Abstract:
The experience of liberalized electricity markets’ ability to allocate scarce energy resources has been mixed. In this paper, we analyze how liberalized markets allocate power in the short and long run through the interaction between the spot and end-user markets. We show that totally inelastic demand in the spot market does not necessarily result in market failure in a shortage situation, as long as price incentives are transferred to the end-user markets. We argue that the market does not have to run optimally to handle a shortage situation, and that problems with short- or long-run allocation of power arise when price restrictions in end-user markets results in a higher demand than that which may sustain the energy situation over time.

Keywords: Liberalized electricity market, allocation of power, demand response

JEL classification: D4, Q4

Address: Torgeir Ericson. E-mail: t.b.ericson@cicero.uio.no
Bente Halvorsen, Statistics Norway, Research Department. E-mail: btl@ssb.no
Discussion Papers comprise research papers intended for international journals or books. A preprint of a Discussion Paper may be longer and more elaborate than a standard journal article, as it may include intermediate calculations and background material etc.

Abstracts with downloadable Discussion Papers in PDF are available on the Internet:
http://www.ssb.no
http://ideas.repec.org/s/ssb/dispap.html

For printed Discussion Papers contact:

Statistics Norway
Sales- and subscription service
NO-2225 Kongsvinger

Telephone: +47 62 88 55 00
Telefax: +47 62 88 55 95
E-mail: Salg-abonnement@ssb.no
1. Introduction
The experience with electricity market deregulation has been mixed. During the winter of 2002/2003, the Norwegian market experienced significant shortages in energy supply as a result of a historically dry autumn. This resulted in a situation where the water reservoirs reached their lowest ever levels before the winter season. In such a predominantly hydro-based power production system, this led to a significant reduction in the supply of energy during the cold winter months. As a result, prices on the Nordic power exchange rose more than fourfold from mid-October until the beginning of January. The price increases reduced consumption sufficiently to secure the supply of energy during the rest of the winter, and the market never failed. In contrast, the Californian market collapsed during the winter of 2000. The Californian market experienced two major supply shocks, one as a result of increased prices of natural gas and the other as a result of a less than normal inflow of water into the reservoirs and reduced river-flow levels. In addition, the Californian market experienced increased demand from both an unusually hot summer and an unusually cold winter season. This resulted in very high wholesale prices and several blackouts during the winter of 2000.1

In both cases, the markets experienced a strained energy situation. The important question is why the Norwegian market was able to handle this whereas the Californian was not. After the Californian crisis, many articles discussed the reasons for the breakdown of the market and possible solutions. As explanations for the failure of the Californian market, some studies pointed to “the problem with market power abuse by a few dominant sellers, poor market design that invites strategic bidding by suppliers, capacity shortage caused by demand growth not matched by new capacity, and thin trading of forward and futures contracts that are critical for price discovery and risk management” (Woo, Lloyd and Tishler, 2003). Most of these arguments point to flaws in the design of the wholesale markets, which resulted in wrong incentives in demand and supply. They are arguments for why the wholesale market did not run at its optimum. However, as we will discuss later, they are not sufficient arguments to explain the failure of the market. Another explanation for the breakdown of the market is the lack of demand response to price spikes in the spot market (Woo, Lloyd and Tishler, 2003, Taylor and Van Doren, 2001, Faruqui and George, 2002, Fraser, 2001). Taylor and Van Doren (2001) argue that if a sufficient percentage of total end-user demand faced hourly real-time prices, this would prevent a similar crisis. According to Hunt (2002), California had to employ rolling blackouts with a shortage of only 300 MW in a system of 50,000 MW, which means that a very small reduction in

1 See, e.g., Borenstein (2002), Woo, Lloyd and Tishler (2002) for a more detailed description of the two events.
demand was needed to avoid the blackouts. However, the Norwegian market was able to handle the extreme energy shortage situation during the winter of 2002/2003 even if only a tiny fraction of consumers faced real-time pricing. An argument that offers an explanation for why the Californian market failed while the Norwegian market was successful is the end-user price cap in the Californian market (see, e.g., Taylor and Van Doren, 2001 for a discussion), which stripped consumers of incentives to reduce consumption during the winter months.

