A survey of capacity mechanisms with lessons for the Swedish electricity market [a]

Pär Holmberg[b] and Thomas Tangerås[c]

March 26, 2022

Abstract

Many electricity markets use capacity mechanisms to support producers. Capacity payments can mitigate imperfections associated with "missing money" in the spot market and solve transitory capacity shortages caused by investment cycles, regulatory changes, or technology shifts. We discuss capacity mechanisms used in electricity markets around the world. We argue that correctly designed strategic reserves are likely to be more efficient than market-wide capacity mechanisms in jurisdictions that rely on substantial amounts of variable renewable energy and hydro power for electricity supply, such as Sweden.

JEL codes: D25, D47, Q40, Q48
Keywords: Capacity mechanism, capacity market, strategic reserve, wholesale electricity market

[a] This work has been financially supported by the Swedish Energy Agency and Svenskt Näringsliv. We are grateful for comments at meetings with Svenskt Näringsliv, Svenska Kraftnät, Energiforsk, and the reference group of the project 46227-1 of the Swedish Energy Agency. In particular, we would like to thank Zarah Andersson, Peter Cramton, Björn Hagman, Lina Håkansdotter, Marie Knutsen-Öy, Robert Ritz, Malgorzata Sadowska, and anonymous referees for help and comments on our work. Finally, we want to thank Christina Lönnblad and Glenn Nielsen, who assisted us by proofreading and editing this paper.
[b] Research Institute of Industrial Economics (IFN), Stockholm. Associate Researcher of the Energy Policy Research Group (EPRG), University of Cambridge. Affiliated with the Program on Energy and Sustainable Development (PESD), Stanford University.
[c] Research Institute of Industrial Economics (IFN), Stockholm. Associate Researcher of the Energy Policy Research Group (EPRG), University of Cambridge. Affiliated with the Program on Energy and Sustainable Development (PESD), Stanford University.

1 Introduction

Restructured wholesale electricity markets have a spot market in which consumers pay producers for the electricity they deliver. In *energy-only* markets, such as Denmark, the Netherlands and Norway, this is the only payment producers earn in the wholesale market. In many other markets, producers receive an additional upfront payment for making capacity available to the power system. In Belgium, Finland, Germany, Sweden and Texas, such *capacity payments* are limited to generation units within a designated *strategic reserve*, which is activated when the market capacity has been exhausted. In Great Britain and in most restructured electricity markets in the US, nearly all generation units on the market receive capacity payments. We refer to such a market-wide capacity mechanism as a *capacity market*.

The energy-only design is efficient. But in practice, politicians and households often find long periods with high electricity prices to be unsustainable. In extreme cases, bankruptcy and complete market failure can be the outcome. This occurred in the 2000-2001 California energy crisis, when the extreme prices drove the California Power Exchange and Pacific Gas & Electric into bankruptcy (Borenstein, 2002; Cramton, 2022). To avoid very high prices, Texas lowered its price cap, the maximum price in the market, after the energy crisis in 2021. The investment incentives of producers then decreased because of lower expected prices.

This paper surveys the literature on capacity mechanisms, which make it possible to lower the price cap without reducing the investment incentives (Cramton et al. 2013; Léautier, 2019). However, they are also used for the purpose of maintaining installed capacity, so that temporary capacity shortages due to investment cycles, new regulations, and technology shifts can be avoided.

Holmberg and Ritz (2020) note that well-designed strategic reserves and capacity markets are as efficient as an energy-only market in a simplified market where capacity is reliable, and the ramping costs are negligible. However, modern markets with non-thermal technologies, such as energy storage, demand response and renewables with intermittent output, are more complicated. For such technologies, it is hard to estimate a generation unit's reliable output (firm capacity), and to give them correct price signals. As an example, a capacity market does not give the right incentives to wind power owners in the choice of location, plant design, maintenance, and preparation for extreme weather conditions. The latter has turned out to be of particular importance in the US, where capacity tends to become unavailable when it is most needed during extreme weather conditions. Another problem is how to give correct price signals to energy storage plants (including hydro power), so that they save optimally for critical days, if prices are capped at a low level. In addition, there is a risk that special interests, short-termism or the excessive risk aversion of political actors controlling capacity mechanisms can lead to inefficiencies. To minimise these distortions, the price cap should be as high as is politically acceptable, so that the capacity mechanisms can be minimised. Well-designed capacity mechanisms can reduce distortions even further.

In comparison to strategic reserves, capacity markets entail advantages for systems that are dominated by thermal power, especially if the market has significant issues with investment cycles or if there are large time variations in fuel prices. Capacity markets are particularly suitable for markets with a centralised day-ahead dispatch as in the US, where the system operator dispatches each individual generation unit. Moreover, capacity markets can be used to implement an engineering-based reliability standard (Aagard and Kleit, 2022).

A main advantage of a strategic reserve is that it is only necessary to define firm capacity and regulate availability for the units within the reserve. This task is straightforward if the reserve mostly consists of thermal peak-load units with high variable costs. Hence, a well-designed strategic reserve is likely to be more efficient than a capacity market if production is dominated by renewables and/or hydro power, as in Sweden.

A problem with strategic reserves in practice is that the regulations for those units tend to be overly restrictive. Sweden has stricter environmental rules for plants inside than for those outside the reserve. In Belgium and Germany, a plant that has entered the reserve can never return to the market. Moreover, in Germany the plant must close after some years in the reserve. We favour an isolated reserve, where the operation of the reserve does not influence spot prices, but in our view Belgium and Germany have exaggerated this isolation.[1]

In the paper, we also discuss how to organise the procurement of capacity. A problem is the limited supply of capacity, which gives suppliers significant market power, especially in capacity markets where the procured volume is large. This problem can be mitigated by introducing an elastic demand for capacity and by procuring capacity further in advance. Other questions to consider are: (i) who will be responsible for procuring the capacity; (ii) how much information should bidders receive during the procurement process; and (iii) what is the best auction design when procuring capacity?

This paper focuses on resource adequacy, i.e., how capacity mechanisms can be used to support the wholesale market. Other mechanisms, such as targeted capacity payments or tenders for new capacity, are useful when solving specific problems, such as supporting renewables, facilitating the closure of coal plants, and strengthening operating reserves.[2]

The paper is outlined as follows. Section 2 gives a quick introduction to electricity markets in the EU and the US.[3] In Section 3, we reiterate well-known results for investments in a competitive energy-only market. In Section 4 we go through capacity markets in detail and how capacity can be procured. In Section 5, we discuss strategic reserves, with a particular focus on the Swedish strategic reserve. We conclude the paper in Section 6.

[1] On the other hand, we would argue that the strategic petroleum reserve in the US was not sufficiently isolated from the market. The size of that reserve was at times used by politicians to stabilise petroleum prices, also when there was no shortage of oil (Bamberger, 2010).

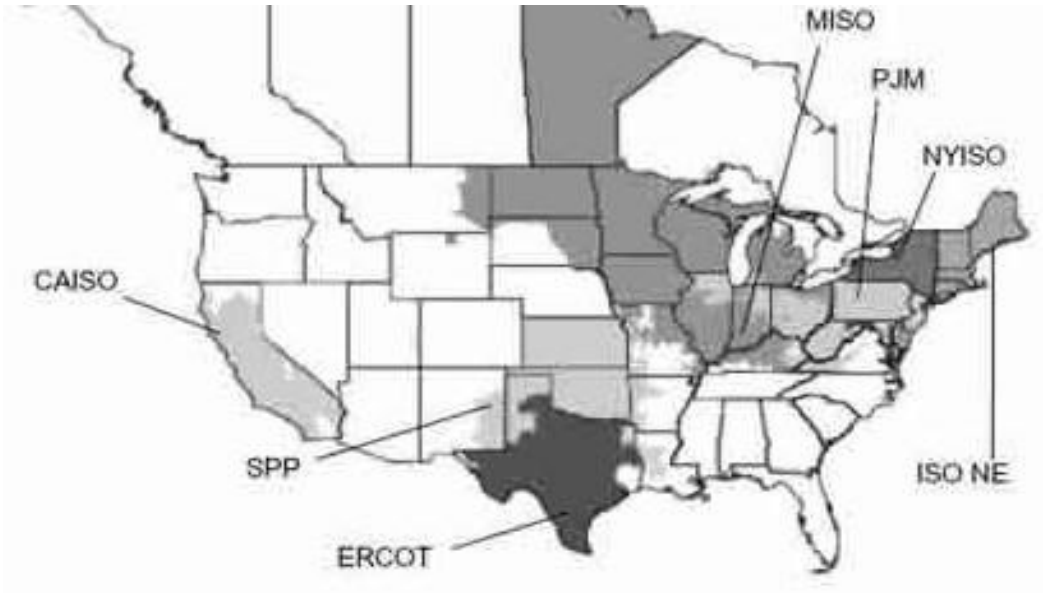
[2] Operating reserves help system operators maintain a reliable electricity system by balancing supply and demand.

[3] A more detailed comparison of US and European electricity markets can be found in Ahlqvist et al. (2022). That report emphasises the advantages and disadvantages of centralised electricity markets.

2 Electricity markets in the EU and the US

The internal electricity market of EU is divided into bidding zones, usually one per country. The amount of consumption and production within each zone is cleared on the spot market, which accounts for the transmission capacity between zones. Intra-zonal constraints are neglected in the spot market.[4] In most EU countries, producers can decide which of their units will produce the contracted amount of electricity within each zone. This is usually called a decentralised or portfolio-oriented spot market. In the EU, a transmission system operator (TSO) often owns the transmission system. Hence, TSOs receive congestion rents when price differences arise between zones. This means that their revenue depends on the spot market outcome. This is one reason why the spot market has often been managed by a market operator (not by the system operator) in EU countries. In European electricity markets, retail is unbundled from distribution.

Figure 1: Seven restructured electricity markets in the US



In the US, all states have substantial autonomy over the design of their electricity markets. For example, several states have decided not to restructure their markets**.** Figure 1 shows the geographical footprint of the seven restructured electricity markets in the US: California (CAISO), Midcontinent Independent System Operator (MISO), New England (ISO NE), New York (NYISO), Pennsylvania-New Jersey-Maryland (PJM), Southwest Power Pool (SPP) and Texas (ERCOT). MISO, ISONE, PJM and SPP each span several states. MISO also includes Manitoba in Canada. CAISO, NYISO and ERCOT cover one state each. ERCOT in Texas has a strategic reserve in the form of 4 GW Emergency Demand Response (Cramton, 2022). The other six electricity markets in the US have capacity markets designed in accordance with the standard model recommended by the Federal Energy Regulatory Commission (FERC). All restructured markets in the US have a centralised spot market in the sense that the system

[4] In practice, intra-zonal congestion is often considered when system operators allocate inter-zonal transmission capacity to the market.

operator decides how much is to be produced by each unit. These decisions are based on the detailed bids at the unit level submitted by generation owners. In the US markets, the system operator is an Independent System Operator (ISO).[5] It neither owns network nor production capacity to generate revenue on the spot market. This independence means that it is appropriate for the ISO to be involved in the procurement of production capacity and the operation of the spot market. In the US, sales and distribution of electricity are often integrated into Load Serving Entities (LSEs). The final customer prices are often regulated.

In this paper, we will also discuss some markets in South America that use capacity markets. Colombia is one particularly interesting example. It should be noted that the regulation of South American markets can differ significantly from EU and US markets. For example, some South American electricity markets are cost-based, meaning that producers are not allowed to submit bids to the market (Ahlqvist et al., 2022).

