International Energy Agency
Secure Sustainable Together

Re-powering Markets

Market design and regulation during the transition to low-carbon power systems

2016

Electricity Market Series

International Energy Agency
Secure
Sustainable
Together

Re-powering Markets

Market design and regulation during the transition to low-carbon power systems

INTERNATIONAL ENERGY AGENCY

The International Energy Agency (IEA), an autonomous agency, was established in November 1974. Its primary mandate was – and is – two-fold: to promote energy security amongst its member countries through collective response to physical disruptions in oil supply, and provide authoritative research and analysis on ways to ensure reliable, affordable and clean energy for its 29 member countries and beyond. The IEA carries out a comprehensive programme of energy co-operation among its member countries, each of which is obliged to hold oil stocks equivalent to 90 days of its net imports. The Agency's aims include the following objectives:

- Secure member countries' access to reliable and ample supplies of all forms of energy; in particular, through maintaining effective emergency response capabilities in case of oil supply disruptions.
- Promote sustainable energy policies that spur economic growth and environmental protection in a global context – particularly in terms of reducing greenhouse-gas emissions that contribute to climate change.
- Improve transparency of international markets through collection and analysis of energy data.
- Support global collaboration on energy technology to secure future energy supplies and mitigate their environmental impact, including through improved energy efficiency and development and deployment of low-carbon technologies.
- Find solutions to global energy challenges through engagement and dialogue with non-member countries, industry, international organisations and other stakeholders.

IEA member countries:

Australia
Austria
Belgium
Canada
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hungary
Ireland
Italy
Japan
Korea
Luxembourg
Netherlands
New Zealand
Norway
Poland
Portugal
Slovak Republic
Spain
Sweden
Switzerland
Turkey
United Kingdom
United States

The European Commission also participates in the work of the IEA.

International Energy Agency
Secure
Sustainable
Together

International Energy Agency
9 rue de la Fédération
75739 Paris Cedex 15, France

www.iea.org

Foreword

The 2015 ministerial meeting of the International Energy Agency's (IEA) member states embraced a new strategic vision to turn the IEA into *the* global clean energy technology hub. Importantly, this new vision is coupled with maintaining and reinforcing the core IEA mission on energy security. There is no field where clean energy and energy security interact more powerfully than electricity regulation and market design. The most powerful image of an energy security problem is a major city in darkness. An interconnected economy with its telecommunications and machine tools is as reliant on the high-quality supply of electricity as its consumers' welfare is on lighting and electric appliances.

For a century, a centralised high-carbon power system kept the lights on. The price has been the generation of over a third of global carbon emissions. Thus, to gain the full social and political support required for decarbonisation, the level of supply security that society has come to expect cannot be compromised. Some countries like Brazil and France have built decarbonised power systems on the basis of large-scale conventional low-carbon technologies, but they are exceptions benefiting from unique natural resources and policies. The most promising technological progress has been seen in wind and solar photovoltaics, which have accounted for the large majority of recent low-carbon deployments. These two variable renewable sources, however, are qualitatively different from a system operation and regulatory point of view. Wind and solar are primarily replacing production from dispatchable capacities in different locations and connection levels; consequently, the transition requires system operation and regulatory reforms.

This should not stop the transition. Previous IEA analysis in *The Power of Transformation* has shown that large shares of variable renewables can be integrated into the power system in a secure and cost-efficient fashion by mobilising flexibility resources. Rapid improvements in low-carbon, demand-response and storage technologies can lead to a smarter, more efficient and more secure system, but achieving their full potential requires new approaches to policy and regulation. While technology is racing ahead, network infrastructure development is lagging behind. Innovation is not only about smart meters; it is also about smart regulation for new flexible business models involving millions of electricity consumers. The old regulatory paradigm designed to deliver kilowatt hours from a centralised system in a unidirectional fashion with meters read only once a year is unlikely to unleash the real-time flexibility that new technologies promise and that the new low-carbon power system will require. If regulatory regimes, market design and system operation end up lagging behind technology deployment, the result may undermine electricity security and, ultimately, the low-carbon transition itself.

Re-powering Markets is the first official publication of the IEA that analyses the electricity market framework for low-carbon power systems. It discusses, for all relevant dimensions of electricity market design, the balance that policy makers must strike between supporting innovation and competition while mobilising capital for the deployment of low-carbon sources. It covers the characteristics of a market design fit for the transition to low-carbon power – one that has proper price signals and the competitive provision of flexibility and adequacy. It addresses in detail the policy and regulatory aspects related to the largest and most complex machine in the world: the electric network, a network which has never been more essential but must nevertheless be transformed. *Re-powering Markets* is the flagship output of the IEA Electricity Security Action Plan and is a key IEA contribution to the post-Paris energy transformation.

Fatih Birol
Executive Director
International Energy Agency

Table of contents

List of figures

List of maps

List of tables

List of boxes

Acknowledgements

This publication was prepared by the Gas, Coal and Power Division (GCP) of the International Energy Agency (IEA), in co-operation with other directorates and offices of the Agency. The publication was designed by Manuel Baritaud under the supervision of Laszlo Varro, Head of Division and Keisuke Sadamori, Director of Energy Markets Security.

The main authors and analysts were Manuel Baritaud (Chapters 1, 2, 3, 4, 6), Sylvia Spruck (Chapters 5, 7, 8, 9), Christelle Verstraeten (Chapters 6, 9), Matthew Wittenstein (Chapter 5), Simon Müller (Chapter 2), Noor Miza Razali (Chapter 9) Jesse Scott (Chapter 2) and Charlotte Vailles (Chapter 4). Special thanks to Jessica Isaacs, Tomi Motoi and Luis Munuera for additional contributions and support. Janet Pape provided consistent and comprehensive formatting as well as administrative support.

This study also benefited from input provided by numerous IEA experts, in particular (in alphabetical order): Marco Baroni, Simon Bennett, Sylvia Beyer, Paolo Frankl, Christina Hood, Cédric Philibert, Keisuke Sadamori, Daniele Sinopoli, Johannes Trüby, Brent Wanner and Dave Wilkinson.

Justin French-Brooks and Therese Walsh were the primary editors of this report. Thanks go to the IEA Communication and Information Office for their editorial and production guidance and to Bertrand Sadin for graphics.

Expert reviewers and contributors

Richard Baron of the Organisation for Economic Co-operation and Development (OECD) served as the official reviewer of the report.

Many senior government officials and international experts provided input and reviewed preliminary drafts of the report. Their comments and suggestions were of great value. They include (in alphabetical order by organisation): Carsten Lehmkoester (Amprion GmbH), Gonzalo Saenz De Miera (Director of Energy Policy at Iberdrola), Stefan Ulreich (E.ON), Bernard De Clercq (Elia System Operator, Belgium), Juan Alba (Endesa, Spain), Juan Perez (EPEX Spot), Philippe Vassilopoulos (EPEX Spot), Tom Howes (European Commission), Manuel Sanchez Jimenez (European Commission), Matti Supponen (European Commission), Matt Wieckowski (European Commission), Richard Lavergne (Ministère de l'Ecologie, du Developpement Durable et de l'Energie, France), Mark Pickup (Ministry of Business, Innovation and Employment, New Zealand), Tomasz Dabrowski (Ministry of Economy, Poland), Doug Arent (National Renewable Energy Laboratory NREL, USA), Jan-Horst Keppler (Nuclear Energy Agency), Craig Glazer (PJM Interconnection), Stephen Woodhouse (Pöyry), Graham Weale (RWE), Katrin Schaber (Stadtwerke München SWM), Elaine O Connell (Department of Energy & Climate Change DECC), Mike Hogan (The Regulatory Assistance Project, United States), Marion Labatut (The Union of the Electricity Industry – EURELECTRIC), Avi Gopstein (U.S. Department of State) and Michael Grubb (University College, London).

The individuals and organisations that contributed to this study are not responsible for any opinions or judgments it contains. All errors and omissions are solely the responsibility of the IEA.

The work could not have been achieved without the support and co-operation provided by many government bodies, organisations and companies worldwide and was sponsored by a voluntary contribution from Japan (Ministry of Economy, Trade and Industry).

Workshops

A number of workshops and meetings under the Electricity Security Advisory Panel (ESAP) were organised in Paris to get essential input to this study.[1] The participants offered valuable new insights, feedback and data for this analysis.

- Expert Workshop I: Scarcity and Flexibility Pricing - 2 July 2014.
- Expert Workshop II: Demand Response - 3 July 2014.
- Expert Workhop III (Joint Workshop with EPRI): Electricity Market Design under Long-term Decarbonisation - 8 October 2014.
- Expert Workshop IV (Joint Workshop with Council of European Energy Regulators CEER): Network Investments and Regulation - 14 January 2015.
- Expert Workshop V (Joint Workshop with CEER): Integrated Resource Adequacy – 15 January 2015.
- Expert Workshop VI: Integrating New Technologies while Maintaining Resource Adequacy – 28 September 2015

Comments and questions are welcome and should be addressed to: Manuel Baritaud (manuel.baritaud@iea.org).

[1] For proceedings from these workshops, see www.iea.org/topics/electricity/subtopics/electricitysecurityadvisorypanel/.

Executive summary

Competitive electricity markets are being challenged by the need to decarbonise electricity production. The Paris Agreement reached at the UNFCCC COP21 conference in December 2015 is expected to give new strength to policies on climate change and the low-carbon energy transition. But the challenge is daunting: according to International Energy Agency (IEA) projections for Organisation for Economic Co-operation and Development (OECD) economies, the average CO_2 intensity of electricity needs to fall from 411 grams per kilowatt hour (g/kWh) in 2015 to 15 g/kWh by 2050 to achieve the goal of limiting the global increase in temperatures to 2°C. While many studies conclude that this is both technically and economically feasible, reaching this goal calls for new power market designs.

This book examines how the design of electricity markets enables the transition to a low-carbon electricity system, at least cost, while maintaining electricity security.

Debates on market design for a low-carbon power system generally present two contrasting policy options: reliance on either wholesale electricity markets with a strong carbon price, or technology-specific policies and regulations. But failures can be observed both in markets and policies. It is increasingly clear that a binary opposition is no longer sufficient to define the market framework.

The transition to a low-carbon power system requires the incorporation of carbon and support policies into a consistent electricity market framework. Competitive markets are an important tool, but they must be supplemented by regulation to ensure an effective transition to low-carbon power at least cost. Table 1.0 provides a high-level overview of such a market framework, i.e. the rules set by governments and regulators and the associated role of competitive markets.

Table 1.0 • Overview of the key dimensions of market frameworks for decarbonisation

Objective	*Policy*	Type of regulation	Competitive markets
Low-carbon investments	*Carbon pricing*	○ Carbon regulation	○ Carbon price (trading scheme) ○ Long-term contracts
	Additional policy: Support schemes	○ Low-C long-term support	○ Auctions set support level ○ Integration in markets
Operational efficiency / Reliability and adequacy	*Short-term energy markets*	○ Market rules ○ Scarcity pricing ○ Reliability standards	○ Energy prices with a high geographical resolution ○ Energy prices with a high temporal resolution ○ Dynamic pricing offers
	Additional policy: Capacity markets	○ Capacity requirements ○ Demand response product definition	○ Capacity prices ○ Demand response participation
Network efficiency	*Regulation*	○ Regional planning ○ Network cost allocation	○ Congestion revenues ○ Transmission auctions
Consumption	*Retail pricing*	○ Network tariff structure ○ Taxation and levies	○ Retail competitive prices ○ Distributed resources

For the longer term, the design of electricity markets hinges on the portfolio of technologies available. There is no definitive answer to the question of what a "perfect" market design will look like once electricity is low-carbon. Instead, improvements in market design are likely to be evolutionary, reflecting interactions between technologies and market rules.

Re-powering Markets **brings together today's best practices in electricity market design,** which can be mostly found in Europe and the United States, and offers insights into possible next steps for the restructuring process in all countries, including those outside of the OECD. It presents the following key findings.

Low-carbon investments: Continuing long-term support while capturing market value

Low-carbon generators need to participate in electricity markets as they can and should earn a high fraction of revenues there. Such participation provides an important market feedback loop, revealing the value of different low-carbon technologies. Low-carbon support should shift away from being the main source of revenues, and investors in low-carbon technologies should be exposed to some degree of electricity price uncertainty. In order to avoid distortions in operational decisions, at times and locations when the value of electricity is negative no incentive to produce should be provided.