To understand why things went wrong in California, we need to understand how the electricity market allocates power resources in both the short and the long run. We will argue that even if a liberalized wholesale market is able to allocate power (that is, energy within an hour), it may have problems allocating power over a period (hereafter referred to as allocation of energy) in a way that prevents market failure; that is, prevents demand from exceeding supply (hereafter referred to as sustainable). This is because the mechanisms for the allocation of power and energy in the wholesale markets are not the same. The allocation of power happens on an hourly basis based on demand and supply in the spot market, whereas the allocation of energy depends on the interaction between the wholesale and end-user markets. Thus, it is not obvious that a liberalized wholesale market will be able to allocate energy if the end-user market is regulated.

The novelty of this paper is that it analyzes explicitly the allocation of power and energy given the connection between the spot and end-user markets. Furthermore, we will show that even if the wholesale market does not function optimally, it will not necessarily collapse in the case of an energy shortage situation, as happened in California. This implies that the market may be able to allocate power and energy in a sustainable way even if, for example, demand is completely price inelastic in the spot market. Based on our analysis, we argue that Taylor and Van Doren’s (2001) conclusion that the introduction of real-time pricing mechanisms for the largest segment of demand is essential to avoid a breakdown of the electricity market is not necessarily true as long as the price cap is removed. Finally, we show that the link between the wholesale and end-user markets is important, not only to secure energy resources but also for good allocation of power in the spot market. Separating the two markets by mechanisms such as a price cap may create problems with the allocation of both power and energy in a shortage situation.
2. The organization of the Norwegian electricity market

In many countries, production and consumption of electricity are organized in different markets: wholesales markets and end-user markets. This is true in both the Californian and Norwegian markets. The organization of the different markets differs in line with variations in their respective liberalization and regulations. Because of these differences, their markets will handle a strained supply of power and energy in different ways.

2.1 The wholesale market

The Norwegian electricity market was deregulated in 1991. Sweden followed in 1996, and a common Norwegian and Swedish Exchange (Nord Pool) was established as the first multinational exchange in the world for trade in power contracts. Finland joined in 1998, western Denmark in 1999, and eastern Denmark in 2000. The Nordic countries are now connected in a common integrated electricity market.

Any producer within the Nordic area can deliver electricity to the common Nordic electricity market. The wholesale market includes power producers, retailers, industry, and other large undertakings. In the wholesale market, the trade of electricity takes place at the Nord Pool exchange and bilaterally between different market players. At the Nord Pool Elspot (Nord Pool’s spot exchange), the next day’s hourly spot prices are based on bids from participators for purchase and sale (a day-ahead market). Each participant submits bids to Nord Pool Elspot on bidding forms, and the bids are aggregated to a demand and supply curve for each of the next day’s 24 hours. The intersection of the demand and supply curves provides the Elspot system price at which trade in the market occurs. The price also determines the obligations for each participant to deliver or take power from the central grid. If there are bottlenecks, the market is divided into pricing areas, and the prices in the surplus areas are lowered while the prices in the deficit area increased until demand and supply balance within each area.2

2.2 The end-user market

The end-user market includes all buyers of electricity for own consumption, such as industry, commercial buildings, households, etc. Most consumers in the end-user market face a price that is fixed during a longer period, from two weeks to several years. During the winter of 2002/2003, most households were on a standard variable price contract, where the price was constant for two weeks or longer. An increasingly popular type of contract is the spot contract, where the price in each month follows the average of the month’s hourly spot prices. In that sense, it is fixed within the month.

2 For more information about the Nordic spot market and regulations, see, e.g., Ericson (2007).
Household consumers are required to report their consumption four to six times a year (but may report more often if they want) and are charged according to their accumulated consumption between the meter reading dates. The price these customers pay is a weighted average over the so-called adjusted load profile from all nonhourly metered customers in the area for the relevant period. Since a single customer has no significant impact on this load profile, it will not receive the entire benefit if it reduces consumption more than other customers do during a high price period. This means substantial dilution of the signal from hourly spot prices. The result is that, for the total bill, it does not matter at what time the consumer uses electricity. The incentive is thus only to save energy for the whole period, independent of the day/week/month when this saving is carried out.