3 Investments in an energy-only market

In an efficient electricity market, electricity is produced at the lowest possible total production and investment cost. The capacity utilisation will be very different for different plants because of fluctuating demand across the day and the year. Some plants run almost all the time, and others are rarely used. Therefore, it is usually efficient to invest in a mixture of different technologies, where the choice of technology for a specific plant depends on how often the plant is to be used. Typically, low variable-cost technologies have a high investment cost, and vice versa. For base load units that are active almost all the time, it is efficient to utilise economies of scale in production. These are technologies with low variable costs and high fixed costs, such as nuclear power and hydro power where available. As the utilisation rate decreases, it becomes more important that the plant does not cost money when not in use. It will then be more economically profitable to use technologies with higher variable costs and lower fixed costs. Some peak-power plants are used only a few times a year. These have very high variable costs and low fixed costs. To optimise investments, the market needs to get both the volume of investments and the right mixture of technologies right.

Investments in electricity markets are often analysed under ideal assumptions, which we refer to as a simplified electricity market.[6]

Definition *A simplified electricity market has perfect competition, free entry/exit of capacity, no network congestion, price-insensitive demand, no risk of an uncontrolled system collapse, risk-neutral and fully informed producers that invest in flexible and enduring production that*

[5] Some system operators in US are Regional Transmission Organizations (RTOs). RTO and ISO essentially mean the same thing, but an RTO typically operates in several states.

[6] These assumptions are standard in the literature on peak-load pricing and screening-curve analysis (Crew and Kleindorfer, 1976; Chao, 1983; Stoft, 2002; Biggar and Hesamzadeh, 2014; Léautier, 2019), but there are studies that analyse investments under more general assumptions. For example, Joskow and Tirole (2007), Tangerås (2018) and Astier and Lambin (2019) allow for elastic demand. Zöttl (2010) considers imperfect competition, and Teirilä and Ritz (2019) simulate strategic investments in the Irish market. Cramton et al. (2021) allow three years for entry/exit, price-sensitive demand and resource non-convexities (startup and no-load costs).

is always available. If there is a shortage of power, the system operator rations demand when the price reaches the price cap in the spot market, $\bar{p}$.

We use the term enduring to denote production that does not run out of fuel. Hydro-power and other energy storage technologies have limited energy that they can deliver and are not enduring. Flexible and enduring production in a competitive market implies that each producer will produce as much as possible so long as the price is above its variable cost. Demand response is growing in importance, but price-insensitive demand is still a good approximation for a large group of consumers. The main reason why these consumers do not respond to electricity prices is that they buy electricity on long-term supply contracts and therefore have little to gain from reducing consumption in resource-constrained situations. Moreover, many consumers lack control of their consumption. Hence, even if they were to face the spot price, they might not respond to it. But this does not mean that customers would be willing to consume electricity at *any* price. At a sufficiently high price, they would prefer to be disconnected by the system operator. *Value of Lost Load* (VOLL) denotes the theoretical price at which an average customer would rather be disconnected than continue to use electricity. We denote this price by p_{VOLL}.

There is a continuum of technologies in which producers can choose to invest, and they can choose the invested volume for each technology. Each technology is indexed by its variable cost *c*. The fixed unit cost of a technology *c* is given by *k(c).* The market is represented by a representative delivery period. To simplify the analysis, we normalise *k* so that it is the fixed cost per delivery period. The fixed cost is assumed to be lower for technologies with a high variable cost. We let $\bar{c}$ be the highest variable cost in which producers choose to invest and $k(\bar{c})$ is the fixed cost for this technology. Note that $\bar{c}$ is endogenous: it is chosen by the market. If demand exceeds the total capacity in the market, there will be a shortage of electricity, and it will be necessary to ration demand to avoid a blackout. Investing in an additional (marginal) plant helps reduce the electricity shortage. The socio-economic value of producing one unit of energy in this plant when there is shortage is given by the difference $p_{VOLL} - \bar{c}$. The probability of power shortage is usually referred to as *Loss of Load Probability* (LOLP). If we state this probability as π_{LOLP}, it is efficient to invest in new capacity if the expected revenue $\pi_{LOLP}(p_{VOLL} - \bar{c})$ for the most expensive plant exceeds its fixed cost $k(\bar{c})$. It follows from the assumptions of a simplified electricity market that the total production and investment costs will be minimised for each plant, taking into account their degree of utilisation.

Conclusion *Investments in a simplified electricity market are efficient*[7] *if the highest variable cost on the market meets the relationship* $\pi_{LOLP}(p_{VOLL} - \bar{c}) = k(\bar{c})$.

It follows that in an efficient market, the highest variable cost on the market is below VOLL. This makes sense as it would be more efficient to ration demand than to use such an

[7] Note that this is under the simplified market assumptions. For example, there would be potential to improve the efficiency if consumers were active and responded to the spot price or if the system operator was informed of each individual consumer's VOLL level and was allowed to ration consumers with the lowest VOLL level first.

expensive technology. From the above conclusion, it is also evident that a positive likelihood of curtailment is efficient so long as consumers' willingness to pay to avoid power shortages is bounded. In other words, it is not economically viable to have such extensive investments that the risk of rolling blackouts (controlled demand rationing) disappears completely. Optimising the duration of blackouts is called the adequacy problem (Cramton et al., 2013). Resource adequacy ensures that there are sufficient resources available to serve electricity demand under all but the most extreme conditions.

In situations of excess demand, the most expensive plant on the market receives the revenue of $\bar{p} - \bar{c}$. It is profitable to invest in additional peak power if the expected profit $\pi_{LOLP}(\bar{p} - \bar{c})$ in the spot market exceeds the fixed cost $k(\bar{c})$. In a market with perfect competition, and free entry, risk-neutral companies will invest until the marginal invested dollar becomes unprofitable. Based on this reasoning and the above conclusion, we can draw the following conclusion:[8]

Conclusion *The investments will be efficient in a simplified electricity market if the price cap is chosen so that* $\bar{p} = p_{VOLL}$.

An energy-only market will thus be efficient and provide the investments the market needs if the system operator rations consumers at p_{VOLL} when there is a shortage of electricity. A simplified market with a price cap at p_{VOLL} does not favour any technology; investments will be efficient for all technologies on the market.

We will use the simplified electricity market as a benchmark in the discussion of capacity mechanisms in Sections 4 and 5. In this discussion, we will relax most of the assumptions of the simplified model one at a time.

4 Capacity markets

4.1 Why capacity markets?

A consequence of the results in the previous section is that an energy-only market will fail to deliver efficient investments in a simplified electricity market if the price cap $\bar{p}$ deviates from p_{VOLL}. If the price cap is set too low, then electricity shortage will occur too often because there will not be enough investment. Capacity markets give producers an extra payment that compensates them for receiving a lower revenue in the spot market.

There are several reasons why the price cap is sometimes set below p_{VOLL}. For one, it is a way of mitigating market power. Holmberg and Newbery (2010) show that a small price cap pushes down the entire supply curve in markets with imperfect competition. Thus, it is not just that price spikes are cut, but the price decreases during all hours of the day. A lower price cap also reduces price risks in the market. This is a particularly important aspect in countries

[8] This is a standard result that can be found in, for example, Stoft (2002), Joskow and Tirole (2007), Léautier (2019) and Willems (2015).

that lack well-developed financial trade. In well-developed markets, on the other hand, producers, electricity traders and consumers have better opportunities to use financial contracts to hedge prices (Tangerås, 2018). Still, in 2021 and 2022, the energy crises in Texas and Europe showed that both consumers and energy traders can struggle, and even go bankrupt, in well-developed markets when electricity prices are high for a long period. In such circumstances, electricity prices can become difficult to handle politically. Political pressure reduces the market's confidence that a formally established price cap will prevail in situations of power shortage. Léautier (2019) argues that, from a political perspective, there may be advantages in setting a price cap below p_{VOLL} and to ensure security of supply through capacity payments.

Conclusion *Market power, social costs of bankruptcy and political considerations can motivate a price cap below p_{VOLL}. Capacity mechanisms compensate producers for a reduced revenue in the spot market.*

A risk of energy-only markets is that a temporary capacity shortage can occur due to investment cycles (Spess et al., 2013). This means that the market will deviate from its equilibrium in the short term and that the risk of electricity shortages might significantly increase. Arango and Larsen (2011) summarise the literature on boom-bust cycles. Those are mainly a problem in capital-intensive industries with long construction delays and can occur when investors are imperfectly informed and act myopically, i.e., they mainly invest when commodity prices are high. Various papers have identified boom-bust cycles in the mining, oil, pulp, paper, chemical and real-estate industries (Arango and Larsen, 2011). Arango and Larsen (2011) study investment cycles in the three electricity markets with the longest history: Chile, Nord Pool and the UK. In the UK and Chile, the variations in the reserve margin are pronounced and consistent with boom-bust cycles. The cyclic pattern is less prevalent in the Nordic countries. Arango and Larsen (2011) argue that the reason for this could be that the dominant producers in the Nordic countries are controlled by the government and that these companies have a longer perspective on investment.

Temporary electricity shortages could also occur due to sudden changes in regulations. Newell et al. (2012) describe an episode where market participants in ERCOT (Texas) were concerned that new environmental legislation could lead to a substantial amount of plant closures. In the end, the market seems to have absorbed this shock. Resource adequacy has not been a significant problem in Texas during the last 10 years (Cramton, 2022). An example from Sweden was the government's proposal for a new tax on waste incineration, which prompted electricity producers to announce the closure of critical electricity production in Stockholm and other metropolitan areas. Another example was safety directives regarding passive cooling of nuclear power plants, which in combination with the tax on nuclear power and low electricity prices, contributed to the decisions to shut down reactors prematurely in Sweden. Deviations from the market equilibrium and the risk of electricity shortages will be lower if regulatory changes are announced well in advance. Disequilibrium could also occur if the market undergoes a rapid transformation in a short time, for example after a technology shift. If a temporary power shortage is a major problem, then volume-based capacity markets should be particularly suitable for controlling market capacity and avoiding abrupt capacity

changes. Moreover, such a procurement would help producers coordinate their investment and disinvestment decisions, as capacity that gets a payment is more likely to enter or stay in the market. Volume-based capacity markets in the US are often used to implement reliability standards (Aggard and Kleit, 2022).

Conclusion *Investment cycles, new regulations and technology shifts can lead to a temporary shortage of production capacity in an energy-only market. A volume-based capacity market could counteract abrupt changes in the production capacity, coordinate entry/exit, and implement a reliability standard.*

In addition, capacity mechanisms (or a price cap above p_{VOLL}) can increase the margin in the system, and thus provide more scope for short-term deviations from the long-term market equilibrium, without leading to a significant risk of electricity shortages during an energy transition.

There are no externalities in a simplified electricity market. Considering positive external effects can motivate a capacity market, or to set the price cap above p_{VOLL}. When demand is so high that there is a shortage of electricity or a significant risk of shortage of electricity, the margins in the electricity system become small. The risk increases of an uncontrolled system collapse whereby all or large parts of the market are shut down for a prolonged period of time (Joskow and Tirole, 2007; Cramton et al., 2013). In Texas it could even take weeks to start up the system from zero (Cramton, 2022). Consumers and producers have much to lose from such a collapse. Security of supply, and more generally, the quality of delivered electricity in terms of low risk of interruption, a stable voltage level and stable frequency is a public good (Abbot, 2001; Amundsen and Bergman, 2007). Hence, each new investment in production capacity has positive external effects that benefit all agents in the electricity system (Fabra, 2018; Llobet and Padilla, 2018). Payments to producers must be increased to obtain optimal investments regarding this type of positive external effect. A higher price cap can be a way of achieving this objective. Another possibility is to increase capacity payments.

