WORKING PAPERS IN ECONOMICS

No. 01/06

EIRIK S. AMUNDSEN AND LARS BERGMAN

PROVISION OF RESERVE CAPACITY ON THE NORDIC ELECTRICITY MARKET: PRINCIPLES AND PRACTICES

Department of Economics

UNIVERSITY OF BERGEN

Provision of reserve capacity on the Nordic electricity market: Principles and practises[1]

By

Eirik S. Amundsen[2] and Lars Bergman[3]

Abstract

System reliability is a key aspect of electricity supply, and the ability to maintain system reliability thus is an important aspect of a liberalised electricity market. But system reliability can be ensured only if there is sufficient reserve capacity at all times. In a liberalised electricity market the provision of reserve capacity is a matter of incentives. The Nordic electricity market, comprising of the integrated Danish, Finnish, Norwegian and Swedish electricity markets, has worked well from a system reliability point of view. A key factor behind this favourable outcome is that the incentives for keeping sufficient reserve capacity have been strong enough. The reason for this is an adequate institutional design point. More precisely the set of markets that is commonly called "the electricity market" includes both regulation and capacity markets, and rules and regulations are such that these markets are well-functioning.

Key words: Nordic electricity market, reserve capacity, peak capacity, capacity market

JEL classification code: L10, L11, L 51, L 94

[2] Professor, Economics Department, University of Bergen. Herman Fossgt. 6, 5007 Bergen, Norway E-mail: eirik.amundsen@econ.uib.no. Scientific Advisor to Institute for research in Economics and Business Administration (SNF). Corresponding author.
[3] President and professor at the Stockholm School of Economics, Box 6501, SE-113 83, Stockholm Sweden, E-mail: lars.bergman@hhs.se.

1. Introduction

During the 1990´s major electricity markets reforms were implemented in the Nordic countries, i.e. Denmark, Finland, Norway and Sweden[4]. The reform process started in Norway 1991, continued in Sweden 1996 and Finland 1998[5], and was completed by Denmark in 2000. During the same period the national electricity markets were opened up for cross-border trade, and a common power exchange, Nord Pool, was established. Along with this development several markets have been established; including the Elspot which is a day ahead spot market, and Elbas, which is an hour ahead spot market used for adjustments (open to Finland, Sweden and recently also to Eastern Denmark). In addition to these markets, and the traditional bilateral market, there is a thriving derivatives market, where futures, forwards and options are traded actively from both inside and outside the Nordic region.

The Nordic wholesale market has, thus, evolved and matured over several years and it seems to be a generally held view that it functions well by now (see von der Fehr et al. 2005). Also, the retail markets seem to be well functioning, at least at a national level (notably in Norway and Sweden), even though a true Nordic retail market with cross border sales has still not emerged (see Amundsen and Bergman, 2005).

However, a well functioning electricity market is also dependent on a good organization of ancillary services and system operation, including provision of reserve capacity to ensure system reliability (see e.g. Hobbs et al., 2001; and Crampton and Stoft, 2005). In particular system reliability must be maintained also in periods with peak demand. But this can be ensured only if there is sufficient generating capacity at all times. Thus, in a liberalised electricity market the system reliability issue is closely linked to the incentives to keep reserve capacity to be used in periods with extreme demand peaks. The peak capacity problem is particularly important in the Nordic countries where instantaneous electricity demand to a large extent depends on temperature, and where extreme demand peaks appear regularly but with extended time intervals.

---

[4] The fifth Nordic country, Iceland has just recently started to liberalize its electricity market and is not included in this analysis.

[5] The Finnish electricity market reform was implemented already in 1995, but it was not until 1998 that Finnish energy companies started to use Nord Pool for power trading.

An often discussed issue is what kind of institutional framework that is suitable for efficient provision of peak capacity in a liberalised electricity market. One alternative is the "energy-only-market" approach, i.e. a market organisation with one or several forward markets and one real-time market for electrical energy. The underlying assumption is that high electricity prices during peak periods would induce generators to provide an efficient amount of peak capacity. The other alternative is the "capacity market" approach, which implies that a specific market for reserve capacity is added to the forward and real-time markets for electricity.

The Nordic countries have each their own approach to handling system reliability and peak capacity issues and each of the countries has been in search of improvements of their existing systems. However, all the four countries have opted for the capacity market alternative. In recent years there has also been a closer contact between the system operators in the Nordic countries with the intension of obtaining a common design and organization of system operations. In this respect a first step was taken in 2002 as a common regulating power market, RKM was introduced. Still there is a long way to go before a common organization of ancillary services and system operations is obtained.

However, in spite of these obstacles on the road to a common system the Nordic electricity market has worked remarkably well from a system reliability point of view. The purpose of this paper is to describe and discuss the institutional framework for managing the system reliability problem that has been adopted in the Nordic countries. To begin with, however, we will briefly discuss the nature of the system reliability problem in a liberalised electricity market and the choice between the "energy-only-market" and the "capacity market" alternatives.

2. The development of generation capacity and electricity consumption

After the deregulation process started the level of installed capacity has shown a rather slow development in the Nordic countries. As can be seen from Fig. 1 the total installed capacity in Sweden has even decreased, whereas the Norwegian generation capacity has only increased by some 3 percent over the last decade. The other Nordic

countries, that started the deregulation process later than Norway and Sweden, have had a somewhat more pronounced expansion of generation capacity.

Fig. 1. Total installed generation capacity in the Nordic countries excluding Iceland (Source: Nordel Annual Reports)
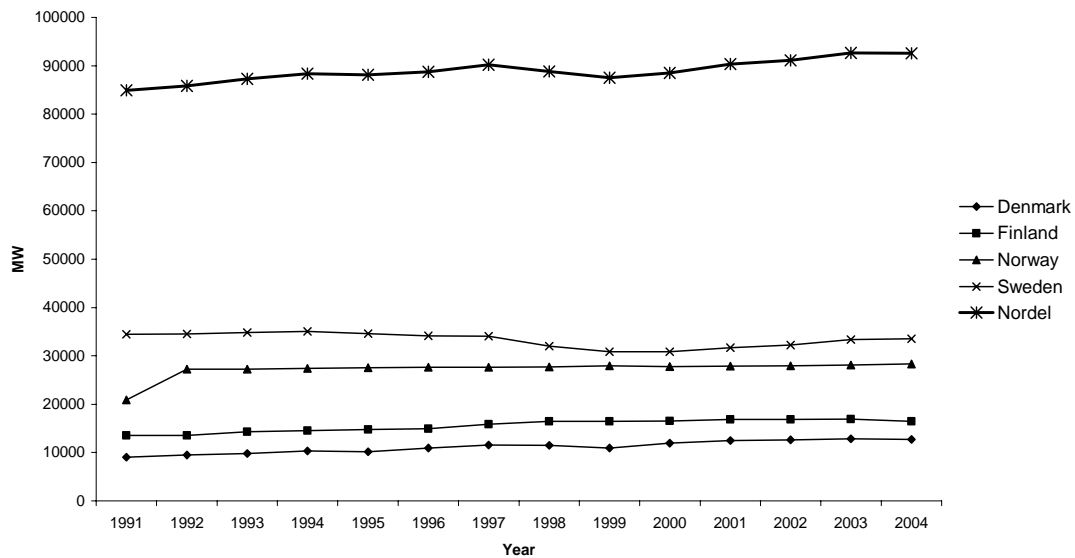
In order to put the development of capacity instalment in perspective, a comparison is made with the development of electricity consumption. Fig. 2 illustrates the relative development of both capacity instalment and electricity consumption in the Nordic countries since 1991. As can be seen from Fig. 2 consumption has increased by some 15 percent whereas installed capacity has increased by about 9 percent. Hence, there is a clear picture of increasing discrepancy between the two features. Thus, one may wonder why capacity instalment is lagging behind. A pertinent question is whether this development in any way can be the result of an ill design of the new electricity markets in the Nordic countries.

