A COMPARATIVE STUDY OF THE ELECTRICITY MARKETS IN UK, SPAIN AND NORD POOL

A Report for Confindustria

Prepared by NERA

27 June 2002
Rome

Graham Shuttleworth
Isabelle McKenzie

NERA
Economic Consulting

Via XXIV Maggio 43
Roma 00187, Italia
Tel: (+39) 06 488 810
Fax: (+39) 06 485 838
Web: www.nera.com

An MMC Company

TABLE OF CONTENTS

1. INTRODUCTION

This report examines the experience of electricity markets in England and Wales, Norway and Spain, drawing occasionally on examples from other countries to illustrate key points. Section 2 provides a brief description of each of the markets considered in this study. However, to summarise recent experience, I have described how each of the markets handles the features of electricity market design that I consider to be most important. These features are:

- The relationship between bilateral contracts and centralised "pools" (or "gross" and "net" pools, as I explain in section 3);
- Rules for setting prices in centralised systems, including imbalance charges (section 4);
- The method of rewarding investments in generation capacity (section 5); and
- The procedures for managing the effect of transmission constraints (section 6).

The importance of these topics is indicated by the lack of consensus about the right way to proceed. Different markets take different approaches, to suit local conditions. However, it is possible to discern some common principles and general lessons.

1.1. Common Principles and Lessons

1.1.1. Contracts and Pools

The discussion of contracts and pools highlights a number of false assumptions that have crept into recent discussions of electricity trading arrangements. It is often suggested that it would be desirable to "abolish compulsory pools", or that the new system in England and Wales "replaces compulsory pools with bilateral contracts". However, all electricity markets require some central system for balancing and a compulsory scheme for putting a price on accidental "imbalances". The only design questions concern the time at which these central elements take over from bilateral trading, and whether the settlement of imbalances requires traders to supply information about their contracts, or not.

1.1.2. Pricing Rules

The discussion of pricing rules leads quickly into a discussion of market power. Evidence about the effect of pricing rules on market power is limited, but provides little to suggest that the exact rule used to put a price on energy has a significant impact on the average *level* of prices.

1.1.3. Payments for Capacity

Different markets adopt different methods of rewarding investment in capacity. Some markets rely entirely on market prices for energy, whereas other markets set up explicit capacity payments. Analysis of the costs and risks of relying on energy prices alone shows why many thermal (hydro-carbon-fired) generation systems use capacity payments rather than market prices. However, there is unlikely to be a universal "first best" solution to this problem and the solution in any market reflects local needs and conditions.

1.1.4. Congestion Management

The same general observation applies to the methods used to deal with transmission constraints: no single solution dominates any other. Although the range of possible solutions is limited, the key decision is whether to split markets at transmission constraints or not. When transmission constraints separate markets, investors see local price signals, which may make investment more efficient. However, some markets span transmission constraints with the aim of increasing liquidity and insulating market prices from the behaviour of individual generators operating within a constrained zone. Again, there is no universal solution, only a range of criteria for designing local solutions.

1.2. Conclusion

In recent years, many countries have set up electricity markets and many providers offer standard systems for operating them. Nevertheless, discussion of electricity markets remains a hot topic. In some instances, the topics for discussion seem to derive from a misunderstanding; some apparently important decisions (gross/net pools and centralised pricing rules) have little impact in practice. On the other hand, a number of important decisions can only be made after taking into account a wide range of local conditions (capacity payments and congestion management).

I hope that this study will help the Italian electricity industry and its consumers to find answers to these questions that meet their requirements.

2. MARKET STRUCTURE

The following sections describe briefly the structure of the industry and of the electricity markets in each of the three study countries. The information provided here is not exhaustive, but is relevant to the discussion of market design issues that follows later.

2.1. England and Wales

The electricity industry of England and Wales was restructured in 1990. A state monopoly responsible for generation and transmission was divided into a transmission company (National Grid Company) and three generating companies (National Power, Powergen and Nuclear Electric). All of these companies were eventually privatised, although the state retained some interest in them until 1996. The 12 distribution companies were privatised in 1991, but the state retained a "golden share" until 1996. Once the golden share lapsed, most of the distribution companies were taken over by other electricity companies from the US, Europe and Scotland.

The three generating companies possessed substantial market power, even though the 1990s saw a rapid expansion of investment by independent companies in gas-fired baseload generation. In 1996, the regulator imposed a price cap and demanded that National Power and Powergen sell of 6 GW of capacity. In 1999 (when National Power and Powergen tried to merge with distribution companies or with their retail arms), the government imposed further asset sales as a condition of approving the merger. As a result, by late 2000, the generation sector was divided among five or six large companies and a large number of smaller ones. From that time onwards, electricity prices fell substantially.

2.1.1. The Electricity Pool

From 1990 to 2001, generators in England and Wales sold all of their physical output to the Electricity Pool and retailers bought all of their physical needs from the Electricity Pool.[1] As explained in section 3, we define this arrangement as a "gross pool". The Electricity Pool was a multilateral contract between generators, retailers and other relevant agents.

By 10:00 on each day, all participating generators offered an amount of electricity for each half-hour settlement period and a set of minimum "offer prices" at which they were willing to generate on the *next* day. The system operator (acting on behalf of Pool Members) compiled all this information into a "day-ahead" production schedule, which matched generator offers against forecast demand. From this production schedule, the Pool derived a day-ahead "Pool Purchase Price" for scheduled output. Separate settlement arrangements handled variations in output between 10:00 and the time of delivery. Retailers paid the

[1] There were only a few exceptions to this rule, for generators connected to low voltage networks or within consumer premises.

"Pool Purchase Price" plus a surcharge to cover the costs of short-term variations in output and other costs of system operation.

Market participants hedged the financial risks associated with the fluctuating pool price by signing financial contracts ("contracts for differences") with agreed prices. These contracts used a cash settlement mechanism to replace the cost of selling power through the Pool (at pool price) with the cost of a direct delivery (at a contract price). They varied in length from short-term products (a day or a month) to long-term contracts (some over 20 years) and allowed market participants to manage risk in many innovative ways.

2.1.2. NETA

"NETA" stands for the "New Electricity Trading Arrangements", which came into force on 27 March 2001. NETA is incorporated in a "Balancing and Settlement Code", to which all generators, traders and retailers must now sign up. Under NETA, the centralised market in England and Wales now operates as a "net pool", ie it only handles trades that fall outside the bilateral contracts agreed between generators, retailers, customers and other traders. (See section 3.)

Electricity generators contract with customers for physical volumes of energy at agreed strike prices. A short time before the time of delivery,[2] participants notify their net contract sales and purchases to the system operator, along with a forecast of production and consumption of electricity and the prices at which they are willing to increase or decrease their production and consumption. The system operator manages the system by calling on these offers and bids through the "Balancing Mechanism". The offers and bids that are accepted then determine the prices that are used to settle imbalances between metered flows and contracts. A low price applies to surplus imbalances and a high price applies to deficit imbalances.

A number of markets have evolved to help sellers find buyers for their contracts. Some markets operate through formal exchanges, including the Automated Power Exchange and the UK Power Exchange. Each exchange offers a variety of contracts (yearly, seasonal, quarterly, monthly, weekly, daily and some for even shorter periods). However, no single exchange has yet attracted the critical mass of trade needed to make markets liquid. A lot of trade is arranged instead by brokers, some of whom publish their own indices (estimates) of market prices.

2.2. Spain

In Spain, generation and distribution networks, as well as part of the 220 kV transmission grid, are mainly owned by four companies: Endesa, Iberdrola, Union Fenosa and

2 Originally, the deadline was 3.5 hours before delivery, but it will shorten to 1 hour from 2 July 2002.

Hidrocantabrico. Other generators participate mostly through a “special regime” which offers a regulated tariff for output from certain types of plant. A separate company, Red Electrica de España, owns most of the 400 kV transmission grid and part of the 220 kV grid.

Competition in the Spanish wholesale electricity market was introduced on 1 January 1998. Broadly speaking, it is organised around a centralized market (a Pool) that coexists with some bilateral physical and financial contracts agreed between producers, resellers and qualified consumers.[3]

The Market Operator (MO) is responsible for coordinating the economic aspects of the electricity market. The System Operator (SO) is responsible for the technical system management processes and for the procurement of ancillary services.

The MO has responsibility for matching offers to sell power against bids to buyer power in daily and “intra-daily” markets, from which the MO determines the final market-clearing prices (“marginal prices”) for each hour. The MO is responsible for settlement of spot trades and imbalances.

The pool is based on a sequence of markets, as summarised below.

2.2.1. The daily market (day-ahead market)

Most physical production is sold on the day-ahead market. All available production units must submit offers in this market for any power not linked to a bilateral contract, in order to be allowed to participate in the intra-day markets. The buyers in this are distributors, resellers, and qualified customers (ie those qualified to act in the market). External agents (holders of capacity on cross-border interconnectors) may also participate as buyers or sellers of electricity.

The day-ahead market matches offers to generate against bids for demand and sets prices for each of the twenty-four hourly periods for the next day. Offers and bids can be submitted to the MO at any time before 10:00 on the day before despatch. International contracts for physical delivery outside the pool must also be notified at that time, while national physical contracts can be notified one hour later.

The MO’s matching process seeks out the lowest cost way to meet demand with available offers for each separate hour, and then makes adjustments to reflect plant ramping constraints and minimum revenue constraints. The day-ahead price for each of the hourly periods of the following day is then defined as the highest price contained in the generation

3 Contracts between generators and resellers will be allowed as of January 2003.

offers accepted in order to meet the demand in that hour.[4] This price is used to settle all energy traded in the day-ahead market in that hour. There is a price cap at 18 €cents / kWh.

In addition to receipts for energy sold in the day-ahead market, generators who were available in the market receive a capacity payment, independently of whether or not the MO accepted their offers. The objective of this payment is to incentivise the building and availability of generation plant. The cost of the capacity payment is recovered from consumers as a surcharge per kWh of consumption. Consumers who have left the regulated tariff pay a lower surcharge than consumers who are supplied under the regulated tariff.

Starting with the production schedule that emerges as the sum of physical contracts and trades in the day-ahead market, the SO resolves any problems due to transmission constraints and contracts for secondary reserve services for the following day. When the SO asks a generator to increase its output because of a transmission constraint, the generator receives its offer price for the extra production. When the SO asks a generator to reduce its output because of a transmission constraint, however, its sales to the day-ahead market are cancelled and the generator receives no compensation. Reserves are paid on the basis of the marginal (highest price) offer that the SO accepts. Agents who wish to modify their program and avoid being charged for imbalances have to “undo” their commitments in the intra-day markets.

2.2.2. Intra-day markets (adjustment markets)

Each day, the intra-day spot market holds six sessions of trade covering all the hours from 2¼ hours ahead up to the end of period included in the last day-ahead market. Each intra-day session is similar to a day-ahead market session, in that agents submit generation offers and demand bids, the MO matches offers against bids, and the highest price of an accepted offer defines a market price for each hour traded in that session of the market. As in the case of the day-ahead market, all trades in intra-day markets are firm commitments.

The intra-day markets provide the agents with the possibility of adjusting their positions as delivery gets closer in time, in order to avoid incurring penalties for imbalances. Sales and purchases for each hour are therefore incremental to trades carried out in previous day-ahead or intra-day market sessions for that same hour.