Most industrial contracts are fixed price contracts where prices are fixed for long periods, some for five years or more. Even though some industries, in particular power intensive industries, have the opportunity to sell power back to the market, this option is not often used because of, for instance, delivery obligations and additional costs of shutting down and starting up production. Thus, a high proportion of consumers have few or no incentives to adapt their consumption to short-term fluctuations in the spot price. Some consumers are measured hourly and have end-user tariffs that change on an hourly basis with the system price on the spot market. Additionally, some large customers (mainly in the private and public service sectors) are present in the spot market. These consumers, however, only account for a few percent of the total consumption of energy.

2.3 The connection between the spot and end-user market

The connection between the hourly spot market and the more long-term end-user electricity market is illustrated in Figure 1. The right figure illustrates the end-user market at a particular point in time and the left figure illustrates the spot market. The supply curve in the spot market \( S \) may be interpreted as the accumulated supply of producers with different marginal and alternative costs. The supply curve is assumed to increase with production. Furthermore, we assume that when production approaches the capacity limit, the supply curve becomes vertical; that is, it is not possible to increase production in the short run outside this limit. In the diagram, reductions in power production capacity will shift the supply curve to the left, while increased capacity will shift the supply curve to the right.

---

3 See also Fraser (2001) for a discussion.
Figure 1: The connection between the spot and the end-user market.

The figure illustrates a situation where the end-user price ($\tilde{p}$) does not change on an hourly basis with the price in the spot market ($p$), but is given in the short run. The demand in the end-user market ($\tilde{D}$) is assumed to be price responsive. However, since the end-user price is fixed in the short run, consumers have no incentives to change their consumption because of short-term changes in the spot price. They choose their consumption quantity ($Q$) given the price $\tilde{p}$. Retailers are obliged to supply all demand at any point in time, and need to buy this quantity in the wholesale market. Retailers thus bid into the spot market for this consumption on behalf of their customers and are forced to pay any price to serve their customers. Thus, in the short run, since demand in the spot market will not change with the spot price, it will appear totally inelastic, as illustrated in the left-hand side of the figure. The price in the spot market will thus fluctuate in the short run as a result of, e.g., the time of day or changes in weather conditions, even if the end-user price does not.

This situation indicates a disconnection between the spot and end-user market, as information about short-term changes in market conditions are not received by consumers and information about consumers’ actual demand response and their willingness to pay for electricity is not reflected in the demand curve in the spot market since end-user prices do not change on an hourly basis.

The situation described in Figure 1 illustrates an extreme case where all end-user prices are fixed in the short run. However, the actual demand curves at the Nord Pool exchange are not as inelastic as they appear in Figure 1. Figure 2 shows an example of the purchase and sales curves at Nord Pool Spot on February 6, 2003, hour 17:00–18:00. In this hour, an extreme shortage in power supply relative to demand occurred. In this hour, the figure illustrates that the supply curve as well as the
demand curve were quite price sensitive in the lower price range and less price sensitive in the upper price range.

**Figure 2: Elspot purchase/sales curves. Hour 17:00–18:00, February 6, 2003.**

![Graph of Elspot purchase/sales curves.](image)

Source: Nord Pool Spot AS.

The demand in the Norwegian spot market is comprised of retailers with and without their own production. An important source of price response in spot market demand and supply is retailers with their own production that shift between being buyers and sellers in this market depending on prices.\(^4\) If the spot price is lower than the production costs at the margin, the retailer will reduce production and buy the power required to meet current demand from end users in the spot market. When the spot price is higher than production costs at the margin, the retailer will increase production and sell the excess power on the spot market. This is one reason why both the supply and demand curves are flexible in the same price region around the marginal costs of retailers with production capacity (see Figure 2). Above this area, the demand curve is much steeper as more and more retailers become sellers, leaving the demand curve to cover the end-user demand. We see from the figure that, even if this was an hour with one of the most strained power supply situations ever recorded, supply and demand curves still intersect since both curves do respond to price changes even at their steep end.

