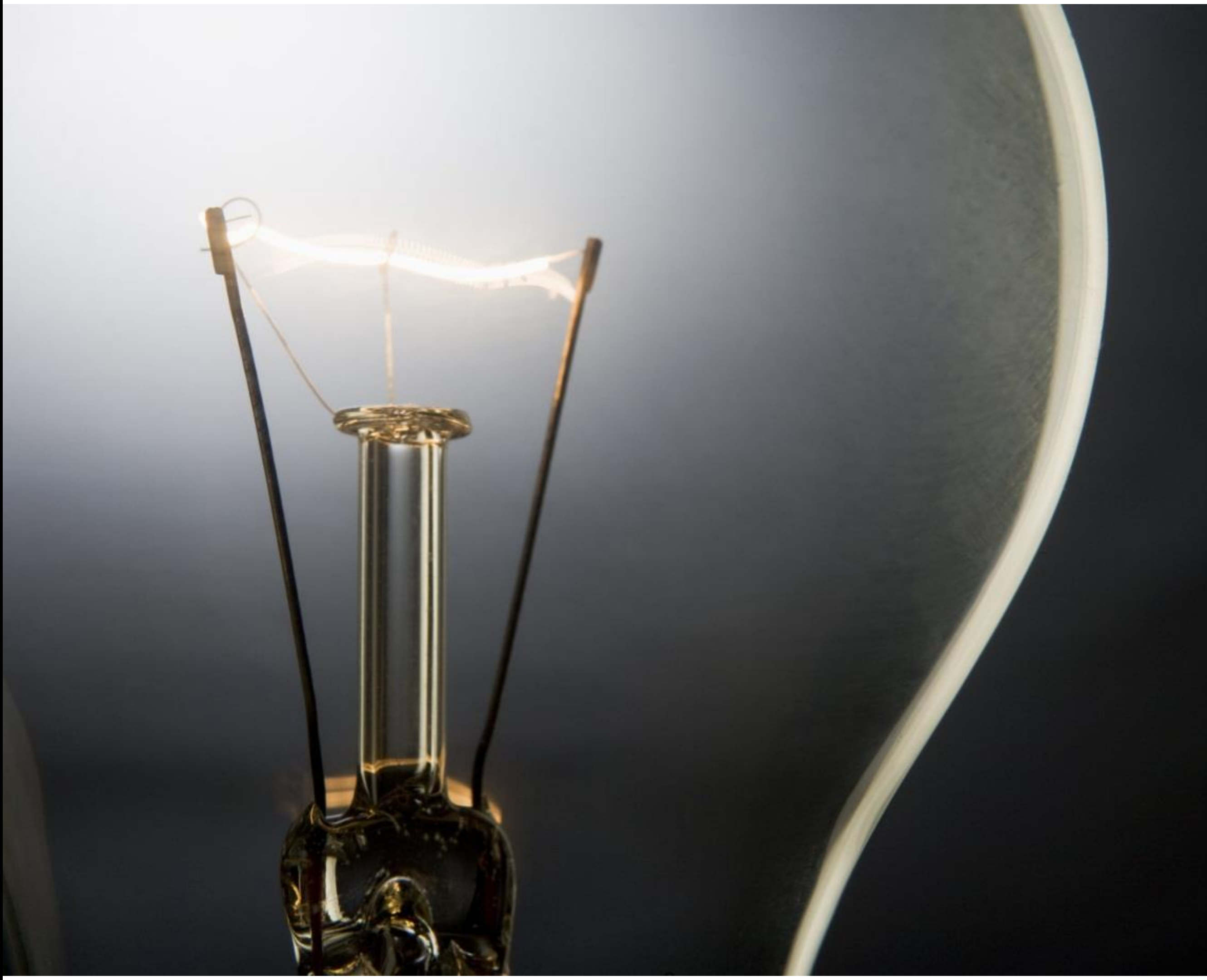


Four market design scenarios for Europe

A NEPP report

March 2015

Table of Contents

Executive Summary

European countries are currently working on the implementation of an internal market for electricity. The main purpose of establishing a common market is to deliver efficient, secure, sustainable and affordable electricity to European consumers. The chosen market design - usually referred to as the Target Model – contemplates the day-ahead and intra-day coupling of national markets, or rather, of "properly defined" bidding zones. Although countries that can justify exceptions are allowed to maintain country-specific market designs, the Target Model contemplates the coupling of energy-only markets. Whether to divide a country into bidding zones to reflect internal congestions is up to the countries themselves, unless cross-border trade is affected.

However, even though Europe is only just gradually starting to implement the first network codes, there is widespread uncertainty as to whether the Target Model - which is strongly influenced by the successful market-coupling of the hydropower-rich Nordic countries - suits current electricity market developments in the European continent. Three issues in particular cause concern. The first issue concerns whether energy-only markets will be capable of delivering credible investment signals to attract enough investment to ensure security of supply, or will additional intervention in the form of capacity remuneration mechanisms be needed? The second issue concerns the Target Model's zonal approach to congestion management, and whether sub-optimally defined bidding zones might not hinder the efficient dispatch of generation and network investment. Should Europe abandon zonal pricing and become a single nodal pricing region? The third issue is a consequence of well-intended interventions. Is Europe giving up on the competitive market model given the challenges of integrating large volumes of subsidized renewable energy? Are we moving back towards central planning? We therefore present in this document four different market design alternatives for Europe: the energy-only Target Model, the Target Model complemented by capacity remuneration mechanisms, a move towards nodal pricing and a return to central planning with a sinlge buyer.

Will energy-only markets be capable of delivering credible investment signals?

In competitive electricity markets, the system marginal price is - according to economic theory - set at the short-run marginal cost of the most expensive plant needed to meet demand, plus an amount to recover long-run marginal costs and guarantee security of supply. This means that in energy-only markets, where generators are only paid when they generate, electricity prices *must* be allowed to peak at very high levels at times of scarcity, otherwise thermal reserve capacity that is seldom needed will not be able to cover its costs. It is important to note that a fundamental characteristic of energy-only markets is that they cannot operate efficiently with price caps.

The share of intermittent, zero-variable cost renewable generation is expected to increase substantially in most European countries in the years to come. Already many countries are witnessing a drop in wholesale prices and capital-intensive thermal generators are finding it increasingly difficult to cover their long-run marginal costs. At the same time, more intermittent generation requires more reserve capacity to generate when the wind does not blow or the sun does not shine, so more thermal generators than today will be expecting to

cover their capital costs in the hours with very high scarcity prices. Sweco's modelling results for energy-only markets show that moving towards 2030 both scarcity prices and zero prices will increase, resulting in a large increase in price volatility in continental Europe.

Are capacity remuneration mechanisms necessary for adequate investment?

If investors believe that very high scarcity prices might be followed by political and regulatory intervention, energy-only markets are likely not to produce sufficient investment, and capacity margins will decrease. Many European governments, including those of Great Britain, France, Italy and Poland have reasoned that in the absence of flexible demand energy-only markets cannot be trusted to produce an adequate capacity margin necessary for security of supply and have decided to introduce national capacity remuneration mechanisms (CRMs) of different designs to reduce risk to investors and, hopefully, the cost of maintaining generation adequacy.

Sweco modelled two different capacity markets scenarios: an Integrated Capacity Market scenario in which capacity markets of similar design are introduced throughout Europe, and a "Patchwork" Capacity Market scenario similar to what we are seeing in Europe today. Modelling results show a distortion between investments in different countries under the "Patchwork" sceniario as investments in countries with CRMs crowd out investments in neighbouring countries without CRMs. Also, non-served demand in countries without CRMs increases when neighbouring countries introduce capacity markets, leaving non-CRM countries with lower capacity margins.

Sweco's modelling also shows that congestion revenues for interconnectors could be impacted when capacity markets are introduced. Congestion revenues drop as price volatility and price spreads between zones are reduced when more generation capacity is installed with capacity markets. We also review some proposed models to enable cross-border generation and interconnector participation in a capacity market in a neighbouring zone.

Will a zonal approach to congestion management hinder an efficient dispatch of generators and an efficient use of the existing network?

Our third market design scenario addresses transmission capacity allocation and congestion management, and opting for Locational Marginal Pricing (LMP) – also referred to as nodal pricing at its most granular form - over the Target Model's zonal pricing. Nodal pricing is often considered not only as more transparent than zonal pricing but also as superior to zonal pricing both to support efficient short-term operations and dispatch, and to facilitate longer-term contracting and investment in new generation, the two main functions of electricity markets.

The Target Model's zonal approach to capacity allocation and congestion management means that market participants may trade freely within a bidding zone regardless of internal congestions, which are dealt with by system operators. Trade between zones, however, requires capacity allocation. To allocate cross-border capacity, the Target Model is introducing flow-based methods, that while more efficient than current capacity allocation methods still leave decisions that will impact markets in the hands of national system operators: the selection of critical transmission lines, for instance, will impact social welfare and day-ahead electricity prices. How to deal with critical branches that are not stable is also important, as it will affect how market participants are able to manage congestion risk. It is therefore important that these choices are made in a transparent way.

Also of concern is the issue of bidding zone definition. At present, zones usually coincide with national borders, not transmission constraints, resulting in significant loop flows. Many countries, however, want to retain the one country one bidding zone definition, arguing instead for investments in internal transmission capacity to reduce loop flows. Only Denmark, Norway and Italy have voluntarily divided the country into bidding zones to reflect internal congestion, however Italy does not expose demand to zonal pricing. In Sweden, bidding zones were only introduced following complaints from Denmark that the system operator was systematically moving congestion to the Danish borders to avoid high costs to countertrade internal congestions and to keep Swedish wholesale prices down.

Given this reluctance to expose market players within a country to different prices it is, at present, hard to envisage a future for central, possibly Europe-wide, dispatch and nodal pricing in Europe. It should be noted that loop flows will continue to be a problem unless Europe becomes a single nodal pricing region, as experience from the United States, where loop flows from adjacent systems are cause for concern in all LMP markets, shows. There are also fears of loss of liquidity under nodal pricing given the very large number of nodes that can make up a market. It should be noted that loss of liquidity is also a concern when discussing the break up of large bidding zones into smaller bidding zones, as larger markets tend to be more liquid and have lower transaction costs. In the United States, hubs have been introduced to facilitate trading.

Single buyer or competitive markets?

During the past years, Europe has been busy creating markets, in the hope they would become competitive. Competition has not always been achieved, and doubts are surfacing as to whether it is possible to combine competitive markets with achieving more recent climate change targets. The reluctance of several European governments to trust energy-only markets and the increasing volumes of technology-biased subsidises to generation stretching far into the future signal that politicians do not trust markets to deliver the volume and type of investment in new generation and transmission capacity needed to achieve a low carbon, secure, and affordable energy system.

Support for government intervention is growing in many countries. For instance government backing of long-term investments is seen as a way to reduce the risk to investors, and if the cost of capital goes down, costs to consumers will (hopefully) go down. And because the share of intermittent renewables will continue to increase, further eroding thermal generation's business case, a single buyer model in which the system operator or a state-controlled entity contracts adequate reserve capacity is preferred by some over Europe's decentralised trading approach, and is presented in this document as the fourth and last scenario for European electricity markets. From there it is not too far to a return to regulated utilities.

1 Introduction

1.1 Background

Throughout the European Union, national electricity markets are implementing significant modifications to their market design in order to align with the common European Target Model upon which the single European electricity market is to be founded. The Target Model reflects the prevailing market design in Europe. Nord Pool is the prime example. However, Europe’s commitment to deliver a progressively high proportion of electricity generation from renewable energy sources means that a very different energy system than the current is emerging. The prevailing market design in Europe does not necessarily deal with high renewable penetration adequately. Renewable energy sources are often intermittent, have low short run marginal costs and are often located far away from load centers.

Intermittency implies that while renewable generators have the potential to generate large amounts of electricity, it is difficult to rely on them, and greater amounts of total installed capacity are needed to meet a given security of supply. However, in a system with large amounts of renewable generation, renewable generation’s low marginal costs will depress the average wholesale price of electricity, making it more difficult for conventional plants to recover their costs. As they will be running considerably less, they will be relying on high but uncertain price spikes to recover their costs. There are growing concerns that this could lead to underinvestment in conventional capacity without some form of intervention.

However, there is still ongoing debate as to how to best respond to this challenge. Will the current market design stand up with only smaller adjustments, or will there be a need for a more fundamental redesign of the market?

In addition to concerns about securing investment both to maintain existing generation and to encourage the development of new sources of capacity, several other issues deserve attention. Large variations in generation over both time and space will further strain the electricity networks, making the efficient expansion and utilization of the grids increasingly important.

Finally the volatility in electricity generation is also likely to lead to volatile electricity prices. Price spikes are likely to be higher under systems with large amounts of renewable generation as conventional generators will have to recover their costs during fewer hours. Public opinion and the media have little or no understanding of this fact. Furthermore, price spikes that significantly exceed the marginal cost of the last generator needed to meet demand can lead to accusations of anticompetitive and manipulative behavior and calls for price caps. Price caps are already in place in several countries.

How electricity markets will evolve will depend on the decisions made by investors and policymakers. Will investors be willing to accept (risky) investments based on electricity prices that risk being more volatile and possibly lower on average, or will they be discouraged to invest in generation capacity? Will politicians (and regulators) rely on the markets even if this result in volatile prices, or will they opt for an interventionist approach with more detailed regulation and central planning? Will the scale of the investment challenge simply force politician to interfere?

1.2 Four market design scenarios

Based on the above, we see four different market design scenarios:

- Energy-only (the Nordic market model for Europe), discussed in Chapter 2.
- Capacity market (addition of a separate capacity market creating income for capacity even if not used), discussed in Chapter 3.
- Locational Marginal Pricing (a way for wholesale electricity prices to reflect the value of energy at different locations, i.e. incorporating the costs for network losses and network congestion), discussed in Chapter 4.
- Detailed regulation (increased central planning and consumer price based on average cost), discussed in Chapter 5.

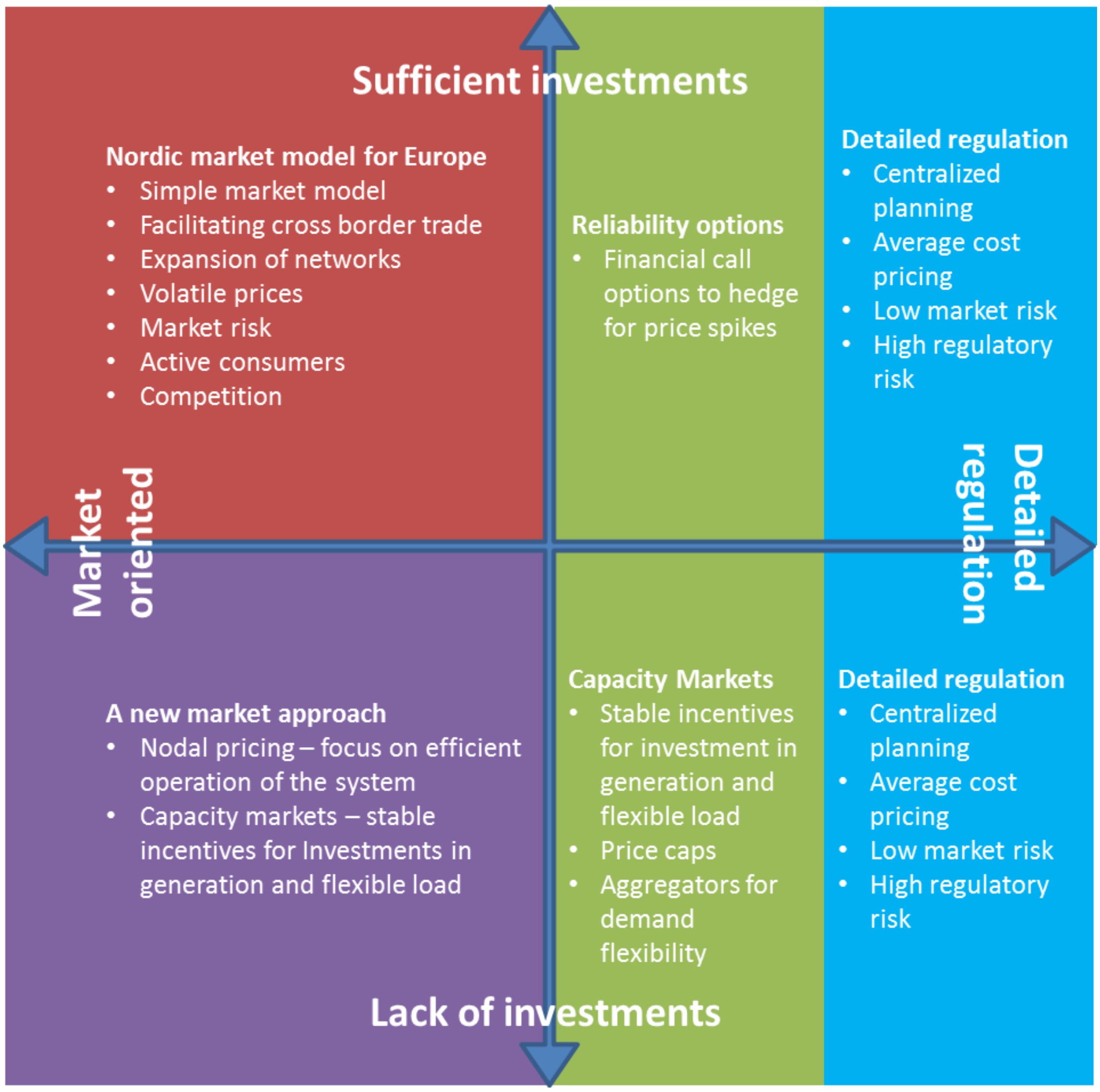
These four scenarios are shown in Figure 1. The energy-only scenario is in line with a continued implementation of the zonal European energy-only target model. The energy-only scenario builds largely on the market model that has been implemented in the Nordic countries. It is the preferred model if it succeeds in attracting sufficient investment in generation capacity.