An argument that is sometimes used to justify capacity markets is that subsidies for renewable electricity production displace thermal electricity production. In part, this type of argument is put forward by producers who have lost revenue from the displacement and who want compensation for their loss. But displacement can also be a welfare economic problem if one wishes to maintain thermal capacity as a complement to renewable production, as in Portugal and Spain (Roques and Verhaeghe, 2015).

4.2 Problems with capacity markets

In practice, procured capacity often tends to be too large, perhaps due to the excessive risk aversion of political decision makers (Aagard and Kleit, 2022). Nelder (2013) argues that the authorities in Australia have been poor at estimating demand. The problem has been noticeable, for example, in Western Australia, where a capacity market delivered large over-investments because of an unexpected fall in electricity demand. Newbery (1997) shows that the authorities systematically overestimated the risk of electricity shortages in England and Wales in the 1990s. Capacity payments were proportional to π_{LOLP} during that period, so the

capacity payments were excessively large. Newbery and Grubb (2014) argue that the new capacity market in the UK is likely to result in an excessive procurement of capacity, mainly because the contribution from interconnectors is neglected. But there are also cases in South America, where the risk of electricity shortage was underestimated (Wolak, 2019).

Cramton and Stoft (2008) argue that the cost of overinvestment in generation capacity does not have to be particularly large. They estimate that installed capacity 10% above the efficient level increases the cost to consumers by about 2% and the cost to society by around 1%. However, estimates by the American Public Power Association (APPA) suggest that the costs could be higher. They estimate that an average household in PJM's area pays $120 extra per year to cover the capacity payments (APPA, 2017).

Availability is another issue with capacity markets. A plant that has received a capacity payment is obliged to provide its firm capacity when called upon by the system operator. But regulations sometimes describe circumstances under which a facility does not have to be available. A potential problem then is that the owner may take advantage of such exemptions. This has been a problem in electricity markets in North and South America, where producers have managed to obtain capacity payments for facilities that were unavailable. Therefore, it is important that the capacity market carefully defines availability, and that the regulations leave minimal room for manipulation of availability. In addition, producers should have particularly strong incentives to supply capacity in situations where one can predict in advance that the demand for electricity will be particularly high (Batlle et al., 2015). Tómasson et al. (2020) make the point that the correct incentives can only be restored when the penalties are as high as the price would have gone to in the absence of the price cap.

Unfortunately, performance is often at its worst under extreme conditions when capacity is needed the most. PJM lost more than 20% of its capacity during a cold period in January 2014 (Rose et al., 2014). A third of the lost capacity in PJM was natural-gas generation with interruptible-gas contracts, which could not get any gas delivered. Texas had similar issues during the winter storm in 2021, where 30 GW of thermal production (mainly natural gas production) was unavailable (Cramton, 2022). Extremely high temperatures can also be a problem for availability, as production efficiency goes down with the ambient temperature. This was illustrated by the heat storm in August 2020, where California lost 1.4-2 GW gas capacity (CAISO, 2020). There was a similar issue with solar power. Nearly 20% of the solar capacity disappeared, which also contributed to the rolling black outs (controlled demand rationing).[9] The extreme weather events illustrate how important it is that natural gas plants receiving capacity payments should be required to have non-interruptible gas contracts and/or being able to switch to a back-up fuel (Rose et al., 2014). After its crisis, PJM has tightened its capacity requirements (Batlle et al., 2015) and introduced pay for performance schemes, where generators that commit to having power available when demand is highest will receive larger payments than those that do not. At the same time, if they do not meet those

[9] This was mentioned in Wolak's presentation "Evidence from California on Challenges Facing Electricity Supply Industries with a Significant Share of Intermittent Renewables" at the meeting of the Swedish Association of Energy Economists (SEEF) on September 24, 2020.

commitments, they will be penalised under a "no excuses" policy. The downside of the measures is that some capacity is excluded from, or disadvantaged in, the capacity market.[10] The consumer side argues that the stricter requirements have led to higher prices in the capacity market (APPA, 2017). Moreover, the regulation of capacity markets is becoming very complex in the US. A concern is that regulators may not have enough competence and resources to regulate this complex market and that industry, system operators and consumers will gain influence over the regulation of capacity markets (Aagard and Kleit, 2022).

Reliability options were launched as a possible solution to parts of the availability problem and are, for example, discussed by Cramton et al. (2013) and Aagard and Kleit (2022). For this type of capacity market, producers must issue options corresponding to the capacity planned to be available in the market. The options have a strike price corresponding to the highest variable cost in the market. If the spot price exceeds the strike price, producers must pay the difference, so that the consumer price will never exceed the strike price. The producers are paid to issue the options, which is partly a compensation for providing the hedge, but it also constitutes a capacity payment. The undertaking is unproblematic for the producer, provided that its plants are available. In case of high spot prices, a producer's revenue from the spot market can be used to pay the consumer, in accordance with the option contract. This gives producers incentives to try to keep capacity available, especially when the price is high, and capacity is needed most. Reliability options have, for instance, been applied in Colombia. A problem there is that the country also introduced a so-called administrative price cap on the strike price of the option. The purpose was to keep prices down and reduce the risk in the market, but an unforeseen side effect was that electricity producers with a dominant position were given stronger incentives to raise the price above the strike price. McRae and Wolak (2019) study this problem empirically and show that the side effect has led to higher electricity prices and reduced availability in Colombia. Colombia's electricity market could be particularly challenging. Shortages of electricity tend to be rare, but persistent once they occur. Still there is a risk that similar problems could arise in Ireland and New England, which use a similar capacity-market design.

Another question is whether different technologies should be procured in the same auction. Experience from PJM suggests that simultaneous procurement increases competition and auction efficiency. However, different technologies often have different availabilities. A collective procurement therefore presupposes that it is possible to calculate firm capacity for technologies with a low availability. However, this is easier said than done, especially as it is not only the physical availability that matters but also the time when a plant is available. Weather-dependent renewable electricity production, for example, is more valuable if it co-varies with demand and delivers as much as possible when the risk of electricity shortage is greatest. The location, design, and maintenance (including ice removal) of a wind power plant can all have a significant impact on how much it can produce when there is an electricity shortage. The problem with derating factors is that they would generally not give the right

[10] Capacity payments have led to many controversies in the US, where generation owners, for example, have argued that they have been disadvantaged by various rule changes. Regulatory changes in the capacity market are often settled through litigation (Spees et al., 2013; APPA, 2017).

incentives to wind-power owners for these, and related, decisions. It is also tricky to define a firm capacity for hydro power. How should one give correct price signals to energy storage (including hydro power), so that they save optimally for critical days, if prices are capped at a level below p_{VOLL}? Cramton and Stoft (2007) argue that firm capacity for hydro power should be defined by how much they can provide during an extremely dry year. Such a design would be robust from a resource adequacy perspective. But it is uncertain whether the design would be efficient. In practice, firm capacity estimates of regulators seem to have been less cautious. Wolak (2019) provides several examples from hydro-power-dominant markets in Latin America where the firm capacity of hydropower was overestimated, so that the procured capacity could not deliver as planned. We conclude that capacity payments are more suitable for thermal production, which is enduring and for which the availability is straightforward to measure, and that capacity payments are less suitable for intermittent renewables and energy storage.

A low price cap reduces consumers' interest in demand response. To a certain extent, this can be compensated if the demand response is rewarded with a capacity payment. But it can be difficult to define firm capacity and regulate the availability for flexible demand. For example, should the consumer be required to always use more than a certain amount of electricity to ensure that there is always enough capacity to reduce consumption?

Conclusion *A problem with low price caps is the difficulty to incentivise availability for plants. For capacity markets it is complex to correctly define the firm capacity of each individual plant, especially for non-thermal technologies such as solar power, wind power, demand response, hydro power, and other energy storage.*

Capacity markets need an external authority to verify firm capacity. It could be the regulator, a system operator or an aggregator. Such a bureaucratic process could hurt households and small industries that have invested in solar panels, wind power, demand response or energy storage. A market-wide capacity market is probably more suitable for centralised markets such as in the US, where all capacity, including demand response and energy storage, must be verified anyway before it can participate in the spot market. Empirical studies show that centralisation of the dispatch decisions typically enhance market efficiency (Ahlqvist et al., 2022). Decentralised energy-only electricity markets, such as in Norway, are less bureaucratic. In such markets, retailers can use historical data to estimate demand response, local production and storage of their customers and bid a corresponding demand curve to the spot market, even if these units and their characteristics have not been registered.

Conclusion *A problem with capacity markets is that the firm capacity of all facilities, small and large, must somehow be individually verified and approved. This bureaucratic process increases the administrative burden, which disadvantages small market participants. This is less of a problem for centralised markets, where the characteristics of generation units are anyway registered to increase efficiency of the day-ahead dispatch.*

A reliability externality (Wolak, 2019 and 2021) arises in markets with low price caps, which are typical for electricity markets with capacity markets. The problem is that consumers of a retailer/LSE that has bought electricity in advance are as likely as consumers of a non-

contracted retailer/LSE to be disconnected in case of an electricity shortage. As contracting does not give priority in such situations, retailers/LSEs will buy too little electricity in the forward market. This effect is particularly noticeable in markets with low price caps, such as in Latin America (Wolak, 2019). If a regulatory intervention is required to fix this problem, then Wolak (2019; 2021) recommends mandatory forward contracting. In some countries in Latin America, retailers/LSEs are required to purchase up to 90% of their clients' planned consumption one or several years in advance.

4.3 How do capacity payments and the price cap affect investment?

The spot price often clears above the variable cost of base-load power. Revenue from the spot market then contributes to covering the large, fixed costs of the base load. As a consequence, a large fraction of the revenue base load receives in excess of its variable cost is independent of the price cap and capacity payments. It is different for peak power. In a competitive market, the spot price covers little else than its variable cost under normal system conditions. However, the electricity price increases to the level of the price cap under a power shortage, during which peak power receives a scarcity rent that can be used to finance the investment costs (Stoft, 2002). For this reason, peak power is particularly dependent on price caps and capacity payments.

In the short term, all plants benefit from a sudden introduction of or increase in capacity payments, but there will be new investments so that the capacity increases. This reduces the risk of electricity shortages, which reduces the scarcity rent for old plants. In a simplified electricity market, the reduced revenue corresponds exactly to the increase in capacity payments (Holmberg and Ritz, 2020). Hence, in the long run nothing happens to the capacity of existing technologies for a simplified market in equilibrium. In the long term, an increase in capacity payments will entirely go to financing investments in new plants, which will have a higher variable cost than the old facilities. The consequence will be the same if the price cap is raised. Conversely, only the capacity of plants with a high variable cost will decrease in the long term if capacity payments or price caps are lowered.

Conclusion *In the long run, only the capacity of plants with a high variable cost is affected by the level of capacity payments and price caps, at least in a simplified electricity market.*

This is also an argument for why new and old plants should receive the same capacity payment. In addition, there is a risk that the agents will try to take ineffective measures if the capacity market discriminates between new and old capacity. There are several such examples from related markets. In Sweden's market for tradable green certificates, only capacity below a certain age was entitled to compensation. Wind power was then scrapped prematurely and old hydropower rebuilt to appear as new (Mauritzen, 2014). PJM, NYISO and ISO-NE make no distinction between new and old capacity, while such differences exist in California (Spees et al., 2013) and have been suggested for the German market (Öko-Institut et al., 2012).