Energy market revenues alone, however, are not enough to attract low-carbon investment at the required scale, in a timely manner and at low cost. Electricity prices in most countries today are too low to recoup the investment costs of any low-carbon technology, including renewables and nuclear. A high and robust carbon price is needed, but introducing one or strengthening existing ones will take time, raise acceptance issues and remain politically contested, creating risks for potential investors. Moreover, reaching decarbonisation objectives by 2030 implies deploying low-carbon technologies faster than existing generation is expected to retire, and this situation will continue to depress prices during the energy transition.

Thus, long-term arrangements backed by governments are still necessary to attract a sufficient amount of new low-carbon power generation. Low-carbon investments are capital-intensive and their cost structure does not fit well with short-term marginal cost pricing. Long-term visibility also needs to be provided to mitigate risks for investors and to keep financing costs low.

A new consistent market framework is needed, which includes carbon pricing and support for low-carbon investments. Risks should be shared among investors, consumers and governments, for instance by modulating the level of support as a function of market prices and partially decreasing support as the carbon price and electricity prices increase.

Auctions can also introduce competitive forces to determine the level of support needed, on top of market revenues. Auctioning procedures allow for better control of the level of capacity deployed, and they reduce information asymmetry about the cost evolution and market value of low-carbon generators, allowing for the discovery of the most competitive low-carbon technologies.

Short-term markets: Increasing price resolution

Short-term markets are pivotal. It is important to make updated price information available during the last few hours before dispatch to incentivise the participation of distributed resources, aggregators and neighbouring markets best able to contribute to system needs. As shares of wind and solar power grow, the need increases for market participants to follow the variations of production, to solve more volatile network congestion and to manage forecast errors. The suite of day-ahead, intraday, real-time (balancing) and ancillary services markets are the place where prices optimise the system in the short run, and reveal the value of electricity (and thus investments in the long run).

A market design with a high temporal and geographical resolution is therefore needed. System operators take many operating decisions to ensure system security and integrate new wind and solar capacity, and an evolution in short-term markets is needed for this to be reflected accurately in electricity prices. Moreover, high-resolution prices need to be as transparent as possible to provide the right incentives on where and when to operate and invest.

In parts of the United States, high geographical resolution pricing already exists in the form of nodal pricing for the day-ahead and real-time markets. It has progressively been adopted in all ISOs and RTOs. Unlike Europe, however, there is no intraday market between the day-ahead and real-time timeframe. One possible change could be to make the evolution of locational marginal prices available and transparent during the intraday timeframe.

In Europe, pricing with a higher geographical resolution has yet to be developed in the day-ahead market. This is due to many reasons including less-congested grids, the lack of competition locally and, more importantly, a political desire to have the same wholesale price apply throughout a given country. That said, electricity systems all obey the same law of physics and most European balancing markets (the closest equivalent of US real-time markets) already provide system operators with the plant-by-plant information needed for managing deviations, resolving congestion and ensuring system security. An evolution in the design of short-term markets is therefore needed to increase transparency of the change in marginal costs during the last few hours before operations, with prices published by location.

Although best practices in existing markets suggest these preliminary ideas, there is no one-size-fits-all solution. Such significant evolutions clearly deserve further and more detailed analysis.

Resource adequacy: Pricing reliability on behalf of consumers

The current level of electricity security is very high in OECD economies, and this plays a vital role for digitalised economies. Recent large-scale blackouts were caused by transmission line losses, and most local power supply interruptions take place at the level of the distribution network. The deployment of wind and solar power, compounded by ageing capacity of the existing power plant stock, sets new challenges for reliability. **Governments should continue to define high reliability standards during the energy transition.**

Scarcity prices remain essential to incentivise the performance of all resources when they are most needed, including demand. **However, prices during hours of capacity shortage cannot be free from regulatory interventions.** Situations of system stress are rare, and market participants often fail to anticipate them. Furthermore, generators enjoy market power during these hours, and, as policy makers usually do not tolerate price spikes, price caps have been set too low in many jurisdictions compared to the level needed to meet high reliability standards.

One possible option is for a regulator to define scarcity prices *ex ante*. Looking to the experience of the National Electricity Market (NEM) in Australia and the Electric Reliability Council of Texas (ERCOT), regulators precisely define the market framework. Scarcity prices could reflect the increasing value and probability of load shedding up to the value of loss of load for one hour (usually in the range of USD 10 000-20 000/MWh). Regulators could also define *ex ante* market power mitigation rules to avoid excess cumulated revenues over a set number of years.

Capacity markets: Creating a safety net

Besides scarcity prices in short-term markets, most restructured electricity markets include some type of capacity mechanism to ensure resource adequacy in the longer run. Capacity mechanisms can provide a safety net in the face of policy uncertainty during the low-carbon transition and insufficient demand response. Capacity markets should be seen as one tool to meet policy-driven long-term reliability goals.

Targeted capacity mechanisms, such as strategic reserves, are a useful fix for short-term security of supply issues. By contracting new capacity or old generation which would otherwise retire, strategic reserves can provide quick and simple solutions. But they do not address investment risk and may incentivise market participants to defer investments until future tenders for new capacity.

System-wide capacity mechanisms such as capacity markets are useful for meeting long-term resource adequacy goals, but they have widespread impacts and need to be well-designed to avoid inefficiency. Capacity markets should be technology-neutral, should include both supply- and demand-side resources, and should be forward looking. Sound penalties can ensure the availability of contracted capacity.

Cross-border participation of capacity in these mechanisms can help to reap the adequacy benefits of regional market integration. Regional resource adequacy forecasts are needed and the definition of capacity products should not conflict with one another. Calculating the maximum contribution from cross-border capacity (including both availability and deliverability) and efficient energy flows during shortages is essential. Inconsistencies between capacity mechanisms can potentially hamper cross-border trade or create distortions in competition.

Demand response: Making the most of dynamic pricing

Another feature that has historically shaped market design is the very low price elasticity of electricity consumers. Until now, price response has mainly been limited to large consumers participating directly in wholesale electricity markets. This situation is changing with decarbonisation and the development of new technologies.

New information and automation technologies allow small consumers to contribute to a more flexible and less costly electricity system, responding to wholesale price variations. This could enable a better coupling of electricity generation with energy services and storage, increasing system flexibility to integrate variable renewables and improve electricity security. Retailers are essential in exploiting this demand response potential, using dynamic pricing options and participating in wholesale markets to source their portfolio of consumption.

A further approach consists of treating demand response as equivalent to generation in energy and capacity markets. This has kick-started a demand response market in certain jurisdictions (for example, PJM in the United States). **But "dispatching" demand response as a generator requires complex market rules.** Demand response can only be assessed according to a baseline consumption levels, which are difficult to define and can lead to hidden subsidies. Setting the right level of remuneration for aggregators has proven to be complex. Instead, dynamic pricing should be encouraged, using new measurement and automation technologies such as smart meters.

Transmission investment: Looking beyond local interests

The electricity grid determines the size of the electricity market and the degree of competition. Despite the increase in distributed energy resources, transmission remains a cost-efficient means to ensure the integration of high shares of wind and solar power. In addition, the transmission grid remains essential to secure electricity supply.

Many transmission projects have benefits exceeding their costs, but are constrained by local acceptance issues. Wind and solar power could develop faster than transmission capacity can be built leading to more frequent congestion. Governments should continue to give a high priority to the development of new lines, particularly across borders. **Proper governance is needed** to look at the welfare of a broader area that includes several jurisdictions.

A promising option for merchant lines is transmission auctioning. Competitive procedures to determine who builds and owns the new transmission assets with likely positive commercial values can be viable options. These lines are still regulated, but transmission auctions can bring in innovation and expose incumbent transmission owners to competition.

Distribution network regulation 2.0

Regulation of distribution networks has to be modernised to accommodate the deployment of distributed energy resources (DER) such as solar photovoltaics (PV), storage, electric vehicles (EVs), heat pumps, micro-turbines and demand response. New distribution models require greater investment in information technology and have higher operating expenses (OPEX) but less capital investment in wires and transformers (CAPEX) than the traditional model. In this context, regulation has to become output-based, enabling many distribution companies to find the efficient level of investment.

The regulatory framework should enable DER to participate in both local and wholesale markets. Several models are currently being proposed, including the traditional model where the distribution system operator integrates distributed energy resources, and the market-based model based on a market platform for distributed resources at the local level (as in New York's "Reforming the Energy Vision"). This evolution requires a modernisation of the regulatory framework of distribution networks.

Retail pricing: Sending the right signals to customers

Reform of retail pricing is urgently needed to better reflect the underlying cost level and structure. Current tariff and taxation structures which do not vary with time can lead to inefficiencies. Investments in distributed resources are not always cost-effective as bill savings do not properly reflect the avoided costs to the electricity system. The significant difference in speed between installing solar PV and small-scale storage and building large-scale power infrastructure can exacerbate this problem.

Retail competition can bring innovative commercial offers and services. Competitive retail rates pass through wholesale electricity prices to final consumers, with the objective of properly reflecting the market value of consumption and investment decisions on the consumer side.

In particular, network tariffs need to be rebalanced towards fixed and capacity components in order to better reflect costs. The structure of retail tariffs should, in addition to providing time-varying prices for energy, give the right signals to consumers and induce efficient investment in and operation of distributed energy resources on the consumer side.

Introduction

The future of the entire energy sector will, to a significant extent, be shaped by the evolution of the electricity sector, which is at the centre of most of the discussions to address the threat of climate change. This should not be surprising: in 2014 the electricity sector accounted for just under 40% of primary energy consumed in member countries of the Organisation for Economic Co-operation and Development (OECD), and 42% of energy-related carbon dioxide (CO_2) emissions. The most significant low-carbon energy technologies, including hydro, nuclear, wind, solar photovoltaics (PV), biomass and carbon capture and storage (CCS), relate to the generation of electricity. Thanks to low-carbon generation technologies already available, and the possibility of electrifying transport and heating, the power sector of OECD countries is at the forefront of climate policies to reach 2050 objectives. The Paris Agreement reached at the UNFCCC COP21 conference in December 2015 can be expected to give new strength to policy signals on climate change and low carbon energy transition – providing greater clarity for investors.

The necessary transformation entails replacing much of the old structure based on fossil fuels and creating a new one based on low-carbon power. Other industries have already experienced such a process of creative destruction. In the transport sector, steam engines were displaced by internal combustion engines in the first part of the 20th century. In the telecommunications sector, the internet and wireless communications are replacing old systems.

The decarbonisation of electricity, however, is complicated by the fact that today's low-carbon generation technologies are not obviously superior to fossil power generation in some respects. Despite continued cost reductions, renewables are often still more expensive than gas and coal, in the absence of a carbon price. Nuclear does not emit CO_2 but concerns over safety have led some countries to phase out this technology and, unlike renewables, the costs of new nuclear have increased. CCS is not yet commercially available at scale. Another complication comes from the fact that core technologies, such as wind and solar power, are weather dependent and therefore "variable", which imposes additional constraints on electricity systems. These topics have received much attention and several studies from the International Energy Agency (IEA) (*Grid Integration of Variable Renewables* [*GIVAR*], *World Energy Outlook* [*WEO*], *Energy Technology Perspectives* [*ETP*]), have analysed this transformation. There is now little doubt that a low-carbon power transformation is technically feasible.

But whether this transformation can actually be delivered now largely depends on suitable market design and regulatory frameworks.

Electricity systems are fragmented and markets are open to competition in many countries. Large electricity systems are usually unbundled between networks and a large number of generation companies in competition with each other. Small project developers also account for the bulk of investments in new renewable plants, reflecting the relatively smaller size of wind, solar and biomass plants. Even households can install their own generation behind the meter.

Historically, electricity sector market arrangements have been introduced with the objective of simultaneously ensuring efficient operations, triggering efficient investments and incentivising the optimal level of reliability. Market arrangements and electricity prices have a key role to play in ensuring the co-ordination of decisions in such a fragmented industry structure. Policy makers, however, have not allowed markets to take on the latter two roles. First, reliability is still heavily regulated. Second, despite its introduction in Europe and other jurisdictions, carbon pricing has not been effective at delivering market-based low-carbon investment.