Fig. 2. Development of installed capacity and total electricity consumption in the Nordic countries relative to year 1991 (Source: Nordel Annual Reports)
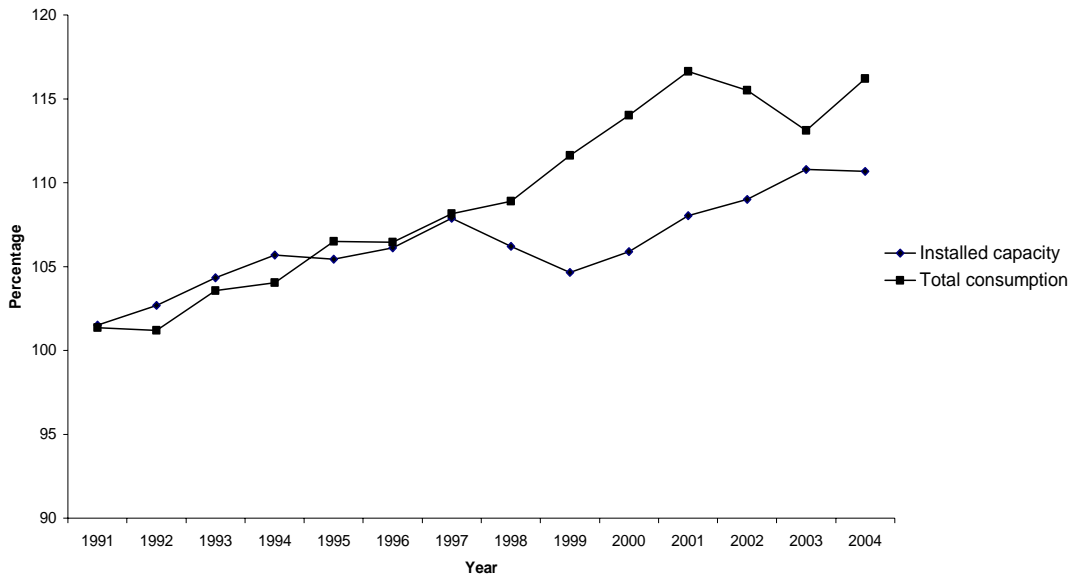
As of now, there does not seem to be a clear answer to this question. However, before concluding that the design of the Nordic electricity market gives insufficient incentives for capacity investment, one should recognize that the Nordic electricity market by the time the deregulation process started was characterized by a sizable excess generation capacity. Hence, one may claim that the relative decline of generation capacity as compared with electricity consumption can be seen as the result of a well functioning market involving a more efficient use of generation resources. Also, one should bear in mind that harsher environmental regulation to a large extent has put constraints on which projects that are acceptable to society and which are not. For instance, expansion of water power projects in Norway and Sweden has more or less come to a halt due to environmental considerations and gas power plants have for a long time been prohibited in Norway for the same reasons[6].

Still the development with respect to new capacity investments does give rise to concern among the Nordic system operators (the SOs). At times demand is close to maximum generation capacity, both within regions of the grid system and nationally. A well known case took place in Sweden on February 5, 2001 when an expected load of 29 GW was projected as compared with a total installation of 31,7 GW. On that

---

[6] Some gas power projects have now been accepted by the authorities in Norway. However, there are strong requirements with respect to the emission of CO_2 and efforts are made to install emission free power plants.

occasion a plea by the SO (Svenska Kraftnät) to customers of reducing consumption probably solved the situation without the controlled "brownouts" that would otherwise take place. With the narrowing gap between maximal load and generation capacity situations like this are likely to appear more frequently and give, thus, rise to problems of sustaining the system. One problem, in particular, is that the spot markets may fail to clear at distinct prices such that the market mechanism ceases to function.

Hence, it is an important challenge to design and implement a good system of reserve capacity and regulating power while taking due account of the interrelationships between the various markets already established. In the following sections we address some basic features and relationships related to this challenge.

3. The nature of electricity as a consumer good

As most other goods, electricity possesses, both a quantity dimension and a quality dimension. In quantity dimension electricity is a so-called *private* good. Basically this means that the electricity that is used by one consumer can not be used by another consumer. This feature is fundamental for the functioning of a market. A necessary condition for a market to emerge is precisely that individual private property rights can be defined and handed over to others. It is only if the good in question is a private good that this condition can be fulfilled.

The quality aspect of electricity has to do with security of supply and is defined by the risk of interruption and by the stability of voltage and frequency. This quality is "produced" through the interaction between aggregate production and consumption in real time (and through the level of maintenance of the transmission and distribution network). If variation of voltage and frequency appear then all consumers connected to the grid will be affected. This relationship implies that security of supply is a so-called *public* good in quality dimension i.e. a good for which individual property rights can not be defined (see Abbott, 2001). Hence, a so-called market failure exists such that it is impossible to organize an ordinary market for quality. Just as for other public goods the responsibility of providing a suitable amount of "quality" will have to be transferred to an agent with a monopoly in performing this task. For the electricity market this agent is the system operator.

The necessity of a continuous balance between production and consumption indicates that all electricity trade takes place in real time on a spot market i.e. on a market for contracts of immediate delivery. From the point of view of the electricity consumers such an organization of the electricity market would, however, imply a high cost of electricity purchase both in terms of time and resources spent. Instead, consumers are primarily interested in ensuring electricity for use and in a necessary amount when the need is there i.e. for lighting, heating, kitchen appliances and so on. For this reason the electricity contracts are predominately of optional character, that is, contracts that give the consumer the right to use electricity within a certain power limit and at a given price when the consumer wishes. The consequence of this is that electricity production at all times will have to be balanced with a demand that will be known with certainty first in the actual hour of consumption.

The task of the system operator is made more complicated by the fact that even the production of electricity is uncertain until the actual hour of delivery. Thus the combination of demand and supply uncertainty and the requirement of continuous balance between supply and demand imply that the system operator must have at his disposal at least some generation plants that may be called upon for regulation purposes. This, thus, implies that a certain spare generation capacity will have to be available at all times to be ramped quickly if demand becomes larger than expected or some generation plant should fail. Conversely, voluntary reductions of demand may be used as an alternative to commanding operation plants.

In the following we apply the term "peak load capacity" to denote the maximum available power capacity[7] of the electricity supply system. Also, the term "reserve capacity" is applied to denote that part of peak load capacity that is not in use but is kept in reserve by the system operator for balancing short term variation of demand. Furthermore, we use the term "regulating power" for power needed in real time for balancing purposes.[8] The market for regulating power is denoted the "regulation market" or the "real time market". With these concepts at hand, the task of the system

---

[7] This is calculated in excess of so called disturbance reserves which is under the direct command of the system operator in e.g. the Swedish electricity system.
[8] The term "balancing power" will be applied to the financial settlement process taking place in the period after real time operations. See discussion of the Nordic regulating market in section 9.

operator may be formulated as, at all times, keeping the sum of reserve capacity and (voluntarily) disruptible consumption at least as large as the sum of the maximum deviance between real and expected consumption that may appear.