Conclusion *An efficient capacity mechanism should not differentiate between new and old capacity.*

4.4 How large should the capacity payments be?

Let π^*_{LOLP} be the efficient probability of curtailment, which corresponds to the efficient total market capacity q^*. Let p^* be the efficient capacity payment that is needed to achieve q^*. Holmberg and Ritz (2020) argue that for a simplified electricity market, it does not matter whether the operator procures the volume q^* or sets a predetermined capacity payment p^*, the outcome will be the same. The efficient capacity payment in a simplified electricity market equals $\pi^*_{LOLP}(p_{VOLL} - \bar{p})$. With this capacity payment, the sum of the capacity payment and the expected scarcity rent will be independent of the price cap and equal to $\pi^*_{LOLP}(p_{VOLL} - \bar{c}^*)$ for the unit with the highest variable cost on the market, which ensures efficient investment incentives (Holmberg and Ritz, 2020).

It is difficult to give producers incentives to run plants for which the marginal cost is above the price cap.[11] Hence, a price cap below $\bar{c}^*$, the highest variable cost for which investment is efficient, would likely lead to inefficiencies. Holmberg and Ritz (2020) estimate that $\bar{c}^*$ is roughly 50-75% of p_{VOLL}. This rough calculation indicates that the price cap should not be set far below p_{VOLL} if one wants to avoid inefficient investment.

Conclusion *In a simplified electricity market, the socially optimal market-wide capacity payment is* $\pi^*_{LOLP}(p_{VOLL} - \bar{p})$.

4.5 Intermittent power and ramp rates

The simplified electricity market assumes that production is flexible and always available. Holmberg and Ritz (2020) study what happens when the share of intermittent renewable electricity increases in an otherwise simplified electricity market. They find that, in the long run, the risk for electricity shortage remains constant if price caps and capacity payments remain unchanged. However, there will be a switch from thermal base load to thermal peak load as more renewables enter the system. In the short term, such a transition can probably cause an increased risk of electricity shortage, for example related to non-coordinated plant closures and investments.

Conclusion *In the long run, additional renewables do not increase the risk of electricity shortages in an otherwise simplified electricity market. However, it changes the optimal technology mix, and this readjustment can lead to temporary capacity shortages.*

If the delivery periods are sufficiently short for fluctuations in electricity prices to internalise the system effects of fluctuations in renewable electricity production, then investments would be efficient also in the presence of intermittent generation and ramping costs in an energy-only market where the price cap is set at p_{VOLL}. But capacity markets would not generally yield efficient investment in such technologies, even if capacity payments were to follow the relation $\pi^*_{LOLP}(p_{VOLL} - \bar{p})$. The problem is that a lower price cap and a higher capacity

[11] Lambin (2020) argues that it would be possible to get efficient outcomes also when the price cap is below this level if tailor-made contracts are used. Tómasson et al. (2020) outline a design with uplift payments to producers so that they effectively face prices above the price cap, while consumer prices are capped.

payment reduce the price fluctuations in the electricity market. This is bad for energy storage and flexible production with short ramping times and low ramping costs that would benefit from price fluctuations. Lower price volatility is good for wind power, which would otherwise suffer more from the cannibalisation effect, namely that prices are high when wind-power output is low and vice versa. To minimise the distortions, the price cap should be as close to VOLL as is politically acceptable.

Conclusion *Capacity markets combined with a price cap reduce price fluctuations. This design favours intermittent production and disfavours flexible production and energy storage, The price cap should be set as close to p_{VOLL} as politically acceptable to maximise efficiency.*

The simplified electricity-market model does not consider the possibility of a system collapse. In practice, intermittent output from renewables is a challenge for the system. Holmberg and Ritz (2020) argue that the risk of a stressed situation evolving into a total system collapse increases with more wind power in the system. In that case, the security-of-supply externality will also increase. Hence, it would be optimal to increase the price cap or extend capacity mechanisms when more renewables enter the system. Newbery (2022) identifies an externality that occurs in power systems with a high wind-power penetration rate, where curtailment of wind power is frequent. Texas was close to a total system collapse in 2021, but this was due to a series of unexpected plant outages (Cramton, 2022), but not intermittent renewables.

4.6 What is the demand for capacity?

In a simplified electricity market, there is perfect competition, participants are fully informed, and producers are risk neutral. In this case, it does not matter if the procurement of capacity is done according to a predetermined price or a predetermined volume if the price and volume ultimately remain the same. In practice, however, competition and information are imperfect, and producers are risk averse. Then the design of the procurement plays a greater role.

Many capacity markets are neither price- nor volume-based, but something in between, where the procurement takes place along a specified demand curve. In Italy, the system operator Terna has designed a value-based demand for capacity that is calculated based on the estimated value of this capacity to consumers.

In a simplified electricity market, the demand for capacity can be derived from the following reasoning. Let q be the production capacity and let $\pi_{LOLP}(q)$ be the loss of load probability for that level of capacity. Adding more capacity would give consumers the marginal benefit $p_{VOLL} \cdot \pi_{LOLP}(q)$. On the spot market, consumers will, on the margin, pay the owner of an additional plant $\bar{p} \cdot \pi_{LOLP}(q)$. In the capacity market, the consumers would be willing to pay the difference between the marginal benefit and what they are paying in the spot market. Hence, we can conclude the following:

Conclusion: In a simplified electricity market, consumers' inverse demand for capacity is $p(q) = (p_{VOLL} - \bar{p}) \cdot \pi_{LOLP}(q)$.

Normally, demand outcomes would be far below market capacity. In such a market, $\pi_{LOLP}(q)$ would not decrease much if new capacity was added. Hence, a value-based demand curve would normally be rather elastic at market capacity.

But it is difficult to estimate p_{VOLL} and accordingly also consumers' valuation of capacity (Aagard and Kleit, 2022). Instead, most US markets have reliability targets and set demand in the capacity market, so that the target is met. Calculating reliability is a difficult task, but one with which engineers have decades of experience, since regulated utilities use essentially the same approach to decide how much capacity to build (Cramton et al., 2013). This is reminiscent of the EU regulation 2019/943, which says that capacity mechanisms are only allowed in EU countries where a resource-adequacy target is not met.

Nevertheless, many markets in the US have chosen to make procured capacity somewhat dependent on capacity payments. These demand curves are not based on consumers' valuation of capacity, but have been designed to, among other things, reduce price volatility and reduce the market participants' opportunities to exercise market power when capacity is being procured.

Price volatility in the capacity market is one of the biggest problems facing suppliers of capacity, according to Spees et al. (2013). To reduce the risk of investors, it has been suggested that the demand curve for capacity should be convex (Hobbs et al., 2005, 2007; Stoft et al., 2004, 2005). By implication, capacity payments would be relatively large even if the supply of capacity was large, and the procured capacity would be relatively large even if the capacity price was high. ISO-NE has a demand curve that is reminiscent of this form, although they allow the procured capacity to be low if the price of capacity is very high. The design of ISO-NE's capacity market stabilises prices, but it also leads to significant inefficiencies (Spees et al., 2013). In addition, consumers do not necessarily value such price stability. If new capacity is willing to enter the market even at small capacity payments, this should have an effect on prices, according to the consumer side (APPA, 2017).

All capacity markets we know of have a reservation price. This means that, in practice, the demand for capacity is partly concave, at least at high prices. This corresponds to the procurer wanting to protect the consumer side against very high prices.

4.7 Which auction format is best?

Procurement of capacity normally uses marginal pricing, pay-as-bid or Vickrey pricing. In a simplified electricity market the outcome and the costs for consumers will be the same regardless which of these price mechanisms is used. This is approximately the case also in practice for electricity spot markets and treasury auctions, where competition is reasonably well functioning (Holmberg and Newbery, 2010). Imperfect competition is often a larger problem when procuring capacity. For a market-wide capacity market, the underlying problem is that if the procurer wants to buy almost all existing capacity in the market, and there is little entry of new capacity, then a seller can push the price to the price cap by withholding a fraction of its volume from the market. If the procurer was fully informed of investors' costs, then it could just set a capacity payment or a very tight cap for the capacity

price. In this case, sellers of capacity could not influence the capacity price, even if they were to have market power. Unfortunately, the procurer is rarely fully informed in practice. Typically, capacity is procured in auctions where bidders have a significant influence on the capacity price. Under these circumstances it matters how the price is set.

Marginal pricing is normally used in electricity spot markets, and sometimes also when procuring capacity. Marginal pricing means that the highest accepted bid sets a market price that is paid to all accepted bids. As long as the bidding equilibrium is well behaved, such a pricing mechanism has several benefits:

1) Bids are not that sensitive to uncertainties in the market.
2) It simplifies the bidding process for small businesses. For them, it is optimal to simply offer the marginal cost of capacity.
3) The market price is well defined.

A well-defined strike price is, for example, an advantage if there is forward trading in the capacity market. The main problem with marginal pricing when procuring capacity is that bidding may not be well behaved. There is the possibility of getting prices at the collusive level. For example, there could be an equilibrium where a dominant firm places a bid at the price cap, and other smaller firms place very low bids. Still, the small firms would sell their entire capacity at a price equal to the price cap. The dominant firm would have to lower its bid substantially to increase sales, which would be unprofitable. The risk that the market ends up in such a high-price equilibrium is particularly large in markets with a dominant firm and when the uncertainty in the market is small, so that it is possible to predict in advance which bid will set the price. The problem was described theoretically by von der Fehr and Harbord (1993) and has also been observed in NYISO's procurement of capacity (Schwenen, 2015). For some scenarios, Teirilä and Ritz (2019) get the high-price equilibrium in their simulations of the Irish capacity market, which has a dominant producer. Corresponding problems arose in Colombia's capacity market both in 2008 and 2011 (Harbord and Pagnozzi, 2014). Consequently, several of these markets have taken measures to reduce the problem. Colombia has tried to make it more difficult to predict in advance which bid will set the price by revealing less information between bidding rounds, although this did not help (Harbord and Pagnozzi, 2014). Harbord and Pagnozzi (2014) have suggested that the demand for capacity should partly be random. As discussed in the previous section, the procurer can make the procured volume sensitive to the price, in which case it would be more difficult to exercise market power. In Section 4.10, we discuss how entry can be increased, which would make bidding more competitive.

Conclusion *Marginal pricing has many advantages, but can sometimes lead to extremely high prices, especially if there is a dominant agent in the market, uncertainty is low, the demand for capacity is inelastic and there is little entry of new capacity.*

Pay-as-bid is sometimes used when capacity is procured. It means that any accepted bid is paid according to the own bid price. Under this setup, all accepted bids set the price for each firm. As many bids are price setting, this constrains the range of possible equilibria. This reduces the risk of outcomes with very high prices (Fabra et al., 2006; Pycia and Woodward,

2019). For example, the high-price equilibrium is not an equilibrium for pay-as-bid auctions. A main problem with pay-as-bid is that each firm has an incentive to raise its bid until it is just accepted. Hence, the bid prices will be very similar, regardless of whether the bidder has high or low costs. Thus, small errors can have major consequences for the outcome. For example, the allocation may be inefficient, such that an agent with a high cost of offering capacity wins the procurement, while an agent who can offer capacity at a lower cost (but bids too high) may not sell any capacity. Anderson et al. (2013) show that this problem is exacerbated by the fact that bidding can become volatile and unpredictable in a procurement that applies pay-as-bid. Furthermore, the agents are more dependent on accurately forecasting the outcome of the auction, so that they can maximise their profit. This increases the costs of participating in a capacity auction, which disadvantages small firms.

Conclusion *Pay-as-bid reduces the risk of high prices, but capacity procurement can be inefficient and disadvantage small firms.*

Power plants are often heterogeneous in terms of ramping rates or their location in the network. If there are significant differences between the plants, it will be inefficient to define a single market price that applies to all procured capacity. It would be theoretically possible to define a market price for each product category, but it may be easier for accepted bids to be paid-as-bid instead. Furthermore, the bidding behaviour changes when the generation units are imperfect substitutes, which should reduce the problems of pay-as-bid pricing.