To be fair, this perception is also due to the failure of regulators themselves to implement well-functioning electricity markets. The structure of electricity markets remains concentrated, with a strong control or oversight over wholesale prices which has prevented efficient pricing during tight system conditions, and there is persistent regulation of retail prices in many jurisdictions.

In any case, policies still need to drive the transformation in the right direction. Carbon pricing can be an efficient approach to internalising the climate externality. A robust global price on carbon, if implemented, would reduce CO_2 emissions, focusing change where it is least costly across industries and countries. After the European Union, and thanks to the efforts of international organisations, a growing number of countries and states are likely to implement a carbon price.

To date, however, the policies implemented have fallen short of performing in practice. Where they are in force, CO_2 prices tend to be relatively low, below USD 20 per tonne of CO_2, due in part to the economic crisis that began in 2008 and also to other climate policies that reduce carbon emissions. Policy issues also concern the distributive effects and affordability of higher carbon prices. The credibility of a strong long-term carbon price has yet to be established. Lowering emissions, however, is only one objective of electricity policies. In particular, a very high level of security of supply is at the foundation of our modern digitalised economies.

While pursuing carbon pricing is crucial, to date this approach has not been sufficient to deliver the actions needed. Nor will existing market designs be sufficient to deliver the right investments in low-carbon technologies.

Market design demands a shift in perspective. Existing markets essentially ensure the least-cost dispatch of conventional, mainly large fossil-fired power plants and were introduced at a time when a new technology, combined-cycle gas turbines (CCGTs), could be deployed at a lower cost than older, less-efficient coal and gas plants. In contrast, future market rules have to be designed so as to enable the efficient deployment of new technologies at the centre of the transition: wind and solar power, demand response, storage, hydro, bioenergy and other renewables, but also nuclear in some countries and potentially CCS.

Looking beyond the usual and simplistic alternative between “free markets” and “utility regulation”, or “decentralised decisions” versus “central planning”, it is increasingly clear that decarbonising the electricity system necessarily involves a combination of instruments.

Yet there is little doubt that electricity markets are needed. First and foremost, market prices allow for the co-ordination of distributed resources locally and over large geographic areas spanning multiple balancing areas. In addition, market prices provide incentives to perform at minimal operational cost and when the system values resources most. What is more, market prices bring transparency and inform collective decisions about the relative value for the system of different resources and, in particular, generation technologies.

The right balance between market arrangements and regulation of power sectors still has to be found to successfully manage the transformation of the power sector.

The electricity sector has always been, and continues to be, heavily regulated. This is the case not only for the grid infrastructure, but also for the choice of generation mix. Nuclear investment has been and remains a policy decision. So are renewables in most cases. The regional integration of electricity markets also results largely from political decisions rather than the spontaneous forces of markets or the natural consolidation of the industry.

In addition, it is not always clear whether investors in competitive electricity markets have performed much better than regulators. During the “dash for gas”, private investors overestimated the electricity demand growth rate, and underestimated the pace of renewable deployment. These costs have not been borne by consumers or ratepayers, but have resulted in overcapacity, stranded assets and low profitability for investors.

The central question of this book is to strike the balance between policies that require a form of regulation, and outcomes that can be left to competitive markets perspective of the transition to low-carbon power systems, with the aim of that transition taking place in an effective manner and at least cost.

One key lesson learned from 30 years of experience of market design and regulation of competitive electricity markets in OECD countries is that there is no "one size fits all" solution. This report identifies best practices in order to set the benchmark for policy makers who are embarking on the transition to a low-carbon power system. The objective is to define a workable solution that balances market arrangements and regulatory instruments.

About this report

This report gathers insights from preparatory IEA work in the field of electricity security and market design, initially endorsed by IEA member countries at the 2011 Ministerial Meeting. Several workshops have been held in 2014 and 2015 within the framework of the IEA Electricity Security Advisory Panel (see IEA ESAP webpages).

The focus is on market design issues as they relate to competitive electricity markets in OECD countries. This report might also be relevant for other markets or countries wishing to develop competitive electricity market, although it does not present a full package for power sector liberalisation. Accordingly, it is assumed that the reader is familiar with the organisation of electricity sectors.

The report covers the key components of all electricity systems:

- Chapter 1 introduces the context of the book and key issues.
- Chapter 2 discusses investment in low-carbon generation.
- Chapter 3 looks at markets for short-term operation of electricity systems.
- Chapter 4 presents the regulation of reliability, adequacy and scarcity pricing.
- Chapter 5 describes the design of capacity markets.
- Chapter 6 analyses demand response.
- Chapter 7 discusses investments in the transmission network and interconnections.
- Chapter 8 deals with the regulation of distribution networks.
- Chapter 9 provides an overview of retail price competition and reform of retail pricing.
- The final section presents conclusions and summarises key recommendations.

Chapter 1 • Re-powering markets: Context and key issues

HIGHLIGHTS

- Electricity generation is at the core of efforts to reduce carbon dioxide (CO_2) emissions.
- Many countries have restructured their electricity markets over the last 30 years, and most of these markets will need to adapt further in order to ensure the decarbonisation of electricity.
- Timing of the low-carbon transition matters. Decarbonisation needs to accelerate, which in practice means reducing electricity generated from coal. The pace of investment in low-carbon generation also needs to increase for decarbonisation to stay on track.
- Decarbonisation cannot be done if security of supply is not ensured. A major security crisis is likely to take priority and delay the achievement of other objectives.
- Efficient markets are needed during the transition and will help to keep bills affordable, as will energy efficiency. While, in the long term, electricity markets could be very different from those we know now, the market framework needs to make low-carbon investment possible and cope with the uncertainties inherent in the transition to low-carbon power.

This chapter provides an introduction to, and overview of, key issues relating to the design of electricity markets suited to the transition to low-carbon power systems.

Competitive electricity markets have been progressively spreading out in an increasing number of jurisdictions, having been introduced in Chile in 1980 and then the United Kingdom in 1990. Japan and Mexico are in the process of reforming their electricity systems and introducing competitive electricity markets. Many lessons have been learnt, which can help in the design of electricity markets in other regions.

Over the last decade, policies to decarbonise the electricity sector have had a major impact on electricity markets. This is likely to remain the case into the foreseeable future. Carbon pricing has been introduced in Europe and certain parts of the United States. But the most important impact stems from renewable support policies. In particular, wind and solar power are reaching a scale where they have to become an integral part of electricity markets. As decarbonisation continues, renewables, nuclear and carbon capture and storage (CCS) are expected to grow.

This chapter begins with a description of the importance of electricity restructuring in countries of the Organisation for Economic Co-operation and Development (OECD) and in non-OECD countries and regions. The second section reviews the key issues that decarbonisation raises for power markets. The final section presents the role that different building blocks of electricity markets can play in meeting the issues raised by decarbonisation.

1.1. Electricity reforms

Industrial organisation of power markets has been restructured in most markets

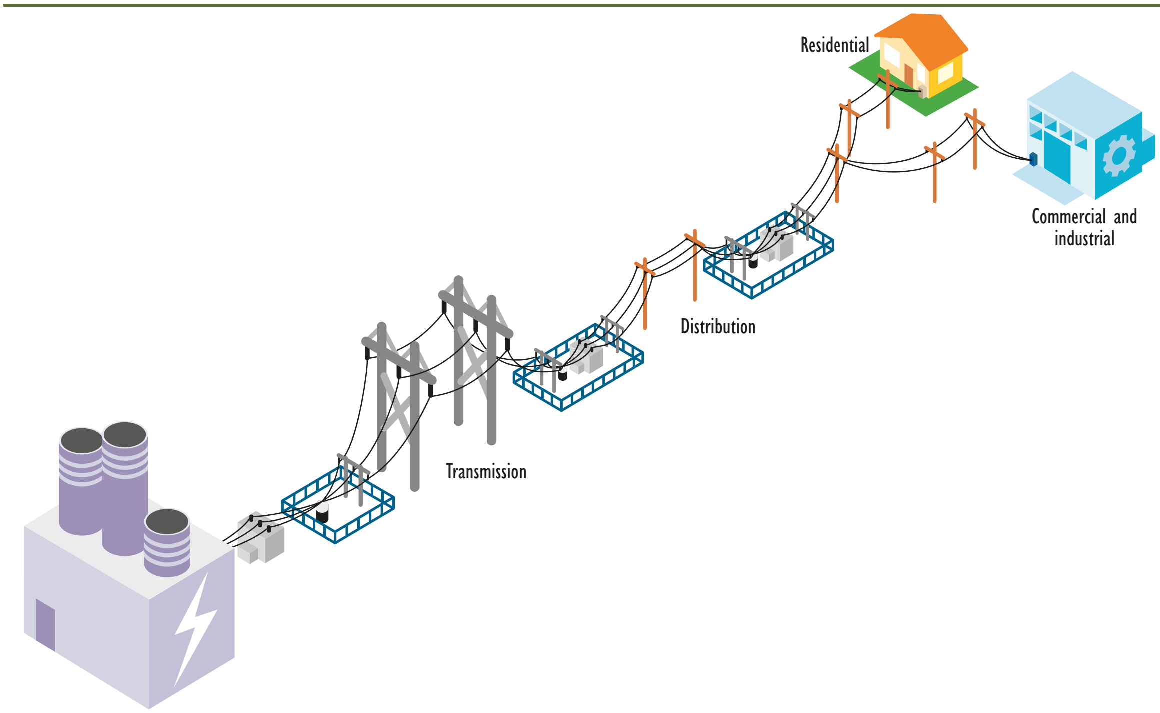
The industrial organisation of electricity systems has profoundly changed in the vast majority of countries over the last 20 years, not only in OECD countries, but also in non-OECD countries and regions (Sioshansi, 2013) (Figure 1.2). This section provides a brief taxonomy of different reforms and the extent to which they have been implemented.

Historically, electricity systems have taken the form of a vertically integrated regulated monopoly, a situation that existed in most countries until the 1990s. This approach can still be found in most African countries and in a number of smaller countries with limited use of electricity. Most such utilities are nationally owned or even part of ministries. Analysis by the International Energy Agency (IEA) has found that this pure monopoly framework represented only 6% of the electricity consumed globally in 2012.

The most basic level of competition is the existence of independent power producers (IPPs) alongside the vertically integrated utility. The IPPs build, own and operate power plants and sell their output at a predefined price to the local utility. In the United States, Congress opened the system to IPPs with the Public Utility Regulatory Policy Act (PURPA), passed in 1978, and this arrangement can still be found in a number of US states. It is also predominant in most Asian countries, including Indonesia and Thailand, and in many countries of the Middle East.

Unbundling represents a further step in market reform. In unbundled systems, vertically integrated utilities are divided into distinct companies, which either own or operate generation assets or the transmission grid and distribution network with related services (Figure 1.1). In large power systems, this approach is often the first step towards introducing a market-based arrangement. The 2002 reform in China, for instance, divided the former State Power Corporation into two grid companies and five generation groups (Andrews-Speed, 2013). It is the largest unbundled network in the world, with total power consumption of 4.326 trillion kilowatt hours in 2012. A further example can be found in India, where the power sector is organised around the Power Grid Corporation of India, which interconnects state electricity boards and several power generation companies (Sen and Jamasb, 2013). In many cases, all the companies are state owned, which in principle reduces problems associated with co-ordination between different unbundled organisations.