4. The need for intervention

As seen from the point of view of a single electricity company, investment in new generation capacity, just as maintenance decisions for existing generation plants, is motivated by the net income that this capacity may give rise to in the future. This income may, in part, stem from contracts with retailing companies (or directly with end users) and in part from the system operator as compensation for using the generation capacity for balancing stochastic variation of demand.

On a well functioning electricity market, the equilibrium price should at all times reflect the marginal cost of generating electricity. In a pure thermal system this implies that the equilibrium price is high under periods of high demand as "short term"[9] generation capacity is in use and low in periods of low demand as "long term" generation capacity is in use. The high prices prevailing under peak load periods, thus ensures the profitability of generation plants that are only in use under such periods. In a hydro power system, where electricity generation easily can be reallocated between periods, this seasonal pattern is generally weaker, but still observable.

If the electricity market is well functioning in this sense, then there is no reason why the system operator or any other party should affect or take responsibility for the provision of peak load capacity. However, there do exist some relationships that imply that the electricity market may function less efficiently. One is the existence of actual or expected regulation of prices. If the government should set a price cap under peak load periods, or the electricity companies believe that this will be the case, then the expected profitability of investments in plants intended for peak periods and thus the supply of peak load plants would be reduced.

---

[9] "Short term" capacity, is to be understood as generation capacity that is intended for use only in peak demand periods e.g. gas turbines. "Long term" capacity is generation capacity that is in use on an annual basis.

Another reason why the electricity market may not function efficiently is market power. If the market structure is such that the market has to rely on a single (or a few) producers to deliver power under peak load demand, this company may be in a position to exercise market power by restraining generation in order to raise price. If this is the case the provision of peak load capacity will be smaller than what is optimal from the point of view of society i.e. that the consumers' marginal willingness to pay for electricity during the peak period is larger than the marginal cost of keeping and generating additional power during this period.

A third reason why the electricity market may not function efficiently has to do with long term stochastic variation of demand in combination with risk aversion on the part of the electricity producers. During years with particularly harsh winters, the utilisation of peak load capacity will be high, in particular in regions with a large proportion of electric heating, and the market will clear at high prices. During years with mild winters the utilisation of peak load capacity will be low as will the price level for electricity. This implies that the investment in peak load capacity only will generate revenue in some of the years and nothing or little in other years.

A risk neutral investor is willing to invest if the expected value of future earnings exceeds the investment cost. A risk averting investor will, however, need a higher rate of return in order to be compensated for the risk implied by the annual stochastic variation of income. This then implies that the combination of risk averting companies and stochastic annual variation of winter temperature may limit the supply of peak load capacity.

Probably the most important reason why the electricity market does not function efficiently has to do with the typical kind of contract that dominates this market. As observed earlier a typical contract gives the consumer the right– at a fixed price settled prior to consumption (and sometimes within a given power constraint) - to use as much electricity as the consumer desires and at whatever time is suitable to the consumer. In addition the settled price is often fixed for a very long time. In practise this implies that the consumer prices during peak load periods are lower, or much lower, than the marginal cost of generating electricity during these periods, while the opposite is true during low load periods.

As compared with an efficiently functioning electricity market this implies that consumption is too high under peak load periods and too low under low load periods. Hence, the diurnal and seasonal variation is increased and the need for peak load capacity becomes larger. Also, the profitability of investing in new peak load capacity, or to keep existing capacity in vigour, becomes smaller. This implies that that the probability of capacity shortage increases and that the periods during the year for which the provision of reserve capacity is too small become more numerous and longer. In extreme cases one may even risk that the electricity market can not clear at any price.

However, it should be observed that the reasons mentioned above as to why an electricity market may function inefficiently are not "market failures" as such and, hence, do not by themselves call for intervention by the system operator or the government. In our opinion there is only one important market failure that calls for intervention in this setting and that is the "public good" aspect of security of supply. As for all public goods the supply will be too small as compared with what is socially optimal if supply decisions are left to private agents. Also, under most circumstances it is not possible for an individual customer to pay a generator for secure uninterrupted supply of electricity as electricity is generally supplied over a jointly used network that is governed by the laws of physics and not by contractual arrangements. [10] Basically the implication of this is that the system operator, on behalf of the customers, will have to demand a sufficient real time reserve capacity to ensure a security of supply that corresponds to the socially optimal level i.e. at the level where the aggregate marginal willingness to pay is equal to the marginal cost of providing that level either by supply or demand responses. The system operator should thus have a position as a monopsonist, though without exercising monopsonistic market power.

It should, however, also be stressed that this does not necessarily mean that the system operator should intervene directly by commanding generators or consumers to act in certain ways to achieve the appropriate level of supply security; or by owning reserve

---

[10] In theory, however, several suggestions exist as to how subscription of capacity may come about; see e.g. Doorman (2000), Hobbs and Iñón (2001) and Vázques et al. (2001).

capacity itself. The system operator may well rely on market mechanisms to ensure a sufficient reserve capacity. However, before addressing such market mechanisms we first consider some factors that in part determine the need for reserve capacity and how an optimal level of supply security may be determined in principle.

5. The need for reserve capacity

The need for regulating power and thus for reserve capacity is in part determined by institutional matters and in particular by the relationship between the forward markets and the real time market. Three features may be pointed to.

The time lag between the clearing of the spot market and real time operation

In planning operations electricity generators normally base their decisions in part on conditions on the spot market and existing long term contracts and in part on prognoses of temperature and other conditions that affect electricity consumption in the given hour of operation. If there is a long time lag between the closing of the spot market and real time operations the conditions that affect electricity use on a short time basis (e.g. temperature) may well change. In general, therefore, the need for up or down regulation and thus the amount of trade on the real time balancing market will increase as the time distance between the clearing of the spot market and real time operations increases. On the Nordic market there is a time lag of 12-36 hours between the closing of the spot market and real time operations. However, in Sweden and Finland (and recently also Western Denmark) companies with a balance requirement have an opportunity to adjust sales or purchases in the so called Elbas market that closes two hours prior to real time. This implies that the demand for reserve capacity will be rather small for these countries. This is contrary to Norway where companies do not have the opportunity to participate in the Elbas market. Consequently, the demand for reserve capacity is, ceteris paribus, larger in Norway.

The incentives for being in balance

An important factor determining the demand for reserve capacity ready to be used in real time, is the companies' incentives of being in balance i.e. to have own production or contracted production that is equal to actual consumption stemming from those contracts that the companies have signed with their customers. In practise a large part of the trade on the spot market is motivated by the companies' desire to be in balance.