2.2.3. Imbalances

In real time, the physical balance between electricity production and consumption is ensured by the SO using ancillary services. Some ancillary services are mandatory (primary regulation and voltage control) and some are voluntary (secondary and tertiary regulation,

[4] The price specified in demand bids cannot set the market price, even if generation capacity is insufficient to meet demand.

additional voltage control capabilities and the service restoration). Mandatory ancillary services must be provided when requested, but are not remunerated.

2.3. NordPool

Norway, Finland, Sweden and Denmark operate a single electricity market for wholesale electricity trading, known as the Nordpool. Nordpool replaced an earlier arrangement in Norway in 1993, and the other countries joined later; Denmark was only incorporated in 2000. Each country retains the rights to regulate and set the market rules and each country had different reasons and timetables for liberalization of their markets.

Within each country, generation and distribution are divided amongst a large number of companies – Norway alone has around 200 distribution companies. The markets appear to behave competitively, although some players are very large – in Norway, Statkraft provides about 30% of total production and Vattenfall's market share in Sweden is larger. However, ownership by the state or by municipalities is widespread and may diminish the incentive for these companies to abuse their market position.

2.3.1. NordPool markets

Nordpool operates two separate markets: the physical market and the financial market. Elspot is the physical commodity market traded for the following 24 hours only and is used to manage transmission and delivery constraints on the network. (In Sweden and Finland, Nordpool also operates markets 2 hours in advance of delivery.) Elspot provides an electricity price for the system as a whole. However, the network of the Nordic region is split by transmission constraints and Nordpool collaborates with system operators to define an "area price" for each part of the network. Sweden, Finland and Denmark each count as one area, but there may be up to 5 areas within Norway. This zonal pricing structure is known as market splitting.

The financial markets cover both futures and options. Contracts can be traded for periods a few days or up to 4 years ahead, whilst financial derivatives products, such as options and volatility products, are also available. Any company (provided they are credit worthy) can participate in the Nordpool markets regardless of whether they own physical assets in the region.

2.3.2. Regulation power markets

The transmission companies in each country are independent of trading activities and act as system operators. They are responsible for real time balancing within their network and each of them operates a real-time market for balancing and for settling imbalances. These real-time markets are known as "regulation power markets".

3. "GROSS" AND "NET" POOLS

In European environments, there has been a lot of discussion lately of the need for and role of power pools, ie, centralised electricity markets. The recent change in England and Wales, from the Pool to NETA, has sometimes been presented in documents and international conferences as if the Pool was "abolished" and replaced by "bilateral contracts", or as if a "compulsory" Pool was replaced by "contractual freedom". Commentators have suggested that "compulsory pools" are undesirable and should be "replaced" by a system of bilateral contracts.

In practice, such descriptions of the UK reform and its implications are uninformative and misleading, as they disguise the true nature of electricity pools and contracts. After all:

- Under the old Electricity Pool, every generator and retailer signed "bilateral contracts"; and
- Under NETA, adherence to the Balancing and Settlement Code is compulsory for a *wider* range of players than the obligation to sign the old Pooling and Settlement Agreement (the contract which established the Pool).

The changes that introduced NETA were therefore subtler than implied by the comments summarised above. Closer inspection reveals that NETA retains many "compulsory" aspects, but that a key aspect of the reform was the switch from a "gross pool" (which is also found in Australia and some US markets) to a "net pool" (as operated in the Nordic region and throughout much of continental Europe). Below, I explain the difference between a gross pool and net pool, and the implications for contracts.

3.1. Basic Structure of an Electricity Market

All electricity markets must balance, minute by minute, to ensure system security and stability. The system moves out of balance when demand or production levels change unexpectedly. Each electricity system must operate at or near a certain frequency (50 cycles per second in Europe). If there is excess consumption, the frequency on the system will drop (or else the system will draw in unplanned imports from neighbouring systems) whereas, if there is excess production, the frequency will rise (or else power will flow out unexpectedly to neighbouring systems). The system operator in each system is responsible for ensuring that the system balances and remains within its operational parameters.

3.1.1. Centralised balancing

In order to achieve a balance, the system operator calls for generators (or despatchable consumers) to increase or decrease production (or consumption) of electricity, until the system is brought back into balance.

The need for minute-by-minute balancing means that, *at some point*, the system operator is always given sole responsibility for controlling individual levels of consumption and production of electricity. The point in time when the system operator takes over, and bilateral trading is no longer possible, is known in the UK as "gate closure". At that time, the "gate" into the bilateral contract market closes and all further adjustments take place through trades involving the system operator (or a central exchange) as one party to the contract.

In some markets, gate closure occurs a day in advance of delivery, in order to allow the system operator time to adjust the "generation schedule", ie the choice of generator to be connected to and synchronised with the system. Day-ahead gate closure is a feature of markets in Spain and Norway (and Germany), and also in England and Wales before March 2001.[5] Each of these systems allows for adjustment to day-ahead positions, but such adjustments require trades in which one party is always the system operator or a central exchange (such as the intra-day markets in Spain).

Some markets, however, leave day-ahead commitment of generation to the decisions of individual traders. The system operator is content to balance the system by adjusting the output of generators already connected to and synchronised with the system (the process of "despatch"). In such systems, gate closure may take place only a few hours or minutes before the start of each (half-)hourly trading period. In Sweden, for instance, bilateral trading continues up to two hours before delivery. In England and Wales, since March 2001, gate closure has been 3.5 hours before delivery; from 2 July 2002, this period will be 1 hour.

3.1.2. Compulsory settlement of "imbalances"

Since system users cannot forecast conditions exactly, temporary unpredictable fluctuations in production and demand lead to "imbalances" between their contracts and actual flows (metered production and consumption). At any time, many individual system users and traders will either be "spilling" power on to the system (in excess of their needs) or drawing power off the transmission system (to meet a deficit). The system operator's control of system balancing ensures that these individual imbalances do not undermine system security and stability. However, the imbalances represent real power flows which must be converted into financial payments. These payments must offer some kind of penalty for running a deficit and some reward for delivering a surplus. If a deficit imbalance does not incur sufficient penalty, all system users would have an incentive to draw cheap power off the network as an imbalance, which would quickly lead to the breakdown of the operational

[5] In Spain there are 6 intraday markets, so technically, "gate closure" does not occur the day before but will happen between 2 ¼ hours and 7 ¼ hours before despatch (there is a range of hours because the interval between intraday markets reaches up to 5 hours). What does happen is that the System Operator can decide to "lock in" certain programmes the day ahead (e.g. for plants which are required because of transmission constraints). Agents are free to adjust their programmes in the intraday markets as long as they have not been locked in by the System Operator. Most programmes are *not* locked in the day ahead.

and contractual system. Hence, anyone using or connected to a network must be obliged to participate in some kind of compulsory settlement agreement, by which they agree to pay for (or to be paid for) imbalances.

3.1.3. Gross and net pools

Given that there will always be a compulsory settlement system for imbalances, the first question that arises is how such imbalances are calculated. A key choice is whether an individual's imbalance is "gross" or "net" of that individual's bilateral contracts.[6]

- **"Net pool"**

A 'net pool', which many people describe as a system based on 'bilateral contracts', measures imbalances as the difference between (1) a user's net contract position (sales minus purchases) and (2) the user's net physical output (production minus consumption). The difference between contract and physical positions is recorded as an imbalance, to be settled at a price that is determined not by a market, but by rules set out in the compulsory balancing and settlement agreement. Of course, system users that manage to abide by their contracts will have no imbalances and will trade electricity only at contract prices.

Such systems must force contract traders to participate in the settlement system, so that the settlement accounts include both parties to every contract. Most traders aim to "close their position" (ie to achieve a zero net balance with their contracts) in order to avoid imbalance charges.

- **"Gross pool"**

A "gross pool" is a settlement system that ignores the bilateral contracts signed by system users and traders. All physical production and consumption is counted as if it were an imbalance, to be settled at prices determined under the compulsory settlement agreement. This arrangement is found in a number of regimes in Latin America, the US and Australia, and applied in England and Wales until March 2001. Spain uses a "gross pool" for domestic generation, [7] although the pool does "net out" international contracts. These regimes are now often described as "compulsory pools".

In such a scheme, traders can still sign bilateral contracts, but they take the form of financial "contracts for differences". The form of these contracts is well understood in such markets

---

[6] See S. Hunt and G. Shuttleworth, Competition and Choice in Electricity, John Wiley and Co, 1996, pp 144-145, for a discussion of the differences between these two methods of settlement.

[7] Bilateral physical contracts are allowed, but when a domestic generator signs a bilateral physical contract with a consumer, the generator loses its capacity payment and the consumer is exempt from the capacity surcharge. (See section 5.) The former is more than the latter, with the result that neither party has any incentive to sign a physical contract.

and allows traders to achieve anything they would wish to do with a "physical bilateral contract". Indeed, many traders found that the introduction of NETA, with its emphasis on physical contracts, prevented them from applying many provisions that had worked in financial contracts.

- **Overview**

Gross pools and net pools have many common features:

1. Producers, consumers and traders can buy and sell power contracts, allowing participants to hedge their risks. In principle, physical and financial contracts have the same hedging abilities, although any set of market rules can limit the possible range of contracts.

2. At some point ("gate closure"), the system operator takes over control of all changes in generation and consumption.

3. Participants can control the level of output from their generation plants by adjusting the bids and offers they submit to the system operator.

4. The ultimate value of any electricity contract depends upon the cost of alternative sources, which ultimately means it depends upon the charges used to settle imbalances (which are available to everyone as a potential source of power).

3.2. Contract Settlement in Gross and Net Pools

The difference between gross and net pools is sometimes presented as a choice between "financial" and "physical" contracts. In practice, no wholesale electricity contract is ever truly "physical", because it is impossible to make power flow physically from a seller to the buyer. The true difference lies in the method of settling contracts under each form of pool.

In a net pool, users and traders "notify" their contracts to the settlement system. In England and Wales under NETA, market participants make this notification at "gate closure" ie before delivery.[8] In Spain, contracts are notified during the day before. In Norway, however, traders are allowed to notify contracts some time after delivery, which permits a wider range of contract forms. Ex post notification (known in the gas industry as "ex post trading") allows traders to use contract forms that depend on real-time information, eg non-firm contracts, or contracts whose volume depends on a meter reading.

[8] This approach means that it is impossible to use "non-firm" contracts, which depend on generation plant being available, or to link contract volumes to actual metered outputs (which are only known much later).

In a net pool, this process of *notifying a contract* constitutes the fulfilment of the power sale arranged in that contract. Whether the contract parties generate or consume electricity is not usually relevant to fulfilment of the contract itself. Of course, if their generation or consumption fails to match their contract volumes, they will be liable for imbalance charges.

In a gross pool, traders use financial contracts for differences, meaning that they fulfil the contracting simply by exchanging an amount of money equal to the contract volume multiplied by the difference between (1) the power price set out in the contract and (2) the "reference price" used to settle flows through the pool. There are many ways to explain why this approach meets the needs of contracting parties. The following may be helpful:

1. the buyer pays for the contract volume (CQ) at the contract price (CP) – total payment by buyer = CQxCP.