Figure 1.1• Organisation of the power sector

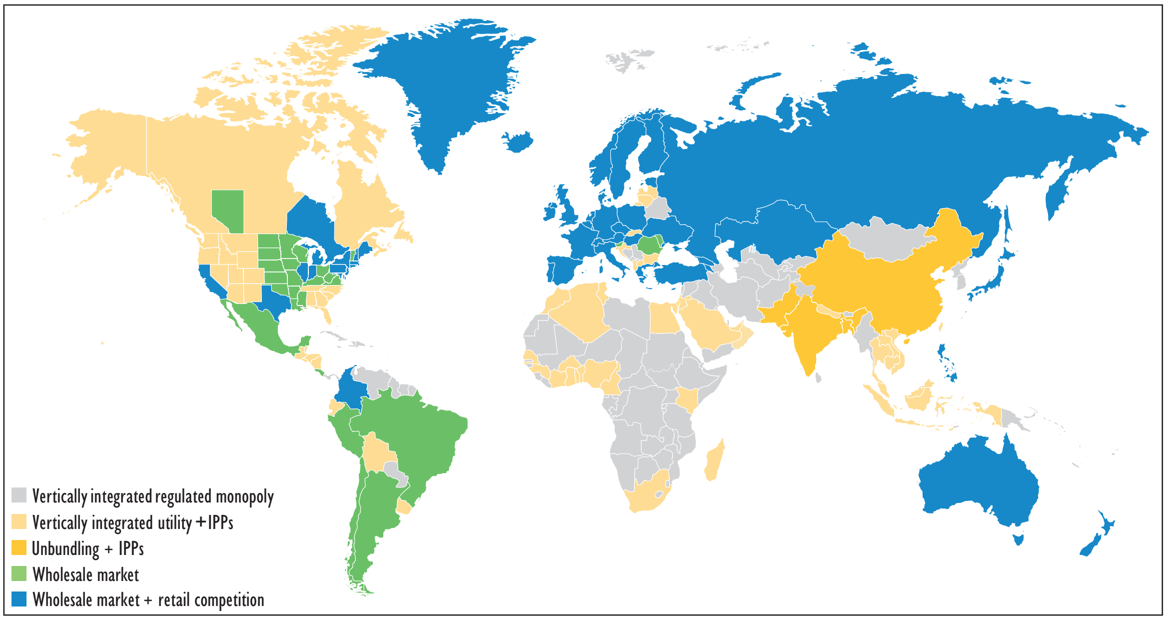


Key point • The traditional vertically-integrated and regulated utilities have been unbundled: while transmission and distribution activities remains regulated monopolies, competition between different generating companies is introduced and consumers can also choose their supplier.

One important objective of restructuring is to better integrate electricity systems into competitive, wholesale power markets over large geographic areas. However, ensuring the efficient co-ordination of large, unbundled electricity systems involves some complexity. By far the most ambitious restructuring process has taken place in OECD countries. In the United States, some but not all states have restructured their electricity sector (Joskow, 2007; Borenstein, 2014). Congress further opened the system to competition in 1992 with the National Energy Policy Act, which allowed power producers to compete for the sale of electricity to utilities.

In certain jurisdictions of the United States, markets have been established with the aim of integrating many small entities in charge of balancing generation and load (balancing areas) into one large wholesale market (IEA, 2014a). An independent system operator (ISO) or a regional transmission organisation (RTO) acts as a central entity that dispatches power plants on the basis of bids, taking into account the technical possibilities of the transmission infrastructure. For instance, PJM in the United States serves a load of 150 gigawatts (GW) across more than 14 states. In addition to the unbundling of networks, generation assets are separated into distinct companies that compete for the provision of electricity. PJM calculates locational marginal prices (LMPs) using the Security Constraint Economic Dispatch (SCED) algorithm based on the bid of the last unit needed to meet demand, and these constitute the uniform remuneration of all the power plants that cleared in the market. ISOs and RTOs represent approximately 60% of the electricity consumed in the country (EIA, 2011).

Figure 1.2 • Map of the status of liberalisation



This map is without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

Sources: IEA and *Renewable Energy and Energy Efficiency Partnership (REEEP) Policy Database 2012-2013*.

Key point • Electricity markets have been restructured in most jurisdictions, with different degrees of competition being introduced

A similar organisation can be found in New Zealand and in Poland, and is being introduced in Mexico. In Australia, the National Electricity Market (NEM) has integrated the previous state organisations into one of the world's largest geographical markets (stretching over 4 000 kilometres). Europe has adopted a different approach to integrating markets across borders (Glachant and Lévêque, 2009). The European Union introduced several directives mandating restructuring in all member states. A first directive on electricity markets was introduced in 1996, followed by a second in 2003 and another in 2009 (IEA, 2014b). Some

European countries decided not to dismantle their national utilities by dividing into different generation companies, and competition is primarily taking place across borders.

In Europe, market coupling has been used as a method for integrating electricity markets across different areas (Glachant, 2010). The market coupling process started in 2010 between France, Germany and Benelux. Under this approach, electricity prices are computed simultaneously for different nationally organised electricity market platforms (also called power exchanges), while taking into account cross-border transmission capacity. The European Internal Electricity Market is built on strong co-operation between transmission system operators and power exchanges from 17 European countries. In 2014, full price coupling of the South Western Europe (SWE) and North Western Europe (NWE) day-ahead electricity markets was achieved.

Adding retail competition to wholesale markets is the ultimate degree of market liberalisation. While it has been introduced in most OECD countries, progress has remained limited to date, both in terms of commercial innovation and market share of new entrants (see Chapter 9).

The varying degree of competition reflects the fact that different countries have different electricity systems and objectives, which shape the organisation of their markets. Some markets have even taken steps back to ensure system adequacy and reliability by having dedicated procurement of capacity alongside an existing wholesale market. This has been the case in California after the California energy crisis of 2001 (Joskow, 2001). Regulatory framework failures had enabled ENRON to manipulate electricity markets, which ultimately caused involuntary load curtailment that imposed high costs on consumers.

Similarly, Brazil liberalised its electricity sector in the 1990s, but new markets did not attract adequate investment. Faced with one of the most serious energy crises in its history in 2001-02, in a context of drought, Brazil resorted to developing an integrated long-term plan for the power sector (Pinguelli et al., 2013). The crisis originated from insufficient hydropower generation during drier years, delays in the commissioning of new generation plants and transmission issues.

Finally in the United Kingdom, recent electricity market reform also marks a step toward a higher degree of regulation (Newbery, 2012). A capacity market was introduced to ensure adequate reliable capacity, while Contracts for Difference (CfD) were introduced to replace the more market-based green certificate scheme and to support investment in nuclear.

Despite these developments, new countries are now also reforming their electricity systems. In recent years, Japan and Mexico have decided to reform their systems and introduce competition.

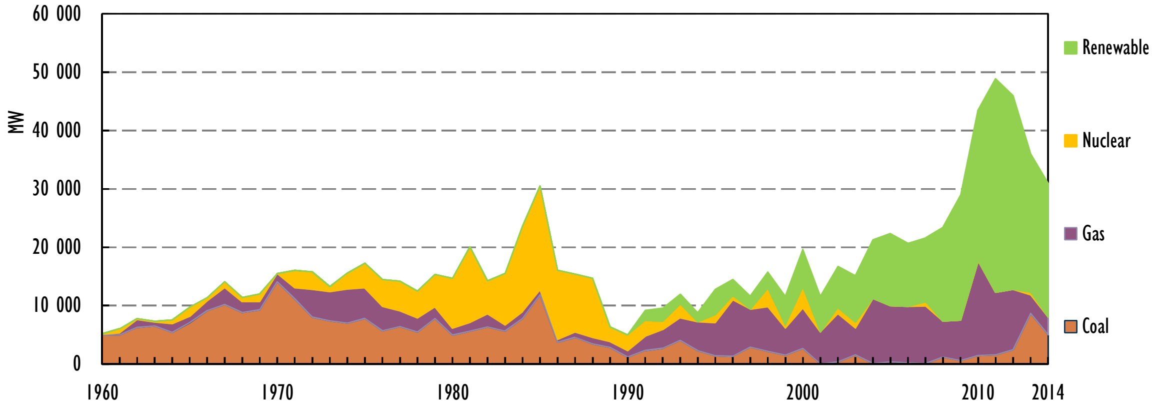
Performance of competitive power markets

Deregulation and restructuring have achieved the initial goal of creating larger markets and promoting trade in electricity to reduce the overall cost of power systems. There is clear empirical evidence that electricity trade has increased both in North America and Europe. In addition, restructuring is also associated with increases in operating efficiency (Davis and Wolfram, 2012), achieved primarily by reducing the frequency and duration of plant outages.

Competitive electricity markets have also triggered a wave of investment in gas-fired power plants, influenced by factors which include the relatively short construction time, the decline in wholesale gas prices and the desire by regional electricity companies to diversify sources. For example, the United Kingdom’s aptly named “Dash for Gas” symbolises the shift by newly privatised electricity companies towards gas-fired plants in the 1990s. An underpinning factor was the development of North Sea gas. In 1990, gas turbine plants made up only 4.5% of the United Kingdom's generating capacity. By 2002, the new combined cycle gas turbine (CCGT) plants comprised 30.9% of UK generating capacity.

In light of climate imperatives, today's policy makers' agenda for electricity has shifted significantly toward decarbonisation. This agenda is profoundly changing the way we look at the role of electricity markets. While markets delivered mainly gas investments (Figure 1.3), very few market-based investments were seen in low-carbon power plants. With the adoption of the renewable energy act in Germany (EEG) in 2000, and the implementation of the Kyoto Protocol in 2005, governments started to introduce emissions reduction policies with meaningful impacts on electricity markets. Carbon prices were introduced in Europe in 2006, and strengthened in the European Union's 20/20/20 climate energy package in 2009.

Figure 1.3 • Capacity addition in OECD Europe by technology, 1960-2014



Note: MW = megawatt.

Key point • Market-based investments have mainly produced gas-fired power plants, while coal and nuclear have been built under a regulated framework, and renewables have been installed with support schemes.

In this context, the bulk of renewable energy investments in the last decade have been policy-driven with support schemes and subsidies. In many countries, renewable deployment is associated with industrial policy, pursuing the objective of creating industrial champions and exporting these technologies.

The timing of the transition implies the rise of renewables capacity in energy markets, according to the IEA *World Energy Outlook* (*WEO*) Special Report, *Energy and Climate Change 2015* (IEA, 2015a). As demand continues to either drop or remain stagnant in many OECD countries, a situation of excess capacity may last into the next decade. Consequently, coal plants and more recent gas plants will have to give way – a higher percentage of stranded assets may become the price to be paid for the push for cleaner power. This raises the question of the pace and the mode of retirement of fossil-fuelled assets.

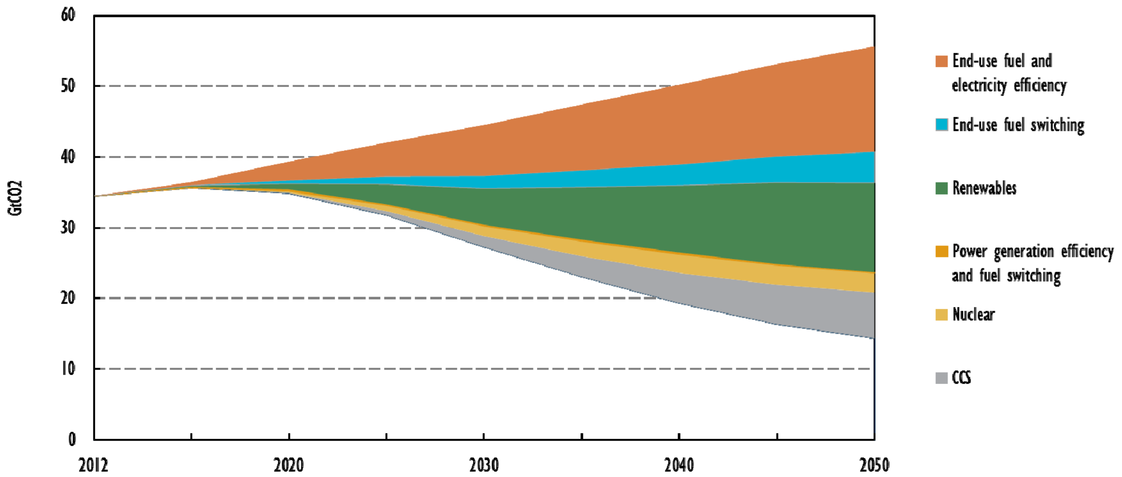
1.2. Challenges facing the power sector during the energy transition

This section considers the three essential challenges facing any energy system: sustainability, security and affordability. First, the electricity system has to reduce dramatically its CO_2 emissions. Second, security of electricity supply has to be safeguarded, and third, efficiency must be ensured to keep the cost of decarbonisation as low as possible.