11

From the point of view of a single company lack of balance implies either a sale or a purchase of regulating power. If regulating power costs more or less the same as electricity bought on the spot market, a company may not make a large effort of being in balance. However, if it is very costly not to be in balance, the company may go at great lengths in trying to assess actual demand and to be able to balance this with own and/or contracted generation. Thus, in general, there is a negative relationship between the companies' cost of not being in balance and the demand for regulating power. In particular, this is the case in Sweden where there is an asymmetric pricing of regulating power. Hence, a company (subject to a balancing requirement) having a deficit during a specific hour as there is a general need of up regulation, will have to pay a higher price of regulating power - the so called up regulation price - than a company with a surplus during this hour. In Norway there is no such asymmetric pricing and the incentives for being in balance are, ceteris paribus, less strong in Norway than in Sweden.

The relationship between scarcity and end user prices

A third factor of importance for the demand for regulating power is the correlation between end user prices and the price of regulating power. The demand for regulating power is, ceteris paribus, larger the smaller is the correlation between these two prices during the actual hour of operation. For Denmark and Finland and to a large part also for Sweden and Norway this correlation is rather weak. This implies that the end users have no economical incentives in adjusting their actual consumption to the level of actual scarcity on the electricity market.

6. Optimal choice of reserve capacity

Disruption of electricity delivery may come about as a result of acute technical problems in the transportation of electricity (e.g. in transmission, distribution, transformer stations etc). This is normally beyond the control of the system operator. However, disruption may also be due to intentional decisions made by the system operator by exercising his right to disconnect selected end users groups in order to

avoid excessive strain electricity system. In this case we are dealing with "blackouts" that are directly dependent on an insufficient generation capacity[11].

Lack of generation capacity during a given hour may come about if the difference between actual and expected demand becomes larger than the actual supply of reserve capacity[12]. It may also come about as a result of random technical problems in generation plants that may thus reduce actual reserves as compared with expected reserves available. Assuming that the transportation system functions satisfactorily and that the spot market is clearing at a finite price, the risk of a "blackout" or a "brownout" will be linked to the availability of reserve capacity. This capacity will in part consist of available generation capacity ($\bar{x}$) and in part of voluntarily agreed disconnection of consumption ($\bar{y}$). The sum of available reserves is denoted $\bar{q}$.

In the following we denote the risk of disruption/blackout due to lacking reserve capacity by ρ and assume the following relationship is valid:

$$\rho = \rho(\bar{q}); \quad \rho' \leq 0$$

Hence, we assume that the risk of disruption/blackout is non-positively related to the amount of available reserve capacity. The inverse of this relationship, i.e.

$$\bar{q} = \bar{q}(\rho); \quad \bar{q}' \leq 0$$

defines the amount of reserve capacity as function of the risk of disruption/blackout.

Disruption, like large variations of frequency and voltage will imply costs and inconveniences for the consumers. The size of these costs and inconveniences vary a lot between the hours of the year. Blackouts, with direct consequences for heating and

---

[11] Lack of generation capacity may concern all of the system or - due to "bottlenecks"- only concern certain regions.
[12] This may in its turn rely on a fundamental imbalance between electricity demand and the electricity generators' ability and be the result of regulations or other factors that imply that the forward markets do not function efficiently. The calamities on the California power market may be an example of this kind of situation.

lighting, will cause great problem a cold winter's evening whereas a similar blackout may hardly be noticed at all a light summer's night.

The costs and inconveniences caused by "blackouts" and "brownouts" define the consumers' marginal willingness to pay for supply security, or quality; i.e. the maximum amount that the consumers are willing to sacrifice in order to reduce the risk, ρ, marginally. From the inverse relationship stated above the marginal willingness to pay for quality can be transformed into a marginal willingness to pay for reserve capacity. In so doing we let MBV_t^j denote consumer j's marginal willingness to pay for reserve capacity during hour t. It seems reasonable to assume that MBV_t^j is decreasing in $\bar{q}$.

The principle of determining the optimal amount of reserve capacity as seen from the point of view of society may be stated in the following way: During each hour of operation t the amount of reserve capacity should be set in such a way that the marginal cost of reserve capacity during this hour is equal to the consumers' aggregate marginal willingness to pay for reserve capacity. In particular it should be observed that the public good aspect of security of supply implies that the marginal willingness to pay should be summed over all affected consumers.

The marginal cost of keeping reserve capacity during a given hour, MK_t, reflects the revenue that the capacity in question otherwise could generate through contracts signed on the forward market. It seems reasonable to assume that MK_t is increasing in $\bar{q}$. As mentioned this cost varies a lot between the various hours of the year. Furthermore, these variations tend to be positively correlated with the variations in the consumers's marginal willingness to pay for reserve capacity. For example, during cold winter days both the opportunity cost of generation capacity and the consumers' imputed value of security of supply are high.

The optimal amount of reserve capacity as seen from the point of view of society may be determined by the following condition:

$$MK_t(\bar{q}_t) = \sum_j MBV_t^j(\bar{q}_t); \quad \forall t$$

However, in practise it is not possible to calculate all relevant MBV_t^j at an acceptable accuracy. Instead various assessments of costs caused by disruption and variation in voltage and frequency will have to be made and acceptable intervals of voltage and frequency, as well as a maximum acceptable probability of blackouts, are determined on the basis of these. Hence, with the notation at hand one may say that the system operator prescribes a maximum risk, ρ^*, of disruption and/or unacceptably large variation in voltage and frequency. This risk may then be translated into a requirement of minimum amount of reserve capacity available, i.e. $\bar{q}^* = \bar{q}(\rho^*)$.

7. Optimal provision of regulating power and reserve capacity in sequential markets

In the following we investigate basic principles of obtaining an efficient amount of reserve capacity i.e. the reserve capacity that minimizes the cost to society of attaining the level of supply security as determined by the system operator. In so doing we focus on the interconnection between the real time market and the forward market. The real time market is the instrument by which the system operator may buy and sell electricity so as to balance generation and consumption in real time.

Given the typical end users contracts electricity demand during a given hour *t*, may be considered a stochastic variable. Hence, denoting the price of contracts on the forward market by *p* and the demand function for forward contracts by *D=D(p)* , real time demand during a given hour *t* may be expressed as $D(p) + \varepsilon$ where ε is the stochastic (negative or positive) part. As total electricity generation in real time will have to be equal to real time demand and as the producers' generation of electricity in general are equal to the expected demand from settled forward contracts, ε will be equal to the demand on the balancing market. Furthermore, as the demand on real time reflects the settled forward price *p*, which is not influenced by the price of regulating power, ε will be completely inelastic with respect to *p*.

In general each producer of electricity has three options available in employing his generation capacity i.e. generate in order to honour forward contracts engaged in, generate electricity for sale on the balancing market and keep capacity idle as reserve. The producer's supply to the balancing market and, in general, the decision of how to

allocate capacity among options, will simultaneously depend on the price on the forward market and the expected price on the balancing market (and on the price on the reserve capacity market, if available). Furthermore, supply to the balancing market may even stem from certain categories of end users, offering to reduce consumption through so called "voluntary disconnection". On the demand side of the balancing market the system operator decides how much to purchase depending on actual real time demand and in accordance with the decided security of supply level. Demand for regulating power is positive if up regulation is needed and negative if down regulation is needed. The general efficiency criteria in the balancing market implies that the price of regulating power should be equal to the marginal cost of generating regulating power which in its turn should be equal to the marginal cost of voluntary disconnection of electricity consumption.