However, even if energy-only markets succeed in attracting sufficient new investment, it is highly probable that electricity prices will be very volatile, and occasionally very high. While this is not a security of supply issue, it may lead to concerns of market power and may be politically unacceptable. If this happens, we may see the end of energy-only markets in Europe. We then see three possible scenarios.

The first alternative to the energy-only scenario is the capacity market scenario. Under this scenario, which is becoming a reality in several European countries, national markets across Europe introduce capacity mechanisms with financial call options that provide a hedge for high prices. This would lead to more stable prices, and support investment in generation capacity. This scenario is still largely market-based, even if it involves more central planning.

The second alternative to the energy-only scenario builds on the capacity market scenario. In the locational marginal pricing scenario, we abandon zonal pricing and introduce nodal pricing to encourage a more efficient dispatch of generation and in the location of new generation investments. The third alternative to the energy-only scenario is Detailed Regulation, which involves increased central planning, long-term contracts for new capacity and possibly a return (or the continuation for several countries) of regulated end-user tariffs. The reason for this approach is not primarily physical security of supply, but rather a reluctance to accept volatile prices, an urge to create a low risk environment for investors and a belief that consumers will gain from regulated average prices rather being exposed to market prices.

Figure 1. An overview of different energy market designs, depending on the level of market orientation and the degree to which there is a belief that there will be sufficient investments in energy generation and distribution



Source: Sweco Energy Markets

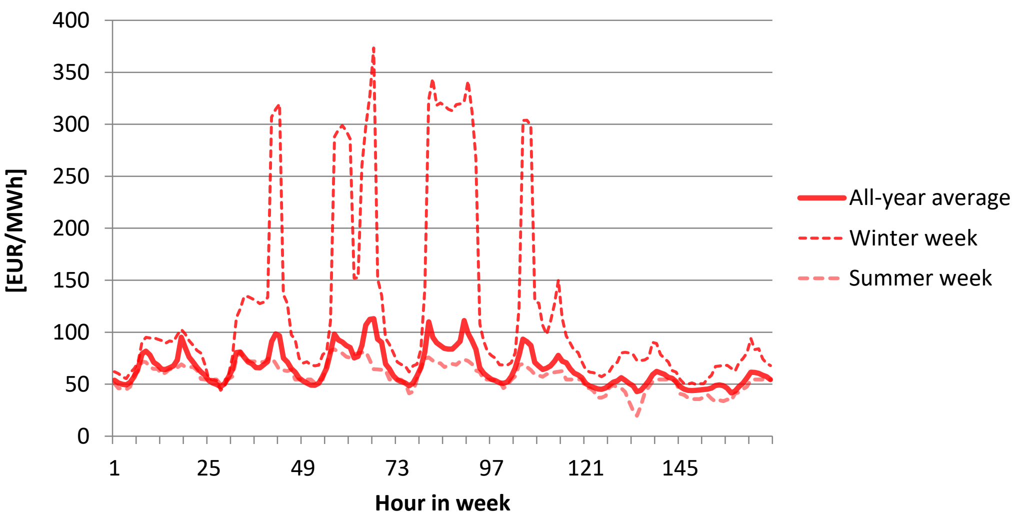
If the current market model is unable to attract sufficient investment, and end in a (perceived or real) security of supply problem, it is likely that the model will be challenged and replaced. Combined with a lack of trust in market solutions, the detailed regulation scenario is a likely outcome. The solution may to a large extent be similar as with the detailed regulation in the upper right quadrant, but the underlying motives differ. In this situation it is more important to handle the physical security of supply. With somewhat more trust in the market we instead expect a situation with some type of capacity mechanism.

We expect that some type of market-based capacity remuneration mechanisms will be adopted in several markets. In addition, with an electricity system under pressure in combination with a trust in markets and price signals, we expect that there will be a willingness to expose market participants (producers and consumers) to the correct local prices. Adopting locational marginal pricing would imply a new market approach for Europe, but one already adopted by a wide range of electricity markets, most notably in the United States.

2 Energy-only market

In the first scenario – an "energy-only market" – generators only get paid when they generate. In competitive energy-only electricity markets, the system marginal price is set at the short-run marginal cost of the most expensive plant needed to meet demand, plus an amount to cover security of supply and long-run marginal costs. Therefore, a typical characteristic for energy-only markets is very high prices at times of scarcity. It is during these hours that peaking plant that is needed only a few hours a year recovers long-run marginal costs. These peaks are necessary to attract necessary investments in peak generation and in demand flexibility. Figure 2 shows an example of price variations for different weeks in a model run for 2030.

Figure 2: Hourly prices in SE4 in representative weeks in 2030: an all-year average week, an average winter week, and an average summer week.

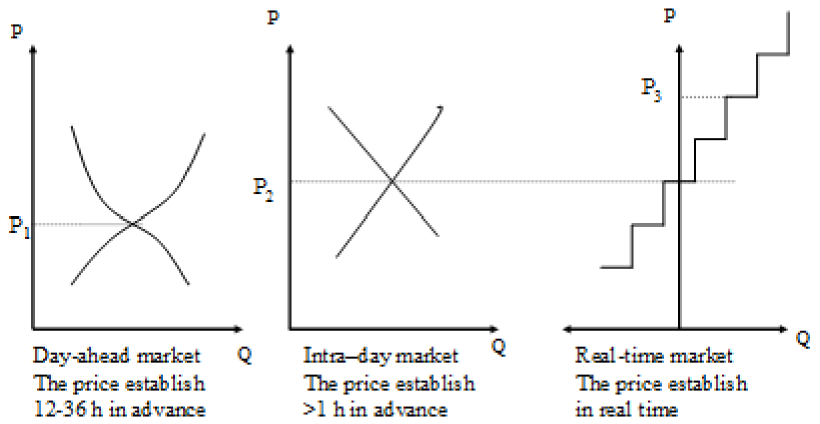


Source: Sweco Energy Markets

2.1 European Target Model

The European Target Model aims to encourage the harmonization of European wholesale market arrangements. It is essentially an "energy-only" market model in which trading takes place in four time-frames: a forwards market, an auction-based day-ahead physical market, an intraday market with continuous trading and a balancing or real-time market run by the Transmission System Operators (TSOs); the latter three time-frames are shown in Figure 3. The Target Model was strongly influenced by the successful market-coupling of the Nordic countries.

Figure 3. Balancing the system in the short run



Source: Sweco Energy Markets

The purpose of the short-term markets is to create incentives and tools for the market players (producers, suppliers, and consumers) to sell and purchase physical wholesale power efficiently through a common platform, and to participate in the balancing of the system. There should be reasonably strong incentives for market players to behave in such a way so as to create a balance between power generated or purchased and power sold. A perfect balance is not possible to achieve and that is the reason why real-time markets operated by the TSOs are needed. Unavoidable imbalances between contracted volumes and actual load and generation are settled ex post.

As mentioned above, generators in an "energy-only market" only get paid when they generate. This is the case for all three markets – day-ahead, intra-day and real-time, although capacity payments are common in real-time markets: to make sure there is a minimum level of reserves in the system for the (TSO) to operate the system safely, generators may get paid for allocating resources (standing ready to generate but not actually generating) through special arrangements administered by the TSO. This is the case in Norway and Denmark. This type of "capacity payment" exists in nearly all markets, even in so-called pure energy-only markets. The important feature of all is that they do not affect the market prices.

The day-ahead markets also have a very specific role in the allocation of cross-border transmission capacity and thus facilitate trade between countries or between price areas within countries. The idea is that power flows between countries shall be a result of decentralized decisions made by generators and consumers. Capacity between countries (or regions) is used to minimize the price differences between the markets. From a practical point of view this is done through the use of a common day-ahead-market (like Nord Pool Spot for the Nordic countries) or through close cooperation between market places often referred to as "market coupling" or "price coupling".

Forward markets are a complement to the short-term physical markets. The forward markets can be based on physical delivery or financial settlement. The forward markets are essentially free to develop as seems fit, while the short-term physical markets must have a common design to be able to facilitate cross-border trade.

2.2 Modelling the European power market: Apollo and the Diversified Supply Technologies scenario

Sweco's European power market model, Apollo, is used to quantitatively analyse various power market scenarios. The model is fundamental and deterministic. It simulates 38 price regions within Europe and establishes trade both within Europe and with seven regions outside of Europe, which are represented as fixed price regions. It optimises system cost on a weekly and hourly level. New investments in thermal generation are input as long as they are profitable and with perfect foresight. Demand response is considered in all countries as virtual plants that can bid into the market at specific price levels and capacity. The model is described further in Appendix A.

The Sweco Diversified Supply Technologies (DST) scenario is akin to the NEPP Green Policy scenario. It is a scenario in which a high quantity of renewable energy sources is introduced across Europe, and whilst the assumptions made are not exactly the same, the two can be considered as proxies. In these scenarios, it is a strong climate focus that drives developments in the European power markets.

2.3 Developments to 2030 in the Diversified Supply Technologies scenario

The power systems in the Nordics and on the European continent ("the Continent") differ somewhat today, in terms of generation mix, power balances, and market design. Whilst some factors will become more aligned in the coming decades as the markets become more integrated, the balance between supply and demand will develop differently as we move from 2015 towards 2030. As the Nordics increase their capacity of renewable energy, they will continue to export to the Continent, and will generate more and export more. Conversely, on the Continent, as more renewables are introduced and the different power markets become more integrated across Europe, certain countries such as Germany will come to import more and generate less.

Price formation on the energy-only market

In the energy-only market, the energy price is set by the marginal bid into the power market. Different plants and technologies bid a certain amount of capacity for a specific price, which in perfect competition is most likely equal to the variable costs of that plant. The baseload of the generation in a market is naturally made up of those plants with the lowest variable costs, traditionally nuclear and coal, and in hours with higher demand it is more likely to be set by those plants with higher variable costs, such as CCGT and gas turbines.

Hydropower that can be regulated, through the use of reservoirs, is different in nature to thermal plants. Such plants can store water to produce electricity now or in the future, and can help to levelise prices. It is however highly dependent on water inflow. If the inflow is too much it must spill water (which could otherwise be used at a later time). If the inflow however is too little it cannot produce any power. Hydropower plants that have little or no storage capacity, on the other hand, will produce when water flows into their plants.

Technologies such as wind and solar have low, or zero, variable costs and hence enter the market whenever they produce, ahead of the more traditional baseload plants. As more renewables enter the system, at times when these plants have optimal weather conditions at the same time, they can push the traditional baseload plants out of the market. As price levels are reduced, as well as running hours, the profitability of traditional thermal generation will be challenged.

Scarcity prices

Scarcity prices occur in hours when supply is insufficient to meet demand in the power market. In these hours, the price will rise to a price cap in the spot market, which is set at different levels in different markets, currently at a maximum value of 3000 EUR/MWh in Europe. With insufficient energy in the power system to meet demand, it is necessary to introduce curtailment, such as brown outs or rolling black outs. As the quantity of intermittent renewable energy increases in the power system across Europe, there will be more price volatility, with more price spikes – as weather patterns change and are not wholly predictable – and also more zero prices – when the weather is such to enable most intermittent generation to operate at close to full capacity in hours when that generation capacity alone is sufficient to meet demand.

On the Continent, there is generally need to invest in more generation capacity as older capacity is decommissioned and power demand grows. However with the increase in zero prices and reduction in running hours for thermal plants, many thermal generators are becoming less profitable. Hence, in the energy-only market there is unlikely to be sufficient investment in generation to avoid some hours of scarcity prices as we move towards 2030, as seen in the modelling results.

Scarcity prices alone do not indicate under-investment from an economic perspective, as curtailment could be the optimal solution. It is, however, fear of these scarcity prices and the number of hours in which they occur that are a main driver in the discussions of market design changes, such as capacity markets. It is also worth noting that it is in the hours with very high prices that capital-intensive thermal capacity can recuperate their capital costs; hence, measures to reduce scarcity prices will directly affect the profitability of thermal generators.

Demand response

Demand response is one way in which scarcity and very high power prices can be reduced. Rather than focussing on ensuring there is sufficient power generation in the system to meet demand, the demand side can be encouraged to reduce their power usage in certain peak hours so that the available power production can meet the new reduced demand.

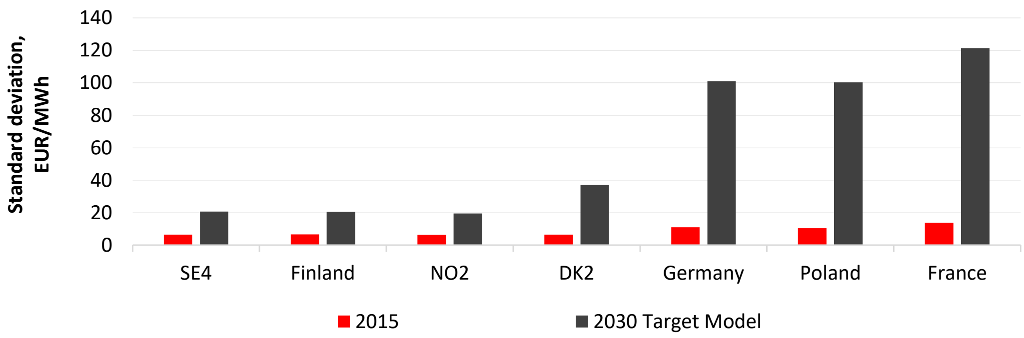
Two main concepts for demand response are firstly a type that would be employed by industry – that is “virtual plants” which can bid to reduce their power consumption by a specific amount in certain hours and receive payment for it – and the second which would typically involve households – that is a kind of demand shifting, in which power demand can be shifted from peak hours to hours with lower prices.

There are challenges and costs associated with the introduction of any type of demand response, but it is an attractive alternative for power systems struggling to encourage investments in new thermal generation capacity.

Price volatility

In 2015 in the Nordics and Baltics, price volatility is expected to be quite small due to hydropower in the Nordic region and pumped storage in Lithuania. Prices in the Nordics are dependent on hydro inflow; in the modelling and results referred to in this section, inflow for a normal hydro year is used, but price volatility could comparatively increase in a dry year. This situation is different to the Continent where the share of hydropower is not so great, and volatility is greater. These effects can be seen in Figure 4.