---

\(^4\) In the theoretical discussions in Figure 1, we assumed that no retailers have their own production.
3. The spot market’s allocation of power and energy

Given the disconnection between the spot and end-user markets, an important question is how well the spot market will handle variations in demand and supply. To discuss this, we need to look at how the spot market allocates both power and energy.

3.1 The allocation of power

Producers bid into the spot market according to their costs of producing power in this period. In a system dominated by hydropower, as in the Norwegian system, the alternative value of production in this period compared with postponing production to a later period depends on the value of the water in the reservoirs. If the reservoirs are full, the water has a lower alternative value. That is, the value of postponing production to future periods is lower than production in this period. If the water level in the reservoirs is low, the opposite may be true. In the first situation, producers will give a low bid into the spot market to avoid the risk of being left with too much water in the reservoirs at the end of the season. The alternative value of the water will increase as the level in the water reservoirs declines, and the supplier will demand a higher price for producing power in this period rather than postponing production to future periods.

How well the spot market handles a strained situation depends on the level of supply, the consumption level, and the short-term price response in demand and supply in that spot market at a particular point in time. Figure 3 describes the supply and demand curves in the spot market in two different cases. The left-hand side shows the spot market in the case where demand is totally insensitive to price changes, whereas the right-hand side shows a case with various degrees of flexibility in demand. From the left-hand side figure, where demand is totally inelastic, we see that the market will clear as long as the demand curve (D) intersects the supply curve at the part with increasing marginal costs. If the demand shifts to D′ and exceeds the production capacity limit, the market will no longer clear and there will emerge a need for power rationing. Thus, if the demand curve is totally inelastic, we may experience a lack of market clearance and the market may fail in allocating power across the system. However, as we see from the right-hand side of the figure, as long as there is some demand response, the market will always clear and rationing will not be necessary. We also see that the more price sensitive demand is, the lower the equilibrium price, in particular at the steep end of the supply curve where the production capacity limit is reached. A more price-sensitive demand will also give less

See Førsund (2007) for a more detailed description of how inflow shortages will affect the alternative value of the water.
volatile prices when demand shifts, e.g., according to seasonal changes in the need for heating/cooling or intradaily demand shifts from peak to off-peak hours.

**Figure 3: The spot market in the case of no price response (left side) and with different degrees of price response in demand (right side).**

![Graph of the spot market](image)

### 3.2 The allocation of energy

The accumulated power over a period gives the market’s allocation of energy. From the discussion above, we know that the market will clear as long as demand and/or supply in the spot market responds to price changes at the margin. However, how the spot market allocates power over a period, that is, how it allocates its available energy resources, is a different question.

**Figure 4: Allocation of energy in the electricity market**

For an efficient allocation of energy, the connection between the end-user and the spot market becomes important. Figure 4 illustrates how the spot market allocates energy in the case of an energy shortage situation. The left-hand side of the figure shows the spot market and the right-hand side
shows the end-user market. To illustrate our point, we show the extreme case of no short-term demand response in spot market demand. We assume that there is sufficiently installed capacity in the system to meet demand before the shortage situation. We compare two situations: one with sufficient inflow of water in the reservoirs, and the other as an energy shortage situation with low water levels.

Assume that the supply curve ($S$) describes supply under normal reservoir inflow levels. Assume now that the autumn rain fails to appear, resulting in water in the reservoirs at very low levels at the beginning of the cold winter season. This was the case in the Norwegian system in the autumn of 2002. In this case, the alternative value of saving water to postpone production will increase, as producers expect prices to rise when demand increases and the level of water in the reservoirs decreases (Førsund, 2007). If producers are rational, they will demand a higher price to produce power in this period rather than postponing production to a later period with a higher price. This is shown in the figure by a positive shift upwards in the supply curve, from $S$ to $S'$. When production approaches the capacity limit, the supply curve is assumed to be vertical. This means that a power- and an energy-scarcity situation will have different impacts on the supply curve, as a shortage of power at a particular point in time will shift the supply curve in the spot market to the left, while an energy shortage situation shifts the supply curve upwards. In Figure 3.2, we see that the shift in supply upwards results in increasing prices in the spot market (from $p$ to $p'$). In the short run, this increase in the spot price will not reduce consumption, as the fixed end-user prices are not affected. However, if retailers expect the situation with high spot prices to last, they will increase end-user prices as soon as possible. This is illustrated in the right-hand side of the figure, with a shift in price from $\bar{p}$ to $\bar{p'}$.