In order to investigate these efficiency principles further and in particular the relationship between the forward market and the balancing market we assume that the system operator is able to determine both the supply of contracts on the forward market and the generation of regulating power. For simplicity we now ignore that the regulating power may be "generated" by voluntary disconnection and we assume that the power industry is characterized by a constant marginal cost, *c*. Furthermore, we assume there is a given generation capacity, *k* available and that the task of the system operator is to determine how much of the generation capacity that should be available as reserves for up regulation during a future hour, *t*. We consider an hour with high demand, i.e. an hour in which all generation capacity could be applied for delivery as contracts on the forward market. Furthermore, we assume that the allocation of generation capacity between the forward market and the real time market is done prior to the closing of the forward market. Therefore, the need of up or down regulation can not be foreseen with certainty. However, we assume that the distribution of the "states of the world" pertaining to this, ε, is known.

For simplicity we assume that only three states of the world are possible i.e.

$$\varepsilon = 0 \qquad \textit{with probability } s^0$$
$$\varepsilon = \varepsilon^L < 0 \qquad \textit{with probability } s^L$$
$$\varepsilon = \varepsilon^H > 0 \qquad \textit{with probability } s^H$$

such that $s^0 + s^L + s^H = 1$. The task of the system operator is to maximize expected social surplus by determining the supply of contracts on the forward market subject to the constraint that sufficient regulating power be available in real time (i.e. so that the decided level of security of supply is satisfied). The amount of contracts on the forward market, z, must be equal to the expected demand that the settled price on the forward market gives rise to. We assume that z is determined in such a way that the demand for regulating power may be satisfied according to the states of the world given above, i.e. such that $x^L = \varepsilon^L$ and $x^H = \varepsilon^H$. Furthermore, the system operator needs to ascertain that the amount of available generation capacity, k-z is at least as large as regulating power needed for up regulation, $x^H \leq k - z$.

The optimization problem faced by the system operator may, thus, be formulated as follows

$$\max \sum_i s^i \int_0^{z+\varepsilon^i} p(r, \varepsilon^i) dr - s^0 cz - s^L c(z + x^L) - s^H c(z + x^H), \quad i = 0, L, H$$

$$s.t.$$
$$x^L = \varepsilon^L$$
$$x^H = \varepsilon^H$$
$$x^H \leq k - z$$
$$z, -x^L, x^H \geq 0$$

This problem may then be formulated as a Lagrangian problem

$$L = \sum_i s^i \int_0^{z+\varepsilon^i} p(r, \varepsilon^i) dr - s^0 cz - s^L c(z + x^L) - s^H c(z + x^H) - \lambda^L (\varepsilon^L - x^L) - \lambda^H (\varepsilon^H - x^H)$$
$$+ \phi(k - z - x^H)$$

In this expression the Lagrangian multiplier λ may be interpreted as a shadow price of regulating power whereas the Lagrangian multiplier ϕ may be interpreted as a shadow price of reserve capacity. The first order necessary conditions for this problem are

$$\frac{\partial L}{\partial z} = E\{p(z;\varepsilon^i)\} - c - \phi = 0$$

$$\frac{\partial L}{\partial x^L} = -s^L c + \lambda^L = 0$$

$$\frac{\partial L}{\partial x^H} = -s^H c + \lambda^H - \phi = 0$$

The first of these conditions may be rewritten as

$$E\{p(z,\varepsilon^i)\} - c = \phi$$

Basically this relationship states that the amount of contracts sold on the forward market should be such that the expected price on the forward market is larger than the marginal cost of generating electricity. This reflects the essential point that a buffer generation capacity is needed in cases of extreme demand in order to ensure a sufficient level of supply security.

It is important to observe that the amount of reserve capacity available during a given hour is determined prior to the demand of regulating power is known with certainty. This means that the probability of full capacity utilisation of reserves is only s^H. If the decided level of supply security is high, the reserves needed to cope with even extreme deviances between expected and actual electricity consumption during a given hour must also be high. This thus implies that s^H is close to zero such that there is normally idle reserve capacity. In spite of this the shadow price of reserve capacity, ϕ, will be positive for all hours as the sum of the reserve capacity that must be kept in cases of extreme demand constrains the supply on the forward market. In our model, the criteria for this to be the case may be expressed as.

$$D(c) + \varepsilon^H \geq k$$

where $D(c)$ is the electricity demand at the lowest possible price i.e. a price equal to the marginal generation cost. The implication of this condition is that the shadow price of reserve capacity is equal to zero only if the available capacity, k, is so large that it both covers maximum demand on the forward market and the most extreme deviance between actual and expected consumption.

The two remaining necessary conditions of an optimum may be written in the following way:

$$\frac{\lambda^L}{s^L} = c$$

$$\frac{\lambda^H}{s^H} = c + \frac{\phi}{s^H}$$

The expressions on the left hand side may be interpreted as the shadow prices of regulating power realized provided that the corresponding "state of the world" is realized. The implication of the first equality is that the shadow price of regulating power is equal to the marginal generation cost. The second equation states that the shadow price of regulating power is larger than the marginal generation cost if there is an extreme (positive) discrepancy between actual and expected consumption during the given hour.

Assuming that the total amount of reserve capacity for a given hour is equal to $k-z=\varepsilon^H$, and that each producer is paid a price for generating regulating power equal to the shadow prices determined above, the expected net income of the producers will be equal to

$$R^e = s^L \left(\frac{\lambda^L}{s^L} - c\right) x^L + s^H \left(\frac{\lambda^H}{s^H} - c\right) x^H = \phi x^H$$

An interpretation of this expression is that the electricity producers are paid compensation equal to ϕ for every unit of regulating power delivered in case of up regulation. In terms of a market outcome this interpretation corresponds to the "energy-only-market" case. Thus, the producers get p per unit of electricity sold on the forward (day-ahead) market, and ϕ for every unit sold on the real-time market.

However, the expression above may also be interpreted in an alternative way. As the generation of regulating power is equal to demand in "all states of the world" and we have assumed that $k-z=\varepsilon^H$, the above expression may be rewritten

$$R^e = \phi(k - z)$$

The interpretation of this expression is that the producers, instead of being compensated for delivered regulating power, should be compensated for the supply of reserve capacity irrespective of whether it is called upon or not. More precisely the producers would be paid ϕ per MW of reserve capacity, and his net income would be $\phi(k\text{-}z)$. In terms of a market outcome this interpretation of the necessary condition for an optimal allocation of resources corresponds to the "capacity market" case.

Within the frame of the optimisation model the two cases are identical, and thus give the producers the same the expected net income. However, due to several factors the possibility of realising the optimal resource allocation as a market outcome differs significantly between the two cases. Thus the choice between the "energy-only-market" and the "capacity market" alternatives has to be based on institutional and other real-world factors.

8. A market for reserve power

The main result stemming from the analysis in the previous section is that two prices may be called for in ensuring an efficient amount of reserve power as seen from the point of view of society. One is the price of regulating power actually delivered and one is the price of the capacity kept in reserve for possible generation of regulating power. The keeping of a buffer reserve may be considered a specific service. Without a specific price on this service a sufficient provision of reserve capacity during peak demand may not be guaranteed.

The price on the delivered regulating power is the same as the equilibrium price on the balancing market. As for the price of keeping capacity for reserve there are two candidates, as indicated by the analysis in the previous section. One is that the price is determined as the difference between the price of regulating power and the marginal cost of regulating power at extreme discrepancies between actual and expected electricity consumption.[13] The other is that a price on reserve capacity is determined.