Figure 4. Price volatility in 2015 and 2030 in the Nordics and Continent



Source: Sweco Energy Markets, 2013

Moving towards 2030, there is a general increase in price volatility across Europe due in part to the increase in intermittent distributed power generation, namely wind and solar power. There is an increase in extreme prices – both scarcity prices and zero prices – and a general increase in price variation, which both lead to the large increase in price volatility in 2030 on the Continent, seen in Figure 4 above.

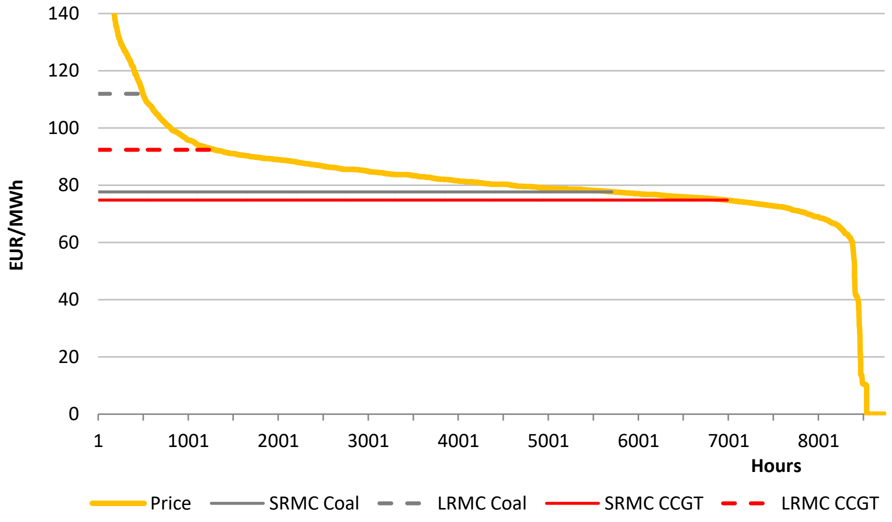
2.4 Incentives for investments in generation

With the increased uncertainty in the power market around changing regulation and large amounts of RES entering the power system, the investment climate on the Continent is changing and questions are arising as to whether the energy-only market is able to ensure sufficient generation capacity to meet demand.

Cost of new capacity compared to electricity prices

In Figure 5, the long-run marginal cost (LRMC) and short-run marginal cost (SRMC) are shown for generation from new condensing coal and CCGT, and compared to the power price in Germany for 2030. In the scenario modelled here, which has a CO_2 price of 68 EUR/MWh, CCGT in fact has lower variable costs, or SRMC, than condensing coal, which is not true for today's power generators.

Figure 5. Price duration curve compared to the SRMC and LRMC of coal and CCGT for Germany in 2030 (truncated at 140 EUR/MWh)



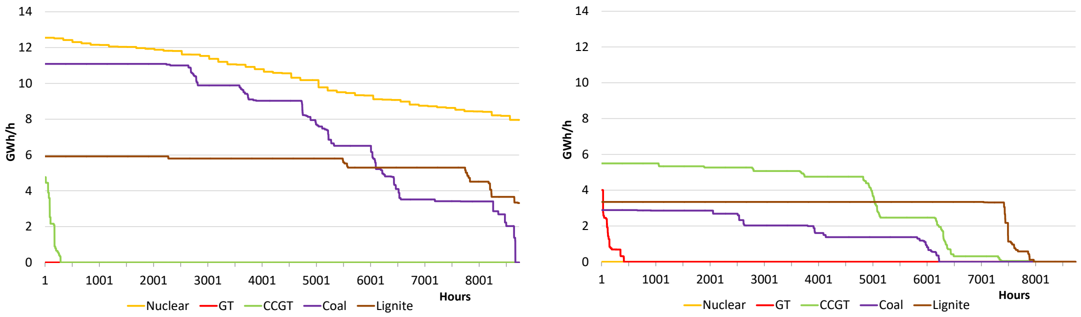
Source: Sweco Energy Markets, 2013

The SRMC represents the variable costs with which power generators are assumed to bid into the power market. When the price is above this value, they can recuperate their capital costs, and when it is below this value, it is preferable to not generate. In the model runnings, this corresponds to around 7000 hours for CCGT, and 5700 hours for coal.

Decrease in running hours of thermal capacity

With the large increase in renewable energy generation capacity, the decommissioning of older plants, and in certain countries the phase-out of nuclear, the generation mix and plant running hours differ substantially between 2015 and 2030. Most striking in the modelling results is the overall reduction in generation from the thermal generators – shown in Figure 6 with Germany as an example – with generation from renewables reducing the amount of thermal generation in the system.

Figure 6. Power generation duration curve for thermal technologies in Germany in 2015 (left) and 2030 (right)



Source: Sweco Energy Markets, 2013

In Figure 6, nuclear, coal and lignite generation have all decreased substantially between 2015 and 2030. By 2030, all nuclear has been phased out in Germany. Older coal and lignite plants have been decommissioned, and the high carbon prices have greatly increased the variable costs of coal and lignite, both factors reducing the total running hours of these technologies. More flexible gas generation enters the system, with a significant increase in generation from CCGT and gas turbines, which are more flexible than coal and lignite in dealing with the intermittent power generation.

Need for new capacity in 2030

Over the coming decades in continental Europe, there will be need for investments in new thermal capacity. Because the increase in RES is expected to be large, thermal capacity will be required to at least back up and complement these intermittent power sources. Certain countries will need new capacity up to 2020, such as the UK, however for the majority, new capacity will be needed in the mid- to later 2020s. Germany, for example, is phasing out all nuclear by the early-2020s, by which time they will need to have brought in sufficient generation capacity to replace it.

Optimal investments

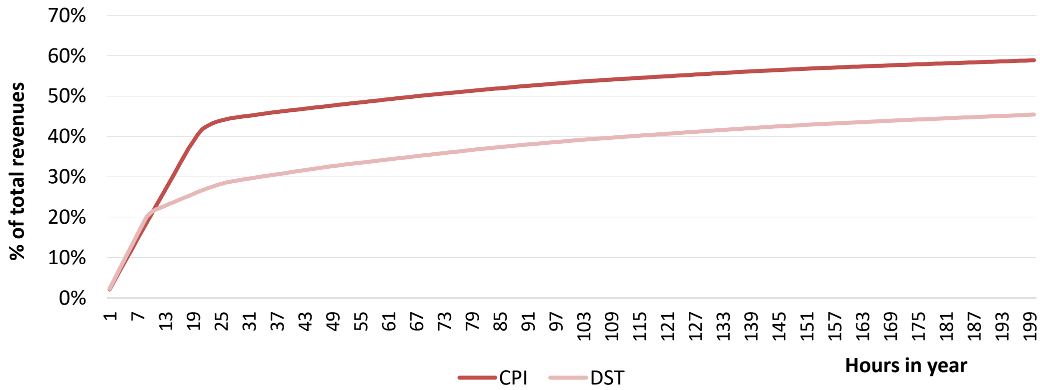
Optimal investments are used when evaluating scenarios in the modelling work – investments are only input if they are profitable, that is if they can cover all of their variable and fixed costs, including cost of capital. This emulates perfect foresight, where revenues in future years are known before investment decisions are made, and is not the setting for the power market in reality.

If investments are more than the optimal level, price levels and the number of hours with very high prices will be lower than in the optimal, and not all generators will be profitable. If on the other hand investments are sub-optimal, there will be an increase in price levels and the number of hours with very high prices. If the social costs of over- or under-investment are also considered in this discussion, with high costs of outages, it is reasonable to expect that the costs of under-investment would be higher than over-investment.

Needed scarcity prices to generate investments

As the number of hours with zero prices increases, scarcity prices become increasingly important for thermal generation to be able to recoup their capital costs. As can be seen in Figure 7, the profitability of new investments can depend on just a small number of hours when there are scarcity or very high prices.

Figure 7. Cumulative share of net revenues for a CCGT plant in Germany in 2030 in two scenarios, one with less renewables – CPI, and one with more – DST (the latter is shown in the other model runnings)



Source: Sweco Energy Markets, 2013

Measures taken to reduce price peaks, or limit the price cap of the power market, all increase the difficulty for capital-intensive thermal generation to be a profitable investment. Also, any investments made will depend even more on the few hours with scarcity, which increases the risk of the investment. An investment based on a small number of very high prices in the year is considerably more risky than an investment based on 2000 hours. Relatively small changes can easily reduce revenues in these hours, decreasing the financial viability of new investments.

3 Capacity markets

In our second scenario for the internal European electricity market, capacity markets have been introduced in some countries, as is happening today. When a capacity market is introduced to a power system, the system is offering two products: one is energy and the other is capacity. Whether a separate (and explicit) capacity payment is necessary in addition to the payment for energy has been discussed for a long time, and the debate is ongoing.

In theory the energy-only market should be able to function by itself, but it requires a very well-functioning market, with, for example, a sufficiently active/flexible demand side and no price cap. Even if this is fulfilled the risk profile may make investments based only on energy prices risky, which may delay investments or lead to investments in less capital-intensive technologies. It also relies on high price spikes to pay for peak load investments. This section discusses some of the theory behind capacity remuneration mechanisms and lends modelling results from Sweco’s Capacity Market Study (2013) to quantitatively add to the discussion.

3.1 What is a capacity remuneration mechanism?

Capacity remuneration mechanisms (CRMs) consist of a payment for capacity, separate to revenue from the energy-only market, regardless of whether the capacity is used or not. They aim to increase available power generation by creating a more stable environment in which to encourage investment in generation, thus reducing the risk of a supply shortage.

Opinions differ on whether, if used, they are a short-term fix until the energy-only market stabilises by itself or other means, or a long-term solution to solve inadequacies with the energy-only design. As described in the Thema Report (February 2014), there are three main types of CRMs: capacity payments, strategic reserves and capacity markets.

A capacity payment is a fixed payment from the market for a certain amount of capacity, with the amount determined by the regulators. In a strategic reserve, specific capacity is kept out of the market and only used in certain stress situations, such as when there are large price spikes in the various markets or a market failure in balancing supply and demand. In a capacity market, a target market capacity is defined and the level of payment determined by the supply and demand of capacity. This is done either by a capacity obligation placed on consumers and load serving entities, by a capacity auction, or by reliability options where certain capacity providers are contracted to pay the difference between the wholesale market price and a pre-set strike price.

Differences between the CRM types and key design features are many, but whichever is chosen, it is essential that these key features be carefully designed to avoid market distortions. Such features include the level of required capacity, a differentiation between technology type, inclusion of interconnectors and demand-side response, and on whom the capacity obligation is imposed.

3.2 Capacity markets in practice

In a capacity market, the required margin is determined by a central party (e.g. a TSO/ISO or a regulator). This margin is typically determined so that a required loss of load probability is upheld. The margins can either be defined for the current situation or be requirements of future margins based on the expected load growth. The latter has become more popular as it facilitates investments in new capacity and increases the competition in the capacity markets, as more parties are able to participate in the market.

The reserve margins are distributed to the load serving entities that are required to meet the requirements either through own resources, bilateral contracting or through a centrally administered capacity market.

Capacity markets typically include all types of capacity, both different types of generation and demand side resources, but the capacity that can be offered as a share of the installed capacity will depend on the technology (the probability that the resource is not available when called upon). It is also possible to only allow certain technologies to take part in the capacity mechanism. If for instance flexible generation/demand is important one could choose to only include technologies that are flexible (e.g. nuclear or intermittent generation would not be allowed in the system). The risk of choosing that path is of course that it may distort investments.

There are several markets globally that have introduced different types of capacity mechanisms, such as most liberalised markets in the United States and South America. In addition, several countries in Europe are introducing or planning to introduce capacity mechanisms, such as the UK, France, Italy, Ireland and Poland. The discussion is also ongoing in e.g. Germany.

3.3 Modelling of energy-only and capacity markets

The rest of this section will discuss the use of a capacity market as a capacity auction, with some reference to the other designs. As mentioned in 2.2, the Sweco European power market model, Apollo was used to model energy-only and capacity markets. This model is further described in Appendix A. The modelling results referred to are based on the Diversified Supply Technologies scenario (mentioned in the previous section of this report) that is very close to the NEPP Green Policy scenario, and the results come from a Sweco Multiclient Study regarding capacity markets (Sweco Energy Markets, 2014).

Four different market designs are considered:

- Target Model – a stylised energy-only market, with no CRMs.

- Integrated Capacity Market – a European-wide capacity market, with target capacity set in relation to national peak demand. Target capacity can be met via domestic or external generation, and external generation is limited by the available transmission capacity.

- Coordinated Policy Scenario (CPS) – Similar in market design to the above Integrated Capacity Market, but limited to the countries which have already or are planning to introduce capacity markets – France, Italy, UK, Spain and Portugal. External generation can still participate in each of these national capacity markets. This scenario also has a sensitivity case, “Case 2”, in which Germany also has a capacity market.

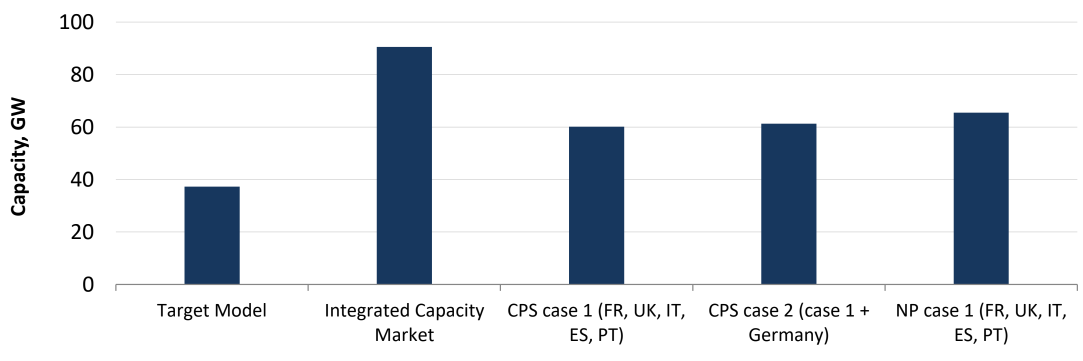
- National Policy (NP) – This scenario is also a “patchwork” design, with the same countries with national capacity markets (France, Italy, UK, Spain and Portugal) as in the CPS, but in these markets external generation cannot participate. There is also a short-term trade uplift placed on exports from the capacity market countries during stress times; although not in line with current thinking, it did represent a concept suggested in the design of the UK capacity market.

3.4 Investments in generation

The introduction of a CRM would be expected to increase the amount of generation capacity installed compared to an energy-only market; with a payment in addition to energy market revenues, there is a more stable environment to make investments in generation and flexible load – returns should appear more secure and the cost of capital should be lower.

The design of the capacity market, the number of countries who introduce them and the degree of homogeneity between them will affect the quantity and location of the investments that will be seen in Europe. As can be seen in Figure 8, most capacity is introduced when capacity markets are in all countries in Europe – the Integrated Capacity Market, and in the “patchwork” scenarios the levels of capacity in Europe are somewhere in between the Target Model and Integrated Capacity Market.

Figure 8. Total new capacity installed in Sweco scenarios by 2030

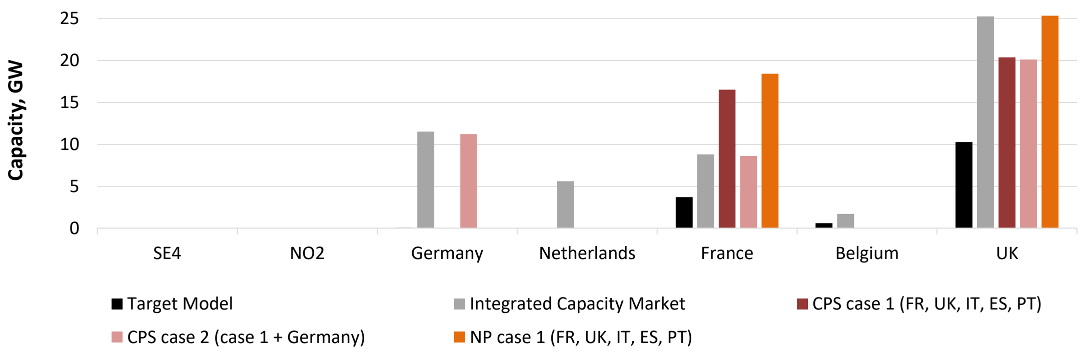


Source: Sweco Energy Markets, 2013

In addition to the total increase in capacity installed, the different capacity market scenarios show a distortion between investments in different countries, see Figure 9. Investments in CRM regions can crowd out investments in neighbouring non-CRM regions. In the deterministic model setting, investments are optimal, that is they are input as long as they appear to cover all costs including the cost of capital.

