* Paper presented at the Research Symposium on European Electricity Markets, The Hague, Sep. 26.
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, Aug. 22.
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, June 10.
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 (Brasil), May 21 - 26, 2000.
PJM Interconnection, L.L.C. (2001) *Reliability Agreement Among Load-Serving Entities in the PJM Control Area, First Revised Rate Schedule*, FERC No. 27.
Schweppe, F.C. (1978) *Power Systems '2000': Hierarchical Control Strategies*, IEEE Spectrum, July: 42-47.
Shuttleworth, G. (1997) *Getting Markets to Clear*, Letter to the Editor in The Electricity Journal 10 (3): 2.
Stoft, S.E. (2002) *Power System Economics: Designing Markets for Electricity*, IEEE Press, Piscataway, NJ.
Willis, K.G. G.D. Garrod (1997) *Electric Supply Reliability, Estimating the Value of Lost Load*, Energy Policy 25 (1): 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 17 (2): 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 7and 8 June 2001: 91-110.