The seller does not then fulfil the contract by handing over a physical volume of power to the buyer:

2. instead, the seller gives the buyer the *value* of the contract volume of power (CQ) at the reference price (RP) – total payment by seller = CQxRP.

Comparing items 1 and 2, it can be seen that the *net* payment from seller to buyer is equal to the contract volume (CQ) multiplied by the difference in prices (RP-CP) (which may be positive or negative).

Hence, neither gross nor net pools really accommodate "physical" wholesale contracts. In a net pool, traders settle contracts by submitting information – a contract notification – to the imbalance settlement system. In a gross pool, traders settle contracts by exchange cash to the value of the power concerned.

3.3. Empirical Evidence on Effects

As discussed above, the difference between gross and net pools lies in the method of settling (1) imbalances and (2) contracts. It is often suggested that this difference in settlement method improves real factors, such as the efficiency of output, the level of prices and the degree of competition. In practice, however, this difference in approach does not seem to have such benefits.

3.3.1. Impact on Efficiency of Output and Emissions

To keep down costs and, ultimately, prices to consumers, any electricity trading system needs to encourage efficient (least-cost) use of generation (and despatchable consumption, ie, load management contracts). The purpose of a centralised balancing mechanism (or pool) is to encourage generators to provide information about the marginal costs of generation, which the system operator uses to determine a least-cost pattern of output. Under a self-dispatch system such as NETA, players trade between themselves to create an efficient (least-cost) dispatch based on their knowledge of their own costs. Both systems are subject to errors and distortions.

One might expect a decentralised, market-based system to produce a more efficient production schedule than centralised decisions taken by a system operator, because more people are able to assess more information. However, the risks created by punitive imbalance prices under NETA affect the behaviour of generators in ways that seem to reduce efficiency and to raise costs. To minimise the risk of incurring the (very large) penalty for a deficit imbalance, market participants are consistently over-generating relative to their contracts (running a surplus imbalance). The system operator must then pay power stations to reduce their output, in order to balance the system.

Furthermore, to minimise the risk of deficits due to unplanned outages, generators are keeping spare generation capacity running at “part load” as protection against the risk of generator failure or unexpected increases in demand. Analysis in Power UK suggests that from 21 September to 6 November 2001, 2.7 GW of surplus generation capacity was “on line”. This surplus equates to 5 part-loaded coal sets (generating units) and reduced efficiency of production. The magazine estimated that such inefficiency increases CO2 emissions by 270,000 tonnes per year.[9]

A comparison of annual statistics on electricity production and C02 emission levels provides further evidence of inefficient dispatch caused by the balancing rules. Demand for electricity has risen around 2 percent per year since 1990, yet from 1990 to 1999 emissions of CO2 fell, as gas-fired generation replaced coal-fired generation. However, in 2000 and 2001, CO2 emissions rose. In 2000, the rise was caused by the low availability of nuclear power stations (due to normal maintenance outages), so that coal plants covered the resultant deficit in output of electricity. In 2001, however, the total amount of hydrocarbon fuel (coal, oil and gas consumption in tonnes of oil equivalent) used for electricity production rose 4 percent compared with 2000, even though total output from coal-, oil- and gas-fired generators rose only 1.8 percent. Total electricity consumption rose 1.65 percent, but emissions of CO2 rose 3.89 percent, despite a recovery in nuclear production.

3.3.2. Effect on prices and competition

In a recent article,[10] John Bower of the Oxford Institute of Energy Studies produces empirical evidence to explain the price reductions observed in the England and Wales electricity market since 1998. Ofgem claims that the 40 percent reduction in wholesale prices were a direct result of the introduction of NETA. Bower compared 18 variables (indicators of costs,

---

9 *“Current NETA operation increases CO2 emissions”*, Power UK, Issue 94, 19 December 2001.

10 J. Bower, *“NETA is no BETTA than the Pool”*, Power UK, [May 2002]. “BETTA” is the British Electricity Transmission and Trading Arrangements, which when introduced will extend the current market arrangements in England and Wales to Scotland, thereby creating a British Electricity market. Here, the author is making a word play on “better”, which sounds the same in English.

market shares, market rules and regulatory environment) with the energy component of pool/market prices over a period April 1990 to March 2002. Using standard regression techniques on monthly data, he found that only seven of these variables had statistically significant coefficients and that the introduction of NETA was not one of them. In recent years (1998-2002), the only significant variables that affected average prices by more than 10% were:

1. the ending of the government ban on authorising new gas-fired generation plants ("gas moratorium"), since the ban had been associated with a price rise of £3.35 per MWh; and
2. the massive reduction in the concentration of ownership among coal-fired plants (de facto, the plants which set market prices), which led to average prices falling by £6.89 per MWh.

Bower concludes:

> "They [international regulators] may conclude that introducing a NETA style reform will make their wholesale electricity markets more competitive. The results discussed here show that this is not the case, and that the industry structure, especially in the mid merit generation sector, is the key determinant of a competitive wholesale electricity market, regardless of the market mechanism employed."

3.3.3. Conclusion

Bower's analysis backs up the investigation of trading rules set out above. The claim that the switch from the Electricity Pool to NETA represented the abolition of "centralised" trading in favour of a "more competitive" system is unfounded. There is no evidence that changing the trading rules per se led to systematically lower prices.

A major change introduced by NETA was the adoption of a "net pool", in which imbalances are calculated *net* of notified contracts. Spain and Norway already operate "net pools", in the sense that imbalances in both countries are calculated net of at least some contracts. However, the system operator retains central control over production from a certain time onwards and participation in the balancing arrangements remains compulsory for generators and retailers. Indeed, under NETA, the obligation to comply with the Balancing and Settlement Code extends to all contract parties, including traders who had previously been exempt from the obligation to become Pool Members.

The choice between "gross" and "net" settlement rules does have some important effects (such as the degree of credit risk borne by centralised markets), but market prices depend on more fundamental aspects of market structure.

4. WHOLESALE PRICING RULES

Whether the power pool is "net" or "gross", the bulk of power trades take place through (or are hedged by) contracts negotiated individually by the buyer and the seller. In the course of such negotiations, the buyer and seller may make use of standard terms and conditions, but the prices always have to be set by the parties.

When, however, traders buy and sell power in advance through a centralised mechanism, such as the day-ahead markets found in England and Wales (until 2001), Spain and Norway, the mechanism must include a rule for setting the price applicable in each (half-)hour, to prevent the market operator from abusing market power in negotiations. When responsibility for power flows is assigned *after* delivery, as for imbalances, a pricing rule is essential, because negotiations are no longer feasible at that time.

The pricing rules for *centralised markets* raise two questions.

- First, the market rules must state whether the market sets a uniform ("market-clearing") price or awards each trade a price that depends upon the offer or bid of the trader concerned ("pay-as-bid").
- Second, in markets that adopt a market-clearing price, the rules must explain what auction process the market operator will use to find the price.

When it comes to *imbalance pricing*, the big question facing market designers is whether to use a single price (which allows simple hedging) or two different prices for surpluses and deficits (to encourage traders to avoid imbalances).

The following sections discuss these questions with respect to the economic theory of pricing and observed experience.

4.1. Pricing Setting Rules

Until 2001, the normal pricing rule for electricity markets was the use of "market-clearing" prices derived from some kind of simultaneous or iterative auction. The justification for this approach lay in its consistency with traditional methods of despatch, in which generators provided information about their own variable *costs*. Proponents of NETA, however, maintained that it would be more efficient to adopt "pay-as-bid" pricing, in which generators offer the *prices* they want to be paid, not least because it would be less conducive to the exercise of market power. Below, we examine these concepts and the claims made about them, first in competitive conditions and second in conditions where generators possess a degree of market power that allows them to raise prices.

4.1.1. Market-Clearing Price

The market-clearing price, sometimes called the "(system) marginal price", is the price at which the volume of offers to sell matches the volume of bids to buy. Markets that set prices at the market-clearing level charge all consumers and pay all producers the same "uniform" price. It is, by definition, the lowest possible price at which producers are prepared to meet demand.

In a competitive environment, settling all trades that the market-clearing price gives producers an incentive to offer their output at their variable ("marginal") cost of production. Given this bidding strategy, the market accepts their offer if and only if the market price exceeds their marginal costs. If producers offer at prices above their marginal cost, they will miss out on some opportunities to generate profitably. Similarly, if they bid below their marginal cost, they may be called on to generate for prices below their costs. Offering a price equal to marginal costs ensures that their offer is accepted whenever they would earn a profit by producing, and is rejected whenever they would lose money from producing.[11]

In such conditions, the resulting pattern of output would be efficient, as the market operator calls on generators in ascending order of marginal cost and satisfies demand at the lowest cost in each time period.

4.1.2. Pay-as-bid

In a "pay-as-bid" system, the market operator declares that producers will receive (and/or buyers will pay) the prices contained in their offer (bid) if it is accepted. In such conditions, producers will try to maximise the prices they receive, but competitive pressure will prevent anyone producer from charging much more than any other. Differences in prices offered and paid will therefore arise out of the difficulty of establishing the competitive market price.

Critiques of marginal pricing suggest that pay-as-bid rules will result in lower prices, but some of these claims are ill-informed. In a marginal pricing system, low-cost producers offer prices below those of the higher cost producers (who actually set the price). Some observers have concluded (rather naively) that a pay-as-bid system would allow them to buy from low cost producers and these lower costs. However, in a pay-as-bid auction, the behaviour of producers changes. Instead of offering to supply at their variable costs, generators will set their prices equal to their estimate of market prices. If each generator correctly anticipates what the market price will be and offers the same price, there will be no difference in

[11] In England and Wales, some plants, mainly nuclear and some Combined Cycle Gas Turbines, offered their output to the Pool at a zero offer price (ie zero minimum price). Such plant had high costs associated with reducing output to zero and starting up again, so that maintaining output was highly valuable to them – as if they had a negative avoidable cost. In Spain, plants offering a zero price are typically (though not only) nuclear plant or run-of-river hydroelectric plants (which have a zero marginal cost).

outcome between pay-as-bid and market-clearing price rules. In practice, however, generators will be prone to forecasting errors, since they must normally submit offers some time in advance of conditions becoming known. In reality, the price offered by a generator with a low variable cost can exceed the price offered by a generator with higher costs, such that the more expensive plant is selected, despatch is inefficient and prices are higher than necessary.[12] Consumers face similar problems in deciding their bid prices.

In practice, however, the debate in the UK about "market-clearing" versus "pay-as-bid" did not focus on potential forecasting errors, but on the implications for exercise of market power. Ironically, as the debate progressed, the degree of market power in generation fell considerably, making much of the debate irrelevant. Nevertheless, public perceptions of the debate – of the supposed behaviour of different pricing rules under less then perfect competition - remain a potent force in electricity market design and deserve closer scrutiny.

4.1.3. Market Power

During the discussions in the UK, very little analysis or evidence emerged on the performance of the difference pricing rules. As explained above, the claim that pay-as-bid would produce lower prices than market-clearing rules failed to take into account the way in which generators would change their bidding behaviour when the rules changed. Criticisms of generators' possible abuse of market power tended to focus on manipulation of the LOLP formula, rather than energy prices. The regulator, Ofgem, (or members of its supervising authority, GEMA) did suggest on occasion that the uniform price rule of the Pool facilitated the exercise of market power by large generators. However, the theoretical and empirical evidence on this question is inconclusive.