Reducing carbon dioxide emissions

The first objective, the low-carbon transformation of the power sector, is primarily a policy-driven process. In principle, climate policy in the form of carbon pricing offers multiple desired outcomes: lower energy demand due to higher prices, a disincentive to new high-carbon investment, an incentive and support for low-carbon generation, and curtailment of the continuing operation of high-carbon emitting assets. But in most jurisdictions, carbon prices – for example, under the EU Emissions Trading System (EU ETS) – are proving too low to have meaningful effects.

Figure 1.4 • Contributions to annual emissions reductions between a 6°C and a 2°C scenario



Note: $GtCO_2$ = gigatonnes of CO_2.
Source: IEA, 2015b.

Key point • Decarbonising energy relies essentially on technologies that generate or consume electricity.

Meanwhile, a broad range of specific measures are being adopted to reproduce the outcome that a high carbon price would have in theory. Such technology-specific policies rely on targets (such as the EU targets for 2030). This approach may not fit well with a market-based approach and may undermine the functioning of the underlying electricity markets. Generally speaking, climate policies have created uncertainty and investors are concerned about the increasing regulatory risk. Across jurisdictions in OECD countries, these policies are coming under strong critique for failing to provide adequate long-term visibility for either operations or investment.

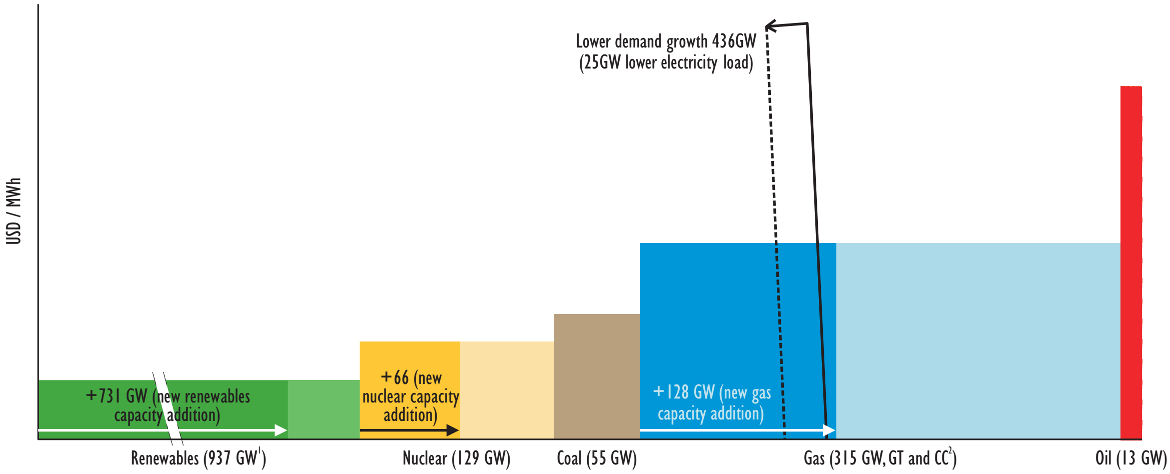
Forecasting electricity demand: Less or more?

Electricity demand has been decreasing in many OECD countries since 2008, after the economic downturn. Corresponding gross electricity production in 2014 in OECD countries (including generation from pumped storage plants) was 10 773 terawatt hours, a decrease of 0.8% on the 2013 level and 1.4% lower than in 2010 (IEA, 2015c). This, in part, reflects efficiency improvements associated with new appliance standards in the building sector and with industrial technologies powered by electricity. Looking ahead, a decoupling of electricity demand from growth in gross domestic product can be foreseen in many OECD countries.

Conversely, with the potential for deep emissions reductions from the electrification of the transport sector and/or space heating and cooling, new drivers may cause an increase in demand. However, with electric cars constituting less than 0.1% of the fleet in 2015 (Cobb, 2015), the prospect for future demand growth remains uncertain (IEA, 2013). The *WEO* assessed that electricity demand could increase by 0.4% per year in OECD Europe under the New Policies Scenario (IEA, 2015b). The annual growth rate is 0.2% under the 450 Scenario,[2] which lowers electricity load by 25 GW on average over the year (Figure 1.5).

[2] The IEA World Energy Outlook 450 Scenario sets out an energy pathway consistent with the goal of limiting the global increase in temperature to 2°C by limiting concentration of greenhouse gases in the atmosphere to around 450 parts per million (ppm) of CO2.

Figure 1.5 • Impact of decarbonisation on the merit order of generation (450 Scenario, EU 2015-40)



Notes: The GW values in parenthesis represent the electrical capacity in 2040 under 450 Scenario. As the average load factor is 29% for wind and 13% for solar photovoltaics, this represents the de-rated capacity that reflects the average contribution of these technologies to the supply slack. CC = combined cycle; GT = gas turbine; USD/MWh = USD per megawatt hour.

Key point • Decarbonising the power system entails a major evolution of the merit order that constitutes the foundation of electricity market prices.

The various factors determining potential demand growth or decline are not well understood yet. In the United States, total electricity use is forecast to grow by an average of less than 1% per year from 2012 to 2040, according to the US Energy Information Administration in a Reference case (EIA, 2014). An alternative Low Electricity Demand case forecasts annual electricity demand in the United States in 2040 reaching a level only slightly higher than that of 2012. In this case, little new capacity is added in the power sector after existing planned capacity additions are completed.

Uncertainty regarding electricity demand growth rates has profound consequences for investment needs. A declining demand scenario would mean little new investment for stagnating power systems. Unlike in fast-growing markets, where excess capacity lasts only one or two years, resorbing this excess capacity could take at least a decade or more, resulting in high economic costs and stranded assets.

But if demand is higher than expected and there is a lack of investment, prices would increase and place a huge economic cost on consumers, not to mention the risk of capacity shortage and security of supply. Governments are inherently conservative and tend to over-invest. Indeed, avoiding "gold plating" and the costs of excess capacity that plagued the power sector in the 1980s were precisely the major drivers for market liberalisation. Based on the experience of the last decade, however, it is not clear whether market participants can do a better job at forecasting demand than governments.

Reducing coal generation

The transition from a high-carbon to a low-carbon generation mix has been much more intensively analysed than demand trends. Policy statements, such as the European Union's goal of cutting emissions by 80% to 95% by 2050, provide a basis for low-carbon investment scenarios. But while there is a clear starting point in today's power portfolio, there are also large uncertainties.

Across OECD countries today, the largest source of high-carbon power generation is coal. Capacity is ageing. In the *WEO* 450 Scenario, the net retirement of coal capacity in OECD Europe is 140 GW, or 73% of the 2013 installed capacity. Power sector investors need to ask the extent to which the current phase-out and retrofit policy arrangements applied to coal generation are consistent with climate change objectives and a transition to a secure decarbonised electricity system. Where the policies seem inconsistent, are they likely to be subject to revision?

In the United States, the Clean Power Plan (CPP) from the Environmental Protection Agency imposes a performance standard limiting total emissions across the generation portfolio of each state's power sector, in effect reducing the market share of high-carbon plants. In the United Kingdom, emission performance standards for new plants have been introduced at a level of 450 grams of CO_2 per kilowatt hour (gCO_2/kWh) for a plant operating at baseload (DECC, 2014). This prevents, in practice, the construction of new coal power plants without CCS. Similarly, in the Netherlands the emissions performance standard is set at 360 gCO_2/kWh. In Germany, an agreement was reached in July 2015 to create a Climate Reserve and close 2.7 GW of lignite power stations (German Energy Blog, 2015).

The future of coal generation raises several issues for the functioning of power markets. Uncertainty over the timing of coal plant retirement adds to the uncertainty surrounding levels of demand in the quest to define future investment needs. While CCS technologies are expected to play a role in decarbonisation, the technology is not yet available at commercial scale, unlike other low-carbon technologies. In addition, it is not clear that CCS will be sufficiently flexible to sit alongside renewables in an electricity system. Finally, despite high CO_2 emissions, coal is a domestic fuel in many countries, contributing to fuel security and, in the absence of a carbon price, offering lower costs than other technologies. Base load coal generation remains important to provide network inertia and stability as well as flexibility when needed. When reducing coal generation, governments should make sure that markets are able to achieve these objectives.

Attracting investment in low-carbon generation

The greatest challenge is to secure finance for massive investment in low-carbon plants during the energy transition. In Europe, for example, the *WEO* projects that the new capacity addition of renewables will reach 731 GW, plus 66 GW of new nuclear capacity, during the period 2015-40 under the 450 Scenario. Variable renewable technologies generate whenever there is wind or sun because their marginal costs are very low (Figure 1.5). The foreseen capacity is likely to reduce wholesale prices.

Will energy markets be capable of delivering low-carbon investment? In theory, an electricity market based on the sale of electrical energy in megawatt hours (MWh) (energy-only market [EOM]) combined with a sufficiently high carbon price could plausibly ensure decarbonisation in the long term. As usual in economic theory, a set of assumptions has to be satisfied, including perfect correction of externalities, separation of efficiency and equity/distributive objectives, convexity of cost functions and perfect competition. In practice however, these assumptions do not all hold and consequently market-based low-carbon investments face a number of challenges.

The first issue is that current market prices for power are too low. In the United States, prices are in the range of 30-40 USD/MWh and in Europe the range is 30-50 EUR/MWh, levels which are insufficient to attract any investment, including low-carbon. If decarbonisation of the power system stays on track by increasing efficiency and deploying low-carbon generation, prices are expected to remain low in the coming years. They would need to come back to the level experienced in 2008-09 (around 80 EUR/MWh) for a long period of time to trigger investment during the transition, but this is not anticipated in the short to medium term.

The second issue is that power markets set prices based on short-term marginal costs. Marginal cost pricing leads to volatile prices and does not guarantee the recovery of the high upfront fixed investment costs of renewables, nuclear and CCS. A very high carbon price, above USD 100 per tonne of CO_2, could restore such high prices in principle, but many governments are concerned about windfall profits, affordability and competitiveness issues, which undermine the credibility of such a high-price scenario. In addition, the rapid deployment of technologies with low short-run costs, such as most renewables, further depresses wholesale electricity prices and can drive them down to zero for some hours under high-renewable scenarios. This issue is analysed further in Chapter 2.

Lastly, governments have specific objectives for the deployment of specific technologies. Letting the market and a carbon price decide the level of decarbonisation and the mix would not be a problem per se, if governments had the single objective of reducing CO_2 emissions. But governments usually have a mix of objectives that determines the selection of, for example, renewables or nuclear that goes beyond CO_2 emissions.

Government policy is seeing the rejection or phasing out of nuclear power in a number of countries. Some countries would also prefer to reduce reliance on gas because it is imported or exposes consumers to long-term gas price risk. Other countries are pro-solar. It should be acknowledged that, to a certain extent, major decisions about the generation mix remain a matter of state energy policy.

To date, existing sources of low-carbon generation have been built under a regulated framework. Nuclear and hydro together represent 80% of low-carbon power in OECD countries, having largely been built before the introduction of competitive electricity markets. The remaining 20% has been subsidised by renewables support schemes. Market-based, unsubsidised low-carbon investments have been negligible.

Support for low-carbon renewables was initially introduced in the years after 2000 as a transitional policy during the inception and take-off phase, with the prospect of becoming close to or fully cost-competitive during the consolidation phase (IEA, 2011). Governments accepted the need to subsidise renewables at the initial stage of deployment, in order to benefit from lower costs subsequently as mass deployment becomes necessary.

In several European countries, onshore wind and solar photovoltaics (PV) have been deployed rapidly and at high cost. These policies have been successful in reducing their associated investment costs. Onshore wind and solar PV are now mature technologies with more than 50 GW of wind and solar power added every year in OECD countries. Several governments, including Spain, Italy and the United Kingdom, have now ceased support, which has stopped new installation as their costs have not fallen sufficiently. These examples illustrate the risks associated with technology-specific support schemes.

Electricity security of supply

Electricity systems in OECD countries deliver power "on demand" with a high level of reliability and at a reasonable price. Decarbonisation is not expected to improve the quality of this basic electricity service; rather it aims to reduce the risk of harmful consequences from the way these services are provided today. The transformation of the power sector involves the retirement of ageing conventional capacity, the deployment of variable renewables and other resources to complement them. The scale and pace of the transformation introduce new challenges, with the need to maintain a high level of reliability receiving much attention.