Naturally, in reality investments are unlikely to be optimal, perhaps with under-investment most probable in the Target Model when there is no second revenue for generators from the capacity market. As noted elsewhere, one of the main risks of distortion with the introduction of CRMs is the risk of encouraging an over-investment in generation capacity.

Figure 9. Total new generation investments in individual countries by 2030



Source: Sweco Energy Markets, 2013

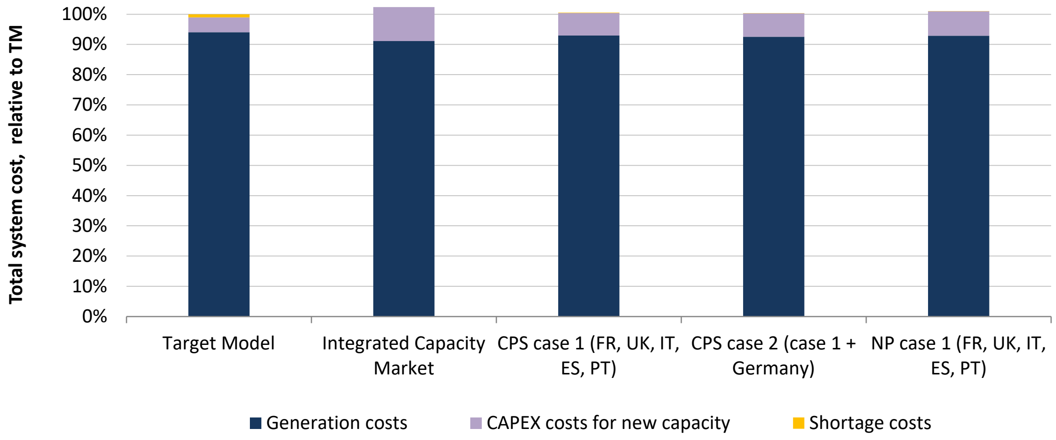
Although a CRM should provide an additional revenue stream for an investor in generation capacity, there will still exist policy uncertainty with the introduction of a CRM; investors must trust in the credibility and longevity of the CRM design if they are to invest now on the basis of receiving CRM revenues in the future.

3.5 System costs

At the basis of the modelling work is the assumption that the different market designs are implemented efficiently. If this is valid, the different market designs have limited impact on the total system cost in Europe, which includes the costs for generation, capital for new thermal capacity (exogenously input into the system after 2012) and shortages. Efficient implementation would result in this in theory; the increased volume of new capacity increases capital costs, but on the other hand reduces variable costs of production, as more technologies with lower variable costs are available.

As can be seen in Figure 10, the system cost is slightly higher when capacity markets are present. The integrated capacity market shows the largest increase in total system cost, approximately 2%, with increased capital costs outweighing reduced generation costs and reduced cost of shortage. The deterministic model setting could, however, favour the Target Model in this result, as in reality the increased costs could be offset by a larger reduction in the risk of electricity shortages than captured here.

Figure 10. Europe system cost in 2030, relative to the Target Model



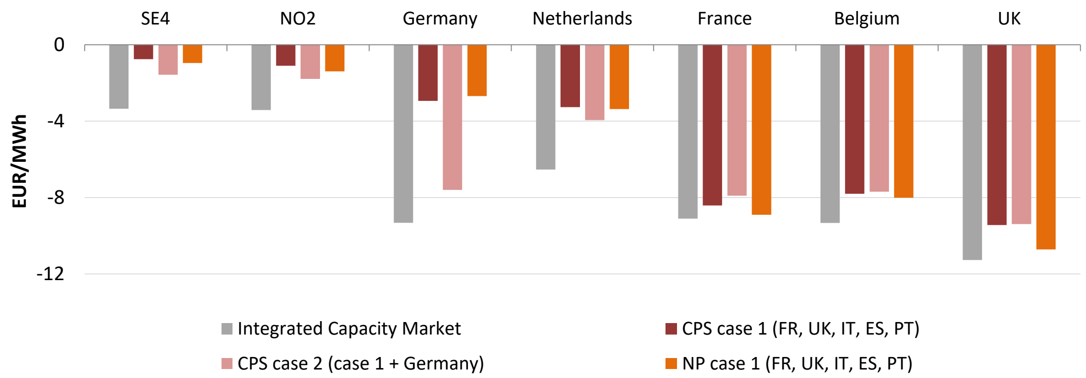
Source: Sweco Energy Markets, 2013

Efficient implementation of a CRM is naturally difficult, and if capacity markets are introduced in several countries, there will likely be differences at a national level that could hinder trade and efficiency. There exists the risk for market failures and regulatory failures in all scenarios; which presents the bigger risk is a different discussion.

3.6 Price effects

As capacity markets encourage more investment in generation, the wholesale power price decreases when they are introduced. The decrease is seen not only in the countries with capacity markets, but all over Europe in all capacity market scenarios; in some cases the spillover effect is substantial, as can be seen in Figure 11.

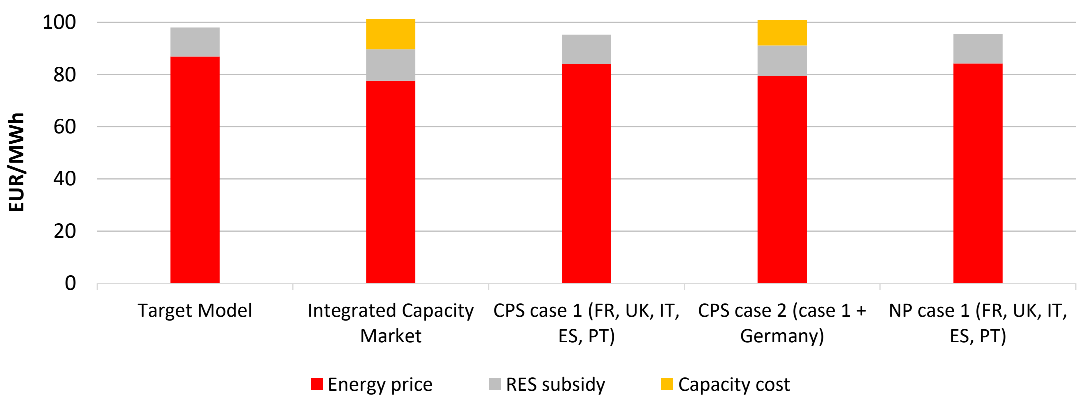
Figure 11. Change in wholesale price relative to the Target Model in 2030



Source: Sweco Energy Markets, 2013

To then consider the impact to consumers, the total customer cost is considered to consist of the combined payments for electricity, renewables subsidies and any capacity payment (excluding grid costs); these cost components are shown in Figure 12, with Germany as an example. Here the bulk of the cost is made up of the wholesale energy price, which decreases in the capacity market scenarios; there is a component to pay for RES subsidies; and in the two scenarios that Germany has a capacity market, there is a component that covers the cost of capacity that clears in the German capacity market.

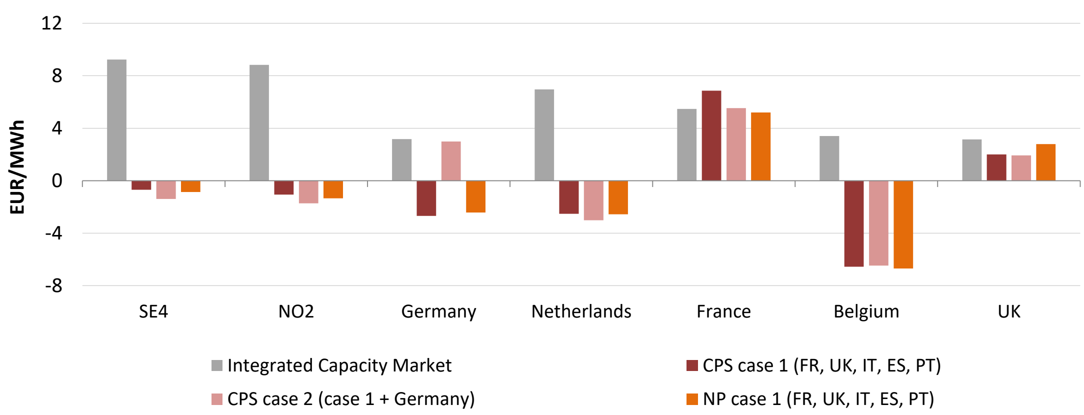
Figure 12. Components of cost to customers in Germany in 2030 in different scenarios



Source: Sweco Energy Markets, 2013

Although wholesale prices are reduced, the cost of paying for the additional capacity more than offsets this price reduction, and in those countries with capacity markets, the customers generally pay more; this can be seen in Figure 13. For those that neighbour the capacity market countries, there can be several spillover effects which these figures show, namely, a decrease in wholesale prices, a decrease in customer cost (without to pay for additional capacity that causes the wholesale price decrease) and a crowding out of investments in that region.

Figure 13. Change in customer cost relative to the Target Model in 2030



Source: Sweco Energy Markets, 2013

An integrated capacity market is generally the most expensive alternative for customers, particularly in the Nordics. Here, the drop in wholesale price is relatively small, but the cost of capacity is much greater than the drop in wholesale price; hence there is a larger customer cost increase here than in other parts of Europe.

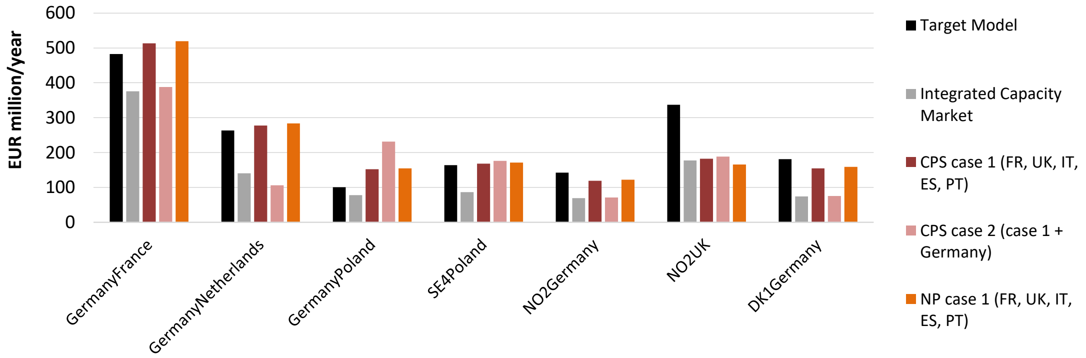
3.7 Impact on interconnectors

A CRM scheme could significantly risk distorting incentives between building interconnector and generation capacity compared to the Target Model. This is brought forward as a key issue in capacity market design in both the Sweco (2014), Elforsk/Thema (2014) and Frontier Economics (2014) reports. Figure 14 shows how congestion revenues for interconnectors could be impacted when capacity markets are introduced. For all interconnectors in the Integrated Capacity Market, these revenues drop as price volatility and price spread between areas are reduced when more generation capacity is installed with capacity markets. In the modelling results presented here, this occurs for all interconnectors when capacity markets are introduced into all of Europe, both in the Integrated Capacity Market and in many of the “patchwork” capacity market designs. Although interconnectors are generally regulated and congestion revenues perhaps do not represent actual profits, it exemplifies the risk posed to interconnector investments if they cannot participate in the capacity market.

Investments in interconnectors are considered in two steps – first the consideration of the addition to national security of supply that the interconnector could provide, and then the possible remuneration from the investment. Lower potential revenues lower the incentive to invest in the interconnector. This can result in more expensive local generation being used instead of increasing interconnector capacity to allow less expensive generation in a neighbouring region to participate, heightening costs to consumers.

Given this risk, the participation of interconnectors and external generation in capacity markets is a key market design question, and an income from the capacity market could perhaps offset the reduction in revenue. Inclusion of these is, however, far from simple and there are several ways this could happen; several potential options are described in some depth in the Elforsk/Thema (14:28, 2014) and Frontier Economics (2014) reports. If included, interconnectors must also be derated according to the addition to generation adequacy that they provide to the capacity market region.

Figure 14. Congestion revenues for selected interconnectors in 2030



Source: Sweco Energy Markets, 2013

The Elforsk/Thema report details three types of models proposed by Frontier Economics and one model proposed by Eurelectric for enabling cross-border generation and interconnector participation in a CRM in a neighbouring region:

- Interconnector models – in these, it is the interconnector that directly participates in the capacity market. The interconnector may or may not bear responsibility for actual delivery in stress periods.

- Generator models – in these, it is the external generator that participates directly in the capacity market. As in the interconnector models, this generator may or may not be responsible for delivering in stress periods.
  Note: In the modelling results discussed in this section it is a generator model that has been used, with no responsibility placed on the external generation for delivery.

- Implicit participation through "corrective" payments. In this alternative, neither generators nor interconnectors participate in the market, but an amount is calculated ex ante or ex post and paid to interconnectors and interconnected generation to compensate them for market distortions caused by the CRM.

- Hybrid generator–interconnector model. In this model, external generators participate in the capacity market but in a separate auction to domestic capacity. The derated interconnector capacity limits the amount of capacity purchased here, and the marginal bid sets the price that generators receive. In order to avoid under-investment into interconnectors, if the price in the domestic capacity market is greater than the price in the separate external market, then the difference in capacity payment should go to the interconnector as a type of "scarcity payment".

The biggest point of discussion on this theoretical level, is around who bears the responsibility for non-delivery. For generator models, interconnected generation needs the interconnector capacity to be available and flowing in the correct direction to be able to deliver to the CRM-market. And for interconnector models, interconnectors need there to be sufficient generation in the external market to be able to deliver to the CRM-market, which is a particular issue when the interconnector is a merchant line as they do not have any control over this. If, however, neither the external generator nor the interconnector bear the responsibility for non-delivery, then there is no incentive for these bodies to deliver at stress times in the CRM-market. Thema and Frontier reach different conclusions but both suggest responsibility should be on the interconnector or generator.

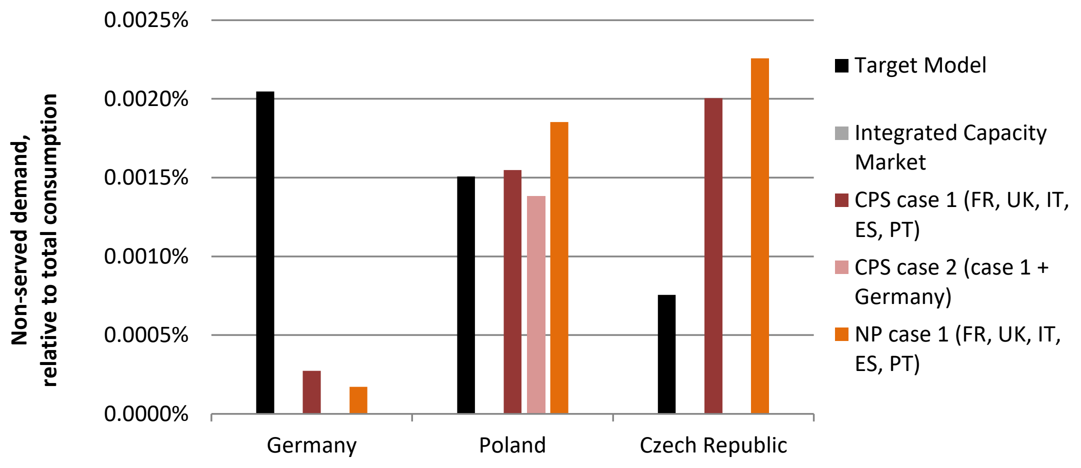
3.8 Generation adequacy

Central to the discussions prompting and involving CRMs is the issue of generation adequacy; by introducing a capacity market, a country is hoping to increase this. This can, however, have positive or negative impact on the security of supply of neighbouring countries.