---

[13] The system that was in use in England and Wales prior to the current system is the best known example of such a system. During periods of high load a capacity compensation was calculated on the basis of "the loss of load probability" (LOLP) and "the value of lost load" (VOLL). This cost element was shifted over to the consumers by adding an "uplift" on the spot price for the actual half hour.

In practice the first alternative, i.e. the "energy-only-market" alternative, implies that the producers are expected to keep given amounts of generation resources that almost never will be applied, and that the system operator is willing to pay a rather high compensation for the use of these resources during rare and short periods of time. Such a system will function efficiently only if firms are risk neutral. Another problem with this alternative is that the system operator does not know in advance how much capacity that actually will be kept in reserve. From the point of view of the system operator it would be better to have a "quantity based" system, i.e. some kind of "capacity market" system, where the amount of reserves is determined in advance.

A market for reserve capacity may be organized as a bilateral contract market. It may also be an organized market like the Nord Pool futures market. In general, the objective of a market for reserve capacity is to ensure that a sufficient amount of power is available so that the settled requirement of supply security and quality may be attained at the least cost to society. At the same time it must function in such a way that regulating power is always generated at lowest cost, which implies that at least the largest customers should have the possibility of participating in the trade.

A market is an institutional arrangement for trade with rights, e.g. the right to administer a certain commodity or to make use of a given service. The right traded on a market for reserve capacity is the right to manage a certain generation capacity during a period and/or the right to disconnect an amount of power during that period. For simplicity we denote this right be "power options". The price on power options is assumed to be determined in an auction where both producers and consumers participate and where the system operator determines in advance how many rights of power reserves that is going to be purchased.

In the following we consider three alternative types of power option contracts between a producer and the system operator. In all of the contracts the producer gets a certain compensation, equal to the price of the power option, for keeping generation capacity in reserve. The difference between these contracts is the rules by which the capacity owner is compensated for regulating power as the system operator is exercising his rights.

a) Compensation based on a predetermined price of regulating power

In this system the owner of the generating capacity system gets a predetermined price per MWh if and when the capacity is used for generating regulating power. Thus, this arrangement has the form of an option with a "strike price". The system operator could, for instance, base the strike price on calculations of the marginal cost of power generated by the capacity lastly employed. As for a rational capacity owner, he will base his decision to issue an option of this kind on mainly three factors. The first is his evaluation of the probability that the capacity in question is actually employed for regulating power generation during the given period. The second is the net income per generated unit if the capacity is called upon for regulating power generation (i.e. the difference between the predetermined price and marginal cost). The third is the net income which the capacity could give rise to if it were used on the forward market instead.

From the point of view of society, however, it is essential that the capacities in generating regulating power are called upon in order of increasing marginal cost. For this reason, the offer from the capacity owner should contain both information on the requested compensation for putting the capacity unit at the system operator's disposal and information on the marginal cost of utilising this capacity unit.

b) Compensation based on reported marginal cost of regulating power

An alternative way of compensation is that the capacity owner reports the marginal cost at which the capacity in question is able to generate regulating power and that the compensation for generating regulating power is based on the reported marginal cost. From the point of view of the system operator this implies that the capacities enter in terms of increasing cost for each hour and that the capacities are used in this order. As the compensation for regulating power is equal to the marginal cost of generating regulating power it is a matter of indifference for the capacity owner whether the capacity unit is used or not. The compensation needed for a rational capacity owner for issuing an option of this kind will therefore depend on the net income that the capacity unit may generate on the forward market.

c) Compensation based on the equilibrium price of regulating power

A third possibility is that a power option in practise is an obligation of participating in submitting supply bids on the real time market. The compensation paid to the capacity owner for the produced regulating power is then based on the equilibrium price on regulating power. For the system operator this implies that the real time market remains a regular auction market. For the capacity owners giving offers to the market for power options the system implies that they will have to evaluate both the probability for their capacities being utilized and the established price on regulating power on these occasions.

If all parties have perfect information on all relevant factors the three alternatives will give the same result. However, in the more realistic case where information is not perfect and, in addition, asymmetrically distributed the differences can be significant. Assume for instance that the capacity owners possess better information than the system operator on the marginal generation cost of the reserve capacity in question. In case a) the capacity owner may have an incentive to report too low a cost in order to increase the probability for being chosen for regulating power generation.

In case b) the capacity owner may, in contrast to case a), have an incentive to report cost that is too high. Clearly, this implies that the probability of being chosen for regulating power generation diminishes, but the net income will be strictly positive if chosen. Finally, in alternative c) the capacity owner may have an incentive of offering power to the real time market at a price that is lower than the marginal cost of generating the power. This increases the probability that the capacity, that will be kept in reserve anyway, will be called upon to generate regulating power. As long as the price on regulating power is determined by power from generation plants that are offered to the real time market at a higher price, this strategy will be beneficial to a rational capacity owner.

9. Provision of regulating power and reserve capacity in practise

As already concluded two prices are needed to ensure system reliability in a liberalized electricity market: A price on regulating power and a price on the capacity kept in reserve to make sure that the regulation market with a certain (high) probability can clear at all times. Moreover, the second price can either be a premium

on the price of regulating power at times with unusually high demand for regulation power, or a capacity price determined on a specific market for "power options" (also denoted a "capacity market"). As indicated in the introductory section the Nordic countries have opted for the capacity market alternative, let alone in different forms. In this section we will describe the organization of regulation and capacity markets in the Nordic countries, and relate our observations in the previous section.

The regulation market, RKM

In September 2002 common rules were established for the use of regulating power for achieving balance of generation and demand in the Nordic countries. Regulating power is now purchased by the TSOs in Finland (Fingrid), Denmark (Elnet), Norway (Statnett) and Sweden (Svenska Kraftnät) from a common price ladder of submitted bids. Clearing between TSOs is done on the basis of a single price (i.e. the price of regulating power) while the system for "day after" financial clearing between single companies and the TSO varies from country to country. In the future the aim is to harmonize the rules for "day after" clearing.

In principle, all participants in the electricity market are responsible for balancing the provision and use of electricity. A "balance responsible" company often possesses generation capacity suitable for regulation (i.e. electricity that may be ramped quickly). This implies that the company may participate in the regulation market. However, physical "own regulation" is not considered beneficial for the balancing of the system, wherefore it is not allowed in Norway and is made uneconomical in some of the other countries (i.e. Sweden). The number of balancing responsible companies varies in the Nordic countries from around ten in Finland to over one hundred in Norway.

The day-ahead market, Elspot, has the important function of providing a common price to the participants and to offer the opportunity for the companies to trade so that they achieve balance in provision and sale of electricity. After the closing of this market (twenty four hours prior to real time), the balance responsible companies submit their plans to the system operator. However, the companies have the additional opportunity to adjust their decisions either through bilateral contracts or by using Elbas that closes one hour prior to real time (This option is only open to companies in

Finland, Sweden and Eastern Denmark). Within the actual hour of operation the system operator ensures balance by using the regulating power market. The day after imbalances of each company are settled, i.e. the companies are either required to pay for or are compensated for the imbalance depending on the direction of the imbalance.