Indeed, security of electricity supply is the first constraint on how the transition develops. In OECD economies, a very high level of security of supply lies at the foundation of our modern digitalised economies. A major electricity security crisis could result in great difficulties for decarbonisation. Reliability should not be taken for granted.

Electricity security encompasses several dimensions: fuel security, system security and adequacy.

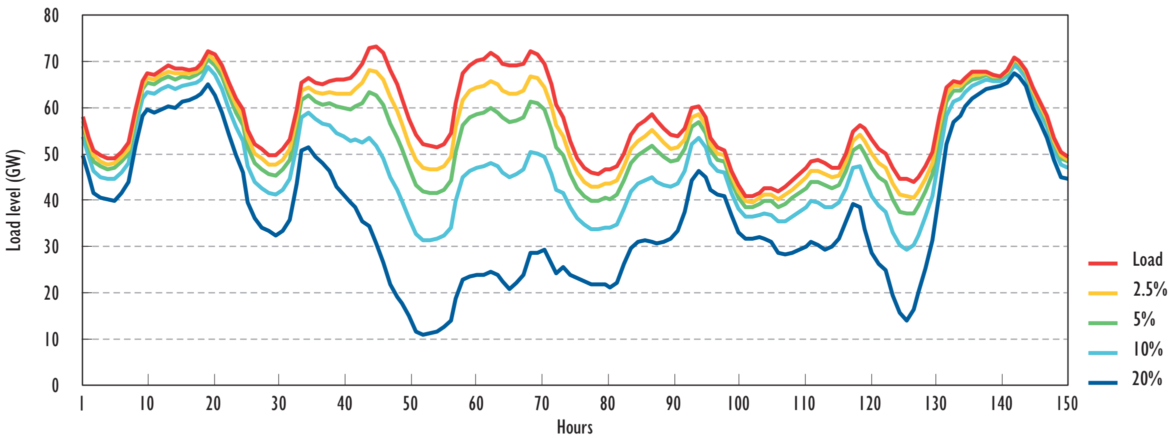
Fuel security

Security of fuel supply is the most intuitive aspect, and from this perspective, low-carbon generation such as renewables can contribute to reducing reliance on imported fuels, in particular gas in European countries, Japan and South Korea. Nuclear can also increase security of fuel supply, because uranium can be stockpiled easily.

System security

System security is essential to maintaining the stability of the electricity system. From this perspective, the integration of wind and solar power raises a number of challenges because their output is weather-related and therefore variable and less predictable (Figure 1.6). Previous IEA work (2014c) has concluded that reaching high shares of wind and solar power is technically feasible. Wind power already constitutes 40% of Denmark's generation mix, while in Spain, wind and solar power together constitute one-third of the generation mix as of 2014. This is already comparable to decarbonisation scenarios with a combined wind and solar power contribution of 31.5% of by 2050.

Figure 1.6 • Evolution of net load for different shares of variable renewables



Note : Load data and wind power data are for Germany from 10 to 16 November 2010. Wind power generation is scaled, actual annual share being 7.3%; scaling may overestimate the impact of variability; for illustration only.

***Key point •** Wind and solar power's variability requires transformation of the power system to ensure security of electricity supply.*

Higher flexibility is needed to accommodate weather-dependent outputs. Technical flexibility stems from all elements of the system in combination, including demand, networks, storage, conventional generation and wind and solar power themselves. Commercial flexibility depends on cost structures and the incentives provided by market revenues. Technologies with very low marginal cost, including wind and solar, tend to be commercially inflexible. Tapping the flexibility potential while maintaining system security might in some cases require improvements to the design of short-term electricity markets (see Chapter 3).

Increasing flexibility requires a more profound transformation of the power system (IEA, 2014c). In the long term, the power system has to be re-optimised. The more wind and solar power in the system, the less baseload power is needed. Conventional gas-plant capacity will be needed in the long run, but will not run or will run at lower output except at times when wind and solar outputs are low. Under the IEA *WEO* 450 Scenario, gas-fired capacity installed in the European Union totals 315 GW by 2040 (of which 128 GW are net capacity addition, Figure 1.5) but runs for 1 081 full load hours, corresponding to an average load factor of 12% (IEA, 2014a).

Adequacy

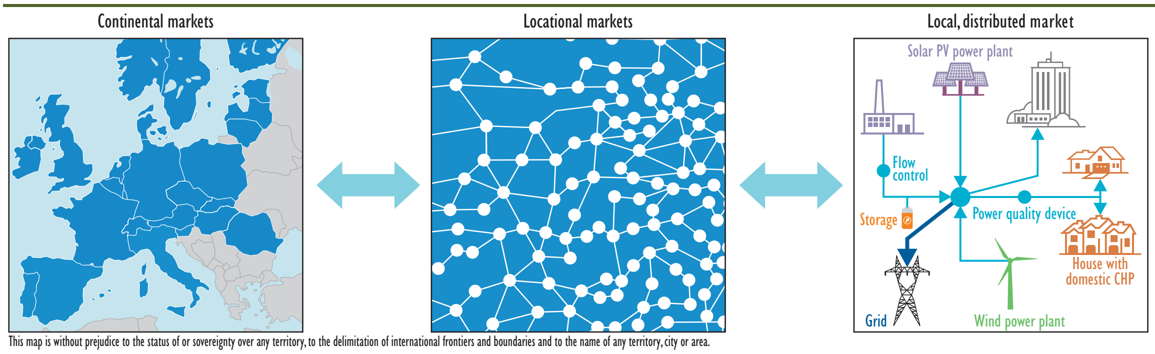
The ability of available generation to meet the predicted system demand remains an issue during the transition to low-carbon power. The power system needs sufficient generating capacity and price-responsive consumers to meet reliability standards. Over the next 25 years, almost 40% of installed capacity is expected to close down; half of all nuclear reactors could reach the end of their technical lifetime, and around 610 GW of coal capacity will be phased out for environmental reasons, according to the *World Energy Outlook Investment Report* 2015 (IEA, 2015a). In addition, the power industry is adjusting capacity by mothballing or retiring recent but uneconomic gas plants. Excess capacity cannot be expected always to be available.

Future conventional plants are expected to run less often and generate less energy due to the fact that they will increasingly complement wind and solar power. In the light of historical experience to date, most market-based investments in CCGT plant were built to run baseload. Whether electricity markets will be able to attract investment for mid-merit and peak requirements remains one of the most debated topics for the design of electricity markets (see Chapter 4 on reliability adequacy and scarcity pricing and Chapter 5 on capacity markets).

Efficiency

The primary objective of competitive electricity markets is to increase the efficiency of power systems. Looking ahead, achieving the transition at least cost necessitates the efficient co-ordination of an increasingly complex, unbundled, large and diversified electricity system. Markets have to be designed at local level to integrate distributed resources (Figure 1.7). Markets also have to be designed in order to ensure the right location of operations and investments. In addition to that, electricity market integration at a continental scale becomes necessary to manage the variability of wind and solar power resources.

Figure 1.7 • Regional, locational and local electricity markets



***Key point** • Efficient co-ordination is needed, from the local level to the continental scale.*

Common to all these geographic scales are networks. Networks will not disappear as a result of distributed energy. On the contrary, they remain the backbone of electricity systems. Network costs represent 30-50% of total costs and their regulation deserves significant attention.

Integration of distributed technologies

Small-scale technologies are playing an increasingly important role in power systems. While gas, coal and nuclear power plants are connected to the transmission network, solar PV and onshore wind are connected to the distribution network. In Germany, for instance, 90% of wind and solar capacity is connected to the medium- and low-voltage grids. New information technology (IT) also enables the remote control of small-scale electrical appliances and storage devices that can increase the flexibility needed to integrate wind and solar power.

Distributed technologies call for a major shift in the design of markets. Electricity consumers have traditionally been passive and perceived as ratepayers, with electricity heavily taxed, not least to finance the cost of renewable policies. As distributed technologies increasingly empower consumers to arbitrage the retail price of electricity, there is a risk of inefficiency if certain active consumers invest in expensive distributed equipment to reduce their bills with no commensurate reduction in cost for the system. Retail pricing of electricity includes regulated costs – levies to recover policy costs and taxes – as well as market price components.

At high shares, distributed resources become important to system and market operations, both for the bulk power system and for the distribution systems themselves. The potential for market-based deployment and co-ordination is an entirely new territory for markets at the local level and is being analysed in New York, California and Hawaii, while many experiments are taking place in Europe and Japan. In any case, small-scale generation, demand response and storage connected at the distribution level will also have to be integrated into wholesale markets.

Locational signals

Whether it is connected to the distribution grid, the transmission grid or elsewhere, the location of new generation matters. New low-carbon electricity is likely to be built in greenfield locations that have to be connected to the existing grid, which developed historically for centralised power. Such network costs can be high, in a range from a few dollars per megawatt hour to about USD 10/MWh (IEA/NEA, 2015). It is important to control both generation and system costs for efficient decarbonisation.

The location of new generation derives both from market design and regulated activities:

- Considering that both renewables and networks are still regulated, detailed integrated planning of investment decisions continues to make sense. Regulators have to decide most of the parameters that influence location decisions, including connection charges and the permitting for new renewables.
- Locational electricity prices reveal the value of electricity at different locations and can provide transparency to guide investment decisions and generation operation decisions.

Consequently, there is a need to strike the right balance between integrated resource planning and market-based investment decisions based on locational signals. Given the scale of the transformation needed, co-ordination between unbundled network and generation activities is needed. Integrated resource planning does not mean, however, that a central entity is in charge. Rather, it has to be seen as an important tool that can bring transparency to ensure informed market-based decisions.

Integrating markets across large balancing areas

The low-carbon transition requires strengthening the integration of markets over large regional areas. This is particularly important in the case of large-scale deployment of wind and solar power. The development of electricity markets is inseparable from regional integration (IEA, 2014c). For instance, the creation of large RTOs, such as PJM and MISO in the United States or the NEM in Australia, is aimed at integrating many small balancing areas into one large market. Similarly in Europe, power markets have largely been designed with the objective of enabling cross-border trade of electricity. The implementation of the so-called Market Coupling in 2014 in 12 countries is a major achievement.

In order to smooth out the variations and forecast errors associated with renewables, market integration has to be deepened. In some regions, larger balancing areas are needed. For instance, the Energy Imbalance Market (EIM) will enable California to share balancing resource across neighbouring jurisdictions. In other cases, some progress has been made to integrate the day-ahead electricity markets, but there is still a need to better integrate short-term intra-day and real-time/balancing markets. This is the case in most European markets. The efficiency of markets over large geographic areas requires strong co-ordination and the consolidation of balancing areas.

Many barriers still stand in the way of regional integration of electricity markets. If the efficiency gains of market integration are important, so are the huge distributive impacts. For example, price increases in some markets can exceed the efficiency gains. When deciding to invest in new interconnections to achieve better integration of markets (Chapter 7), regulators have to look beyond the interests of domestic consumers and consider the broader implications for integrated markets. Regional markets require regional governance.

Efficient network investments

The transition also creates challenges for network activities. With the exception of some merchant lines, networks continue largely to be viewed as natural monopolies that need to be regulated. In the 1990s, regulation underwent modernisation with the creation of independent economic regulators in Europe and the introduction of incentive-based regulation aimed at replicating the discipline of markets. Still, regulators and governments have not always fully adapted the regulatory framework to be fit for decarbonisation.

As previously mentioned, regulation has to look beyond borders in order to fully reap the benefits of regional integration of electricity markets. This is yet to be the case.

Regulation also has to keep pace with technological progress: new possibilities offered by distributed resources and smart grids have to be efficiently deployed and integrated. For instance, active management of renewable resources connected to distribution networks can help reduce or delay distribution network investments. Failure to do so can result in inefficiencies, requiring upgrades to the distribution network capacity to feed in renewables, while seeing less energy being consumed, which reduces the billing base. And as illustrated by the experience of Australia, new distribution network investments can significantly increase electricity bills (Chapter 8).

Affordability

Efficient power markets should remain a first priority to keep bills affordable during the transition to low-carbon power. In some countries, the price of electricity for households has increased significantly over the period 2005-14, reaching 300 EUR/MWh in Denmark and Germany, compared to around 150 USD/MWh in the United States, for example (Chapter 9). Even in countries with the most ambitious low-carbon policies, affordability issues matter.