A third alternative is to combine marginal pricing and pay-as-bid pricing. The design has been studied theoretically by Ruddell et al. (2017) and Woodward (2019). New Zealand had plans to introduce the combination in the spot market for electricity, but it has never been tested in practice. For example, 80% of the payment could be according to the marginal price and 20% according to the bid. In this case, every accepted bid would be partly price setting. In theory, this should rule out the high-price equilibrium outlined by von der Fehr and Harbord (1993). Once this bad equilibrium has been ruled out, bidding should be similar to well-behaved bids in an auction with marginal pricing, which would lower the risk of inefficient outcomes (a potential problem of pay-as-bid auctions). Hence, the combination would seemingly avoid the worst possible outcomes, high prices or inefficient allocations. Another advantage is that it is easy to adjust the shares for the two pricing methods if any problems should arise after all.

Auctions with a Vickrey–Clarke–Groves (VCG) design are another option (Ausubel and Milgrom, 2006). In this setup, firms are paid the procurer's opportunity cost. This is the additional amount the procurer would have to pay if the firm did not participate in the auction. Bidding the true marginal cost is a weakly-dominating strategy under this mechanism. For this equilibrium, there will be no mark-ups on the bids, not even from agents with market power. The allocation of capacity will then be efficient. However, the transaction prices will still be above the marginal cost, as for other auction designs. Similar to marginal pricing, there can also be equilibria with prices at the collusive level (Blume and Heidhues, 2004; Blume et al., 2009). Another challenge is that there will be different prices for different bidders. This problem also exists with pay-as-bid, but for VCG the price is systematically higher for producers with market power because these firms require higher compensation to

make truthful offers. This can be perceived as unfair and in the long run can also encourage smaller producers to merge, which worsens the competition in the market. Problems can also arise if a firm prefers competitors to earn as little as possible from the auction and is able to roughly predict which bids will be accepted. The firm can then strategically choose its own non-accepted bids in such a way that the competitors are paid a low price for their capacity while the bidder itself loses nothing. This problem occurred in European spectrum auctions from 2010 to 2012 (Fanebust and von der Fehr, 2013). The consequence was that different firms had to pay very different prices for similar licenses.

Conclusion *VCG auctions are efficient in theory, but prices can be very high in practice. Small firms are systematically paid less than large firms.*

4.8 Who should procure capacity?

System operators are independent in the US and the UK, and it is therefore appropriate that they handle the procurement of capacity. This arrangement is often referred to as the central-buyer approach. Things are more complicated in the EU because the system operators also own the transmission network and are therefore affected by the prices in the spot market. In these countries, it may be better to let another party handle the procurement of production capacity. This is especially true for a market-wide capacity market, where large volumes are procured which will affect spot prices and the system operator's rents.

In some US markets, CAISO and SPP, and in France, retailers purchase capacity on behalf of their customers. This decentralised capacity market is said to be of the *capacity obligation* type. In the EU, the approach is often referred to as decentral obligation. This arrangement could be suitable for countries where the system operator owns the grid. Another potential advantage with decentralisation is that retailers might be better at making use of the demand response, which reduces the need for capacity (Neuhoff et al., 2016). Yet, the synergies of an integrated market may be lost under decentralisation: the peak loads of different retailers do generally not coincide perfectly, meaning that the overall peak load is somewhat smaller than the sum of the peak loads of all retailers (Neuhoff et al., 2016).

4.9 What information should firms receive during/after procurement?

Under the simplified market assumption, producers would be fully informed. In this case, the information structure of the procurement process does not matter. In reality, investment costs and future electricity prices are uncertain when capacity is procured. These uncertainties are largely common to investors, at least to those that consider investing in similar technologies. In the latter case, both investment costs and the spot market revenues of investors can be expected to be strongly positively correlated. All investments of the same type will therefore have roughly the same profitability, although it is uncertain just how profitable they will be. Under these conditions, there are advantages to investors learning from each other. With better information about competitors' bidding, each bidder can make a better estimate of its own investment costs and better estimate the future revenue in the spot market. This will reduce the winner's curse, so that each supplier of capacity can bid more confidently (Aagard

and Kleit, 2022). One way of revealing such information is through short delivery periods for capacity and frequent procurements, so that firms learn from the outcome of each procurement. Another way is that procurement is preceded by trading in financial instruments, where agents are given the opportunity to secure a future capacity price. The price of such a financial product is based on the information gathered in the market, and it is thus informative for investors. A third way is to arrange a dynamic procurement with several bidding rounds. After each bidding round, information is disclosed to the agents. Usually only aggregate information is revealed. Agents do not receive any detailed information about individual competitors' bids. To improve the exchange of information, Cramton and Stoft (2007) recommend that bid data from the previous round should be reported by production technology. Ideally, the increased information will lead to bids that better reflect costs, so that the most suitable capacity wins the auction. Under favourable conditions, more public information improves competition (Milgrom and Weber, 1982; Holmberg and Wolak, 2018).

Harbord and Pagnozzi (2014) are critical of dynamic auctions in the capacity market. They claim that firms investing in different technologies cannot learn much from each other. Based on interviews with bidders in Colombia's capacity market, they argue that the bidders' lowest acceptable price was not affected by the information they received during the bidding process. Moreover, there are disadvantages associated with improving the information flow during an auction. When ISO-NE procures capacity, the auction can last for five days with eight bidding rounds per day. This means that the procurement process will be very time-consuming and costly, both for ISO-NE and for the market participants. It becomes particularly complicated for smaller agents to participate in such an auction. In addition, there is a larger risk that firms coordinate their bids in a dynamic auction, which in that case leads to higher prices.

Conclusion *Frequent procurement of capacity, or procurement in a dynamic auction, makes bidders more informed and can provide more efficient outcomes. However, the procurement process becomes more costly and the risk of collusion increases.*

4.10 How long in advance should capacity be procured?

Under the simplified market assumption, capacity can freely enter and exit without any delay. But, in practice, it can take years to build new production capacity. This imperfection matters for the decision of how long in advance to procure capacity. With a long time frame, suppliers will have the time to build new capacity or upgrade existing capacity, after the auction has ended and before capacity is to be delivered. By implication, the supply of capacity will be more flexible and competition more intense (Chao and Wilson, 2004). It will also be easier to coordinate investments if they are procured well in advance. At PJM and ISO-NE, capacity is traded up to three years ahead of delivery (Spees et al., 2013). Cramton (2006) argues that in the ideal case, the procurement should be made so far in advance that all technologies have the chance to build new capacity. In practice, however, it is difficult to procure technologies with different construction times in the same auction (Batlle et al., 2015). In South America, there are often three different auctions with one-, three- and up to 20-year deadlines for delivery of capacity, where the latter is targeted at new hydro power (Harbord and Pagnozzi, 2008; Batlle et al., 2015). PJM and ISO-NE have auctions where the procured capacity can be

corrected every year (Spees et al., 2013). California, MISO and NYISO have instead made the assessment that it is sufficient that capacity is procured a few months or days before delivery (Spees et al., 2013). Harvey et al. (2013) believe that this has contributed to poor competition in NYISO's procurement of capacity. However, an advantage of purchasing capacity just before delivery is that it becomes easier to estimate the availability of plants.

A related issue is the length of the delivery period, i.e., for how long a plant should commit to be available. In PJM, ISO-NE and MISO, the delivery period for capacity is one month and the procured capacity is the same for each month. In California, the delivery period is also one month, but the procured capacity changes every month. The delivery periods are much longer in South America. There, the delivery period is often several years and can be up to 30 years (Batlle, 2015). A problem with long delivery periods and capacity procured long in advance is that it becomes difficult to estimate the firm capacity for hydro power and demand response.

Conclusion *Procuring capacity long in advance improves the competition in the auction, but makes it difficult to estimate the firm capacity of hydro power and demand response.*

5 The strategic reserve

Belgium, Finland, Germany, Sweden and Texas, use a *strategic reserve* instead of market-wide capacity payments. The motives for strategic reserves are roughly the same as for capacity markets; to ensure the security of supply while limiting the price risk, or to provide robustness to temporary deviations from the long-term market equilibrium.

Holmberg and Ritz (2020) show how an efficient strategic reserve should be designed for a simplified electricity market. Like a capacity market, plants in the strategic reserve should receive a fixed support in proportion to the capacity of the plant. When procuring capacity, bids should be accepted from those who are willing to offer capacity at the lowest price per MW (irrespective of their marginal cost). For an efficient reserve, the efficient capacity payment is $p^* = \pi^*_{LOLP}(p_{VOLL} - \bar{p})$, just as for market-wide capacity payments. Hence, if the price cap $\bar{p}$ has been set at p_{VOLL}, then there is no need for a strategic reserve in a simplified electricity market. For a simplified electricity market, it also does not matter whether the capacity price is predetermined to p^*, or if an amount $\Delta q = q^* - q^0$ is procured to the reserve, where q^0 is the market capacity that is active outside the reserve and q^* is the efficient total capacity. Hence, implicitly the expression for the capacity payment determines the optimal size of the reserve. Consumers' marginal value of capacity can be calculated as in Section 4.6.

The strategic reserve should only be used when other production is insufficient, and there is a threat of power shortage. In that situation, the spot price should be set at the price cap.[12] Thus, for a given price cap and capacity outside the reserve, the price for plants outside the reserve

[12] Neuhoff et al. (2016) argue that there should be a trigger price, below the price cap, where the strategic reserve should be activated. The motivation is that it is important to avoid price spikes to increase the social acceptance of the electricity market design. We think that it would be inefficient to have such a trigger price below the price cap. If social acceptance of the design is a major problem, it would be better to lower the price cap (when possible), instead of introducing a trigger price.

is independent of the size of the strategic reserve. It does not matter for the price of other plants if the reserve is able to meet the rest of the demand or if there is a shortage of electricity. In this way, the rest of the market is isolated from the strategic reserve. For plants outside the reserve, the electricity market essentially works as an energy-only market.

Holmberg and Ritz (2020) advocate that units in the power reserve should be paid according to a market price for reserve power when they are used. That is, all production in the reserve is paid a price that is set by the running plant in the reserve that has the highest variable cost. Plants outside the reserve do not receive any capacity payment. On the other hand, they benefit from a high spot price, at the price cap whenever the power reserve is used. Holmberg and Ritz (2020) show that in the market equilibrium of a simplified electricity market, the expected value of this extra income is equal to the capacity payment. This means that the power reserve is equivalent to a capacity market, at least for a simplified electricity market.

Conclusion *In a simplified electricity market, the strategic reserve is efficient if the following conditions are met: 1) the price cap $\bar{p}$ is higher than the efficient cutoff $\bar{c}^*$ and below p_{VOLL}; 2) units in the reserve receive the efficient capacity payment $\pi^*_{LOLP}(p_{VOLL} - \bar{p})$; 3) the reserve is used only when the capacity in the rest of the market is exhausted; 4) the spot price is set at the price cap as soon as the power reserve is used, and 5) energy produced by the reserve is paid the marginal price of reserve power.*

It may be interesting to compare the strategic reserve with a capacity market in more detail. Consider a strategic reserve and capacity market with the same price cap and the same level of payment per capacity unit (plants outside the strategic reserve do not receive this payment). For a simplified electricity market, the total capacity and the technology mix will be the same in both markets. The total capacity payments are higher in a capacity market because all generation units receive a capacity payment. This is exactly mimicked by the fact that producers' total revenues in the spot market are higher in a market with a strategic reserve. The spot price is the same in both markets provided there is sufficient capacity outside the reserve or if there is a shortage of electricity. The difference in the spot price arises when the reserve has been activated without a power shortage. In that situation, the spot price for production outside the reserve is set at the price cap, while a market with market-wide capacity payments has a spot price that is equal to the marginal price of the plants in the reserve.