John Bower's 2002 study of the UK market confirms that the change in market rules had no impact on the energy component of wholesale prices.[13] From his analysis, he concluded:

> "The first, and most important conclusion that can be immediately drawn from this analysis is that RETA (Review of Electricity Trading Arrangements)[14] inquiry, and NETA itself, had no detectable effect on system marginal price / reference price data.

---

12 Similar errors can arise in markets where wholesale prices are based on system marginal prices, if generators have to submit simplified information on their costs (eg a simple price per MWh, when the average cost per MWh depends on the level and duration of output, which is not known at the time). The potential for such errors is reduced, if generators are allowed to adjust their programme, either internally (by adjusting the output among their own portfolio of plant) or in short-term ("intra-day") markets. Efficiency may also be improved if generators submit complex offers that specify the costs associated with different running regimes (although this makes it more difficult for the market operator to identify a unique "least-cost" solution).

13 J. Bower, op cit.

14 RETA was the previous investigation to the introduction of NETA (new electricity trading arrangements).

> In other words, out of a total of 18 variables tested, RETA and NETA, were quickly eliminated because they were statistically insignificant."

Hence, Bower's study did not find any evidence that the adoption (or, more correctly, the expansion) of pay-as-bid pricing rules had reduced prices by weakening market power or by any other means.

Under market-clearing price rules, the abuse of market power and tacit collusion are relatively simple to identify by comparing offer prices with marginal costs. (Regulatory authorities cannot presume that the relationship between offer prices and marginal costs will remain fixed, but offer price rises that are unrelated to changes in marginal costs merit close examination.) It is difficult for companies to raise prices to excessive levels without regulatory scrutiny. Both the Spanish and England & Wales regulators have started inquiries into the bidding practices of the generators which focused on the offers made by owners of the "mid-merit" (or marginal) plant that effectively sets prices. Each investigation was prompted by "anomalous" bidding behaviour as well as "anomalous" market prices. In each case, it was clear which companies were influencing the marginal price.

Under pay-as-bid rules, all parties depart from their own marginal costs when submitting offers and bids. All parties submit prices that are close to the market-clearing price and, in some cases, cheap plant will inadvertently find itself the most expensive plant accepted in the market. It is therefore more difficult to estimate whether each generator's offer prices were 'following' the market or were collusive or abusive. Hence, scrutinising market behaviour becomes more difficult.

Furthermore, if large incumbents collude to raise prices, all generators benefit from higher prices. These elevated prices encourage new entry into the market, thereby eroding the existing participants' market power. Under pay-as-bid rules, new entrants may not anticipate that they can achieve the same prices as incumbents, because their information about market conditions will be inferior. A study commissioned by the California Power Exchange examined the choice of pricing rule and concluded:

> "Under the uniform price rule, competitors prosper or fail on the basis of their relative generating efficiencies alone; that is not only a consequence but also a prerequisite of an effectively competitive market. Under pay-as-bid, their profitability depends heavily also on their successful forecasting [of the prices charged by competitors]."[15]

This finding has important implications for efficiency and for competition. The system marginal price system is a relatively simple concept for small companies to manage. Small

[15] Kahn, A, Cramton, P, Porter, R, Tabors, R (2001) "Pricing in the California Power Exchange Electricity Market: Should California Switch from Uniform Pricing to Pay-as-bid Pricing?" p7

companies (both independent generators and single consumers) need to understand only their own cost structure for short-term bidding in the Pool, and to form a view on long-term market prices for planning purposes. Under pay-as-bid, for every period, small companies must forecast the market price – ie the offer price of the marginal plant - or they will forego a significant amount of profit. This forecasting activity is disproportionately intensive for small companies. It can constitute a competitive disadvantage and become a barrier to investment and entry by independents, which tends to enhance the position of incumbent producers.

4.1.4. Conclusions

The choice between "market-clearing" and "pay-as-bid" pricing rules consumed a lot of time and effort during the discussion of NETA in the UK, and yet the debate proved rather uninformative in the end. No evidence emerged to support the notion that "pay-as-bid" rules would systematically produce lower prices and, indeed, the greater burden of forecasting seems fall disproportionately on small traders, including independent generators and individual consumers. In such conditions, incumbent generating companies might actually find it easier to exert market power. However, in practice, there is little evidence to suggest that the choice of pricing rule has any major impact on overall price levels.

4.2. Price Identification Procedures

To apply the market-clearing price rule, an electricity market needs a procedure for finding or identifying what the market-clearing price is. Broadly speaking, there are two alternative approaches and the procedure used to find market-clearing prices for electricity depend upon the type of plant that predominates within the system – hydro or thermal.

4.2.1. Pricing procedures for hydro systems

In hydro dominated systems, such as Nordpool, the cost of fuel is zero, but water stored in the dams has an opportunity cost that varies according to perceptions about future supply and demand. Generators tend to be highly responsive, being able to start and stop production within a matter of minutes, or even seconds. In such markets, it is possible to apply (and, indeed, the Norwegian system operator does apply) the kind of pricing rule known in economics as "tatonnement", in which the system operator acts as a kind of "blind auctioneer" who raises and lowers the market price until supply balances with demand.

Within Norway, generators and consumers notify their position at the start of a day, and submit offers to expand production (reduce consumption) and bids to expand consumption (reduce production). During each hour, Statnett runs a "regulation power market", in which the system operator continually declares a market price in real time, starting with the market-clearing price from the day-ahead (Elspot) market for that hour. Producers (and consumers) evaluate this price and decide what quantity of energy they are willing to generate (or consume) at that price. The system operator evaluates the resulting balance of generation and demand by reviewing the flow over interconnections with neighbouring countries. If Norway is *exporting* more than agreed, production is too high; the system operator lowers the price on offer (by accepting and announcing the next lowest offer price submitted by traders). If Norway is *importing* more than agreed, production is too low; the system operator raises the price on offer (by accepting and announcing the next highest offer price submitted by traders). Traders then adjust their production and consumption in response. This process continues until supply and demand are in balance within Norway. Statnett sets the market-clearing price as the highest accepted offer price (if Norway was short overall) or the lowest accepted bid price (if Norway was in surplus overall).

4.2.2. Pricing procedures for thermal systems

In systems dominated by thermal generation (ie fired by hydro-carbon and other combustible fuels), each power station has a specific avoidable cost of generation, covering fuel costs and variable operations and maintenance costs. Thermal plants also have technical constraints that rule out certain flexible modes of operation, eg rapid changes in output. As a result, the system operator usually retains more control over the physical level of output, as seen in the despatch methods practice in England and Wales and in Spain. In such systems, market-clearing prices must be derived from the pattern of output imposed by the system operator (rather than prices determining output as in hydro systems).

Generators submit information on their production costs (eg the cost of starting the plant, plus the variable cost per MWh generated), and technical parameters such as ramp rates (ie the rate at which they can change their output, in MW per minute). The system operator uses a complex computer programme to find the lowest cost pattern of output from each generator (over some period such as a day) based on the costs and technical data. The market uses this output data to derive a market-clearing price by the following means:

1. Establish the least-cost pattern of output from each generator sufficient to meet demand;
2. Identify the cost in each (half-)hour of running each generator;
3. Find the most expensive generator running in each (half-)hour;
4. Set the market-clearing price equal to the cost of the most expensive generator running in each half-hour.

This four-step approach applies in Spain's daily and "intra-day" markets, and also lay behind the selection of the System Marginal Price in the old Electricity Pool of England and Wales. The last step in this process has been criticised as uneconomic, or inconsistent with competitive markets, because it finds a "maximum price", instead of the "minimum price" that buyers are used to seeking. However, this criticism reflects a misunderstanding. Given the least-cost pattern of output, the price identified in the last step is in fact the *lowest* uniform price that will call forth the level of production needed to meet demand.[16] This outcome matches that in any other competitive market.

4.2.3. Implications for Italy

The Italian system is largely based on thermal generation. Consequently, it is to be expected that any market-clearing price will have to be derived from information on costs and technical data submitted by the power station operators.

4.3. Single or Dual Imbalance Pricing Rules

Every electricity market must have a compulsory system for settling the power flows associated with imbalances between contracts and actual flows. The prices put on these imbalances have important consequences for incentives. In the long-term, the level of these prices determines the ultimate value of power, and the incentive to build and maintain generation capacity. In the short-term, the value of imbalances affects incentives to generate and consume electricity and to sign power contracts.

One of the main elements of NETA is the adoption to a "dual imbalance price regime" – lower prices for surplus imbalances than for deficit imbalances. The declared intention of this system is to provide an incentive for market participants to balance. (Ofgem was convinced that using the same price for surplus and deficit imbalances would encourage market participants to tolerate imbalances.) However, the dual price system has other implications for risks and incentives. In the following, we consider the relative merits of the single and dual price systems.

4.3.1. Single price systems

In England and Wales, Norway and Spain, participants pay for (or are paid for) deviations from contracted (or notified) levels of generation (and sometimes consumption). The rules for pricing such imbalances are based on one of two pricing philosophies: the market-clearing price ; or a punitive price.

- In England and Wales, under the Pool until 2001, generators received the Pool Purchase Price (System Marginal Price plus LOLP.VOLL element) for all scheduled generation. If a generator failed to produce the scheduled volume due to an outage, the generator had pay the Pool Purchase Price for the missing volume of energy.[17] On the other hand, consumers paid for their consumption at the Pool Purchase Price (plus a small surcharge to cover various costs of system operation). Hence, short-

---

16 If the price were any lower, at least one generator would be operating at a loss and would be unwilling to generate.

17 Different prices – primarily the plant's own offer price – applied to variations in output that were the result of instructions from the system operator.

term variations in production or consumption (compared with contract volumes) used to be priced at Pool Purchase Price.

- In Spain, generators are charged the marginal price of the generation offers accepted to maintain the balance between generation and consumption in real time, plus a surcharge based on the costs of secondary reserve (allocated pro-rata to the absolute value of each trader's deviation from the programme).

- In Norway, generators sell most of their output through contracts (in which prices are negotiated) and they sell some of their output in the day-ahead (Elspot) market (which sets a single market-clearing price). Any last minute variations in output or consumption are settled at a market-clearing price derived from Statnett's "regulation power market".

In all these cases, a single charge for energy is or was used for imbalances, based on the derived market-clearing price for energy (or a proxy for such a price).

In the old Pool, and in Spain, this system is combined with the adoption of a "gross" (or "very largely gross") pool, in which all (or nearly all) physical output flows through the centralised market as if it were any imbalance. It is therefore hard to say whether traders are/were relying "too heavily" on the imbalance market. However, both markets operated centralised markets from the day-ahead stage. As a result, the system operators had a great deal of time to adjust the pattern of output and were not unduly inconvenienced by anticipated imbalances.

In Norway, the system operator effectively has control only within real-time (ie within each hourly settlement period). However, Norway's hydro generators can respond to system problems within seconds, so imbalances do not create operational problems.