On the regulating market, bids from companies are collected in a list that is available to all the Nordic system operators in a common information system, NOIS (Nordic Operational Information System). The system operators use this list to keep the load frequency by jointly making decisions on up- or down regulation. In practice local imbalance may to a large extent cancel against each other so that the need for regulating power for the system operator is less than the sum of the needs of the balance responsible companies. On the regulating market the system operator receives bids from companies that are willing to either up regulate or down regulate their generation or demand. A bid on up regulation involves submitting a price at which the company is willing to increase generation or reduce its consumption in a certain quantity; likewise for down regulation. Within the hour of operation the system operator uses the bids in order of increasing asking price, unless bottlenecks or other considerations of supply security make this impossible.

In Norway, Sweden and Finland the participants in the regulating market are all paid the highest asking price of the accepted bids within the hour of operation in case of up regulation. In case of down regulation all are paid the lowest asking price of the accepted bids. In Denmark participants are paid the asking price if their bids are accepted. The prices of up- and down regulation determine the regulation price that is used for settlement between the system operators for the broader regions. If there are no bottlenecks the regulating price is the same in all regions.

In the period after the real time market has been cleared and the regulating price has been established the process of balance settlement between the system operators and the involved parties commences. The purpose of this is to distribute income and expenses for regulating the market. A basic principle applied is that the participant that causes an imbalance has to compensate the system operator for its expenses of re-establishing the balance. The settlement within the regions between the system operator and the participants varies from country to country. Sweden and Denmark

has a system of settling separate imbalances for production, trade and consumption whereas Norway and Finland only settle an aggregate balance. Also, the principles of settlement pricing vary from country to country. Norway uses a "single price" system where the same price is used for both purchase and sale. In the other countries there is a "two price" system where the price charged for purchase or the price paid for sale depends on whether the purchase or the sale supported or counteracted the aggregate balance in the hour of operation. Hence, by this system of asymmetric pricing the incentives of being in balance in real time becomes very strong.

Hence, in general one can say that the Nordic balancing system aims at making the trade on the real time market as small as possible. In practise a large part of the responsibility for balancing is delegated to the balancing responsible companies. For that reason a large part of what would otherwise be trade on the regulating market is shifted over to Elspot and Elbas. Therefore, the discussion of how to ensure the provision of reserve capacity has involved a discussion of Elspot and Elbas.

There may be reason to question the rationality of having each and every balancing company carrying a large part of the responsibility of balancing generation and consumption in real time. Clearly, it is the system operator that with his overview over the development of demand, the magnitude of reserve power and geographical distribution, as well as of the pressure on the "narrow" parts of the transmission network, that should have the best conditions of minimizing the cost to society of managing the system responsibility. On the other hand one can claim that it is the companies that have the best information on relevant generation costs. Therefore, provided that the system operator's pricing of transmission services are efficient i.e. that the prices correctly reflect network losses and scarcity of transmission capacity, decentralized decisions would lead to the most efficient allocation of available resources.

However, if a decentralized solution is to be efficient at least two conditions will have to be fulfilled. Firstly, the closing of the "last" forward market, i.e. Elbas should be as close as possible to real time operations so that big differences can be avoided. Clearly, this condition is fulfilled as it is quite exceptional that demand change significantly within an hours. The second condition is that Elspot and Elbas always

clear at distinct prices. Even though this is normally the case, situations emerge every now and then where Elspot, due to extreme cold weather and thus a continuous high demand, is close to not clearing at a distinct price. One reason for this is that the demand on the spot market is very price inelastic. Another reason is that the provision of peak load capacity is not sufficiently large. Hence, as has been discussed earlier, it is an important task to provide a sufficient buffer of reserve capacity that may handle situations of occasionally high demand.

The Norwegian capacity market

A common system of reserve capacity provision is still not in place on the Nordic electricity market and the countries have all their own arrangements. For instance, in Sweden, the problem of sufficient capacity provision is now being analysed with a view to introduce a new arrangement within a few years. Meanwhile, Svenska Kraftnät, has signed bilateral contracts with power companies that have agreed to keep a certain amount of reserve capacity at Svenska Kraftnät's disposal.[14]   In Norway, however, a whole new power reserve market run by Statnett, was established on November 1. 2000. This is the so called RKOM that will be described further in the following.

The RKOM is run by Statnett and comes in addition to the other existing markets. This market bears some resemblances with the category of reserve capacity markets discussed under "alternative c)" in the previous section. This means that the system operator decides how much reserve power is needed during a given period and then solicits options for reserve capacity from producers and consumers to fulfill this need. Hence, along these lines the RKOM invites participation of both producers and consumers (e.g. large consumers in the paper and smelting industry) to place bids on the RKOM. The bids are based on size (minimum 25 MW) and an asking price (option price). Bids are now weekly but may be presented for 8 weeks at a time. The bids for next week may be adjusted until gate closes on Thursday at noon. On Thursday between noon and 2 pm Statnett is systematizing and analyzing the amount to be purchased.

---

[14] This comes in addition to the contracts of the specific "disturbance" reserves that involve compensation for keeping instantaneous reserves at Svenska Kraftnät's disposal.

The price on the RKOM is determined in an auction where the asking price of the last accepted offer is paid for all accepted offers. The offers are, however, distributed, among three geographical areas as determined by the grid system and normal bottleneck situations. Inside each area it is, in principle, the asking price that determines whether an offer is accepted or not. However, Statnett has a certain policy of achieving balance between the producer and consumer side that may imply a departure from the strict price ladder of acceptance. Otherwise, it should be noted that no account is taken of the real cost of power generation of the accepted bids, only the asking price matters.

The basic idea is that suppliers accepted on the RKOM market must guarantee to bid the accepted volumes of power reserves into the regulating power market, RKM. Accepted capacities must be available at 15 minutes warning with full activation for at least 1 hour, all days of the week from 06.00 am to 10 pm. There exist sanction rules for not complying. In this way Statnett[15] is guaranteed to have a sufficient power reserve to draw from as the real time of operation is approaching. An accepted supplier called upon to deliver regulating power on RKM, thus gets two kinds of remuneration; the option price for capacity and the regulating price for power delivered. The payment for regulating power generated is thus determined by the rules on the RKM and is the same as regulating power generated by suppliers that are not committed on the RKOM. Hence, there is no direct coupling between the RKOM and the RKM.

Thus far, experiences from running the RKOM seem to be positive (see Nilssen and Walther, 2001). In particular, - as is apparent from Fig. 3, the system has to a large extent succeeded in involving the demand side. Also, it seems that the RKOM has managed to bring forth a sufficient amount of power to the RKM thus relieving the pressure that would otherwise have been on the RKM. It is for instance interesting to note that "down regulation" to a large extent also takes place in the "option period" from mid November to late April (almost 30 percent of the time during the season 2004-2005). Price and quantities on the RKOM are reported in Fig. 4. As for the

---

[15] Presently, Statnett considers that about 2000 MW of fast operating reserve is needed in the RKM for Norwegian purposes. Statnett has also some bilateral long term contracts with durations of 5-10 years.

recent development of this market RKOM has now been made open to other Nordic countries with a common bid ladder. Up until now there has been some participation from Western Denmark price area.