Within the Target Model, the modelling results show that there will be insufficient capacity to avoid some hours of physical shortage. This result would be expected, but given the deterministic nature of the model and the optimal level of investments input, the number of hours of shortage could be more in reality. Whilst curtailment in itself can be the optimal solution in a given system, the fear of too many hours of physical shortage is what lies at the heart of this discussion.

In the modelling of the capacity market scenarios, there are no hours of physical shortage in the countries that have capacity markets; this was a part of the modelling process, and as discussed earlier the reality could be different. As seen before, countries that neighbour a capacity market without having one themselves, are able to benefit in terms of investment levels and system cost. There is, however, also the risk that they experience a lower security of supply; this can be seen in Figure 15 for both Poland and the Czech Republic, their non-served demand increases when other countries have introduced capacity markets.

Figure 15. Non-served demand as a percentage of consumption in 2030



Source: Sweco Energy Markets, 2013

3.9 Further considerations for capacity mechanisms

With the likely failings of the energy-only market in some countries, a capacity mechanism could be beneficial. A quantity-based mechanism will ensure a certain reserve margin, which is likely to lead to increased generation adequacy. Investment risks could be reduced, leading to reduced capital return requirements and thus lowering the cost of investments. In addition it could allow for more capital-intensive technologies. With increased capacity margins, price spikes are likely to be reduced and the possibilities to exert market power in shortage situations will diminish. Experience from markets with capacity mechanisms also indicate that they are fairly successful in attracting demand side resources, such as in the PJM market, but the design of the mechanism is crucial.

On the negative side, capacity mechanisms typically require quite detailed regulation. It is naturally necessary to define reserve margins (or the capacity payment if it is a price-based mechanism) and it is necessary to define what technologies can be included and in what way. It can be debated whether it is necessary to control the actual physical backing of the resources or if a financial backing is sufficient, but most (if not all) markets that have such mechanisms have opted for a physical backing, that needs to be controlled. In a quantity-based mechanism the prices can become very volatile, which was the case in PJM before the introduction of the existing mechanism. Support to investment is also facilitated if the mechanism is forward looking, which requires the regulator or TSO to have load forecasts for different load serving entities. There is thus a risk of “micro management” of the sector. The outcome is also highly dependent on the details of the design, which implies a high risk of regulatory failure.

4 Locational Marginal Pricing

In our third market design scenario we look at congestion management. Europe's Target Model to integrate national markets and manage transmission congestion in the day-ahead and intraday timeframes is based on the introduction of flow-based market coupling to clear, in a coordinated manner, trades between "properly defined" bidding zones. This zonal approach is sometimes criticised for not promoting efficient dispatch and transmission usage, and for lacking efficient signals for investment in both generation and transmission in particular when bidding zones are not properly defined. In this chapter, we discuss some pros and cons of adopting a nodal approach to congestion management and energy pricing. For the sake of completeness we begin this chapter with some background on Europe's Target Model for congestion management.

4.1 The Target Model: a zonal approach to capacity allocation and congestion management

A bidding zone is an aggregation of injection and exit nodes, defined in the Network Code on Capacity Allocation and Congestion Management (NC CACM) as *the largest geographical area within which market participants may trade without transmission capacity allocation*. This means that congestion *within* bidding zones is treated separately from congestion *between* bidding zones:

Congestion between bidding zones: flow-based market coupling

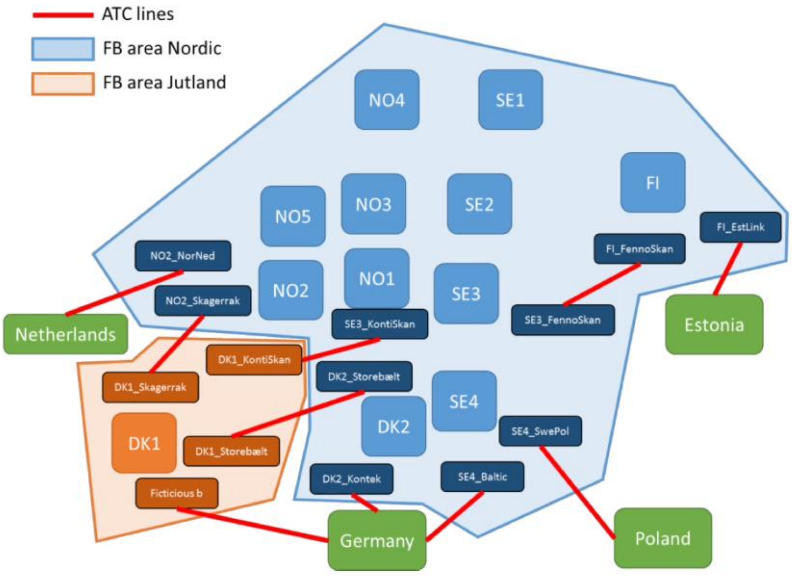
Because an efficient use of interconnectors is vital for the success of the internal market, the NC CACM is introducing significant changes to the way transmission capacities for day-ahead trading *between* bidding zones are determined and allocated. At present, system operators decide in advance, border-by-border, how much transmission capacity should be made available for trading between bidding zones. These capacities, along with market participants' bids and offers, are the only input to the power exchanges' day-ahead market clearing algorithms: the actual, unscheduled physical flows that the cleared trades may give rise to in real time[1] - known as *transit flows* - are not considered, and have to be dealt with by system operators. System operators therefore allocate capacity conservatively, reducing the amount of possible commercial trades between bidding zones.

The NC CACM foresees the introduction of flow-based methods as the preferred approach to determine and allocate cross-border transmission capacities between bidding zones. Central to flow-based methods is the inclusion, in the market coupling algorithm, of a sophisticated grid model that roughly represents the underlying network (*see Appendix B From nodal to zonal*). Information about how changes in generation and consumption in individual nodes affects flows across interconnectors provided by system operators is used by the market-coupling algorithm to predict the physical flows resulting from all cross-zonal commercial bids and offers, allowing the algorithm to prioritize those flows that resolve

[1] Because electricity flows over the path of least resistance, a trade between two nodes will result in an actual physical flow that will be different from the contracted flow: power will move in parallel along all paths between the injection and the exit node, impacting the feasibility of any other transaction and conditioning access to the transmission network.

congestions in the most economically efficient way, for the whole region. As a result, transit flows are dealt with in the algorithm[2]: lines are used more efficiently, less congestion is expected in real time, and more cross-zonal capacity can be allocated for day-ahead trading. Also, capacity will be allocated to those trades that maximize *global* social welfare, across the entire market-coupled region, as the objective of the internal market is to increase pan-European social welfare. This can lead to controversial distributional effects, and unintuitive results. The flow-based setup of the Nordic countries is shown in **Fel! Hittar inte referenskälla.**.

Figure 16. Flow-based setup of the Nordic countries used for market simulations in EUPHEMIA



Source: Nordic TSOs

Congestion within bidding zones: re-dispatching and countertrading

As mentioned before, commercial transactions within bidding zones are not limited. TSOs resolve internal congestions by resorting to remedial actions like re-dispatching and countertrading, and socialise the costs within the bidding zone. In general, this is not considered optimal from a transparency perspective, as re-dispatching costs are not always made public. Also, resources for re-dispatching and countertrading are procured without the use of market-based methods and within the bidding zone, even though cheaper resources could be available in neighbouring zones.

How bidding zones are defined is therefore very significant. A strong internal network is a prerequisite for efficient zonal markets, not only to reduce remedial actions but also because unplanned physical flows resulting from commercial transactions within bidding zones with considerable internal congestion will reduce the amount of cross-zonal capacity. The volume

[2] Transit flows, however, are not expected to be reduced equally across all borders, and may in some instances increase once flow-based market coupling is implemented across Europe. See ENTSO-E Technical Report, Bidding zones review process, January 2014

of unplanned (or unscheduled) flows - known as *internal flows* when they are confined to the zone in which they originate, and as *loop flows* when they spread into adjacent bidding zones - will depend on whether bidding zone borders match the underlying physical network to reflect transmission constraints.

At present in Europe bidding zones are defined along national borders, with a few exceptions: Germany, Austria and Luxemburg constitute a single bidding zone, while Denmark, Italy, Norway and Sweden have been divided into two or more bidding zones to reflect internal congestion. Sweden was divided into bidding zones following a formal complaint from Danish energy companies to DG Competition that the Swedish TSO was restricting the competition in the internal market by “moving congestions to the border”, i.e. curtailing interconnector capacity to Denmark to lower its countertrading costs and keep wholesale prices in Sweden down, at the expense of consumers in eastern Denmark.

The Swedish example shows that in general, countries in Europe have been reluctant to redefine bidding zones to reflect the underlying network. Arguments for large, single-country bidding zones include less complexity, increased liquidity, and a desire to let generators and consumption face a single, national price. In Italy, demand faces a single national price. Investment in transmission capacity is seen as a better alternative to bidding zone reconfiguration given the magnitude of the consequences for all stakeholders of bidding zone re-configurations, and the uncertainties regarding the possible benefits of introducing smaller zones.

Because central Europe is highly meshed, however, the existence of large bidding zones - in particular the very large Germany/Austria/Luxemburg bidding zone with substantial generation-demand imbalances and limited transmission capacity - is nevertheless controversial. Calls for splitting the German-Austrian bidding zone are frequent, with Poland and the Czech Republic having long complained that loop flows caused by internal transactions in the Germany/Austria/Luxemburg bidding zone have a discriminatory impact on market participants in Poland and the Czech Republic.

Because bidding zones should evolve over time to reflect changing market conditions, the NC CACM foresees a periodic bidding zone review, to potentially modify how bidding zones are defined. There are however, fears that this process will not achieve optimal results given national interests and the fact that there is no simple and objective definition of what an optimal bidding zone configuration is, in particular given the impact of increasingly large shares of less predictable generation on power flows.

Critics of zonal congestion management also point out that regardless of how bidding zones are defined some market participants always get subsidized at the expense of others, and that re-defining bidding zones may is merely an exercise in re-defining who gets subsidised. These critics would often like to see Europe moving on to Locational Marginal Pricing (LMP), often also referred to as nodal pricing.

4.2 Locational Marginal Pricing (LMP): a nodal approach to capacity allocation and congestion management

Nodal LMP has been implemented in many electricity markets around the world, including liberalised electricity markets in the United States. LMP was the mechanism proposed by the Federal Energy Regulatory Commission (FERC) to manage congestion in its 2002 attempt to establish a common standard market design for all electricity markets across the United States[3]. Also in the discussions leading up to the Target Model, the European Regulators Group for Electricity and Gas (ERGEG) favoured the nodal approach, describing it as "*the ultimate goal and (technically and economically) optimal solution*[4]". Zonal pricing was, however, more in line with the philosophy of the decentralised market design most European countries had adopted following liberalisation, when facilitating competition was the main challenge.

Managing congestion by the node

Central to LMP markets is the use of a full network model to simultaneously establish dispatch volumes and prices at each generator injection node and demand exit node on the transmission network, taking into account not only market participant's bids and offers (and scheduled bilateral transactions) but also the effect of the resulting power flows.

Prices in LMP markets have three components:

- A *system energy price*, which representing optimal dispatch ignoring transmission congestion and losses. It is determined by the marginal generator and is equal for all nodes.
- A *transmission congestion cost*, which will vary by location in the presence of congestion: prices in nodes with a generation surplus will fall, and prices in nodes with a generation deficit will rise, reflecting the increased cost to deliver lower cost energy to locations behind constraints.
- A *cost for losses*, which will vary by location and reflects the marginal losses of transmitting electricity from one node to other nodes; each generator will see a price for losses that exactly reflects the incremental cost of transmission arising from its contribution to power flows. Generally, transmission losses increase as power is transferred over longer distances, at higher volumes, and over lower-voltage lines. In markets in which generation is located far away from demand, transmission losses and thus price differences between nodes can be significant even without congestion.

Under full nodal pricing, generators are paid the price at the location (node) where they inject, and loads pay the price at the location (node) where they withdraw electricity. This

[3] In 2002, the FERC issued a Notice of Proposed Rulemaking (NOPR) setting out a proposal for a standard wholesale electricity market design to be implemented across the United States. The NORP was later withdrawn as several entities were against a national one size fits all approach, and because many of the proposed changed were already being implemented on a voluntary basis.
[4] impact assessment of the Framework Guideline on Capacity Allocation and Congestion Management

means that both generators and load face the actual cost of using the transmission network, and the costs for managing grid constraints are no longer socialized.

In LMP markets in the United States, Independent System Operators (ISOs) calculate prices day-ahead and at 5-minute intervals (real time). Because of this, LMP is sometimes considered LMP is often considered difficult and expensive to implement. Computing nodal prices at thousands of nodes in real time requires detailed and up-to-date models of loads, power plants, and transmission and distribution grids. Advocates of LMP argue that the information needed to compute nodal prices is the same information that system operators already use to plan the operation of the system, so no new data collection is required. With clever algorithms and powerful computers, the computations become feasible.

Short- and long term benefits of LMP

Because nodal prices reflect both the temporal and the locational value of electricity, it is often argued that nodal pricing encourages, in the short term, a more efficient dispatch of generators than zonal pricing, and a more efficient use of the existing network. However, in the absence of demand response and with inelastic demand, nodal prices are generally not considered particularly effective at providing efficient signals for consumption. This is because most consumers, with the exception of large customers, do not see the nodal prices directly. In several nodal markets, only generation faces nodal prices, with demand facing zonal prices, usually the load-weighted average of the prices at exit nodes within the zone.

In the longer term, nodal pricing is said to deliver more efficient signals for the location and timing of investment in generation, demand response and/or transmission. These longer-run benefits are, however, not as evident as short-term benefits. Concerning investment decisions, it is doubtful that nodal price differences by themselves are sufficient to have a significant impact on where to invest, as potential investors may regard transmission constraints as temporary and likely to be relieved by transmission investment in the future. In such cases, only the impact of losses on nodal prices will be taken into account, bringing – if possible – new generation closer to demand. Experience from markets with nodal pricing seems to suggest that forecasts of nodal price differentials do not play a major role on investment decisions, and that investment decisions are mostly driven by other factors such as permits, renewable subsidies, capital costs, expected operating costs and fuel availability.

It should be noted that many nodal markets have resorted to special arrangements to encourage investment in peak generation. For instance, nodal markets in the United States have introduced locational CRMs to sustain a reliable level of investment. There is a locational aspect to the capacity auction: capacity that is offered at nodes that have traditionally experienced deficits are valued higher than capacity offered at other nodes. New Zealand, Australia and Singapore have introduced scarcity pricing arrangements.

4.3 Congestion risk and congestion-risk liquidity

Because congestion can result in unexpected and significant differences in prices at different locations, congestion-related price risk is a primary concern in both zonal and nodal markets. In nodal markets, market participants can hedge this risk by purchasing Financial

Transmission Rights (FTR) that offset the effects of nodal price differences. Appendix B describes how congestion risk is managed in Europe and how FTRs (obligations and options) operate.

A major argument in Europe against LMP (and smaller bidding zones) is that nodal markets (and smaller bidding zones) are inherently less liquid than zonal markets, particularly if bidding zones are large with many participants and various generation types. The Nordic market is usually cited as an example of lack of liquidity following the introduction of smaller bidding zones in Sweden. In the Nordic market, market players first reduce their exposure to price volatility in the region as a whole by entering into forward agreements that pay the difference between the Nordic system price and an agreed strike price, and, on a second step, purchase Contracts for Difference (CfDs) that pay the difference between the system price and the zonal price. Theoretically, these two contracts should provide a full hedge against congestion risk. CfD contracts written on the difference between the system prices and prices in SE4 (southern Sweden, a predominantly import area) are, however, expensive and illiquid. This is partly explained by the fact that because CfDs are issued by generators and not TSOs (as FTRs in the United States are), they are therefore not linked to the physical capacity of the network, so a lack of CfDs can be expected in import areas especially if CfDs are predominantly offered by a dominant generator, as is the case in SE4.

LMP, with a potentially very large number of nodes is often regarded in Europe as too complex to support liquidity in long-term contracting. In the United States, nodal complexity is simplified by aggregating nodes into trading hubs. Hub prices are calculated by averaging the individual nodal prices at the nodes that make up the hub. The general view in the United States appears to be that it is not market design but efficient use of transmission, and diverse ownership of generation and retailers that drives liquidity.