4.3.2. Dual price systems

Under NETA, both generators, traders, retailers and some consumers in England and Wales pay for (and are paid for) imbalances calculated as the difference between metered generation (or consumption) and notified contract sales (or purchases).[18] Participants who run a deficit must pay the System Buy Price (SBP). SBP is a relatively high price, derived from the offer prices of expensive generators that the system operator instructs to increase their output through the Balancing Mechanism. When participants run a surplus, they receive the System Sell Price (SSP). The SSP is the average of the prices received by the system operator for its sales through the Balancing Mechanism, and tends to be rather low (sometimes even negative).

---

18 Generation (and its net contract sales) must be held in a separate account from consumption (and its net contract purchases), so that traders cannot net off the imbalance risks for generation and consumption.

The dual price system can be regarded as a penalty system, in the following sense. Suppose that in any period there is a market price, PM, which lies between SBP and SSP. Then it is possible to regard the two imbalance prices as a market price plus or minus a penalty:

- System Sell Price (for surpluses) = PM – (Penalty 1)
- System Buy Price (for deficits) = PM + (Penalty 2)

In the Netherlands, for example, the dual prices for imbalances are calculated in precisely this way, as the market price (emerging from the Amsterdam Power Exchange) plus or minus a fixed penalty (ie, Penalty 1 = Penalty 2). However, under NETA, neither of the imbalance prices is a true market price. Instead, they reflect a restricted set of production conditions and the associated penalties are implicit and highly unpredictable. Both imbalance prices are highly influenced by the behaviour of the system operator.

These features of the imbalance prices would not matter if market participants could take steps to avoid imbalances. However, many market participants are unable to control, or to predict exactly, their actual sales to consumers (particularly in the case of small retailers) or generation (particularly for intermittent generators such as wind turbines). The dual pricing system therefore creates unhedgeable risks. Market participants are unable to predict what the System Buy Price and System Sell Price will be, or even which price will apply to their imbalance. As a result, they have not found ways to share imbalance risks with other market participants, through hedging agreements. Instead, market participants are managing these risks (1) by horizontal (and vertical) integration, and (2) by continuously "going long" and (3) by choosing to operate plant flexibly, even if that means operating is less efficiently.

1. Horizontal integration allows market participants to spread their risks over a wider portfolio of production and sales, so that their imbalances vary less relative to their overall revenues. This approach is only available, of course, to large companies, and acts as a barrier to entry. In Germany, the Federal Cartel Office has criticised similar dual price schemes operating in some areas of Germany for similar reasons.

2. The second tactic, "going long", means that traders to try to run a surplus in all conditions. This approach suits risk averse traders, because the penalties implicit in the System Sell Price (for surpluses) have been consistently less volatile and less expensive than the penalties implicit in the System Buy Price (for deficits). If market participants consistent enter the market committed to a "long" position, ie having committed excess generation, the system operator must continually use the Balancing Mechanism to pay generators to reduce their production.

3. Finally, generators are operating their plant in a less risky, but less efficient, manner, eg by part-loading a lot of capacity, rather than taking some capacity off-line overnight. (When generating capacity is off-line, generators face a bigger risk that it will not return to operation on time, than if the plant is part-loaded.)

The Balancing and Settlement Panel is currently considering a modification to the market rules to eliminate the dual price regime and to replace it with a single price. This modification has been inspired by the risks and problems facing small companies, which are unable to benefit from horizontal integration. However, the system operator has responded with a counter-proposal that maintains the dual pricing system, but narrows the gap between them.

4.3.3. Conclusions

Although all electricity has implicitly the same value at any one time, no matter where it is coming from or where it is going, several electricity markets apply two different prices to imbalances, a lower one for surpluses and a higher one for deficits. The reasoning behind this approach is the desire to encourage individual market participants to balance, ie to keep their net production (production minus consumption) in line with their net contracts (sales minus purchases).

The creation of two imbalance prices (and the abolition of any notional "market price") was one of the design priorities for NETA. The stated aim was to promote trade, ie to encourage traders to balance their own portfolios and not to rely on the system operator. In practice, however, the outcome has been slightly different: fearful of the high penalty associated with deficits, market participants have reacted by trying to run a permanent surplus. As a result, the system operator has consistently had to intervene to reduce production from the plants committed by market participants. Such a process does not necessarily produce an efficient pattern of despatch, or reduce the role of the system operator. Moreover, the unpredictable nature of the implicit penalties has created unhedgeable risks and financial difficulties for independent (ie small) companies operating a small number of generators, or serving a small number of customers. These problems act as a barrier to entry.

In a system such as the UK's, where production is divided among a sufficient number of generators to allow competition, the discouragement to independent companies may not be sufficient. In Germany, however, where production is highly concentrated (at least within regional control areas), the discouragement of new entry is serious enough to merit investigation by the Federal Cartel Office.

Spain's market applies a single market price to both surplus and deficit imbalances, but then allocates the costs of ancillary services to absolute imbalances (surpluses and deficits), to give an incentive for market participants to abide by the programmes emerging from their trading activities. A similar approach is found in the Netherlands. Using a single market price allows traders to manage risks more effectively. For instance, if demand rises, traders know that the system operator will call for extra output from generators. The highest price paid for this extra output determines both the revenue to the generator and the charge to the retailers whose demand increases. The generator and the retailer therefore face similar, hedgeable risks, since their net revenues depend on the same "basis price". Adding a small charge for any imbalances does not create any major additional "basis risk", if the size of the charge is relatively stable.

Nord Pool operates with a single imbalance price, which tends to be fairly predictable and can be hedged. Statnett appears to have few concerns about difficulties caused by imbalances, not least because Norway's hydro generation can respond very quickly at reasonable cost, so imbalances do not create operational problems.

The emphasis placed on the need for market participants to balance (ie to avoid imbalances), and the penalties placed on the absolute value of imbalances, therefore seems to depend upon a number of factors.

- The system operator may want market participants to balance, if imbalances are expensive to counteract, or threaten system security. Neither of these conditions seems to hold in Norway, for instance, which may explain the adoption of a single price. Such risks also diminish if the system operator has control over output from the day-ahead stage.
- The penalties for imbalances tend to hit small generators and retailers hardest, because they cannot spread imbalance risks over a portfolio of production and consumption. In markets where new entry (ie small companies) are important, this effect is considered a major disadvantage.

The choice of single versus dual imbalance price therefore depends heavily on whether the system operator can ensure system security at low cost by other means, and whether facilitating entry by small companies is a priority or not. Furthermore, even if it is considered desirable to penalise imbalances, such penalties should be predictable and limited in scope. NETA shows that large, unpredictable penalties for imbalances create risks that cannot be hedged, and that unnecessarily penalise small players. Some reform of this aspect of NETA is therefore likely.

5. CAPACITY PAYMENTS

5.1. Economics of Investment in Generation

The electricity industry is characterised by large investments in long-lived assets such as power stations and transmission lines. These assets last between 20 and 40 years and such investments are irreversible. The fixed cost of buying and building these assets is therefore a "sunk cost" that cannot be avoided. To operate these assets, the owners incur "avoidable costs". Some avoidable costs are "fixed" (ie incurred just to keep the asset available) and some are "variable" (ie linked to the level of output from the asset). Investors hope to recover sunk costs over the life of the asset during periods when prices lie above the "avoidable" costs of operating the asset.

In competitive markets, prices will be driven by the marginal cost of meeting demand. In most periods, generation capacity will exceed demand and the marginal cost of meeting demand is the variable cost of a generator. These variable costs - the fuel costs of power production – form a relatively low share of total costs. Hence, for as long as prices reflect only variable costs, few generators, if any, will be able to cover their sunk or fixed costs. If this situation persisted, it would not only drive all generators into bankruptcy, it would also discourage any investor from building new generation capacity.

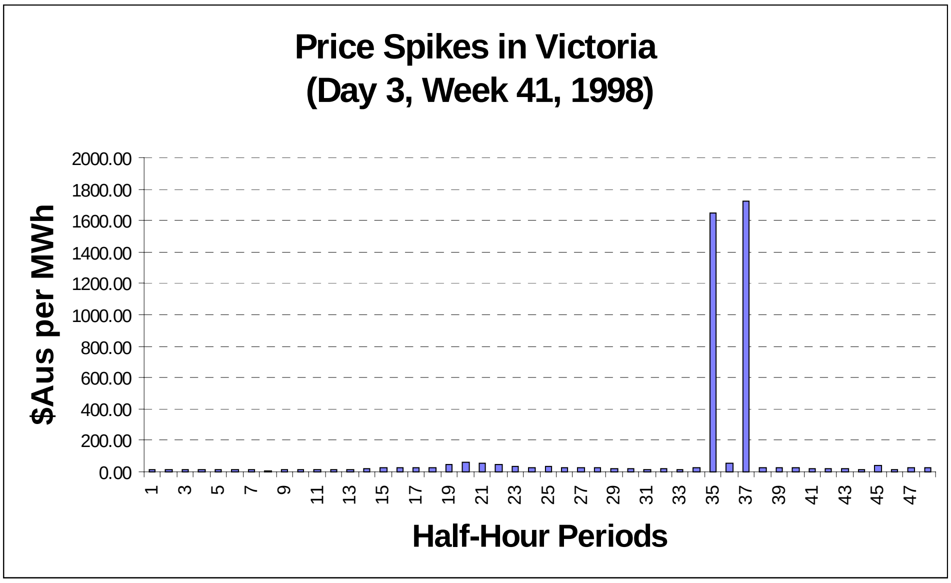
However, at some point, the excess generation capacity will disappear, either because old plant costs down or because demand rises. When demand actually exceeds generation capacity, the competitive price of electricity will no longer reflect only the variable costs of production. Instead, the competitive market price of electricity will rise to the level necessary to ration the available supplies, by discouraging demand.

If demand is highly sensitive to price (price elastic), the resulting increase in prices may be relatively small. However, in most electricity systems, a large volume of demand is price inelastic and prices rise dramatically when capacity is short. These high prices signal to the market that new investment is required and new power stations are built.

5.2. Implications for Pricing

The way in which these high prices arise depends upon the time at which prices are set. In "real-time" markets – the centralised pools (gross or net) that set a price for each (half-)hour with respect to actual conditions at the time – prices oscillate between very low levels (fuel costs of generation) and very high levels (rationing prices when demand exceeds capacity). Experience in the electricity market of Victoria, Australia demonstrates how a shortage of generating capacity can produce very short "price spikes", in between sustained periods of low prices. (See Figure 5.1.)

Figure 5.1
Electricity Prices in Victoria, Australia



The purpose of such real-time pricing mechanisms is to ration demand, but consumers (or their retailers, acting on their behalf) may find it difficult to curtail their demand in response to prices, if they do not find out until too late what the market price actually was. Even in periods where capacity is short, real-time market prices will not rise much above fuel costs. Market prices only hit the extreme levels seen in Figure 5.1 when load is actually lost (ie when consumers are cut off involuntarily). In a thermal system, where prices are derived from the pattern of production, such prices are only posted after the hour is over, by which time it is too late for consumers to respond.[19] Demand response then relies on consumers' (or their retailers') ability to anticipate large but unpredictable changes in spot market prices.