Conclusion *In a simplified electricity market with a strategic reserve, producers receive larger total revenues from the spot market and smaller total capacity revenues compared to a market-wide capacity mechanism that has the same price cap and capacity payment. Producers' total revenues, and consumers' total costs, will be equal in both markets.*

For the simplified electricity market, we can also compare efficient strategic reserves and capacity markets with an efficient energy-only market. In both capacity mechanisms, the spot price is lower than the energy-only market when there is electricity shortage (loss of load). In capacity markets, this is compensated by a capacity payment. A similar payment is given to plants in the strategic reserve, whereas plants outside the reserve get an extra high spot price,

higher than in the energy-only market, when the reserve has been activated and there is no loss of load.

5.1 Advantages of a strategic reserve versus a capacity market

Strategic reserves feature several of the same advantages and disadvantages as we observed in our discussion of capacity markets in Section 4. There are also notable differences. One is that capacity payments are paid only to a fraction of the generation units in a strategic reserve. Only those units with the highest variable costs and the lowest utilisation rate would participate in the reserve, as it is only used when all other capacity is insufficient. Reserve capacity will mostly consist of thermal peak power. Hence, it will be straightforward to define firm capacity for a large part of the units that will receive a capacity payment. Demand response is another technology that is activated at peak electricity prices. A problem of including demand response in the strategic reserve is the complicated task of defining and verifying firm capacity, the same as for market-wide capacity mechanisms. The demand response is energy limited in the sense that reducing demand for long periods of time is inefficient. Moreover, for consumers in Sweden, it has been difficult to commit demand response long in advance and for long periods of time. Hence, we would argue that the demand response is unsuitable for inclusion in a strategic reserve.

Normally, renewable electricity production, hydro power and energy storage have a relatively low variable cost, or opportunity cost.[13] Thus, it is rarely interesting for these technologies to participate in the reserve. An advantage compared to capacity markets is that there is no need to define a firm capacity and to incentivise the availability of these technologies. They can operate outside the reserve under essentially the same conditions as in an energy-only market.

Conclusion *In an electricity market with a strategic reserve, it is relatively easy to define firm capacity for units that receive capacity payments, unless the demand response is included in the strategic reserve.*

A consequence of this result is that price signals will be more precise for individual plants inside and outside the strategic reserve, compared to a capacity market. Moreover, plants outside the strategic reserve will have economic incentives to be available when capacity is most needed. A potential advantage of procuring a smaller volume is that even if the procurement is primarily targeted at plants with high variable costs, the large supply of other capacity should, in theory, ensure viable competition in the procurement.

An additional potential advantage of the reserve is that it is isolated from the spot market if the market price is set at the price cap whenever there is a supply shortage in the spot market. The size of a reserve will then have no effect on spot prices in the short or long run. Such minimal market impact is an advantage if the TSO is responsible for procuring the reserve.

[13] The opportunity cost of energy storage can occasionally be very high, but on average it is not high during a period of several months.

Under a strategic reserve, market participants are still responsible for forecasting demand and choosing an appropriate investment portfolio. The reserve provides a margin, reduces the risk of electricity shortages, but in theory it should not affect other investments, at least not in the long run.

Neuhoff et al. (2016) argue that distortions occurring because of special interests, short-termism or excessive risk aversion of political actors can occur in all types of capacity mechanisms. Such concerns appear to be particularly relevant for capacity markets because of their large volumes.

5.2 Disadvantages of a strategic reserve versus a capacity market

A strategic reserve has some potential drawbacks as compared to a capacity market. One is that there can be situations where it would have been efficient to use a plant in the reserve before a plant outside the reserve (Bublitz et al., 2019). This cannot occur under the simplified-market assumption in Holmberg and Ritz (2020), but it can happen under more realistic assumptions. For example, if the two plants have different locations and/or different flexibility, it may depend on the market conditions which plant it would be efficient to dispatch first. The merit order of units may also change due to changing fuel prices or fluctuating opportunity costs. This is a problem if the units that are most suitable for the reserve vary during the period when the reserve is active.

For the same price cap, the reserve has a higher spot price (when the reserve is used) compared to a capacity market. This means that, for the same price cap, the strategic reserve does not have the same mitigating effect on risks. A higher risk means that investments are likely to shift towards low-capital investments with short lead times (Bublitz et al., 2019).

Arguably, a strategic reserve is less effective in mitigating market power compared to a capacity market. However, the strategic reserve may have relatively smaller problems with market power in the procurement phase. If risk or market power is a problem for a market with a strategic reserve, then it can partly be mitigated by lowering the price cap on the spot market and increasing the capacity payment and procured volume of the reserve.

In some markets authorities set a level of reliability in the system and procure a capacity corresponding to that level. This will also reduce the investment cycles (Bublitz et al., 2019). Yet, authorities do not have the same control over the total capacity in an electricity market with a strategic reserve. In effect, a spot market complemented by a strategic reserve is an intermediary between an energy-only market and a capacity market.

Ideally, investments outside the reserve will be the same regardless of the size of the strategic reserve. In practice, however, there is long-run uncertainty about the size of the reserve, which adds investment risk. The advantages and disadvantages of the strategic reserve can be summarised as follows:

Conclusion *A strategic reserve yields fewer distortions for intermittent generation, and energy storage compared to a capacity market that has price caps and capacity payments at*

the same level, unless the market has significant issues with risk, investment cycles or large time variations in fuel costs and opportunity costs of plants.

5.3 Strategic reserves in an integrated electricity market

Analyses of capacity mechanisms are usually conducted under the assumption of a national electricity market. In an integrated multinational electricity market, the level of strategic reserves chosen in one country could have effects abroad through the effects on prices and resource constraints. On the one hand, a larger domestic reserve has a positive effect abroad because it increases the overall reserve capacity. On the other hand, the domestic reserve can distort market prices and investments abroad. Tangerås (2018) shows that strategic reserves can be too large or too small in equilibrium depending on which effect dominates. That paper also demonstrates that the international externalities associated with strategic reserves will lead to underinvestment in network reliability even if network investments are coordinated across national borders. Underinvestment is exacerbated if the countries invest in network capacity based on domestic considerations.

If the EU sets a common price cap for all member states, then the price distortions of the strategic reserves disappear. This should leave only the positive effect associated with capacity reserves being available for uses abroad. By implication, strategic reserves in EU are likely to be too small. EU Regulation 2019:943 states that countries are only allowed to introduce a strategic reserve if it can be justified based on a detailed analysis of the risk of power shortages. This supra-national approach to capacity reserves makes sense because of their cross-border effects. The welfare effect of strategic reserves is larger if their utilisation is coordinated across borders, compared to the case where they are entirely used for national purposes. Neuhoff et al. (2016) argue that such coordination is both beneficial and feasible.

5.4 The strategic reserve in Sweden

The deregulation of the electricity market in Sweden in 1996 caused electricity consumption to increase and unprofitable production to shut down (Swedish Government, 2009). On top of that, the government decided to close the remaining nuclear power plant in Barsebäck. The Swedish TSO, Svenska Kraftnät (SvK), was instructed to procure a strategic reserve to cope with the shut-down of nuclear power without increasing the risk of power shortage. At the outset, this was a small preliminary reserve of 400-600 MW thermal power (SvK, 2013). The reserve became statutory in 2003, after which the size increased to 2 000 MW. From the beginning, the idea was that the reserve should constitute an intermediary solution until 2008, as the market was expected to catch up in the long run. However, the deadline of the capacity reserve has been extended on three occasions, most recently until 2025. Between 2011 and 2017, the reserve was reduced to 750 MW. Due to the phase-out of additional nuclear-power plants in southern Sweden and domestic network congestion, SvK currently only procures capacity in southern Sweden. The reserve was activated on approximately ten occasions

during the years 2009-13. On these occasions, up to 826 MW of the reserve were used (SvK, 2013).[14] Curtailment has never occurred in Sweden.

The government's long-run ambition is to phase out the reserve and leave capacity decisions entirely to the market. According to the Swedish regulator, the Energy Markets Inspectorate, this energy-only market could be achieved with additional bidding zones and improved market integration (EMI, 2008). SvK (2013) estimates that the market reform in 2011, when the number of bidding zones in Sweden increased from one to four, increased the trade with neighbouring countries, which indeed reduced the need for a strategic reserve. EMI (2008) also proposed to increase the share of the demand response in the strategic reserve. In the short term, this would increase the procurement cost, but the change should have long-term benefits by stimulating demand flexibility (Swedish Government, 2009). According to Government Regulation 2010:2004 on strategic reserves, demand flexibility in the reserve was expected to increase from 25% to 100% during 2017/2018. However, it was impossible to achieve this objective, and in 2014 the regulation was relaxed so that at least 25% of the reserve should be demand response.

An issue with demand response is that it typically has limited endurance. It can be feasible to reduce electricity consumption for a few hours, but it will often be very costly to shut down industrial production for a prolonged period of time. An advantage of demand response is that activation is fast (SvK, 2013). For thermal production capacity in the reserve, it is usually the opposite, this capacity is enduring but needs to be notified well in advance (SvK, 2013).

All reserve capacity is required to be available 95% of the time, or else the capacity payment is reduced. For demand response, this means that a consumer must commit to using more than the sold capacity 95 % of the time. For energy-intensive industries, it has proven difficult to make such a capacity commitment six months in advance, when the reserve was procured. Hence, SvK has moved the procurement of the demand response closer to the delivery period to facilitate the participation by the electricity-intensive industry. Moreover, the demand response was allowed to leave the reserve temporarily if it is active on the spot market, but it will not receive any capacity payment for such periods. There is some concern whether this arrangement is consistent with the new EU regulation for strategic reserves (EMI, 2020), so SvK has suspended the procurement of the demand response to the strategic reserve.

The procured capacity is reserved only for the winter period. There is only one bidding round, and the auction is executed approximately six months before the start of the winter period. Each generation unit in the reserve receives a fixed compensation equal to its own bid. Generation units that participate in the procurement auction are heterogeneous, and the bids are therefore ranked according to a scoring rule. The scoring rule admits bids from all types of generation units, but it assigns a low score to units that fail to meet all technical performance and environmental specifications. The scoring rule also penalises plants that ask for a high variable compensation. In addition, variable compensation is paid upon activation. For

[14] In some cases, it would have been possible to use power outside the reserve (SvK, 2013). SvK estimates that a reserve of 400 MW would probably have been enough to avoid a power shortage during 2009-13.

production capacity, this is done according to the bids (pay as bid). Energy from the demand response is usually accepted via the balancing market[15] and is paid the market price of that market.

Under the current design, the spot price is set at the price cap as soon as there is excess demand in the day-ahead market. This means that the spot price is not affected by the size of the reserve, at least not in the short term. This procedure is in line with the design of the strategic reserve advocated by Holmberg and Ritz (2020).

In 2016, the government decided that SvK should take environmental aspects into account in the procurement of the strategic reserve (Proposition 2015/16:117). The entire reserve must now consist of renewable generation capacity.[16] By implication, most units that previously participated in the reserve must be converted into biofuel plants. According to one study, this would increase the variable costs by roughly 30%, and the total costs by 10-20% (Ceije, 2016). A concern has also been that the increased environmental requirements will reduce the supply of reserve capacity. These problems may possibly be a partial explanation for the recent lack of competition in the procurement auction. For example, a supplementary procurement had to be suspended during the winter of 2019/2020 when only one firm submitted a bid. We also note that in recent years, all capacity has been procured from one specific plant, *Karlshamnverket*, since the competing *Mälarenergi* and *Stenungsund* power plants have been closed.