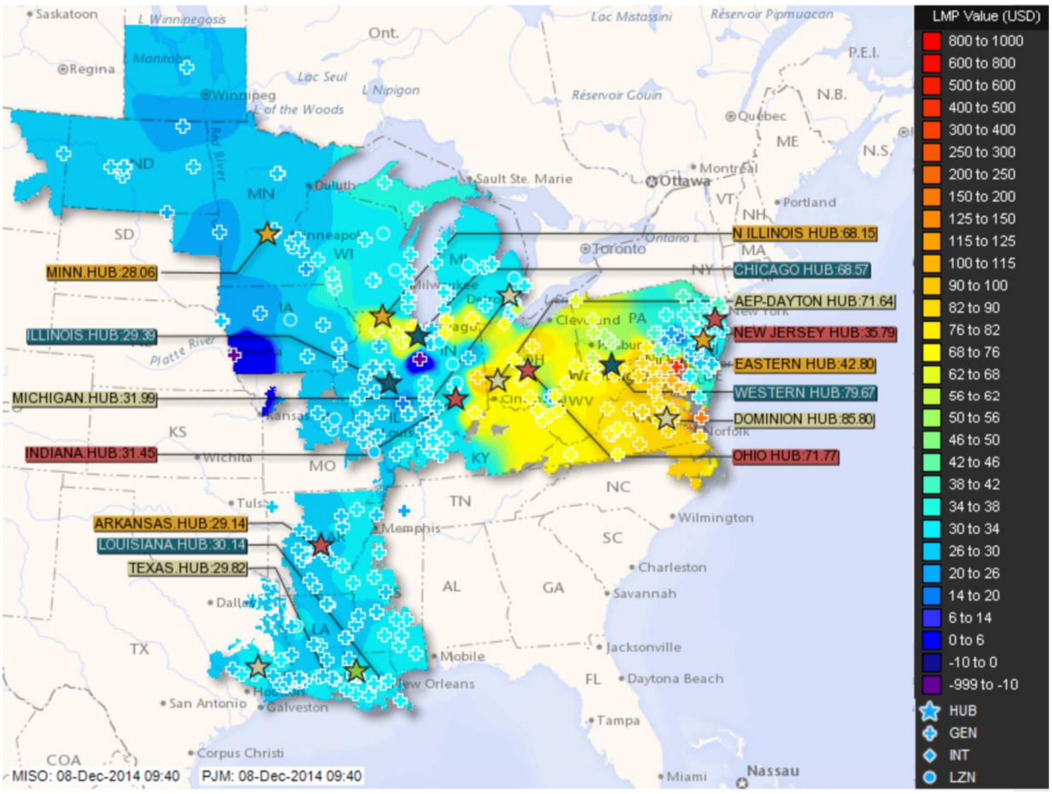
It should also be noted that in most LMP markets in the United States, wholesale demand is settled at a zonal price[5]. Reasons for having zonal prices for demand include a desire to have the same price over a certain region or state and to remove some decrease the exposure of independent retailers to congestion price volatility. Zonal prices are calculated as the load-weighted average price of all load nodes in the zone. **Fel! Hittar inte referenskälla.** Figure 14shows the LMP contour map of the PJM/MISO Joint and Common Market (JCM), showing selected commercial pricing nodes and their respective LMP values.

System operator’s FTR auctions are overwhelmingly regarded as very competitive. The resulting FTR contracts may either be traded at system operator’s FTR markets (e.g. reconfiguration auctions) or may be converted, after initial positions have been established, into equivalent financial contracts. In general, FTRs are fairly standard, facilitating liquidity in secondary markets. FTR equivalents are traded in different clearinghouses like ICE Futures, the NY Mercantile Exchange and the Nodal Exchange. The Nodal Exchange has the largest market share and offers locational futures contracts at over 1800 commercially significant hubs, zones and nodes across all six nodal markets. Contracts are financially settled using the monthly average of the relevant hourly LMPs.

[5] Exceptions with full nodal pricing include New Zealand, see Appendix D

In general, nodes have a lower price risk but a higher liquidity risk, while hubs have a high price risk but a lower liquidity risk. Zones have low liquidity risk and - in general – a lower price risk than hubs. Market participants do not always hedge at the node, as the hedge price for hub or zonal hedging may be more attractive. Currently, the Nodal Exchange does not offer contracts on individual load nodes, which often lack liquidity.

Figure 17. PJM/MISO Joint & Common Market (JCM) - LMP contour map of the JCM showing selected commercial pricing nodes, with their respective LMP values.[6]



Source: http://www.jointandcommon.com

Finally, it should be noted that large bidding zones, although inherently more liquid are likely to have higher countertrading and re-dispatching costs than smaller bidding zones or nodal markets, and that this is likely to increase with larger volumes of wind generation.

Market power

LMP has been criticized for being vulnerable to market power. In some cases such criticism is based on empirical studies of trading patterns in LMP markets and claims of possible market power are not backed up by an underlying theory. In other cases, a theoretical argument is put forward that LMP segments large markets into smaller markets and therefore gives market power to dominant generators in deficit nodes. Here the counter-

[6] Prices are updated every five minutes. Contour maps are used to visualise LMPs and transmission constraints in real time. Significant colour changes mark changes in LMPs across a geographic region and can usually be associated to transmission constraints from one location to another.

argument is that market power is not an effect of LMP per se, but a consequence of real grid constraints. LMP only serves to bring this market power out into the open. The only thing a non-LMP system achieves is to spread the pain of market power among a larger set of consumers. This may be positive for consumers that are close to the generators with market power, as they no longer have to face these generators alone, but it also hides the market power to some extent. When the pain is shared, the incentive to respond to market power by reducing demand is reduced. The incentive to enter into long-term contracts with new entrants into the generating market is also reduced.

Even though it is unclear to which degree LMP is vulnerable to market power or not, the mere possibility of market power has led to the introduction of price caps in highly constrained areas in some LMP markets, notably in the US.

4.4 Is there a case for nodal pricing in Europe?

Most critics agree that probably not, if the internal transmission network of the largest bidding zones are reinforced. The United States is often said to have been driven to nodal pricing by weak transmission networks with significant congestion, and by gaming. LMP is, however, regarded as best practice, and experience from the United States shows it can work. But it is unlikely that Europe would become a single nodal market, and in adjacent nodal markets problems with loop flows would remain.

5 Increased central planning

In the fourth and final scenario, governments return to central planning in order to reach the goals they have set out for the electricity sector. In this scenario, governments have given up on the ability of market forces to deliver the volume of investment in new generation and transmission capacity needed to achieve a low carbon, secure, and affordable energy system, and have conceded that the transformation must be driven by authorities.

While governments are not likely to turn on their heads and undo the privatization of the electricity industry in this scenario, governments have nevertheless accepted that introducing competition has been more complicated than expected, and that the decarbonisation objective they wish to pursue is a public policy objective that requires not only non-market viable investments, but also significant investment in plant to back up non-market viable, non-manageable generation. Governments have finally conceded that relying on commercial decision-making serves no useful purpose to the climate change agenda and merely increases risk, and therefore the costs to be borne by the public.

Governments will therefore resort, under this scenario, to far-reaching measures to steer the electricity sector in the direction they want. The old premise of deregulation, that market forces should be allowed to guide investments, is abandoned in favour of a system where investment is driven by central planning and subsidies. While there will still be commercial actors, their behaviour will be largely controlled by governments, resembling times past when governments instructed their utilities to install a preferred portfolio mix.

At the same time, governments will want to ensure that the rules that are laid down will be in force for long periods of time, so that investors that are contemplating long term investments in the electricity sector can assume that the conditions are true when an investment decision is made will continue to remain true for as long as is required to make the investment profitable. Changes to regulation will always be accompanied by grandfathering clauses to protect existing investments.

The overall goal is to create an environment where the risks associated with investments in generation and transmission are reduced to a point where investors will feel comfortable to make the investments that will ensure security of supply and a decarbonized electricity sector. Another goal is that the transformation should be carried out in a manner so that costs to consumers are kept to a minimum.

5.1 Examples of increased central planning in Europe today

Already we are seeing clear attempts at market interference, for different reasons. In France, where nuclear generation enjoys a very high scarcity rent as a result of interconnections to Germany, regulated electricity tariffs that pass on the cost-advantage of France’s ambitious nuclear program to consumers are still widespread, and have impeded the development of competitive retail markets. To remedy this situation, the French government has passed a law to reform the electricity market. The Law regulates the price of a significant volume of France’s nuclear output by granting EDF’s rivals supplying end-consumers the right to buy electricity generated by EDF’s nuclear power plants at a regulated tariff set by the Government. In addition, the Law hints that future costs to replace the nuclear fleet, i.e. to build new plant, will be covered by a regulated item included in end-consumers electricity tariffs. It can therefore be argued that only investment in peak generation will be left to market forces. The Law also opens the door for the establishment of a capacity market.

Also Great Britain, who pioneered electricity market liberalisation, is intervening to de-risk investment by transferring some investment risk to electricity customers. The British government has concluded that current electricity market arrangements do not provide the long-term market certainty for the large volumes of capital-intensive low carbon generation sources, including nuclear power, needed to decarbonize Britain, ensure security of supply, and make sure that costs to consumers are kept reasonably low. Unlike the Nordic countries, the Britain’s share of renewables is low, and demand is expected to increase as the heating sector switches from natural gas to electricity.

The British Government has therefore decided on a set of reforms. The key elements are long-term contracts in the form of technology differentiated feed-in tariffs, long-term price signals implemented though a carbon floor price, a capacity mechanism to ensure an adequate security of supply, and stricter environmental legislation to make sure that new carbon-based power plants without CCS (Carbon Capture and Storage) are not built.

The argument for a carbon floor price is that the carbon price of the EU ETS is too low, too volatile, and lacks longer-term credibility. The British government does not believe that an EU-wide tightening of the EU ETS will come about quickly enough, so it has decided to use taxation to bring the price of carbon in the Britain to what it describes as “sensible” levels - required to make low-carbon alternatives more attractive compared to the carbon-based alternatives. In addition, since the wholesale electricity price is set by coal or gas plus the cost of carbon, the effect will be to drive up the wholesale price, so investment in nuclear can happen without easily seen and possibly controversial subsidies.

As the French Government, the British Government is also planning to help nuclear energy into the market. A higher wholesale price and carbon taxation is not long-term enough - taxes that currently make an investment look attractive can be rolled back at any time, thus destroying the conditions on which the investment depends for its profitability. Instead, the government intends to offer long-term, legally enforceable contracts to new low carbon generation projects. The intent is that such long-term contracts that promise regular revenue streams over long time periods will make investments in low carbon alternatives attractive.

The long-term contracts being proposed by the UK Government are feed-in-tariffs based on Contracts for Difference. If the wholesale electricity price is below the price agreed in the contract, the generator will receive a top-up payment to make up the difference. If the wholesale price is above the contract price, the generator pays the surplus back. Because feed-in-tariffs will be technology differentiated, it can be argued that market decision-making is being replaced by the decisions of a central agency. A reminder seems appropriate at this point that a major argument for liberalization was that decentralized investment decisions would deliver the best outcome. However the proposed British system resembles the single buyer model in which a single buyer in the market holds periodic tenders for new capacity, with the winners signing long-term power purchase agreements with the single buyer.

Given the centralized, long-term contracting of all new generation (either through feed-in-tariffs or competitive tenders) the wholesale market price is no longer used as a signal for investment under this scenario. Achieving retail competition is no longer a major goal, and consumers can once again pay regulated tariffs. In a further step, abandoning decentralized contracting in favour of centralized pools might be considered, as several commentators argue that centralized dispatch is better suited to intermittent generation. Gains are made in simplicity, and transparent market prices make it easier to evaluate investment opportunities. However, this would mean abandoning the European Target Model.

5.2 Possible design of a single buyer model in a European country

A market design with detailed regulation would probably be set up as a single buyer model, where a state-controlled entity acts as the single buyer.

Investment decisions

The single buyer contracts generation capacity through long-term contracts. The contracted amount is based on a forecasted demand. Existing capacity could be contracted for a shorter period, e.g. 3-5 years, while new capacity would be contracted for longer periods, e.g. 15-40 years depending on technology, to mitigate the investment risk.

The capacity could be contracted through an auctioning process. Some technologies with few credible vendors, e.g. nuclear and off-shore wind, might not be suitable for an auctioning process. Here the contracted price could be an outcome of negotiations.

The long term contracts are typically be based on two types of payments:

- Capacity charges: cover the fixed costs of the plant, including all capital costs and fixed operations and maintained costs. This payment is made so long as the plant is available to be dispatched.

- Energy charges: cover the variable costs of the plant including fuel costs and all variable operations and maintenance costs. This payment is made only if the plant is actually dispatched.

Varying cost such as fuel cost will typically be passed through. Those payments could be constructed so that the generator has incentives to use the fuel efficiently and source it at lower cost.

Allocation of risk

Long term contracts have a typical allocation of risk:

- Demand risk is borne by the buyer. In other words, the buyer contracts to purchase a certain amount of electricity, independent of whether demand actually exists for this electricity. This allocation of risk makes sense because the private party has no control over whether or not their generation asset is dispatched

- Fuel price risk is almost always borne by the buyer. This means that changes in fuel costs—such as coal or gas—are passed through to the buyer under the contract.

- Inflation and foreign exchange risk. This risk is almost always borne by the buyer through escalation provisions in the contract

- Technical generation risks are always borne by seller, who is best able to control whether the plant is appropriately designed and functioning properly.

With a single buyer model most of the risks are shifted to the single buyer, and in the end to the consumers/tax payers. This results in a lower risk for then generator and should therefore expect a lower return. This could on theory lead to lower wholesale prices of electricity.

Dispatch

Dispatch will be centrally controlled. In most cases generation will usually be dispatched in merit order according to the contracted energy price. Other considerations could however be taken such as environmental concerns or grid stability issues. Hydropower will be optimized separately to minimize the cost of thermal generation.

Trade

Trade with neighbouring countries or regions can be hampered by the single buyer model. Cross border trade could be handled through bilateral trade where the single buyers share the profit of the trade.

6 Concluding remarks

Four different potential market design structures have been presented in this report. These market scenarios all have different advantages and disadvantages.

The energy-only market has the advantage that it is market-oriented, and – if functioning – using resources efficient. It also facilitates electricity exchange and further integration between countries. As the model result for the energy-only market shows spot prices are highly volatile, with very low prices at times and occasional high price peaks. This is especially true for markets that lack regulatory power such as hydro power. In order for some power producers (those with high marginal costs and fewer running hours per year) to cover their costs, prices will need to peak at times. If the spot price is allowed to vary freely, there will be incentives for demand response and further power generation investments. However, there might be a public opinion against volatile energy prices and price caps may be introduced. If so, the incentives for demand response and for power generation investments are decreased. Then other market models might be needed in order to create incentives for investments in power generation in order to ensure energy supply.

Capacity markets create incentives to provide capacity. The model results also show lowered wholesale power price, not only in the countries with capacity markets, but all over Europe. However, the cost of paying for the additional capacity more than offsets this price reduction, and in those countries with capacity markets, the customers generally pay more. It may also sub-optimal investments, if a capacity market is located next to for example an energy-only market. Power generation will then be built in the country with additional payments for capacity, even if the other location would be slightly more profitable (if it was not for the capacity market). The model results also showed a small increase (2%) in system costs for capacity markets. However, the increased costs may to some extent be offset by a larger reduction in the risk of electricity shortages. Since price volatility and price spread between areas are reduced when more generation capacity is installed with capacity markets, congestion revenues for interconnectors are reduced. Hence, capacity markets decrease incentives for market integration.