Because it is difficult to anticipate (and respond to) such events, experience like that seen in Victoria often leads to calls for alternative pricing mechanisms. In Spain, market rules prevent prices from rising above the highest offer price submitted by any generator. In Norway, the regulator has considered forcing the system operator to sign contracts with peaking generators. The market in Argentina, to look further afield, has always set prices at the *start* of each hour, based on forecast conditions. However, several markets offer additional payments to generators, in the hope that the resulting investment will avoid the need for price spikes to occur. Such mechanisms must nevertheless provide a reasonable prospect of recovering sunk and fixed costs, as well as the variable costs of generation.

[19] As discussed in section , traders can watch prices emerge from a "tatonnement" process in Norway's electricity market, but such processes are limited to predominantly hydro systems.

5.3. Centrally Administered Capacity Payments

Several electricity market designs try to eliminate the huge fluctuation in prices that occur if generators can only recover their sunk costs in the few hours when demand exceeds capacity.

- **England and Wales: Probabilistic Capacity Payments**

The Electricity Pool in England and Wales provided a "capacity payment", which spread the value of peak-time capacity over a number of periods, in proportion to the *probability* (calculated on the previous day) of a capacity shortage. When the probability of a capacity shortage rose (because of a high demand forecast or an anticipated lack of capacity), generators received higher capacity payments.[20] The purpose of the payment was two-fold:

- First, to provide short term signals to the market so that maintenance decisions and short-term availability for stations reflected the requirements of the market.
- Second, to provide long-term signals for new capacity construction, thereby maintaining sufficient plant margin to ensure system security.

The scheme worked well during the early and mid 1990s, when it matched the daily planning cycle of coal-fired generators. By the late 1990s, the formula was beginning to offer increasingly inappropriate signals, because it failed to keep abreast of short-term changes in generator availability due to (1) more flexible management at coal-fired stages and (2) short-term arbitrage between electricity and gas markets. However, the abolition of the capacity payments under NETA has raised concern about incentives for investment in new generation capacity.

- **Spain: Fixed Surcharge per kWh of Consumption**

The Spanish Electricity Pool collects a charge on each kWh of consumption and divides the resulting revenue among available generation capacity as a reward for making capacity available: the more plant is available, the lower the payment to each generator. This scheme therefore mimics some aspects of the formula in England and Wales although, while the Spanish formula does not attempt to reflect consumers' valuation of unserved energy, as in England and Wales. In Spain, the system seems intended to allow generators to recover the costs of reserve (peaking) plant. However, in practice the level of the charge paid by consumers is defined by the government, which has progressively reduced it from an average level of 0.78 €cents / kWh to 0.48 €cents / kWh.

---

[20] See Appendix A for a detailed explanation of the formulae.

"Free" consumers (ie consumers who have exercised their right to leave regulated retail tariff) pay an average of 0.18 €cents/kWh. "Captive" consumers (ie who have not exercised this right) pay whatever amount is necessary to bring the average contribution up to 0.48 €cents/kWh (currently more than 0.60 €cents/kWh). Thus, the more consumers exercise their right to leave the tariff, the more captive consumers have to pay. The methodology for defining the level of the charge levied on consumers and paid to generators in Spain has not been made public, so it provides no long-term guarantee of cost recovery. Perhaps not surprisingly, there are increasing concerns in Spain about the risk to investment (and the cost of relying on generators constructed under the "special regime").

- **Appraisal and comparisons with real-time pricing**

Both these capacity payment systems allow generators to spread the recovery of their sunk costs over a larger number of hours and a larger volume of sales, so that prices fluctuate less than in a "pure" real-time pricing system. In practice, the main benefit may be the avoidance of extremely high prices *and the associated tendency for governments to intervene in markets*. Consumers can protect themselves against high prices by signing long-term contracts with generators; such behaviour would encourage efficient plant construction. However, consumers may believe it is cheaper not to sign long-term contracts and instead to demand government intervention if and when electricity prices actually rise. Such behaviour is an example of "market failure" (ie the lack of a market, in this case for long-term contracts) that may prevent efficient construction of generation capacity.

Nord Pool does not offer any similar capacity payments. However, the reason seems to be the relative stability of power prices in a system dominated by hydro-generation, in which prices are determined by energy shortages (not capacity shortages) and by arbitrage between periods (which allows the reward for capacity to be spread). Although the Norwegian market has seen high prices during seasons when water was short (notably in 1996). However, the price rise was small relative to price rises observed in thermal systems (about four times normal prices, rather than a hundred times) and was limited by the ability of many consumers to reduce their demand. The regulator, NVE, has investigated the market at various times, but has not undermined the basic freedom of generators to raise prices in times of shortage.

In other words, real-time pricing may offer economically efficient pricing signals in theory, but in practice these signals may be unsustainable, because they invite government intervention to change the market rules, or to impose caps on prices. Such interventions undermine the prospect of recovering generators' investment costs and so discourage investment. As a result, despite their shortcomings, capacity payments may offer a more stable and *credible* basis for rewarding investments in capacity, than real-time pricing.

5.4. Conclusions

Security of supply remains a critical issue within the electricity industry. Relatively inflexible demand (particularly close to real-time), combined with high capital costs and long lead-times for investment, makes the electricity market prone to large swings in market price. If prices are allowed to follow marginal price signals, electricity markets are prone to long periods of low prices followed by extreme pricing levels. Consumers can protect themselves through bilateral contracts, through which they undertake to pay the fixed costs associated with generation. Alternatively, consumers can rely on political pressure to override or prevent high prices. The latter strategy often seems attractive to consumers, but in the long term it prevents generators from recovering their fixed costs, undermines investment signals and raises costs and prices.

The Victoria market in Australia demonstrates extreme prices at times of capacity constraint, whilst the California market demonstrates the political problems caused by sustained high prices. To avoid such pressures, a number of pools have implemented a centrally administered capacity payment that "smoothes" electricity prices, ie to spread the reward for capacity over more periods. Consumer groups have sometimes complained about generators manipulating capacity rules in ways that raise prices. The spreading of capacity payments may make it difficult for interruptible consumers to avoid high price periods. However, the problems with real-time pricing (including the difficulty for consumers of anticipating when high prices will occur) mean that centralised, administered capacity payments may continue to offer some advantages. On this point, the solution is likely to be driven by local production conditions and concerns.

6. CONGESTION MANAGEMENT

Transmission constraints within a network segment the electricity market. As a result, if production costs differ from one area to another, electricity may have markedly different values at different points on the network. Signalling the local market values of electricity may be important for encouraging efficient decisions about the production and consumption of electricity, and about the location of generation capacity.

On the other hand, market segmentation may limit the effective number of competitors within some areas, which creates fears of a reduction in market liquidity and even possible abuse of market power. The prospect of electricity prices differing by a large amount within one country may create political problems. Many electricity markets therefore make no allowance for transmission constraints when setting prices - even though, of course, the system operator must take transmission constraints into account when deciding the pattern of output.

The treatment of transmission constraints, and the resulting congestion of the network, therefore represents a major design issue. The following section sets out three different approaches, beginning with the one that most closely reflects the physical reality.

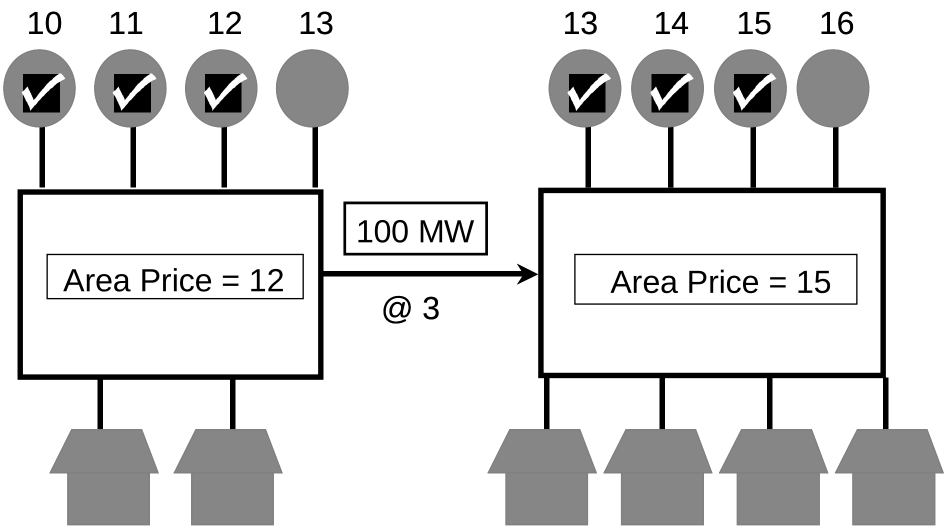
6.1. Market Splitting

Market splitting creates separate markets in each area bounded by transmission constraints. The Norwegian market applies the principle of market splitting, as illustrated in Figure 6.1 below. This example shows two zones linked by one interconnector, but the economic principles apply equally to multiple zones and interconnectors.

In the figure, each "ball" represents 100 MW of generation and each "house" represents 100 MW of demand. Associated with each generator is an offer price in øre per kWh. Demand is 200 MW in the left hand zone, and 400 MW in the right hand zone, making 600 MW in total. If the system faced no transmission constraints, the market (or the system operator) would select the six cheapest generators, ie those offering prices from 10 to 14 øre per kWh. The "system price" would equal the highest accepted offer price, ie 14 øre per kWh.

However, selecting generation on this basis would create a surplus of 200 MW in the left hand zone, which would overload the 100 MW link between the zones. The system operator must therefore leave one generator in the left hand zone unused – and efficiency demands that the most expensive generator, priced at 13 øre per kWh, is not despatched. To replace this plant's output, the system operator despatches the next most expensive plant in the right hand zone, at a cost of 15 øre per kWh. The generators required for an efficient despatch in accordance with transmission constraints are shown with "ticks".

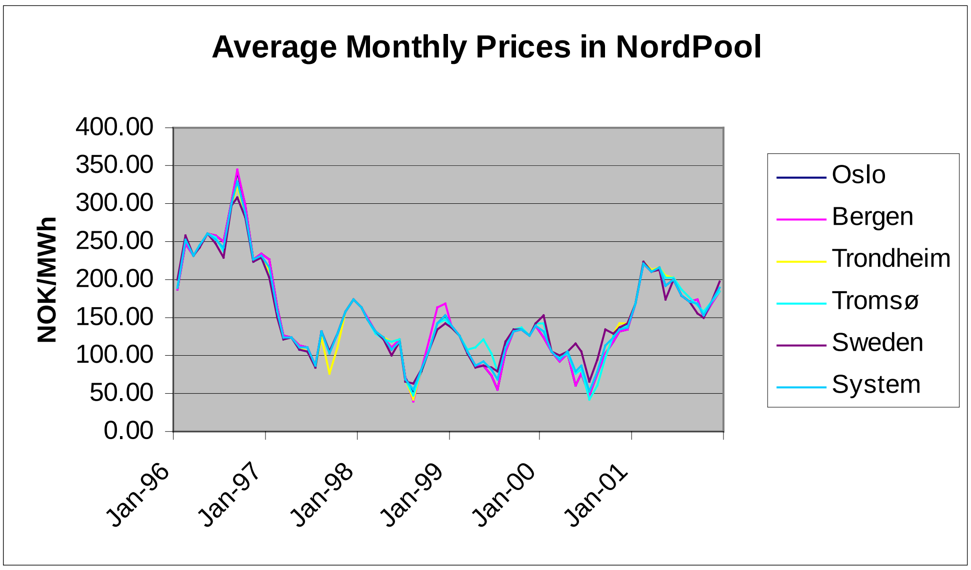
Figure 6.1
Zonal Pricing in Norway
(Prices in øre/kWh)



The market-clearing price now differs between the zones. On the left, the "area price" is 12 øre per kWh, whilst on the right the "area price" is 15 øre per kWh. The flow over the interconnector is 100 MW and earns a profit of 3 øre per kWh, which the system operator retains as a contribution towards the costs of the main transmission grid.