One key issue is to determine whether the cost of the energy transition can and should be supported by taxpayers or entirely by bill payers. In the United States, renewables are deployed with subsidies that take the form of tax credits, reducing fiscal revenues, while in Europe, the cost of renewable policies is entirely paid by electricity consumers, albeit not all consumers contribute proportionately to their consumption. Under the European approach, electricity consumers pay the full cost of electricity, but this also includes the cost of industrial policies or fuel security policies associated with renewables. Increasing electricity prices, however, discourages the electrification of transport and heat that are also needed to reduce emissions from the entire energy sector.

Lastly, policy makers are not only looking at efficiency but are also often sensitive to issues of social justice. In principle, high CO_2 or electricity prices create distributive impacts that should be addressed by means of general taxation and redistribution, not by distorting prices. In practice, however, policy makers tend to prefer to keep prices relatively low and affordable, in particular for public services such as electricity.

The future of generators in competitive markets

Utilities in many OECD countries have experienced a change of fortune: from being healthy, profitable and low-risk investments during the 2000s, they are today regarded by the financial community as high risk and unprofitable. In Europe, the power sector has become "uninvestable" (Financial Times, 2014) and has lost half a trillion Euros in value (The Economist, 2013). After spending billions on new plants and acquisitions in the 2000s, many utilities cannot invest more without being further downgraded by credit rating agencies.

This situation largely results from a range of causes including the economic crisis, declining demand and the divestment of regulated assets such as the transmission networks. It can also be linked to decarbonisation, for instance the rise of renewable and distributed resources. It also has consequences for the market design discussion. The traditional role of large diversified utilities is, in principle, to manage the risks associated with electricity markets. If they are decreasingly in a position to fulfil that role, low-carbon investments will increasingly have to come from other investors with different risk appetites.

1.3. Re-powering markets

Wholesale markets are pivotal

Broadly speaking, the discussion about future market design for a low-carbon power system often features two opposing camps. "Market purists" want to remove all policy intervention that distorts market prices and internalise the climate externality with a strong carbon price. The "climate change planners" want to minimise the financing cost of low-carbon generation investments by insulating investors from market risk, introducing procurement auctions for power purchase agreements for low-carbon generation projects. Ultimately, following this logic would lead to the abandonment of competitive markets.

Is there a problem with the competitive power market itself, and not only the (lack of a) carbon price? If energy-only power markets will not be fit for purpose in a decarbonised power system in 2050, then should the market design based on marginal cost pricing be reconsidered? This question is receiving increasing attention.

The main concern is that electricity markets are inherently volatile, while low-carbon technologies have high upfront fixed costs. Even with a high and robust carbon price, exposing low-carbon generators to the long-term uncertainty of gas prices does not provide any certainty that the investment costs can be recouped. Consequently, market risks will increase the cost of capital considerably. This issue is reinforced in scenarios with high shares of wind and solar, which further depress wholesale prices when the wind and sun are plentiful. The cost of decarbonisation in an energy-only electricity market would therefore be higher and this might jeopardise decarbonisation targets.

Despite these investment-focused discussions, there is no doubt that wholesale energy markets are essential and are needed more than ever for the best functioning of large and complex power systems with an increasing number of participants. Wholesale energy markets can:

- Ensure co-ordination of millions of distributed resources locally (including demand response and storage) and co-ordination across large geographic areas spanning multiple control areas.
- Provide incentives to perform, i.e. minimise operation costs and be available when the system values the resources most.
- Bring transparency and inform collective decisions about the relative value to the system of different resources and in particular renewable generation technologies.
- Incentivise innovation in the power system.

These are now discussed further in turn.

The first reason why markets are needed is to ensure the co-ordination of many resources. Decarbonisation implies more diverse technologies and that more resources become distributed or located far from consumption centres. This increases the frequency of network congestion. Letting markets facilitate dynamic trade of electricity is the most straightforward way to ensure efficient operation and to minimise the cost of such low-carbon electricity systems.

Without market arrangements to co-ordinate the dispatch of resources, vertically integrated monopolies used to perform this task relatively efficiently, but only for a few dozen power plants within their control area. Now, with an increasing share of decentralised resources, dispatch would also have to be done for thousands, or even millions, of distributed resources (demand response, back-up generators, etc.).

Furthermore, reaping the benefits of large electricity systems requires sizeable regional markets. Enabling the trade of electricity across highly fragmented local balancing areas was precisely the primary objective of ISOs and RTOs in North America. Similarly in Europe, the gains associated with cross-border trade of electricity were the single most important driver for the creation of the internal electricity market. Large markets smooth the variability and lack of predictability of renewables, tap the potential of the windiest and sunniest places, ensure least-cost dispatch and keep the cost of decarbonisation as low as possible. Agreeing on exchange schedules that change every hour or every 15 minutes would be an almost impossible task without transparent markets.

The second reason why electricity markets are needed is the operating efficiency of power plants. Many empirical studies have found that markets increase efficiency. Exposing market participants to electricity prices is an effective way of ensuring that power plants and demand-response resources are given an incentive to be available when their value to the system is highest. The owner of a power plant stands to lose a significant amount of revenue if its plant is not available when prices are high, but under a regulated cost-recovery regime is guaranteed its income under any circumstances.

The third reason why markets are needed is that they send investment signals. For reasons discussed further in Chapter 2, markets might not be sufficient to incentivise low-carbon investments, due to the uncertainty associated with carbon pricing and other policies interacting with electricity markets. Markets, however, are necessary to reveal the value of low-carbon investments to the system. Even at high shares of wind and solar deployment, market revenues can represent a significant fraction of the total revenue needed to recoup the investment costs. The higher the market revenues, the higher the efficiency of the investment.

Assume, for instance, that the average market revenue of new wind turbines is very low, say below 20 USD/MWh, while the average wholesale market price is 50 USD/MWh. This is an indication that new wind turbines generate mainly when there is already a lot of wind and the additional value of new wind capacity is therefore low. This would be a signal that other low-carbon technologies with a different generating profile might be preferable, for instance solar PV plant with market revenue of 50 USD/MWh, even if its costs remain higher than wind on a levelised cost of energy basis.

Although these revenues might not be sufficient to ensure full cost recovery, such differences in value have to be properly factored in when deciding to invest.

The final reason why electricity markets are essential is that they foster innovation in the electricity sector. New entrants can select novel low-carbon or demand-side technologies. Also, without markets, vertically integrated monopolies may be likely to attempt to protect their assets from becoming stranded, slowing down the pace of decarbonisation.

Markets incentivise innovations as much as innovations shape markets. Innovation may even change how electricity markets operate by 2050: storage and demand response have the potential to transform traditional electricity markets. If batteries can store electricity and consumer demand response decreases or increases load in line with supply, then prices will be less volatile and simple market arrangements are more likely to function well. In addition, while there is a strong focus on wind and solar today, bioenergy plays a notable role in most scenarios.

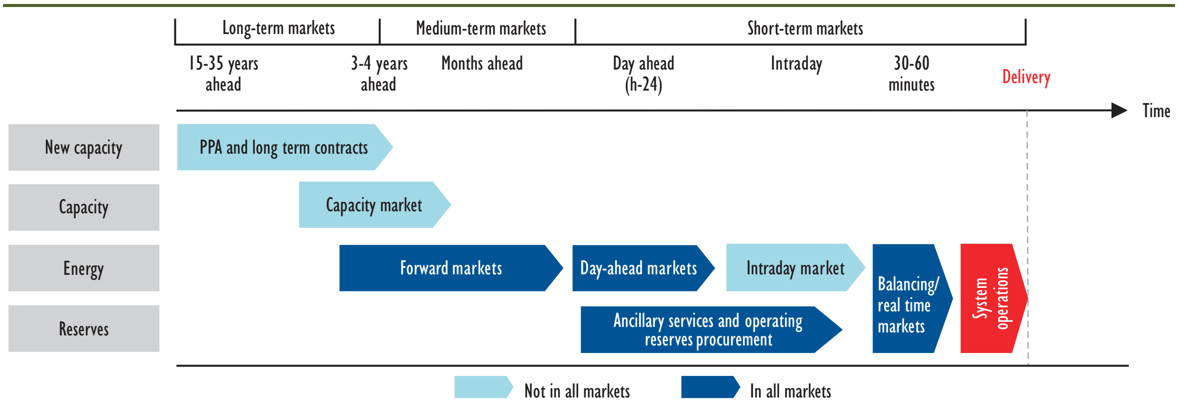
Defining wholesale electricity market

To help address the issues mentioned above, electricity markets are typically structured around three durations: short term, medium term and long term (Cramton, 2015).

Short-term markets (minutes to hours)

Short-term markets provide the foundation to all electricity markets. In most cases, they consist of two main markets: the day-ahead market and the real-time market. In the day-ahead market, participants bid for energy and the market clears and sets hourly prices for each hour of the next day. Generating units are committed according to these prices. Then, during the day, adjustments have to be made to balance supply and demand, which are continuously updated. This is done either by system operators or by generators. In Europe, participants can exchange electricity blocks on an intra-day market platform before system operators set balancing energy prices that clear the balancing market (Figure 1.8). In North America, system operators calculate real-time prices every five-minute.

Figure 1.8 • Overview of different building blocks of electricity markets



Key point: A suite of interrelated markets is used to match generation and load in the short, medium and long term.

System operators must also have enough flexibility to balance the system permanently in case of an unexpected outage of a generator or a transmission line. The system operators procure a number of ancillary services, including operating reserves, to restore the frequency. While the denomination differ in every market, operating reserves typically include frequency response reserve or primary reserve, spinning and non-spinning reserve and replacement reserves. These reserves provide the capability to balance the system second by second after a failure occurs. They usually represent a smaller fraction of costs (see Figure 1.9). The exact definition of balancing services is complex and varies from one country to another.

Short-term prices have a locational dimension, in that prices depend on the location of the generating unit or the consumption in the network. The United States has Locational Marginal Prices, while in Europe prices are uniform for large zones, often defined by the borders between countries. In a system without congestion or energy losses, prices are equal at different locations, but when the transmission line is congested or losses are taken into account, prices are differentiated. The locational dimension of prices sends signals for supply and demand on the marginal costs at different nodes of the network.

Short-term markets play a key role in mobilising the flexibility of the power system, and the detail of their design affects the level of integration of renewables that can be reached. These markets are also essential for the integration of power systems over large market areas. The prices constitute the references against which other medium- and long-term prices are set, and they motivate participants both in the short and long run.

Given their relatively smaller importance in terms of volume, and their high complexity and diversity, this report does not discuss certain technical ancillary services, such as voltage regulation, black start facilities and primary frequency control reserves.

Medium-term markets (month to three years)

Medium-term markets allow price risk to be better managed by consumers. In well-functioning markets, most energy is traded before the short-term markets, from a few months in advance up to three or four years. The medium-term market may be a formal, organised market with future and forward standard products traded bilaterally over the counter, or it may be informal, with variable quantities traded by traders or retailers. In liquid European markets, roughly 90% of energy is traded on these medium-term markets. Short-term spot markets play an essential role in settling the deviation between energy contracted on medium-term markets but not consumed, or buying energy not contracted in advance.

Long-term investment market (3 to 25 years)

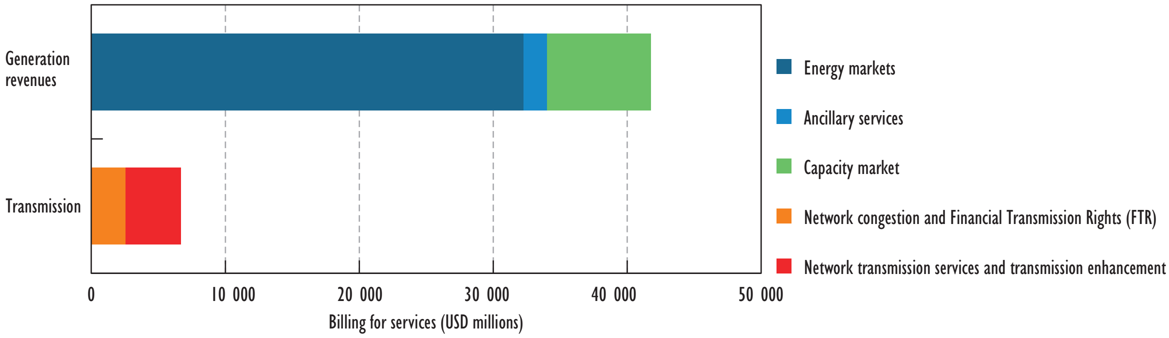
Long-term investment typically involves taking decisions on long-lived assets that will operate well beyond the three years of most forward markets. Beyond these time horizons, investors have to make reasonable long-term assumptions regarding the evolution of demand growth, the evolution of the capacity mix and fuel prices, and all the other fundamentals of electricity prices.