Critics of zonal pricing point out that some market participants get subsidized at the expense of others, and argue that a nodal approach (LMP) would be better. Advocates of LMP also point out the difficulties associated with loop flows, generation shift keys, and identifying critical network elements make nodal pricing more transparent and fair than zonal pricing. Since nodal prices reflect both the temporal and the locational value of electricity, it is often argued that nodal pricing encourages, in the short term, a more efficient dispatch of generators than zonal pricing, and a more efficient use of the existing network. One should note though that most consumers, with the exception of large customers, do not see the nodal prices directly. Nodal pricing is said to deliver more efficient signals for the location and timing of investment in generation, demand response and/or transmission. However, potential investors may hesitate to make decisions based on transmission constraints, as these might be temporary. Furthermore, in Europe there is a political unwillingness in some countries to let consumers and/or generators face different electricity prices, which is the case in a nodal approach. There are also concerns in Europe about the potentially very large number of nodes that can make up a market and the impact on market liquidity. LMP has also been criticized for being difficult and expensive to implement even though some argue that this is not the case. LMP has also been criticized for being vulnerable to market power, based on both empirical studies and the theoretical argument that LMP segments large markets into smaller markets and therefore gives market power to dominant generators in deficit nodes.

In the Increased central planning scenario, the government has control over the amount and type of new generation capacity. Governments will want to ensure that the rules that are laid down will be in force for long periods of time, so that investors that are contemplating long term investments in the electricity sector can assume that the conditions are true when an investment decision is made will continue to remain true for as long as is required to make the investment profitable. In this way it can to a higher extent ensure security of supply, and by using for example power specific subsidies enable a shift towards renewables. The spot prices can to a larger extent be controlled and price peaks be avoided. However, end prices as a whole may increase due to costs of subsidies. This is a step away from market approach (and hence also the European Target Model), since Increased Central planning means commercial actors and their behavior are largely controlled by governments. The use of resources might be suboptimal in comparison to a more market based approach. Another disadvantage is that projections of future demands have to be made, and there is a risk that these differ from actual outcome.

The aim of the study is not to validate which one of these that is most suitable for Europe's future energy market design, but rather to present each model's characteristics, advantages and disadvantages. Time will tell whether today's market design can evolve and meet future demands or if a more fundamental redesign of the market is needed.

Appendix A

Sweco's Apollo model

The modelling of the energy-only market and the capacity market scenarios was carried out with Sweco's power market model, Apollo. This model is a fundamental and deterministic power market model used for the analysis of the power market on the long-term horizon. In this work, it simulates 38 price regions within Europe and establishes trade both within Europe and with seven regions outside of Europe, represented as fixed price regions. It optimises system cost on a weekly and hourly level. Demand response is considered in all countries as virtual plants which can bid into the market at specific price levels and capacity. The model itself has been programmed in-house at Sweco; the model engine uses C++ and the application itself uses Java.

Fuel prices

Fuel prices are generally based on the Diversified Supply technologies scenario in the European Commission's Energy Roadmap 2050. Coal and oil prices, as globally traded commodities, differ slightly to the figures in the Roadmap (since those figures are driven by changes to global energy policy, unlike European policy as is here). Carbon prices are also based on this Roadmap for 2030, but are held closer to market prices today for the price in 2020.

Table 1. Assumed fuel prices in the scenarios

Fuel prices	2020	2030
Coal [€/MWh]	9.9	10.1
Natural gas [€/MWh]	26.5	24.5
Carbon [€/tonne]	5.0	68.3

Source: Sweco Energy Markets

Generation capacity

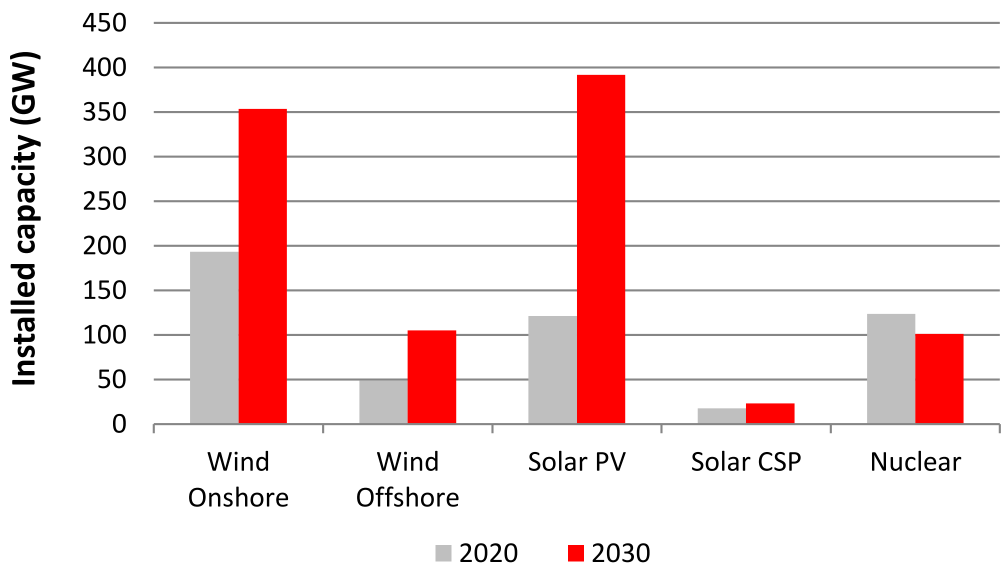
The model currently uses 381 different generic technologies, with power plants aggregated according to technology and efficiency. Users have the possibility to add further technologies if necessary or even add a specific power plant as a unique technology. The input side includes the following base technologies:

- Hydro (Reservoir, Run of River, Pumped Storage)
- Condensing (including Carbon Capture and Storage)
- Extraction (condensing CHP)

- Must-run (CHP and RES)

Availabilities are also set for each technology and time step. New investments in thermal generation are input as long as they are profitable and with perfect foresight. The assumptions regarding installed capacity in wind and solar technologies are shown in the figure below – most technologies increase significantly in capacity between 2020 and 2030, alongside the assumptions for installed capacity in nuclear – which decreases over time.

Figure 18. European installed capacity for wind, solar, and nuclear in scenarios



Source: Sweco Energy Markets

Cost of new capacity

New investments in thermal capacity were determined via an iterative process, with certain cost assumptions assumed. Expected annual revenues for the new investments are output from the model and compared with the annualized cost assumptions. If expected revenues are much higher than the cost assumptions then more capacity is input into the model in the second step of iteration, and vice versa if revenues are much lower. The assumptions for capital expenditure figures are shown in the table below.

Table 2. Cost assumptions for new thermal investments in the scenarios

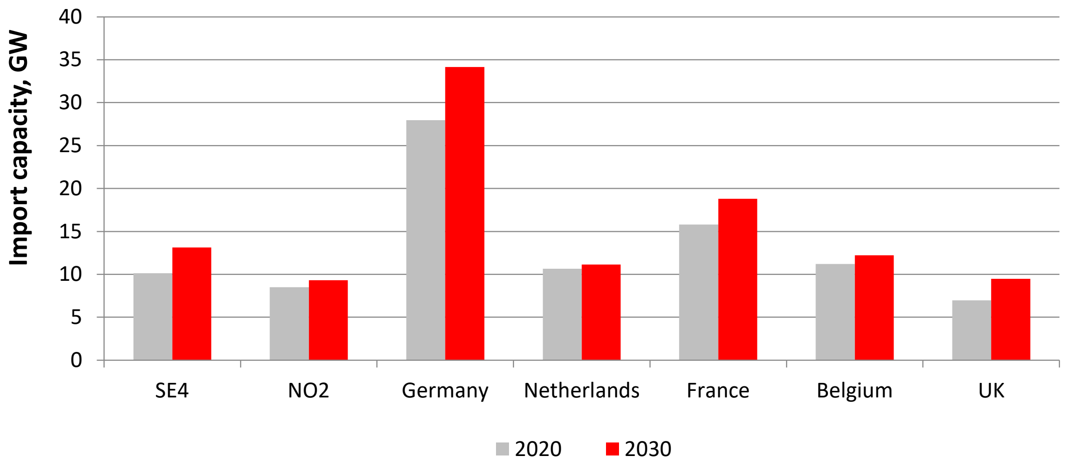
	CAPEX, €/kW/year	OPEX, €/kW/year
Gas turbine	59	9
CCGT	99	20
Coal condensing	224	20

Source: Sweco Energy Markets

Interconnector capacity

Interconnector capacity is based on up-to-date values stated by ENTSO-E, national development plans and the TYNDP2012 for investments into the future. With the large increase in RES generation, the need for additional interconnector capacity above and beyond these plans was also evaluated based on hourly price differences and input if deemed profitable in the results. In the modelling, individual price areas are assumed to be single nodes with no internal congestion. A summary of the total import capacities is shown for specific countries in the figure below, with capacity increasing in all regions.

Figure 19. Interconnector import capacity in 2020 and 2030 in selected countries



Source: Sweco Energy Markets

Capacity market assumptions

The modelling of a capacity market in each bidding region is done in a module separate to the power market model. The market is designed somewhat in line with a previous GB proposal, insomuch as the form of the demand curve.

The demand curve is determined by two figures: the net cost of new entry (net-CONE), taken to be that of a CCGT plant, and a target demand. This target demand is set at a value equal to peak demand in each region, though it varies for some regions and some policy designs. It is a sloping demand curve, and its cap is twice the value of net-CONE.

As for the supply curve, a plant bids into this market at a price equal to its missing money – that is the difference between the plant's costs and the expected revenues it can earn from the energy-only market.

The quantity of capacity that a specific plant can bid into the capacity market is defined according to an assumed availability. In this market, it is possible for certain RES to participate, but – except for hydro – this is generally at a very low availability.

In the Integrated Capacity Market, all countries have a capacity market and can freely trade between regions – limited only by available transmission capacity.

In the Coordinated Policy Scenario (CPS), there are two cases in which countries have capacity markets:

- Case 1: France, UK, Italy, Spain, and Portugal
- Case 2: Same as case 1, plus Germany

In all three configurations, external capacity can bid into the capacity market, even if the region in which it is physically located has no capacity market itself.

In the National Policy, there is only one case considered – that of case 1 in the CPS. In these markets there is also the limitation that only domestic capacity can bid into each capacity market.

In these markets, investments are input to a level that is dependent on two conditions – that there it is accepted into the market until some plant is marginal, and that there are no scarcity prices – classed as those above 1000 EUR/MWh – in the regions with the capacity markets.

Appendix B

From nodal to zonal: PTDF matrixes

The day-ahead market coupling algorithm to be used in the internal European electricity market is called *Euphemia*.

Starting two days before delivery, system operators calculate a series of inputs to Euphemia. These include a carefully chosen set of transmission lines, transmission capacities, and information on how incremental changes to generation located in the different biding zones will affect flows across the entire network.

The transmission lines, sometimes called *Critical Network Elements (CNEs)***,** are those lines that constitute commercially significant transmission constraints. The capacity on a CNE that is given to Euphemia for cross-zonal trade is known as the *Remaining Available Margin (RAM)*.

The information on how incremental changes to generation located in the different biding zones will affect flows across the entire network is encoded in *zone-to-CNE Power Transfer Distribution Factor (PTDF) matrixes* that are computed for every hour. A zone-to-CNE PTDF matrix has a row per CNE, and a column per bidding zone. Below is an example of what such a matrix could look like given three zones A, B and C and three two-directional CNEs between the bidding zones. C is the hub zone, where withdrawal takes place.

	Impact from bidding zone (transaction)		
CNE	**A** (A→C)	**B** (B→C)	**C** (C→C)
A → B	0,33	0,33	0
B → C	0,33	0,67	0
A → C	0,67	0,33	0
B → A	-0,33	0,33	0
C → B	-0,33	-0,67	0
C → A	-0,67	-0,33	0

Cell values represent the impact a 1 MW commercial exchange between a zone and the hub will have on the different CNBs. For instance, a trade from bidding zone A to hub zone C will lead to 0,33 MW of additional physical flow on CNE B → A. Negative signs indicate that the flow opposite to the designated direction. The hub zone is equivalent to a reference hub, and its choice is arbitrary: it will not change the results of the computations. From this matrix it is easy to compute a larger “transfer” PTDF matrix with all combinations of transactions between each pair of zones.

Zone-to-CNE PTDF matrixes are not created from scratch: they can be derived from significantly larger *node-to-line PTDF matrices***.** A node-to-line PTDF matrix has a row for every transmission line in the network, and a column per node.

	Nodes (transaction)			
Line	1 (1→1)	2 (2→1)	3 (3→1)	4 (4→1)
1-2	0	-0,75	-0,5	-0,25
2-3	0	0,25	-0,5	-0,25
3-4	0	0,25	0,5	-0,25
4-1	0	0,25	0,5	0,75

Cell values represent the impact on the different transmission lines of injecting 1 MW in one node and withdrawing it at the designated reference node. The choice of reference node is arbitrary and will not change the results of the computations. For instance, a 1 MW injection in node 2 and withdrawing at the reference node 1 will result in a 0.25 MW additional flow on the transmission line between nodes 3 and 4. As in the case of zone-to-CNE PTDF matrices, transfer matrices can be easily computed.

Node-to-line PTDF matrices are determined from measurable physical properties of the grid. In LMP markets, transactions are charged based on their PTDF factor.

Node-to-CNE PTDF matrixes are computed from *base cases* that are developed by system operators for each hour of the delivery day. Typically, transmission lines are included as CNEs if its maximum PTDF-factor is higher than a certain threshold.

To build a base case, system operators use the best information available and make assumptions (best estimates) of grid topology, net positions (either excess generation or excess consumption in a bidding zone) and corresponding power flows in each hour of operation on a nodal level for the delivery period. The base case is created one day before the delivery period (D-1), and is computed using data from two days before the delivery period (D-2), or a combination of D-2 data and forecasts for generation and consumption.

From node-to-line to zone-to-CNE PTDF matrixes

A zone-to-CNE PTDF aggregates nodes into zones and transmission lines between nodes into CNEs between zones. A problem that arises is that the nodes that are aggregated into a bidding zone will have varying influences on the CNEs and should therefore be assigned different weights, which also change over time. This leads to uncertainty in the capacity calculation, so the capacities that are given to the market are reduced by a *Flow Reliability Margin (FRM)* factor. Also, it is important to note the choice of CNEs will determine the grid constraints in Euphemia, and will therefore have an impact on welfare and day-ahead market prices.

The cell values in zone-to-CNE PTDF matrices are computed using *Generation Shift Keys (GSK)*. GSKs translate the incremental changes in the bidding zones into an increase of generation in the specific nodes, i.e. system operators “pretend” there is a change in the bidding zone and assign this change to the different nodes. The nodal changes are then used to compute how transmission lines are affected by these hypothetical changes in the aggregated nodes. GSKs thus reflect a system operator’s expectations regarding generation patterns in its bidding zone. There are many different types or “strategies” of GSKs, for generation, for load, for different generation types, etc.

Finally, it is important to note that system operators will have a very important role under flow-based market coupling, as their selection of critical transmission lines will impact social welfare and day-ahead electricity prices, and how to deal with critical branches that are not stable will affect how market participants are able to manage congestion risk. It is therefore very important that these choices are made in a transparent way.

Appendix C

Managing congestion risk in the internal European market

For the long-term allocation of cross-zonal capacity in the upcoming internal European market, the Network Code on Forward Capacity Allocation (NC FCA) foresees the use of either Financial Transmission Rights (FTRs), options or obligations, or Physical Transmission Rights (PTRs) subject to a Use-It-Or-Sell-It condition, however both FTRs and PTRs on the same border will no be accepted. The NC FCA also calls for harmonisation of PTRs and FTRs but does not develop further, other than at least PTRs of duration one month and one year should be offered. It should be noted, however, that the adoption of flow-based market coupling means that PTRs and FTRs will likely be flow-based. The implications of this should be studied further.

PTRs are the traditional approach for hedging congestion risk between national markets in continental Europe, PTRs are issued and auctioned (at present mostly explicitly) by system operators. If day-ahead markets at both sides of the interconnector are coupled, a Use-It-Or-Sell-It PTR is an option that gives its holder the right to physically schedule (or not schedule) a power flow (MW) over an interconnector between two locations A and B, in a given direction, and over a certain period of time, thereby capturing the associated congestion revenues. If a PTR holder does not nominate the PTR before day-ahead gate closure, it loses its right to the transmission path and the system operator will resell the capacity on day-ahead market and give the revenues to the PTR holder.

It is important to note that even with use-it-or-lose-it conditions, PTRs offer fewer opportunities for trading than FTRs in a multi-zone market. Unlike FTRs, PTRs are only valid over a specific interconnector. This means that market participants will likely need more than one PTR for bilateral trades between non-adjacent zones. Accordning to CASC.EU, , system operators at present in the CWE region auction PTRs via a central auction house, but no portfolios of PTRs are offered.