When the price increases, consumers will respond by reducing their consumption from $Q$ to $Q'$. Thus, in the next period, retailers will take account of this reduction when bidding into the spot market, and the demand curve in the spot market will shift toward the left, as shown in the left-hand side of the figure. This will go on until the energy situation is no longer stressed. Thus, even if the wholesale and end-user markets are not fully integrated and/or the allocation is not optimal (e.g., as a result of market imperfections), the development in the consumption of power and energy may still be sustainable; that is, strained resources can be allocated without market failure.

The disconnection of the two markets will make the wholesale price differ from the end-user price in the short run, but as time goes by, we can expect the average price levels in the two markets to approach each other. The lag before changes in the spot price affecting the end-user price is illustrated

---

In Figure 4, we assume that the inflow shortages do not affect the production capacity limit in the short run, only the alternative value of the water in the basins.
by Figure 5, which shows the hourly spot prices in the Oslo pricing area during the winter of 2002/2003, along with the prices offered through a standard variable contract from the dominant supplier in the Oslo region.

**Figure 5: Hourly spot prices in the Oslo region and the price offered from a supplier through a standard variable price contract in the winter 2002/2003.**

As seen in the figure, the spot price rose to very high levels during December 2002 and the beginning of January 2003. The standard price facing customers, however, was lower in parts of this period, sometimes only about half of the system price in the spot market. Furthermore, from mid-January until May, the customer price was well above the market price, sometimes more than twice as high. We can see that the standard price did not bring the energy scarcity price signal to customers at the point in time when the market considered the situation to be constrained. Moreover, the standard price signal did not inform customers when the market considered this situation to be over. Also important is the price spike on February 6, 2003, when the peak price signaled a power shortage situation (see Figure 2 for the relevant demand and supply curves at the Elspot at this hour).

This figure illustrates that there is a lag before the price signals from the spot market enter the end-user market. However, after a period, the end-user prices follow the spot prices. The higher the proportion of the consumption associated with fixed contracts and the longer the contracts are fixed, the more disconnected the two markets are, and the longer the period before an increase in the spot price affects end-user prices and demand.

Source: Nord Pool Spot AS and Hafslund.
4. Why did the Californian market fail?

Can this analysis help us understand why the Californian market failed in allocating power and energy during the winter of 2000? Figure 6 illustrates what happened in the Californian case by showing the effect on the allocation of energy and power in the spot market when introducing a price cap on the retail price. We assume that the price cap is set at $\hat{p}_0$; that is, the end-user price cannot exceed this price no matter how high the system price in the spot market.

Figure 6: Allocation of energy in the electricity markets in the situation with a price cap on the retail price.

In the Californian situation, marginal production costs increased as a result of increased prices of natural gas, and the alternative value of water increased as a result of low precipitation. This led to an upward shift in the supply curve to $S'$, which increased the spot price from $p$ to $p'$. If there is no price cap in the retail market, this would lead to an increase in the retail price from $\tilde{p}$ to $\tilde{p}'$, which shifts consumption to $Q'$, where consumption is sustainable. However, since this new retail price exceeds the price cap, the retail price is not allowed to go this high. Thus, the demand curve $\tilde{D}D^0$ appears to be totally inelastic above the consumption level $Q^0$. This means that, for this demand curve in the retail market, the demand curve in the spot market cannot be lower than $Q^0$ (while the level without price cap would have been $Q'$). We see that in the short run, the spot market still clears since there is enough power in the system to supply demand in the short term. Suppose now that end-user demand shifts to the right, from $\tilde{D}D^0$ to $\tilde{D}'\tilde{D}'^0$, as a result of an increased need for electricity during a hotter than normal summer and/or colder than normal winter. This will increase the consumption of electricity from $Q^0$ to $Q''$. If this demand shift is sufficiently large, there may not be enough power in the system to meet demand at a particular point in time, and there will be rationing of power or blackouts.
A price cap in the end-user market makes the market vulnerable to demand and supply shifts since the price cap prevents consumers from obtaining the price signals that makes consumption sustainable with the available supply of power and energy in the system. If the price mechanism is allowed to work freely, signals of a shortage will reach consumers and the situation may be avoided. The problem with the price cap is that it separates completely the end-user and the spot market in a strained situation, with the effect that the spot market can no longer allocate power and energy in a sustainable way, as demand ($Q'_0$) exceeds sustainable demand ($Q'$). The price cap also makes the demand in the spot market virtually inelastic at the margin, which results in market failure in a power shortage situation and a need for rationing of power (see the discussion in Figure 3).