Fig. 3. Accepted quantity bids (MW) in the Norwegian market for capacity options (RKOM) from week 47, 2004 to week 16, 2005 in the NOB area (Source: Statnett)
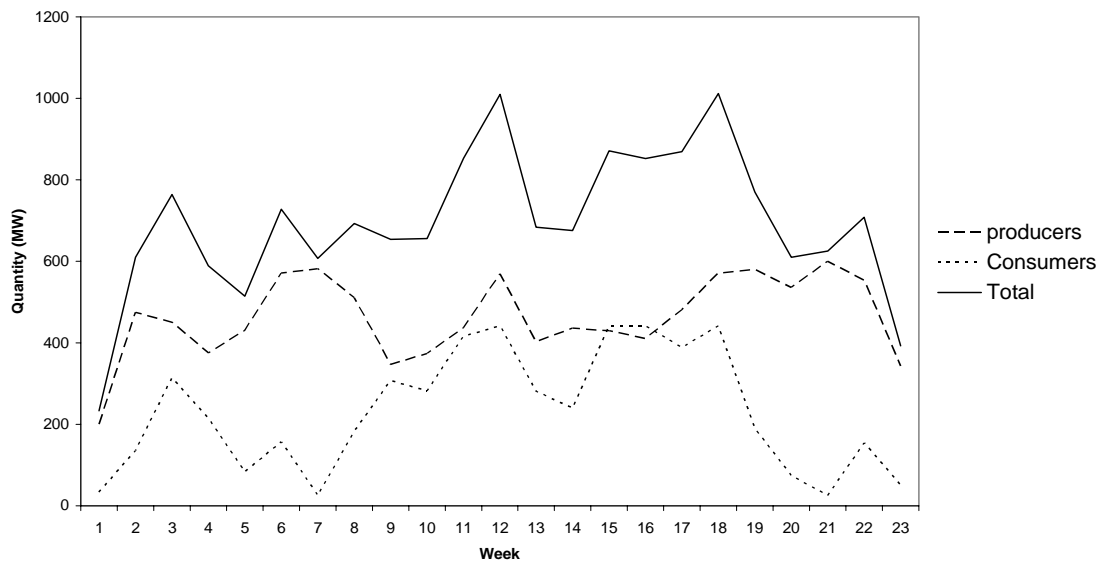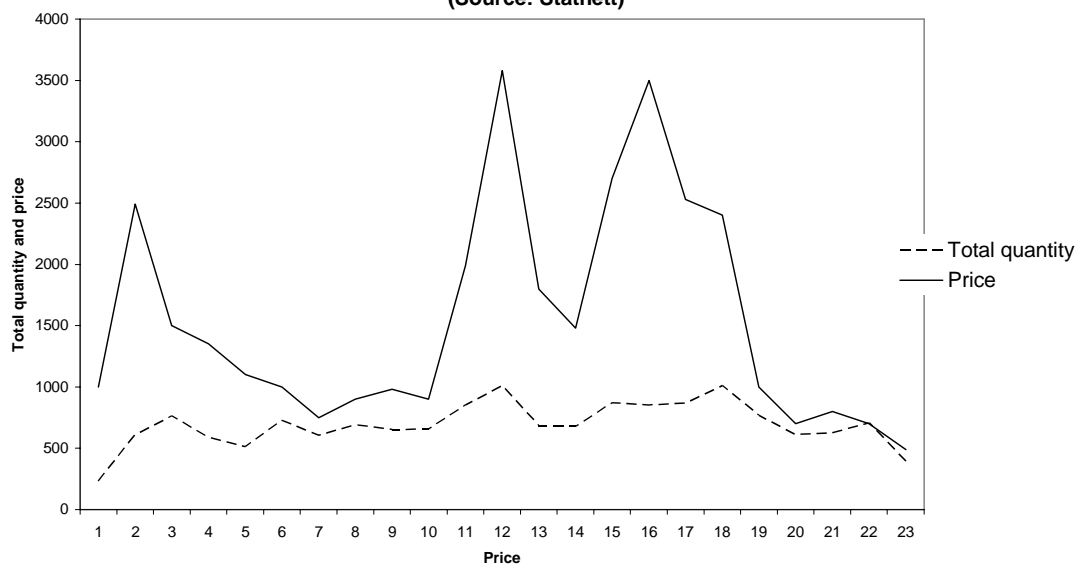

Fig. 4. Accepted quantity bids (MW) and price (NOK/MW) in the Norwegian market for capacity options (RKOM) from week 47, 2004 to week 16, 2005 in the NOB area (Source: Statnett)

10. Conclusion

The Nordic electricity market has worked well from a system reliability point of view; the lights have stayed on even during periods with unusual demand peaks. A key factor behind this favourable outcome is that the incentives for keeping sufficient reserve capacity have been strong enough. From an institutional point of view the reason is that the set of markets that is commonly called "the electricity market" includes both regulation and capacity markets, and that these markets function sufficiently well.

Yet the institutional design of the Nordic electricity market can be improved. In particular the remaining inter-country differences in terms rules and regulations can be reduced or even eliminated. Moreover the demand for regulation power could be reduced, possibly significantly reduced, if the deviations between the price of regulation power and end user electricity prices could be reduced. In practice this means that the standard electricity contracts have to be redesigned, so that the marginal end user prices better than now reflect the relevant marginal cost of electricity. Key to success is also a sizable demand side participation in the emerging capacity markets.

Literature

Abbott, M.(2001) "Is the security of supply a public good?", *The Electricity Journal*, August/September 2001, pp. 31-33.

Amundsen, E.S. and L. Bergman (2005) "Why has the Nordic electricity market worked so well?", Working Papers in Economics, Department of Economics, University of Bergen, No. 18/05

Crampton, P. and S. Stoft (2005) "A Capacity Market that Makes Sense", *The Electricity Journal*, 18, 43-54.

Doorman, G. (2000) "Securing Peak-Load Capacity in Restructured Power Markets", Ph.D. Thesis, NTNU, Norway.

Hobbs, B.F. and J. Iñón (2001) "Issues concerning ICAP and alternative approaches for power capacity markets" paper presented at the ELFORSK:s et al. conference "Methods to Secure Peak Load Capacity on Deregulated Electricity Markets" i Saltsjöbaden, 7-8 june 2001.

Hobbs, B.F., J. Iñón and S. Stoft (2001) "Installed capacity requirements and price caps: oil on water, or fuel on fire? *The Electricity Journal*, 14, 23-34.

Nilsson, G. and B. Walther, (2001)"Market-based Power Reserves Acquirement", presented at the ELFORSK conference "Methods to Secure Peak Load Capacity on Deregulated Electricity Markets" in Saltsjöbaden, June 7-8, 2001.

Vázques, C., M. Rivier and I.J. Pérez-Arriaga (2001) "A Market Approach to Long-term Security of Supply" paper presented at the ELFORSK:s et al. conference "Methods to Secure Peak Load Capacity on Deregulated Electricity Markets" i Saltsjöbaden, 7-8 june 2001.

Von der Fehr, N.H.M., E.S. Amundsen and L. Bergman (2005) "The Nordic Market: Signs of Stress?" *The Energy Journal*, Special Issue, 71-98.

Nordel, "Årsstatistikk 1991-2004" (Annual reports 1991-2004). http://www.nordel.org

Department of Economics
University of Bergen
Fosswinckels gate 6
N-5007 Bergen, Norway
Phone: +47 55 58 92 00
Telefax: +47 55 58 92 10
http://www.svf.uib.no/econ