In Europe, FTRs have to date only been implemented internally in Italy and in the integrated Spanish-Portuguese market (MIBEL) (*see separate box Current financial arrangements in Europe*). The Nordic market has also implemented Contracts for Difference (CfDs). Common for these markets is that they manage internal day-ahead congestion by means of market splitting, and so FTRs' transmission capacities can be allocated efficiently in the day-ahead market, guaranteeing FTR holders the congestion revenue generated by the FTR. The successful introduction of market coupling is therefore a prerequisite for the continued introduction of FTRs in Europe, as if markets are not coupled and in the absence of a reliable reference prices, capacity allocation for zone-to-zone FTRs cannot be implemented.

The models used for flow-based market coupling model electricity systems less accurately than market splitting or LMP models. Forecasting how much transmission capacity may be available for use by transmission customers in the not-so-near future will therefore be a significant challenge, if flow-based methods are adopted. As is the case in regional LMP markets in the United States, how well the capacities system operators allocate match the amount of transmission capacity available day-ahead will determine whether the congestion and auction revenues collected by system operators will be enough to fully fund the transmission rights offered. But not even in LMP markets is forecasting capacities a simple task, in spite of DC models that model all nodes in the network or even complex AC network models that directly model losses and the effects of voltage limits. Changing market conditions in the month(s) leading up to dispatch, weather patterns, loop flows, outages and even market design issues contribute to FTR revenue inadequacy.

In Europe, system operators have to ensure the availability of auctioned capacities, exposing them to firmness risk. There is therefore a risk that system operators will be very conservative when auctioning longer-term capacities. This is not a problem in the United States, where the different markets have chosen other ways to allocate underfunding (or overfunding) of transmission rights. It is important to note that it is a very complicated issue that also touches on many issues including incentives to system operators, inter-TSO compensation, allocation of transmission costs, transmission tariffs, how information between system operators is shared and how bidding zones are defined. Many of these issues are often in the hands of national regulators.

Current financial arrangements in Europe

FTR options in the MIBEL market for the interconnection Portugal-Spain

The first joint auction of FTR options that provide a hedge for hourly day-ahead price differences across the Spanish-Portuguese interconnector took place in December 2013. Two independent products are auctioned, one for each direction of flow:

- FTR PT-ES (flow Spain->Portugal, i.e. $P_{ES} < P_{PT}$): the holder has a right to the price difference between Spain and Portugal (P_{PT} - P_{ES}) in the hours that the flow is from Spain to Portugal.
- FTR ES-PT (flow Portugal->Spain, i.e. $P_{PT} < P_{ES}$): the holder has a right to the price difference between Spain and Portugal (P_{ES} - P_{PT}) in the hours that the flow is from Portugal to Spain.

The Spanish and Portuguese system operators each issue one half of the FTRs auctioned in each direction. Subsequent trading takes place via OMIClear, the Iberian clearing platform for energy products.

FTR obligations in Italy

Italy is divided into price zones that reflect transmission constraints. Generators get the zonal price (P_Z) of the zone of injection, while load pays a single national price known as the PUN (*Prezzo Unico Nazionale*). FTR obligations known as CCCs (*Coperture dal rischio di volatilità del Corrispettivo di assegnazione della Capacità di trasporto*) allow generators to hedge the volatility of a congestion fee that generators pay (or receive) based on the difference between P_Z and the PUN.

The system operator organises annual and monthly auctions in which generators may purchase CCCs. The terms of a CCC are as follows:

- In the hours that P_Z < PUN generators pay the congestion fee. Holders of CCCs will receive (PUN - P_Z)*(number of CCCs) from the system operator
- In the hours that P_Z > PUN generators receive the congestion fee. Holders of CCCs will pay (P_Z - PUN)*(number of CCCs) from the system operator.
- There are separate markets for baseload and peakload hours and the duration of CCCs is either in a one-month or 12 months.

Contracts for Difference (CfDs) in the Nordic market

Nordic CfDs are issued by generators: market players first reduce their exposure to price volatility in the region as a whole by entering into forward agreements that pay the difference between the Nordic system price and an agreed strike price, and, on a second step, purchase CfDs that pay the difference between the system price and the zonal price.

Managing congestion risk in a LMP market

Because congestion can result in unexpected and significant differences in prices at different locations, congestion-related price risk is a primary concern in both zonal and nodal markets. In nodal markets, market participants can hedge this risk by purchasing Financial Transmission Rights (FTR) that offset the effects of nodal price differences.

Financial Transmission Rights (FTRs) in LMP markets
FTRs are tradable financial instruments defined by a quantity (MW), injection and exit locations (A and B) and a time period during which the terms of the FTR are valid. There are two main types of FTR products:

- *FTR obligations* that entitle its holder to a positive or negative credit based on the difference in day-ahead congestion price (i.e. the congestion revenues collected by system operators) between A and B during the relevant trading period.
    - *Positive credit if "prevailing flow" FTR:* if the path specified in the FTR is in the same direction as the congested flow (i.e. if the flow is from A to B, day-ahead congestion prices at the B are higher than at A), the economic value of the FTR for the relevant trading period will be positive.
    - *Negative credit if "counterflow FTR":* if the path specified in the FTR is in the opposite direction as the congested flow (i.e. if the flow is from A to B, day-ahead congestion prices at the B are lower than at A), the economic value will be negative, and the holder of the "counterflow FTR" will pay charges. Charges may be offset by delivering energy along the designated path.

- *FTR options* that entitle its holder to a positive credit if congestion is in the direction specified in the FTR, but no charges are required if congestion is in the opposite direction.

FTRs were first implemented in regional transmission systems in the United States following the introduction of competition in electricity markets. FTRs replaced the physical rights that had been used to guarantee the long-term delivery of cheaper generation to load. Physical rights were considered difficult to define and use-it-or-lose-it conditions difficult to enforce, therefore not suitable to the new LMP market design that relied on economic dispatch for an efficient use of the transmission system. Financial instruments would impose no constraint on the economic dispatch and are independent of actual transmission capacity.

Another perceived advantage of FTRs was that they provided a mechanism to allocate the congestion charges associated with LMP market design. FTRs could be used to compensate firm transmission service customers (typically load serving entities[7]) from congestion charges, as recognition of the fact that these customers had historically paid for, and would continue to pay for the fixed costs of the transmission system. Without compensation, these customers would end up paying for congestion twice. FTRs could also be awarded to those that invest in transmission.

At present, FTRs are offered in all LMP markets under different names: Congestion Revenue Rights (Texas-ERCOT, California ISO), Transmission Congestion Contracts (NY ISO) and Financial Transmission Rights (PJM RTO, Midwest ISO and New England ISO). System operators organise FTR markets to auction FTRs on a periodical basis, typically annually or seasonally prior to the beginning of summer and winter. FTRs may be of different durations: longer-term, one year, six months or one month. Sometimes there are different markets for peak and off-peak hours. All markets conduct monthly reconfiguration auctions.

In three markets (PJM, New England ISO and Midwest ISO), financial Auction Revenue Rights (ARRs) have taken over the role of FTRs. ARRs are directly allocated to firm transmission customers that request them, giving then a right to a share of the revenue generated in annual FTR auctions, or alternatively to convert ARRs to equivalent FTRs. ARRs allow firm transmission customers to receive the equivalent of a firm transmission service, i.e. compensation equal to the level of congestion. This means that in these markets, the role of FTRs is slightly different: FTRs are auctioned off to market participants that purchase them voluntarily, at a market value determined at auction, to hedge positions. ARRs are allocated annually, prior to the annual FTR auction.

The amount of ARRS/FTRs allocated at auctions is based on the system operator's estimates of the amount of transmission capacity that will be available at the time of dispatch. How well the auctioned FTR capacities match the amount of transmission capacity available at the time of dispatch will determine whether the collected day-ahead congestion revenues are enough to fully fund FTRs. In theory, according to Harvey 1997, if the auctioned FTR capacities are simultaneously feasible then economic dispatch with locational

[7] Load serving entities may be paying transmission charges to transmission owners, or may be utilities with transmission assets. In both cases, the cost of transmission services would be included in retail rates charged to customers.

pricing will achieve revenue adequacy (full funding). Annual ARR/FTR allocations are subject to simultaneous feasibility testing but simultaneous feasibility of ARR/FTR capacities is difficult to guarantee for many different reasons that are more or less relevant in the different markets. To start with, the FTR model needs to accurately model the network's topology. It is also very important to account for loop flows from adjacent systems when calculating capacities. Loop flows are a significant source of FTR underfunding, as they are not reflected in LMP prices. Loop flows, however, are difficult to predict, particularly for longer-term auctions. Unmodelled events like unforeseen plant or transmission outages will also affect the amount of transmission capacity available between two locations at the time of dispatch.

An important feature of FTR market design is therefore how to address what happens if a system operator allocates ARRs/FTRs for more or less transmission capacity that is actually available at the time of dispatch. While some markets allocate the shortfall to different actors and guarantee full hedging regardless of the level of FTR funding, other markets scale back the FTR payout on a pro rata basis, and FTR holders will not be fully hedged from congestion price differences on the affected pathway.

On such market is PJM. While FTR products are not designed to achieve full funding[8], underfunding has lead to much dissatisfaction among market participants. Underfunding has lead to lower FTR clearing prices, reflecting the expected lower value of a given FTR, but market participants complain that this makes PJM FTRs less useful to hedge congestion as it is unclear whether it is expectations of underfunding or less congestion that is the main FTR price driver. Other markets, like NYISO, recognise that the initial allocation of transmission capacities will not be simultaneously feasible at dispatch, and always pay the full FTR day-ahead market value, even in case of underfunding. The funding deficit is covered by transmission owners, who share all congestion charge surplus/shortfall as well as FTR auction revenues.

Simultaneous feasibility also means that in LMP markets in the United states, two-way FTR obligations are the norm, as one-way FTR options that do not support counterflows result in a smaller feasible set of FTRs, as cancellation of FTRs in the opposite direction (netting) is not possible. Some markets offer both, but offering both is complex and expensive to implement.

[8] PJM also allocates congestion imbalances back to FTRs, increasing underfunding

Appendix D

The case of New Zealand

New Zealand is one of the few markets globally that has implemented full nodal pricing, i.e. both generators and load face nodal prices at wholesale level.

System characteristics and challenges

New Zealand has a small, isolated electricity system with a long and slender transmission network. The bulk of demand is located in the North Island while the bulk of hydro generation, which on average accounts for approximately 60% of total generation, is located in the South Island. High hydro reliance and limited hydro storage capacity means that New Zealand is vulnerable to dry years, and reliant on flexible thermal back-up generation to ensure security of supply.

Generation adequacy

New Zealand's approach to 'keeping the lights on' has been to temporarily reduce demand by exhorting the public to reduce consumption to avoid the risk of brownouts and blackouts.

This approach risks suppressing the price spikes that are expected during shortages to signal that generation is scarce, thus undermining the incentive for generators to invest in last-resort generation and/or suppliers to invest in demand response capability. Because of this, New Zealand has introduced scarcity pricing arrangements in the form of cap and a floor wholesale spot prices to be applied in rare, sudden and short-lived situations when generation is so scarce that the system operator imposes emergency load shedding. The price cap addresses concerns that spot prices could settle well above the levels expected in a workably competitive market during instances of load shedding.

Full nodal pricing based on marginal costs

Transmissions constraints were not a major concern in New Zealand at the time nodal pricing was introduced, but reflecting the cost of transmitting electricity over long distances was. Because of the long distances involved, price differences between nodes can be significant even without congestion.

Nodal pricing creates a surplus of funds that in New Zealand is known as *loss and constraints excess* (LCE). It should be noted that since nodal pricing reflects short-run marginal costs rather than long-run marginal costs, the LCE does not fully recover the fixed costs of providing transmission network services.

Trading arrangements

Energy and reserves markets are cleared simultaneously. Generation offers, reserve offers, and demand bids are submitted to a centralised wholesale information and trading system (WITS) managed by the New Zealand Stock Exchange (NZX) at least 71 trading periods

(35,5 hours) before the beginning of the trading period to which the offer applies. All generation over 10MW connected to the transmission grid must submit offers.

Spot prices are calculated at 52 generator nodes (grid injection points) and 196 demand nodes (grid exit points) for each half-hour trading period. Unlike offers and bids in European day-ahead markets, offers and bids in New Zealand are not binding and may be modified until gate closure, currently two hours prior to real time. Intermittent and embedded generation may revise offered quantities up to 30 minutes before the beginning of the trading period.

Price forecasts, dispatching schedules and ex-post settlement

Final spot prices are not published until at least two business days after real time. To enable market participants to manage their resources, price forecasts and dispatching schedules for every node are published over a range of different time horizons leading up to real time:

- every 24 hours for the next coming days
- every two hours for the next 71 trading periods
- every half hour for the next 7 trading periods
- real-time prices are calculated every five minutes for the immediately preceding five minute period.

Improving the accuracy of forecast prices

Advance schedules using demand forecasts made by wholesale market buyers were found not to provide good advance indication of nodal prices. To benefit from the fact that demand can generally be more accurately forecasted centrally, grid exit nodes have been classified into conforming or non-conforming nodes. Conforming exit nodes are those nodes that have a predictable demand pattern. The system operator is responsible for forecasting demand at these nodes. Nodes that do not follow a predictable pattern are called non-conforming. Wholesale buyers only need to submit forecasts for non-conforming nodes.

Buyers at conforming nodes may submit ‘difference bids’ on a voluntary basis. The purpose of difference bids is to signal that some load at the node may vary depending on real-time expectations of price. Bids at non-conforming exit nodes are called “nominated bids” and are inelastic. Since differential bidding was introduced in 2012 the TSO has been publishing price-responsive schedules (PRS) that assume elastic response to prices, and non-responsive schedules (NRS) that assume inelastic load. In this way, market participants have better information about how price-responsive bids affect schedules.

Because the short-run marginal cost of providing demand response is typically higher than for generation, demand response is financially more attractive at times of system stress, when prices are higher. Under normal conditions, real-time prices are quite reliable to predict final prices. At times of system stress, however, real-time and final prices may diverge significantly, making demand response risky, as it cannot rely on price signals.

Managing locational price risk

FTRs covering the biggest source of locational price risk - between the North and the South islands - were introduced in 2013. The introduction of inter-island FTRs followed an electricity market review that put in evidence a lack of retail competition in some regions. The review concluded that locational price risk was being managed through vertical integration. Because it was not possible to hedge nodal price differences, retailers were reluctant to enter regions in which they did not own local generation assets, as the financial consequences of certain (infrequent) transmission constraints could be severe.

An Energy Market Services, a division of the system operator Transpower New Zealand Limited was appointed as FTR Manager to establish and operate the FTR market. Two types of FTR products are offered by the FTR Manager through uniform price auctions: FTR options and FTR obligations. There are two auctions per month: a primary auction and a variation (reconfiguration) auction. The FTR period is one month.

Unlike FTRs in the United States, FTRs in New Zealand provide a hedge for locational price risk, including transmission losses. Payments to FTR holders are funded by auction revenues and congestion rents. A full hedge is not guaranteed: in case of revenue inadequacy payments are scaled back as required.

Three new hubs to manage intra-island locational price risk will be added shortly.

About Sweco

Sweco's experts are working together to develop total solutions that contribute to the creation of a sustainable society. We call it sustainable engineering and design. We make it possible for our clients to carry out their projects not only with high quality and good economy but also with the best possible conditions for sustainable development.

With around 9,000 employees, Sweco is among the largest players in Europe and a leader in several market segments in the Nordic region and Central and Eastern Europe.

Sweco Energy Markets delivers value to our clients through deep insights on energy markets. We work with market design, regulation and market analysis. We support a continuous development of the market and help our clients to effectively participate on the energy markets.

Insights. Delivered.

Sweco
Gjörwellsgatan 22
P.O. Box 34044
SE-100 26 Stockholm, Sweden
Telephone +46 8 695 60 00
Fax +46 8 695 60 90

Sweco Energuide AB
Reg. No. 556007-5573
Stockholm
Member of the Sweco group