5. Conclusions

In this paper, we have shown that even if demand is totally inelastic in the spot market in the short run, consumers will change their consumption when more permanent changes in spot prices are carried through to end-user prices. Thus, a liberalized energy market will normally be able to allocate energy and power. Essential to this argument is that the price signals in the spot market are transferred to the end users by increased end-user prices. That is, there is no interference in the price mechanism. If the price mechanism is not allowed to work freely, e.g., if all end-user prices are fixed or there is a price cap attached to the end-user price, consumers have no incentive to reduce consumption in a scarcity situation when spot prices are rising, and the market will no longer be able to allocate either energy or power in a sustainable way. This means that the larger the share of end users in variable price contracts, and the shorter the period prices are fixed, the better the spot market will be at allocating power both in the short and long run. In this paper, we have focused on whether the spot market is able to allocate energy in a sustainable way and not on whether the allocation is optimal. An increased proportion of consumers with time variable price contracts will also increase the efficiency of the allocation of power and energy in the spot market.

For politicians, the problem with the free market solution is that it may have consequences in conflict with other political goals, in particular with perceptions of a fair distribution of resources in society, as a strained power and/or energy situation may result in long periods with very high prices on the necessity good electricity. The political pressure to reduce negative welfare effects by introducing a price cap may be hard to resist. During the winter of 2002/2003, the Norwegian ministry of energy was under strong pressure to introduce a price cap on the end-user price to reduce the welfare loss in the most vulnerable households, and the minister was confronted on television with tales of elderly people who had frozen to death in their apartments. The minister did not cave in to this pressure, and
insisted on solving these problems with income-policy instruments. This was probably an important reason why the Norwegian market survived that winter.

What this exercise tells us is that it is not sufficient to liberalize the wholesale electricity market to enable it to allocate power and energy in a more efficient way. This is because the allocation is done both in the wholesale and end-user markets. To allocate the resources in the system in an efficient way, both the end-user and the wholesale markets must be liberalized. The more integrated the two markets, the more robust the system will be during shocks in demand or supply. That is, the larger the share of energy consumption on short-term price contracts that follow the spot price, the more flexible will be demand and the less volatile will be prices. This will reduce the strain on the system resulting from supply and demand shocks.

However, the Norwegian experience also tells us that markets do not need to allocate resources optimally to avoid a collapse in the system. Even if most end users do not have price contracts that change with the spot price by the hour, the system may very well be able to allocate its resources in a sustainable way without the risk of market collapse. It only takes a little longer, and the prices will be more volatile than if both markets were running at their optimum. As long as the two markets are not totally separated, e.g., by a price cap, the market will in most cases be able to handle a strained situation. It has been argued that the reason for the Californian collapse was the lack of short-term price response in demand (Taylor and Van Doren, 2001). The Norwegian experience tells us that this is not a sufficient explanation for the Californian collapse. It is true that more flexible demand in the short term, e.g., by a substantial share of respondents on real-time prices, may have prevented the crisis by reducing the price spikes, thus avoiding making the price cap in the retail market effective. However, given the existing demand structure, the Californian system probably would not have failed if it had disposed of the price cap.

As we see it, the most important lesson from the above accounts of the Nordic and Californian markets is that it is eminently wise not to interfere with the pricing mechanism to avoid negative welfare effects of increasing prices if the goal is to secure sustainable availability of electricity in a liberalized electricity market.
References