5.5 Strategic reserves in other countries

Finland, Belgium and Germany introduced strategic reserves in 2011, 2014 and 2020.[17] The reserves in Belgium and Germany were partly motivated by a phase out of nuclear power. Similar to the Swedish case, reserves are only procured for the winter season, and are only activated if there is a shortage of electricity, i.e., when the spot price reaches the price cap. All reserves have, or at least allow for, both thermal production and demand response. There are also differences relative to the Swedish design. In both Belgium and Germany, a plant cannot return to the market after it has been included in the reserve.[18] One purpose is to isolate the reserve from the rest of the market. In addition, German plants can only stay in the reserve for a couple of years before they must close. This policy seems overly restrictive, at least from a resource adequacy perspective. The size of the reserve should be predictable, but this does not mean that the plants in the reserve should be fixed. We would argue that restrictions of this type would make it unattractive to enter the reserve. This could reduce market efficiency and

[15] The system operator (SvK) uses the balancing market to get supply and demand in balance when electricity is delivered.

[16] SvK has the possibility to deviate from the environmental requirement if this drastically reduces the cost of procurement. However, the new EU regulation also contains environmental requirements. In the future, production in the strategic reserve may emit at most 550 g of CO_2 per kWh of electricity produced (EMI, 2020).

[17] In Germany, the procurement of capacity took place in 2019, but the delivery of capacity started in 2020.

[18] Germany has also introduced two other reserves. One is targeted to retiring coal plants, and the other deals with redispatch. Bolton and Claussen (2019) discuss the three reserves in Germany and the political process behind them. The reserve in Belgium is briefly discussed by Höschle and de Vos (2016).

worsen competition when capacity is procured. Moreover, it should be noted that if a strategic reserve is well designed, then it will (under the simplified market assumption) result in efficient investments for each technology, including for plants in the reserve. Plants in the reserve are not paid a subsidy. They are just compensated for a smaller price cap. From this perspective, it makes sense that plants in the reserve should be free to be active on the market once their commitment to the reserve has ended.

The Emergency Response Service (ERS) is a strategic reserve in Texas. It includes two types of products. For the 10-minute [30-minute] ERS, demand must be curtailed within 10 [30] minutes after notification. The procured capacity for each product changes by time of day and season. Demand resources enrolled under ERS are dispatchable by ERCOT during an emergency but cannot be called outside their contracted hours and cannot be called for more than twelve hours total per season.

5.6 Can Sweden's strategic reserve be improved?

In Section 4, we discussed how capacity markets should be designed. Many of the results also apply to strategic reserves. It is an advantage that Sweden's strategic reserve mainly contains enduring thermal production capacity that can provide energy for a long period of time. It is useful to have a fast demand response as a complement. But, at least for Sweden, we think that it would be better to procure non-enduring technologies as operating reserves. An advantage with such reserves is that the procurement of capacity is done close to delivery and delivery periods are short, so that it becomes easier for consumers to estimate their demand-response capacity. In Sweden, the demand response within the strategic reserve is already activated in the balancing market (one of the markets for operating reserves), so our suggestion should not be a big step. Moreover, such a change should mean that the demand response could continue to be active in the spot market without risking that the Swedish design would violate EU law.

In this paper, we argue that the size of the strategic reserve should be chosen such that the capacity payment is equal to $p^* = \pi^*_{LOLP}(p_{VOLL} - \bar{p})$, i.e., the capacity payment should compensate for a small price cap. This is not quite how it works in Sweden. The size of the reserve is regulated in the law and is based on resource adequacy studies by SvK and the energy regulator (EMI, 2008). The current plan is to phase out the strategic reserve by 2025. The Swedish energy regulator recommends that the loss of load expectation (LOLE) should be less than 0.99 hours per year (EMI, 2021) in Sweden. Moreover, their view, which is in line with the EU regulation 2019/943, is that keeping the strategic reserve beyond 2025 would only be possible if the resource adequacy target cannot be met without the reserve.

Competition is improved if the demand for reserves is price sensitive. There is legal room for this demand feature. SvK has the possibility to procure less capacity than the law prescribes if the price for capacity becomes too high, which introduces some elasticity in the demand for capacity. The supply of reserve capacity would probably increase, and competition improve, if reserves were procured further in advance. This would give winners more time to modify their plants.

Marginal pricing has many advantages, and it is used in the design that Holmberg and Ritz (2020) outline for a simplified electricity market. But it is likely that other designs such as pay-as-bid would also be efficient for a simplified electricity market. Experience from capacity mechanisms outside Sweden suggests that marginal pricing can be problematic when procuring capacity. Another issue with marginal pricing is that it is probably inefficient to use a single market price for all plants when plants have somewhat different properties and locations in the network. Thus, it is not obvious that marginal pricing would be better than pay-as-bid, which is currently used by SvK.

For a simplified electricity market, it would be efficient to procure the volume such that plants asking for the lowest capacity price are accepted first and to disregard the level of variable compensation that a plant asks for in case of activation. But this result applies to situations with perfect competition, where the reported variable compensation is truthful. Under imperfect competition, it can probably be a good idea to give priority to bids with a low variable cost, which is the case today in Sweden.

The strategic reserve is only procured for the winter months. Hence, a corresponding energy-only market would have a high price cap during the winter and a lower cap during the summer. Most houses in Sweden use electric heating, so it might very well be the case that the value of lost load is higher when it is cold outdoors. If so, it would make some sense that the design changes during the winter.

Environmental requirements are stricter for production within the reserve compared to production outside. But reserves are used so infrequently that the environment would hardly be affected if the government dropped the requirement that 100% of the reserve must be renewable. From an environmental perspective, it would make more sense to have stricter rules for plants used on a regular basis (Neuhoff et al., 2016). In addition, the increased environmental requirements for the strategic reserve seem to have worsened competition since many facilities fail to meet those requirements. Relaxing the regulation would increase the supply of capacity and make the procurement more competitive.

The Swedish strategic reserve has contributed to the ability of the electricity system to absorb various unexpected political decisions. There are probably short-term benefits to this flexibility, but the energy policy would perhaps have been more predictable if there had been no reserve to fall back on. There has generally been a great deal of experimenting with the reserve regarding its size and duration, the proportion of demand response and the environmental requirements. This experimentation has led to increased political uncertainty and increased investment costs.

6 Conclusions

Energy-only is the most efficient electricity market design. In practice, many markets use complementary capacity mechanisms. One reason is that it takes a long time to build new capacity, so temporary capacity shortages can occur due to investment cycles, new regulations, and technology shifts. Investment cycles, simultaneous closure of ageing power

plants, combined with stricter environmental regulation, were part of the reason why the UK decided to introduce a capacity market (Bolton and Clausen, 2019). The strategic reserves in Belgium, Germany and Sweden have facilitated the phase-out of nuclear power. Investments in renewable technologies, mainly driven by support policies, have reduced electricity prices. Portugal and Spain have introduced capacity mechanisms to prevent thermal production from being pushed out of the market. A problem with US market designs has been that price caps tend to be low, partly for political reasons. Capacity payments have been used to compensate for the resulting missing money problem. Moreover, capacity markets in the US are often used to implement reliability standards. Low price caps reduce price volatility, which is useful in countries where financial markets are less developed, such as South America. A small price cap has reduced the risk of bankruptcy among market participants during energy crises in the EU and the US.

This paper has focused on two types of capacity mechanisms, capacity markets with market-wide capacity payments, and strategic reserves for which capacity payments are targeted to a few selected plants. For an idealised benchmark market, well-designed strategic reserves and capacity markets would both be as efficient as an energy-only market. But, in practice, capacity mechanisms introduce distortions. Defining firm capacity is straightforward for thermal production, but it is hard to define firm capacities that give correct price signals for each individual unit with intermittent production, demand response, and energy storage (including hydro power). Another issue with capacity markets is the bureaucratic process of verifying and approving capacity. This is mainly a problem for small plants in decentralised markets. In centralised markets, plants and their properties are registered anyway. Moreover, distortions can occur due to special interests, short-termism or excessive risk aversion of political actors.

An advantage of a strategic reserve is that only plants high up in the merit order, normally thermal plants, would be interested in contributing to a strategic reserve. For such plants, it is straightforward to define firm capacity. Moreover, the bureaucratic process of verifying and approving capacity is reduced to a small number of generation units. If the price cap is set at a federal level in the EU, then strategic reserves should not cause any price distortions in neighbouring countries.

This paper has discussed the strategic reserve in Sweden in detail. The main problem is how to manage the demand response in the reserve. It has been difficult for consumers to commit their demand response for a long period of time. To give them more flexibility, they have effectively been allowed to temporarily leave the reserve and instead be active on the spot market, which may violate EU regulations. Our suggestion for Sweden is to instead transfer the demand response to the operating reserves. The demand response is fast, but not enduring, so it should be more suitable for such reserves. This would also mean that the demand response would be procured closer to delivery and have shorter delivery periods. This should make it easier for consumers to predict how much and for how long they can reduce their consumption, i.e., to determine firm capacity. Moreover, the demand response could then continue to be active in the spot market without risking that the Swedish market design would violate EU law.

Imperfect competition is a major problem when procuring capacity. In theory, this problem should be smaller for strategic reserves, which only constitute a small fraction of the total capacity in the market. But many countries place additional constraints on plants in the reserve, which seems inefficient from a resource adequacy perspective, and it stifles competition. The Swedish market has faced problems of limited competition when procuring thermal capacity for the strategic reserve. These problems seem to have become worse after the requirement that all capacity in the reserve must be 100% renewable was implemented. We suggest that this rule be relaxed, and that thermal capacity is procured longer in advance to increase the supply.

Capacity markets have advantages if the merit order of units is changing due to changing fuel prices or fluctuating opportunity costs. Which plants are most suitable for a strategic reserve, then varies over time. In addition, volume-based capacity markets are better at stabilising the capacity in the market and at counteracting investment cycles. Finally, it should be stressed that capacity mechanisms have many flaws. The price cap in the spot market should be as high as is politically acceptable to minimise the size of the capacity mechanism.

References

Aagard, T.S. and A.N. Kleit (2022). *Electricity Capacity Markets*. Cambridge University Press.

ACER (2019). Market Monitoring Report 2018 -- Electricity Wholesale Markets Volume. Agency for the Cooperation of Energy Regulators, 11 November 2019.

Ahlqvist, V., P. Holmberg and T. Tangerås (2022). A survey comparing centralized and decentralized electricity markets. *Energy Strategy Reviews*, 40, 100812.

Amundsen, E.S., and L. Bergman (2007). Provision of operating reserve capacity: Principles and practices on the Nordic Electricity Market. *Competition and Regulation in Network Industries*, 2(1), 73-98.

APPA – American Public Power Association (2017): RTO capacity markets and their impacts on consumers and public power, Issue Brief, July.

Anderson, E.J., P. Holmberg and A.B. Philpott (2013): Mixed strategies in discriminatory divisible-good auctions. *RAND Journal of Economics* 44(1), 1-32.

Arango, S., and E. Larsen, (2011). Cycles in deregulated electricity markets: Empirical evidence from two decades. *Energy policy* 39(5), 2457-2466.

Armstrong, M., S. Cowan and J. Vickers (1994): *Regulatory Reform: Economic Analysis and British Experience*. MIT Press.

Astier, Nicolas, and Xavier Lambin (2019). Ensuring capacity adequacy in liberalised electricity markets. *The Energy Journal* 40, no. 3.