The Nordpool market as a whole has used market splitting for a number of different areas within Norway and, more recently, for Sweden, Finland and Denmark. Figure 6.2 demonstrates how prices varied in the Norwegian and Swedish areas since 1996.

Figure 6.2
Average Monthly Prices in the NordPool from 1996 to 2001



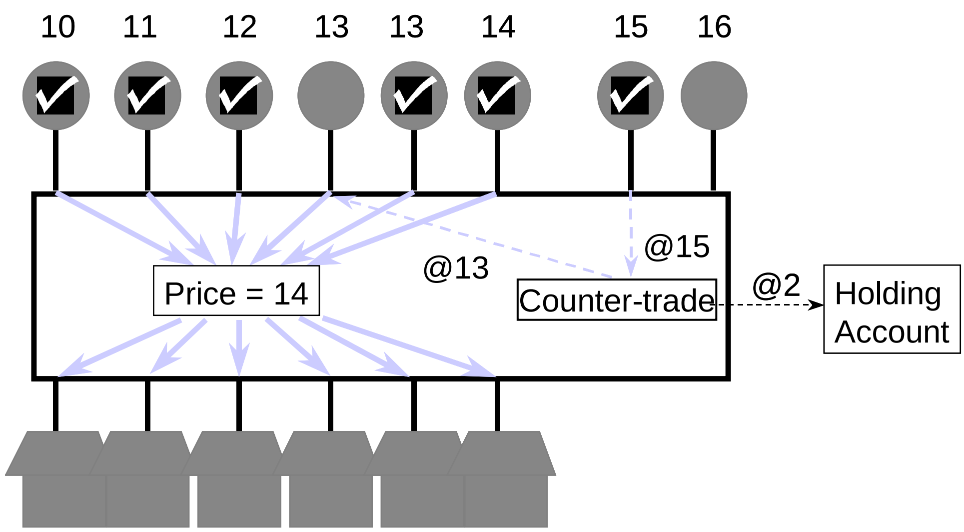
Market splitting works effectively when there are liquid spot markets on either side of the constraint and when the location of constraints is relatively stable. Market splitting provides clear signals of the value of the transmission connections and informs decisions regarding transmission investments. Market splitting can be problematic if the market becomes too fragmented and zones are illiquid. Competition in the price-setting segment of the market for each area is the most effective method of curbing generator market power.

However, it is not always simple to assess the state of competition within one area. Sometimes, due to the physics of electricity transmission, the price that a generator can command in one area depends on production conditions in other areas. In New Zealand and the PJM market in the US, prices are calculated using an integrated model of the whole transmission system, in order to capture this effect in pool prices. NordPool relies on individual traders to understand such interactions, but in any case the generation sector is not concentrated enough for competition problems to be a major concern to government.

6.2. Counter-trades

Several electricity markets ignore the physical state of the network and define markets that span one or more transmission constraints. Within such markets, there is only one market price and traders select generators without regard to constraints. The system operator has to manage constraints by arranging a matching pair of buy and sell trades at specific prices. Figure 6.3 shows the general approach used in such markets (or, indeed, when a constraint occurs within an area of a split market), for the same set of production and consumption data as in Figure 6.1.

Figure 6.3
Transmission Constraints Managed by a "Counter-Trade"
(Prices in øre/kWh)



As before, demand is 6 blocks of 100 MW each, requiring traders to buy 100 MW from each of 6 generators. As before, they select the cheapest generators and the market-clearing price is 14 øre per kWh. However, as before, the system operator must instruct the generator priced at 13 øre per kWh not to generate. The system operator does this by selling power to the generator at its offer price, at which the generator is indifferent between producing or buying power to fulfil its sale to the market. To meet demand, the system operator also has to buy power from the next most expensive generator in the importing (right hand) zone, at a price of 15 øre per kWh.

Overall, therefore, the system operator buys power (in the importing zone) for 15 øre per kWh and sells it (in the exporting zone) for 13 øre per kWh, making a loss of 2 øre per kWh on the net trade. (The trade is known as a "counter-trade" because it flows in the opposite direction from the actual flow over the interconnector.) The loss is transferred to some kind of holding account and must be recovered from system users through an increase in transmission charges, or a service charge levied by the market operator.

In England and Wales, the system operator adjusts the output of generators (both up and down) to deal with transmission constraints. Under both the Pool and NETA, the price for these adjustments was the offer prices submitted by the generators themselves.[21] The resulting costs were recovered through a charge per kWh of consumption (under the Pool) or generation and consumption (under NETA), but such costs are now small, since transmission constraints are relatively uncommon.[22]

In Spain, the system operator adjusts the output of generation plants as shown in Figure 6.3, but the pricing rules are slightly different. High cost plant receives its offer price when "constrained on" (ie when instructed by the system operator to increase its output), as in England and Wales. However, low cost plant does not pay its offer price to buy power when "constrained off" (ie, instructed by the system operator to reduce its output). Instead, its sales to the national market are cancelled, which is financially equivalent to buying in power at the national market price. This approach has, in the past, caused generators to distort their offer prices, in order to avoid being "constrained off" and losing potential profits.

Constraining plant by the system operator works effectively when constraints are short-lived and vary from place to place. However, persistent constraints that incur substantial costs for counter-trades might merit alternative solutions.

---

21 Under the Pool, generators submitted offer prices for all their output; these prices applied to all instructed deviations from the output schedule emerging from the day-ahead market. Under NETA, generators notify the system operator of their intended output level and submit offers to increase and bids to decrease their output; consumers can offer similar facilities if they can control their net consumption.

22 Constraints arise in England and Wales during some periods of low demand which coincide with either line outages or generator outages.

6.3. Congestion Pricing

The third method of dealing with constraints on the network is to put an explicit price on the use of a congested part of the network. The outcome is similar to the market-splitting shown in Figure 6.1, but instead of defining separate market prices by area, the transmission company defines an explicit charge for moving power over a constraint to the general market. In Figure 6.1, for instance, the right hand zone might represent the general market, with a market-clearing price of 15 øre per kWh. The transmission company would then charge 3 øre per kWh for each unit transmitted from the left hand zone to the right hand zone, such that all power produced and consumed in the left hand zone was worth 12 øre per kWh (= market price of 15 øre per kWh minus transmission charge of 3 øre per kWh).

For the most part, this approach is limited to transmission capacity at national borders (eg around the Netherlands). However, *in practice*, Statnett applies the market-splitting system in Norway by imposing a "bottleneck fee"[23] on all generation and consumption, rather than by defining area prices explicitly. The bottleneck fee depends upon the area within which power enters or leaves the transmission system. It is calculated as the difference between the system price and market-clearing area prices. Traders value electricity at the system price *net of* the bottleneck fee, ie at the area price, even though settlement procedures do not use area prices explicitly.

Ofgem has also proposed a similar approach for the system in England and Wales, entailing auctions for rights over transmission capacity on the national grid. However, it has proven impossible to agree how or where such capacity would be defined. The relatively low cost of transmission constraints also makes such schemes redundant.

6.4. Review of Schemes

Transmission constraints arise on transmission networks from time to time, due to changing patterns and costs in generation and demand. Electricity markets can handle constraints (1) by market splitting, (2) by "counter-trades" ("constraining plant on or off") or (3) by putting an explicit price on congested parts of the network.

Market splitting indicates the local value of electricity when there are liquid and competitive spot markets within each zone or area. Such signals are most valuable when constraints occur frequently at the same location, as in Norway. If the location of constraints is unpredictable, however, it is difficult to draw boundaries between zones. Some markets, such as New Zealand and the PJM, have adopted generalised versions that define a price for electricity at each "node" of the transmission system, rather than for wider zones.

[23] This is the informal name for the element of Statnett's tariff known as the "capacity charge". It is defined on an hourly basis in øre per kWh.

An alternative is to defining wider markets and to use “counter-trades” to deal with constraints within them. This method has proven broadly satisfactory in England and Wales, as well as in Sweden and some other markets. (Spain applies a similar approach, but its attempt to minimise the associated costs by adopting special pricing rules has cause other, unnecessary distortions to generators offer prices.) Counter-trading works well in systems where constraints are few or transitory, since it avoids the need for complex (and perhaps unpredictable) pricing systems. Regulators also favour this approach when area or zone markets would suffer from a lack of competition or liquidity. However, redefining the market does not remove the market power of large generators within small areas; such generators may still be able to raise their offer prices, but their offer prices will have little or no impact on market prices.

Congestion pricing works best when there is competition for the congested link, ie when markets on both sides of the constraint are liquid and competitive. Norway uses a real-time variant of congestion pricing to split its markets but, within the EU, the scheme is more commonly found at international borders.

The main difference between these schemes is the nature of the cost signals given to system users. Under market-splitting and congestion pricing, system users pay the cost of congestion, in that the value of electricity in a (constrained) exporting zone is less than the value of electricity in a (constrained) importing zone. Under a system of counter-trades, only the system operator sees the cost of individual constraints. (System users only see the charge required to recover the total costs.) These costs offer a variety of incentives:

1. to locate new generation capacity in high price zones and (possibly) to close down generation capacity in low price zones;
2. to locate new (industrial) demand in low price zones and (possibly) to close down demand in high price zones;
3. to pay for increased transmission capacity between low price and high price zones.

It is therefore important whether users or the system operator bears these costs in the first instance. The location of generation and demand may be more efficient if system users bear the cost of constraints (which implies the adoption of market-splitting or congestion pricing). However, in practice, individual investors may be unable or unwilling to respond to these incentives, if other factors predominate. For instance, it may be impossible to build additional transmission capacity without the cooperation of the existing monopoly transmission network owner. Efficiency may be greater if the transmission company (the system operator) to bear some or all of the costs, to create an incentive to reduce congestion.

Furthermore, decisions to reallocate the cost of congestion will affect the value of investments in generation and power-consuming facilities, such that market designers may not have a free hand. The best allocation of such costs – and the choice of the appropriate scheme for dealing with congestion - therefore depends upon a variety of local factors.

7. CONCLUSION

There is no such thing as a simple electricity market. The Pool Rules that applied in England and Wales from 1990 to 2001 covered about 900 pages (leaving aside the agreement that gave them force). A commonly stated objective for the new arrangements was to offer a simpler system, but the Balancing and Settlement Code, which replaced the Pool Rules, runs to over 1000 pages.

Electricity markets cannot be simple, because the physical electricity system (generation and networks) is not simple to operate. Nevertheless, it is possible to identify a few key questions that affect how the market affects those involved with it:

- How the elements that are administered centrally interact with bilateral contracts;
- How the centrally administered elements fix prices for markets, balancing and imbalances;
- How investors are rewarded for building essential generation capacity; and
- How the market manages the physical constraints within the transmission network.