Ausubel, L. and P. Milgrom (2006): The lovely but lonely Vickrey auction. Chapter 1 in P. Cramton, Y. Shoham, and R. Steinberg (red.): *Combinatorial Auctions*. MIT Press.

Batlle, C., P. Mastropietro, P. Rodilla and I.J. Pérez-Arriaga (2015). The system adequacy problem: Lessons learned from the American continent. In: L. Hancher, A. de Houteclocque and M. Sadowska (red.). *Capacity Mechanisms in the EU Energy Market: Law, Policy, and Economics*. Oxford University Press.

Bamberger, R. (2010). *Strategic Petroleum Reserve: History, Perspectives, and Issues.* DIANE Publishing.

Bergman, L. and C. Le Coq (2019): Blowing in the wind. Energiforsk 2019:587.

Biggar, D.R. and M.R. Hesamzadeh (2014): *The Economics of Electricity Markets*. IEEE Press.

Blume, A. and P. Heidhues (2004): All equilibria of the Vickrey auction, *Journal of Economic Theory* 114, 170-177.

Blume, A., P. Heidhues, J. Lafky, J. Münster, and M. Zhang (2009). All equilibria of the multi-unit Vickrey auction. *Games and Economic Behavior* 66(2), 729-741.

Bolton, R. and T. Clausen (2019): Ambiguous interventions: Capacity remuneration mechanisms and the politics of energy transition in Britain and Germany. *Energy Research & Social Science* 56, 101218.

Borenstein, S. (2002). The trouble with electricity markets: understanding California's restructuring disaster. *Journal of Economic Perspectives*, 16(1), 191-211.

Brattle Group (2014). Estimating the economically optimal reserve margin in ERCOT. Report prepared for the Public Utility Commission of Texas

Bublitz, A., D. Keles, F. Zimmermann, C. Fraunholz and W. Fichtner (2019): A survey on electricity market design: Insights from theory and real-world implementations of capacity remuneration mechanisms. *Energy Economics* 80, 1059-1078.

CAISO (2020): Preliminary Root Cause Analysis – Mid-August 2020 Heat Storm. October 6, 2020.

Ceije, J. (2016): En konsekvensanalys till förordning om effektreserv. Unpublished manuscript, Ministry of Environment and Energy.

Chao, H.-P. (1983). Peak load pricing and capacity planning with demand and supply uncertainty. *Bell Journal of Economics* 14(1), 179--190.

Chao, H.-P. and R. Wilson (2004). Resource adequacy and market power mitigation via option contracts. EPRI report, Palo Alto, CA.

Cramton, P. (2006). New England's forward capacity auction. Unpublished manuscript, University of Maryland.

Cramton, P. (2022). Fostering resiliency with good market design: Lessons from Texas, Discussion paper No. 145, ECONtribute, University of Bonn and University of Cologne, Germany.

Cramton, P., E. Bobbio, D. Malec, and P. Sujarittanonta (2021). Electricity markets in transition: A multi-decade micro-model of entry and exit in advanced wholesale markets," Working Paper, University of Cologne.

Cramton, P. and S. Stoft (2007). Colombia's firm energy market. Unpublished manuscript.

Cramton, P. and S. Stoft (2008): Forward reliability markets: Less risk, less market power, more efficiency. *Utilities Policy* 16(3), 194-201.

Cramton, P., A. Ockenfels and S. Stoft (2013): Capacity market fundamentals. *Economics of Energy & Environmental Policy* 2(2), 27-46.

Crew, M. and P. Kleindorfer (1976). Peak load pricing with a diverse technology. *Bell Journal of Economics* 7(1): 207--231.

EMI (2008). Effektfrågan – Behövs en centralt upphandlad effektreserv? Rapport 2008:15.

EMI (2020). Ren energi inom EU – Ett genomförande av fem rättsakter. Rapport R2020:02.

EMI (2021). Ei:s förslag till tillförlitlighetsnorm för Sverige. Rapport R2021:05.

ENTSO-E (2015). Statistical Factsheet 2015. www.entsoe.eu/publications/statistics-and-data/#statistical-factsheet

Fabra, N. (2018). A primer on capacity mechanisms. *Energy Economics* 75, 323-335.

Fabra, N., N.H. von der Fehr and D. Harbord (2006). Designing electricity auctions. *RAND Journal of Economics* 37(1), 23-46.

Fanebust, E. and N-H. von der Fehr (2013). The combinatorial clock auction: A critique. Unpublished manuscript, University of Oslo.

Fehr, N. H. and D. Harbord (1993). Spot market competition in the UK electricity industry. *Economic Journal* 103(418), pp. 531-546.

Harbord, D. and M. Pagnozzi (2008). Review of Colombian auctions for firm energy. Report commissioned by the Colombian Comisión de Regulación de Energıa y Gas.

Harbord, D. and M. Pagnozzi (2014). Britain's electricity capacity auctions: lessons from Colombia and New England. *The Electricity Journal* 27(5), 54-62.

Harvey, S.M., W.W. Hogan and S.L. Pope (2013). Evaluation of the New York Capacity Market. *Cambridge, MA.*

Hobbs, B.F. (2005). Affidavit of Benjamin F. Hobbs on Behalf of PJM Interconnection, LLC. Filed before the Federal Energy Regulatory Commission. August 5, 2005.

Hobbs, B.F., M.-C. Hu, J.G. Iñón, S.E. Stoft and M.P. Bhavaraju (2007). A dynamic analysis of a demand curve-based capacity market proposal: The PJM reliability pricing model. *IEEE Transactions on Power Systems* 22(1), 3-14.

Höschle, H. and K. de Vos (2016): Implementation of a strategic reserve in Belgium: Product design and market results. Report C5-216, Cigré Meeting.

Holmberg, P. and D. Newbery (2010). The supply function equilibrium and its policy implications for wholesale electricity auctions. *Utilities Policy* 18(4), 209-226.

Holmberg, P. and R. Ritz (2020): Optimal capacity mechanisms for competitive electricity markets. The Energy Journal, 41(Special Issue).

Holmberg, P. and F.A. Wolak (2018): Comparing auction designs where suppliers have uncertain costs and uncertain pivotal status. *RAND Journal of Economics* 49(4), 995-1027.

Rose, J., S. Muthiah, F. Brock, J. Karp, and T. Sakya (2014): Addendum to polar vortex energy pricing implications— Commercial opportunities and system reliability, White Paper, ICF International, Inc.

Joskow, P. and J. Tirole (2007): Reliability and competitive electricity markets. *RAND Journal of Economics* 38(1), 60-84.

Lambin, X. (2020). The integration of demand response in capacity mechanisms, *Utilities Policy*. Forthcoming.

Llobet, G. and J. Padilla (2018). Conventional power plants in liberalized electricity markets with renewable entry. *The Energy Journal*, 39(3), 69-91

Léautier, T.O. (2019). *Imperfect Markets and Imperfect Regulation: An Introduction to the Microeconomics and Political Economy of Power Markets*. MIT Press.

Mauritzen, J. (2014). Scrapping a wind turbine: Policy changes, scrapping incentives and why wind turbines in good locations get scrapped first. *The Energy Journal* 35(2), 157-181.

McRae, S.D. and F.A. Wolak, (2019). Market power and incentive-based capacity payment mechanisms. Unpublished manuscript, Stanford University.

Milgrom, P. R. and R. J. Weber (1982). A theory of auctions and competitive bidding. *Econometrica*, 1089-1122.

Nelder, C. (2013). The perils of electricity capacity markets, September 4, 2013, *Greentechmedia*.

Neuhoff, K., J. Diekmann, F. Kunz, S. Rüster, W-P. Schill, and S. Schwenen (2016). A coordinated strategic reserve to safeguard the European energy transition. *Utilities Policy* 41, 252-263.

Newbery, D. (1997). Pool reform and competition in electricity. Unpublished manuscript, University of Cambridge.

Newbery, D. and Grubb, M. (2014). The final hurdle? Security of supply, the capacity mechanism and the role of interconnectors. Cambridge Working Papers in Economics 1433, Faculty of Economics, University of Cambridge.

Newbery, D. (2022). High renewable electricity penetration: a new "tragedy of the commons", University of Cambridge.

Newell, S.A., K. Spees, J.P. Pfeifenberger, R.S. Mudge, M. DeLucia, and R. Carlton (2012). ERCOT investment incentives and resource adequacy. Consultancy report by The Brattle Group.

Öko-Institut, LBD, Raue (2012). Ein neues Marktdesign für den Übergang zu einem neuen Energiesystem. Studie für die Umweltstiftung WWF. Berlin, 8 October 2012.

Pycia, M., and K. Woodward (2019). Auctions of homogeneous goods: A case for pay-as-bid. Working paper.

Roques, F. and C. Verhaeghe (2015): Different approaches for capacity mechanisms in Europe: Rationale and potential for coordination? In L. Hancher, A. de Houteclocque and M. Sadowska (ed.). *Capacity Mechanisms in the EU Energy Market: Law, Policy, and Economics*. Oxford University Press.

Ruddell, K., A.B. Philpott and A. Downward (2017). Supply function equilibrium with taxed benefits. *Operations Research*, *65*(1), 1-18.

Schwenen, S. (2015). Strategic bidding in multi-unit auctions with capacity constrained bidders: the New York capacity market. *RAND Journal of Economics*, 730-750.

Spees, K., S.A. Newell, and J.P. Pfeifenberger (2013). Capacity markets—Lessons learned from the first decade. *Economics of Energy & Environmental Policy* 2(2), 1-26.

Stoft, S. (2002): *Power System Economics: Designing Markets for Electricity*. IEEE Press.

Stoft, S.E. (2004): Prepared direct testimony on behalf of ISO New England. Submitted before the FERC.

Stoft, S.E. (2005): Prepared rebuttal testimony on behalf of ISO New England. Submitted before the FERC.

Svenska Kraftnät (2013). Effektreserven - En uppföljning och analys av avvecklingen av den svenska effektreserven. Report in Swedish.

Swedish Government (2009). Effektreserven i framtiden. Proposition 2009/10:113.

Swedish Government (2015). Effektreserv 2020–2025. Proposition 2015/16:117.

Tangerås, T. (2018): Equilibrium supply security in a multinational electricity market with renewable generation. *Energy Economics* 72, 416-435.

Teirilä, J. and R.A. Ritz (2019). Strategic behaviour in a capacity market? The New Irish Electricity Market Design. *The Energy Journal*, forthcoming.

Tómasson, E., M. Reza Hesamzadeh, L. Söder, D. R. Biggar (2020), An incentive mechanism for generation capacity investment in a price-capped wholesale power market, *Electric Power Systems Research* 189, 106708.

Willems, B., (2015). The generation mix, price caps and capacity payments. EU energy capacity mechanisms In: L. Hancher, A. de Houteclocque and M. Sadowska (ed.). *Capacity Mechanisms in the EU Energy Market: Law, Policy, and Economics*. Oxford University Press, 140-156.

Wolak, F.A. (2004). What's wrong with capacity markets? Unpublished manuscript, Stanford University.

Wolak, F.A. (2019). Wholesale electricity market design. Unpublished manuscript, Stanford University.

Wolak, F.A. (2021). Market Design in an intermittent renewable future – Cost recovery with zero-marginal-cost resources, *IEEE Power & Energy Magazine*, January/February.

Woodward, K. (2019). Hybrid mechanisms for the sale of divisible goods. Mimeo. University of North Carolina at Chapel Hill.

Zöttl, G. (2010). A framework of peak load pricing with strategic firms. *Operations Research* 58(6), 1637-1649.