Different markets have adopted different solutions to these problems, because there is no universal "first-best" solution. Instead, market designers have to choose from several options, each of which has several advantages and disadvantages. Their task is to make rational choices, by identifying which option is most likely to meet local needs and preferences.

If this report helps to inform a rational debate of the choices facing the Italian electricity industry and its consumers, it will have fulfilled its aims.

APPENDIX A. CAPACITY PAYMENTS IN ENGLAND & WALES

From 1990 until 2001, the Electricity Pool of England and Wales offered all generators a contribution towards the cost of capacity, provided their plant was available to run.

A.1. Pool Purchase Price

The capacity payment used in England and Wales was based on the day-ahead production schedule derived under the Pool Rules, which acted like a forward market where the price was set equal to the *expected* value of energy.

For any future time period, there is a small probability, equal to the Loss of Load Probability (LOLP), that demand will be greater than capacity, in which case demand will have to be rationed and the value of energy rises to the Value of Lost Load (VOLL). In all other conditions, demand can be satisfied by scheduling and despatch available capacity and the value of energy is the real-time System Marginal Price (SMP). SMP is the cost of meeting additional demand from generation capacity; the Pool calculated it by calculating a least-cost production schedule (based on prices and quantities offered by generators) and then finding the price offered by the most expensive generator included in this schedule. The price at which the Pool purchase power in the day-ahead, least-cost production schedule was the probability weighted average of these two prices:

Pool Purchase Price = {(1-LOLP) x SMP] + [LOLP x VOLL]

The "LOLP-VOLL" element of this formula constituted an additional reward for making capacity available.

A.2. Availability Payment

In order to ensure that incentives were balanced, the Pool also paid a similar reward to capacity that was not in its day-ahead production schedule, but which was available and which might be required.

Under the Pool Rules, a generator would receive its offer price for output generated "out of merit", ie when it had not been called in the least-cost day-ahead production schedule. The system operator would call generators "out of merit" to deal with transmission constraints, generator outages and unforeseen increases in demand.

Most of the time, a generator's offer price would be a reasonable estimate of the value of its output when it ran; system marginal costs at such times would exceed the plant's offer price by only a small amount (ε). Occasionally, however, load would be lost, in which case each MWh of output from such a generator would have a value equal to VOLL, ie significantly more than its offer price. The anticipated value of the plant's output is therefore:

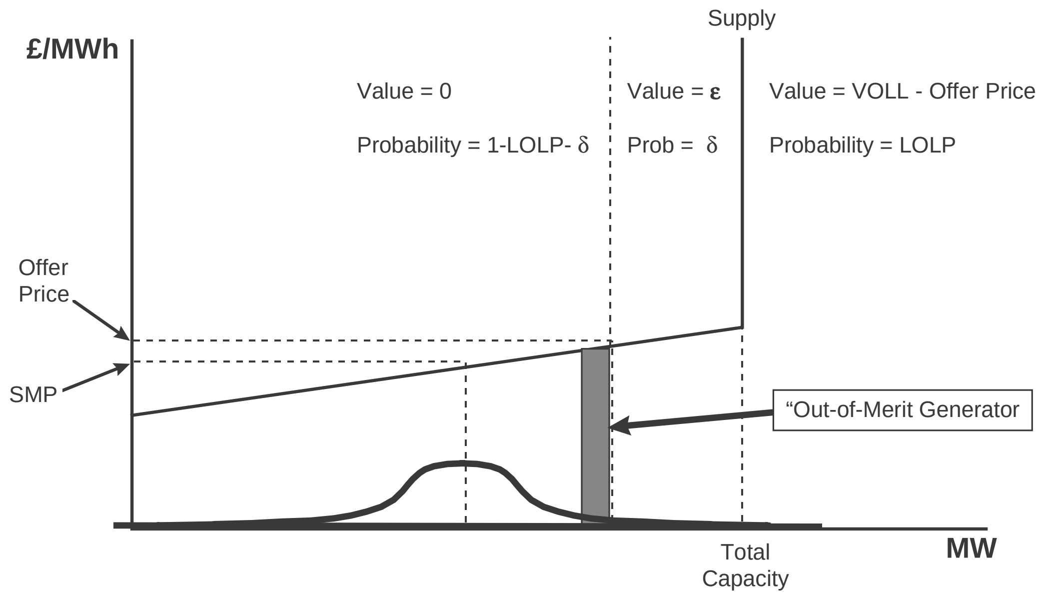
Case 1 - assume the plant does not run:
anticipated value of output = 0

Case 2 - assume the plant does run, and load is not lost:
anticipated value of output = offer price + ε

Case 3 - assume the plant does run, and load *is* lost:
anticipated value of output = VOLL

Now consider the anticipated value of the plant's availability to the system. Figure A.1 shows the three cases, by comparing a typical curve of generator offers (up to the level of total capacity), a typical demand forecast (which defines the marginal plant and SMP), and a typical probability distribution for demand (which indicates that and "out of merit" generator might be required.

Figure A.1
Availability Payments to Unscheduled Capacity



In case 1, to the left of the generator's position in the merit order, the generator's availability has no value, since the market value lies below its offer price and it's capacity is not needed to meet demand.

In case 2, demand is high enough to require the generator, but not high enough to exceed total capacity. The typical generator's output is equal in value to the marginal cost of some even more expensive generator, which is at the margin. The Pool will in any case pay for the plant's actual output at its offer price (under the arrangements for plant that is despatched "out of merit"). The value of the plant's availability (ε) is the difference between the value of its output at the time and the plant's own offer price. In practice, the marginal cost of the marginal generator is normally not much higher than other generators' own offer prices, so ε is small. In addition, the probability of this case occurring is also relatively small. Hence, day-ahead probability-weighted value of the plant's availability in case 2 would most likely give a negligible figure, even though the plant is actually running.

In case 3, when load is being lost because demand exceeds capacity, generation capacity is most valuable. We already know that this case has a probability of LOLP. The value of the plant's output at such times is VOLL, but it will receive its offer price for unscheduled output under the arrangements for plant despatched "out of merit". The *net* value of the plant's availability is therefore (VOLL - Offer Price) per MWh of capacity.

The total expected value of the plant's availability can therefore be estimated using the mathematical formula for an unbiased estimate, ie a probability weighted average (the sum of probabilities times values):

	0	*times*	0	(case 1)
plus	negligible (δ)	*times*	negligible (ε)	(case 2)
plus	LOLP	*times*	(VOLL - Offer Price)	(case 3)

equals 0 + negligible + LOLP × (VOLL – Offer Price)

equals LOLP × (VOLL – Offer Price) *approximately*

The total sum gives the anticipated value, one day ahead, of a MWh of available generator capacity. This value varies half-hourly. A higher probability of a capacity shortage means a higher payment for capacity made available, which therefore encourages generators to make plant available at times when it is needed most.

APPENDIX B. CAPACITY PAYMENTS IN SPAIN

Since its introduction in 1998, the Spanish Electricity Pool has been offering generators an additional payment as incentive for investments in generation capacity with a view to guaranteeing supply to all customers.

The methodology behind the capacity payment methodology in Spain has not been made public.

B.1. Remuneration of the Generation Activity

Following Article 16 of the Electricity Act (Law 54/1997), the wholesale energy price is equal to the price in generation offer accepted to meet demand. There are also additional payments for the ancillary services processes in which generators are involved, and whose provision is not considered compulsory. Finally, there is a capacity payment (known as "Capacity Guarantee Payment") paid to generators who are available to supply energy.

At the time of the introduction of the pool in 1998, the capacity payment value was set at an average of 0.78 c€/kWh. It has since been reduced twice down to 0.48 c€/kWh in spite of the sustained reduction in the generation capacity margin.

B.2. The Capacity Payment

B.2.1. The Capacity Payment for Producers

According to the Electricity Act, the capacity payment component is an amount defined by taking into account the long term capacity needs of the system. Therefore, the payment must reflect the verified availability of generation and their assets life.

Spanish legislation distinguishes between two different types of generation units that have the right to receive the capacity payment; the Ordinary Regime, which is composed by those plants over 50 MW of installed capacity, and the Special Regime, which is composed of those plants under 50 MW of installed capacity and plants over 50 MW not included in the Ordinary Regime because of the primary energy source they use (eg wind power, solar, biomass).

In the Ordinary Regime, generators must submit generation offers to the day ahead market in order to be entitled to receiving the capacity payment. Since there may be generators who are rarely required to generate, provisions for ensuring the operating conditions of the plants are also defined. In 1998, plants had to demonstrate that they were reliable, and thus eligible to receive the capacity payment, by having run a minimum of 100 hours at full capacity (or equivalent) over the previous five years. This requirement was subsequently modified, so that now thermal plants have to run a minimum of 480 hours at full capacity (or equivalent) *each* year. Hydroelectric plants have to run 480 hours a year over the previous five years in order to receive the capacity payment.

The requirement currently imposed on thermal plants distorts despatch, since peaking plants who would otherwise run less than the required 480 hours have incentives to offer their generation below their variable cost of production in order to fulfill the requirement. This displaces more economic forms of generation and increases the cost of meeting demand. In addition, the losses in which some plants incurr may exceed the capacity payment revenues, which would mean that some reserve plants would prefer to close rather than stay in the market and run 480 hours.

The capacity payment to Ordinary Regime plants is defined on a per-kW basis, affected by an availability coefficient in the case of thermal plants (defined as the percentage of peak hours in which the plant was available) and by a “producibility” coefficient in the case of hydro plants or fuel constrained thermal plants (defined on the basis of the historic average level of production).

The per-kW payment is obtained by subtracting from the amount paid by consumers for capacity (which is defined by the government) the amount paid to other forms of generation (see below). Thus, the more capacity is available in the system, the lower the per-kW payment will be, and vice versa. Similarly, as demand grows the amount of money to be paid to generators also increases.

Special Regime generators are not obliged to participate in the pool, at least during a transition period. Special Regime generators who do not participate in the pool are paid a regulated tariff. Special Regime generators who present offers in the pool are paid predefined levels as capacity payment on a per-kWh basis.

Generation units under physical bilateral contracts and electricity imports cannot receive the capacity payment.

B.2.2. The Capacity Payment for Distributors, Eligible Customers, Retailers and Foreign Agents

In 1998, all agents paid the same amount of money as capacity payment. However, in order to incentivise the exercise of eligibility, in 1999 the government decided to reduce the amount of payment to be contributed by customers who exercised their eligibility. The government did not, however, reduce the amount payable to generators, which means that the reduction in capacity payments by customers who exercise their eligibility is financed through an increase in the capacity payments paid by other customers.

Currently there are two different groups that are charged the capacity payment in the pool. These groups are the following:

- Agents subject to the access tariff regime: eligible customers, retailers and foreign agents
- Agents subject to the full-service tariff regime: distributors (on behalf of regulated tariff customers)

The capacity payment is time differentiated and charged on a per-kWh basis on the basis of the number of periods in the tariff paid by agents subject to the access tariff regime.

The total amount of money collected as capacity payment from agents subject to the access tariff regime depends on their actual consumption. The capacity payment by these agents averages 0.18 c€/kWh.

The amount of money payable by distributors is defined as that which is required to bring the average level of payment to 0.48 c€/kWh. Thus, the amount paid by distributors depends on the amount paid by agents subject to the access tariff regime.