Consequently, there are two long-term markets: the first for capacity and the second for long-term contracts for the off-take of electricity at a predefined price.

Capacity markets are typically mechanisms where a system operator procures or imposes capacity requirements (in MW), often three to four years in advance. In capacity auctions, different capacity resources such as generators, demand response, and in some cases storage and interconnections, bid a price for providing available generation capacity. In PJM, capacity markets represent around 20-30% of generators' revenues (Figure 1.9). The capacity does not have to actually produce electricity, but only to be available in case of need.

Long-term contracts for off-take of electricity include power purchase agreements or feed-in tariffs. The contract duration can vary between 10 and 35 years, for long-lived investments such as nuclear power plants. Such agreements can be bilateral contracts between a utility and an independent power producer. Very often, however, they involve government intervention aimed at promoting new investment, either via an obligation or a regulated price. These long-term contracts can be the result of procurement mechanisms, such as auctions.

Figure 1.9 • PJM billing for services (2014)



Source : PJM (2015)

Key point: Generators' revenues come mainly from the energy markets and, where they exist, capacity markets.

Regulation

Regulated activities primarily include the transmission and distribution networks. Tariffs that generators and consumers pay for using the grid are subject to economic regulation and cost recovery, including return on investment. Regulated prices include connection charges, planning and network investment.

Regulators also usually have to approve market rules, in particular for the short-term and long-term markets. The detailed technical rules for balancing and ancillary services can increase the cost of regulated activities and are subject to regulatory approval. The introduction of a capacity market, for instance, is not spontaneous, but is decided by governments and regulators by law or direction. Regulators also mitigate market power on the different markets.

The price of long-term contracts is also usually regulated in order to ensure new investment, and is set according to the investment cost of the technologies being developed. To date, long-term prices, such as feed-in tariffs or power purchase agreements signed to meet a renewables obligation, have been above market prices. The additional costs are passed through to final consumers in the form of a surcharge calculated by the regulator.

Retail market

In the retail market, consumers' bills cover all the costs arising from the previous markets and regulated activities. This includes energy costs, capacity costs, network costs, the cost of different obligations, in particular renewables, as well as taxes. With the exception of large industrial users, consumers are not active directly in electricity markets, but buy their electricity from retailers. Very few consumers directly experience the variations in short-term markets. Retailers usually offer simpler retail tariff structures with a limited number of price components.

Retail competition has been introduced in Europe, Australia and some US states. Under this approach, retailers compete to sell electricity and make commercial offers that the final consumer can choose. The commercial offers can differ in the nature of the electricity provided (e.g. green or not), the average price, tariff structure and time differentiation. Some retailers also integrate services to manage energy consumption in their offer, or even to generate or store electricity behind the meter.

Interactions between low-carbon policies and electricity markets

The transition to low-carbon power is a major challenge for market design and the regulatory framework. Where carbon markets have been introduced, they are reflected in short-term prices, medium-term prices, and to some extent in long-term contracts. Governments, however, also promote energy efficiency, renewables and in some cases nuclear and CCS. Policy proliferation significantly increases the complexity of an already complex set of markets.

The interaction of low-carbon policies and electricity market design can have unintended consequences. While a carbon price increases prices in electricity markets, renewables policies and energy efficiency policies can have the opposite effect of reducing wholesale electricity prices. This makes it more difficult for markets to incentivise other low-carbon investment in nuclear or CCS, or even renewables.

Similarly, renewables deployment increases low-carbon generation, which makes it easier to meet the cap on emissions in Europe and might tend to reduce carbon prices. Electricity markets can then dispatch coal power plants that are cheap and displace gas power plant which would emit less but are more expensive. This would create a paradoxical situation where growth in renewables does not reduce CO_2 emissions. The history of carbon price collapse in the EU Emissions Trading System, however, reflects several other factors including the impact of economic crisis and inflows of international carbon credits. Carbon market design can also address these risks: the adoption of a new Market Stability Reserve mechanism within the EU ETS aims to stabilise the carbon market against risks of exogenous events and of policy interactions.

Renewable support policies have been successful in deploying renewables, but in certain countries renewables have significantly increased electricity bills. The rapid deployment of wind and solar power has also made a limited contribution to meeting peak demand, while displacing conventional fossil-fired capacity. This argument, discussed in Chapter 2 and Chapter 4, is often presented as a justification for the introduction of capacity markets.

Conversely, capacity markets are another example of potential misalignment between different instruments. Many analysts consider that the introduction of capacity markets provides additional revenues for coal plants in order to keep them available, while coal generation should be reduced. They argue that, as they are available, coal plants are likely to run for a longer period of time and thus increase CO_2 emissions or require a higher carbon price in the future. While capacity remunerating mechanisms are not meant to reduce CO_2 emissions, they are a political construct that has been used to serve multiple purposes.

It is clear that different policies are interacting with one another, sometimes in unexpected ways that are not always aligned with the intended transformation of the power sector (OECD, 2015).

References

Andrews-Speed P. (2013), "Reform Postponed: The Evolution of China's Electricity Markets" in Sioshansi P. Ed., *Evolution of Global Electricity Markets*, Elsevier.

Borenstein (2014), "The U.S. Electricity Industry after 20 Years of Restructuring", https://ei.haas.berkeley.edu/research/papers/WP252.pdf.

Cobb J. (2015), "One Million Global Plug-In Sales Milestone Reached". HybridCars.com, www.hybridcars.com/one-million-global-plug-in-sales-milestone-reached/.

Cramton P. (2015), *Market Design in Energy and Communications* www.cramton.umd.edu/papers2015-2019/cramton-market-design-in-energy-and-communications.pdf.

Davis L. W. and Wolfram, C. (2012), "Deregulation, Consolidation, and Efficiency: Evidence from U.S. Nuclear Power", *American Economic Journal: Applied Economics*, Vol. 4, No. 4.

DECC (Department of Energy and Climate Change) (2014), "Implementing the Emissions Performance Standard: Further Interpretation and Monitoring and Enforcement Arrangements in England and Wales", DECC, United Kingdom, www.gov.uk/government/uploads/system/uploads/attachment_data/file/357217/implementing_emissions_performance_standard.pdf.

EIA (Energy Information Administration) (EIA)(2014), *Annual Energy Outlook for 2014*, EIA, United States.

The Financial Times (FT) (2014), "Lobbyist's take on renewables causes it to lose friends", www.ft.com/intl/cms/s/0/9b05ad2a-5ab0-11e4-b449-00144feab7de.html#axzz3qLtGE1dY

German Energy Blog (2015), "Climate Reserve Agreement on Closing 2.7 GW of German Lignite Fired Power Stations – Expected Closure Costs EUR 1.61 Billion", www.germanenergyblog.de/?p=19505.

Glachant, J.-M. (2010), "The achievement of the EU electricity internal market through market coupling", European University Institute, EUI Working Paper RSCAS 2010/87.

Glachant, J.-M. and F. Lévêque (eds.) (2009), *Electricity reform in Europe*, Edward Elgar.

IEA (International Energy Agency) (2015a) *World Energy Outlook* (*WEO*) *Special Report*, *Energy and Climate Change 2015*, OECD/IEA, Paris.

IEA (2015b), *Energy Technology Perspectives*, OECD/IEA, Paris.

IEA (2015c), *Electricity Information* www.iea.org/publications/freepublications/publication/Electricitytrends.pdf.

IEA (2015d), *World Energy Outlook 2015*, IEA/OECD, Paris.

IEA (2014a) *Energy policy of IEA countries - the European Union,* OECD/IEA, Paris, www.iea.org/bookshop/486-Energy_Policies_of_IEA_Countries_-_The_European_Union.

IEA (2014b), *World Energy Outlook 2014*, OECD/IEA, Paris.

IEA (2014c), *Seamless Power Markets: Regional Integration of Electricity Markets in IEA Member Countries*, Insights paper, OECD/IEA, Paris, www.iea.org/publications/freepublications/publication/SEAMLESSPOWERMARKETS.pdf.

IEA (2013), *The Global EV Outlook,* OECD/IEA, Paris, www.iea.org/publications/globalevoutlook_2013.pdf.

IEA (2011), *Deploying renewables*, OECD/IEA, Paris.

IEA/NEA (Nuclear Energy Agency) (2015), *The Projected Cost of Generating Electricity*, OECD Publishing, Paris, www.iea.org/bookshop/711-Projected_Costs_of_Generating_Electricity.

Joskow, Paul L. (2001), "California's Electricity Crisis", *Oxford Review of Economic Policy*, Oxford University Press, Vol. 17, No. 3.

Joskow P. J. (2007), "Lessons learned from electricity market liberalisation", http://economics.mit.edu/files/2093.

Newbery, D.M. (2012), "Reforming Competitive Electricity Markets to Meet Environmental Targets, *Economics of Energy and Environmental Policy*, Vol. 1, No. 1, pp 69-82.

OECD (2015), Aligning Policies for the Transition for a Low-Carbon Economy, OECD publishing, Paris.

PJM (2015), Annual report 2014 - Financial report, PJM, http://www.pjm.com/~/media/about-pjm/newsroom/annual-reports/2014-financial-report.ashx.

Pinguelli Rosa L. et al. (2013), “The Evolution of Brazilian Electricity Market”, in Sioshansi P. Ed., *Evolution of Global Electricity Markets*, Elsevier.

REEEP (Renewable Energy and Energy Efficiency Partnership) (2013), “Policy Database 2012-2013”, www.reeep.org/policy-database.

Sen A. and T. Jamasb (2013), “Not Seeing the Wood for the Trees? Electricity Market Reform in India”, in Sioshansi P. Ed., *Evolution of Global Electricity Markets*, Elsevier.

Sioshansi F. P. (ed.), (2013), *Evolution of Global Electricity Markets*, Elsevier.

The Economist (2013), “How to lose half a trillion euros”, www.economist.com/news/briefing/21587782-europes-electricity-providers-face-existential-threat-how-lose-half-trillion-euros.

Chapter 2 • Low-carbon generation investments

HIGHLIGHTS

- A strong carbon price should be introduced to attract new investment in low-carbon generation, as this is an efficient approach to internalising the climate externality. However, governments must recognise that this is likely to take time and to increase potential investors' perception of market risk.
- By 2050, a market based on energy prices (energy-only market) with a carbon price could drive the transition to a low-carbon power system under certain scenarios. This might be the case if demand response continues to progress and storage costs fall, or if carbon and gas prices drive wholesale prices to a level high enough to recoup low-carbon investment costs, including a return.
- Other scenarios challenge energy-only market design in a more fundamental way. For example, scenarios featuring very high shares of wind and solar could drive down wholesale electricity prices and impede the recovery of high upfront investment costs. Similarly, if distributed generation dominates new investment, the role of retail pricing would increase markedly.
- In any case, market prices can provide a very important market feedback loop on the relative value of different low-carbon technologies. Low-carbon generators can earn a significant fraction of their revenues from markets. Fully integrating low-carbon generation into markets can provide the incentive for low-carbon projects to maximise their value to the system.
- But energy market revenues alone may not deliver low-carbon investments at the required speed and scale. At the time of writing, wholesale electricity prices in the Europe and North America are in the range of USD 30-50 per megawatt hour (MWh). These are far too low to recoup investment costs and could remain low for most of the transition period if rapid low-carbon investment leads to prolonged excess capacity.
- During the transition, government intervention is necessary to promote long-term arrangements. Low-carbon investments are capital intensive and this cost structure does not fit well with short-term marginal costs due to carbon price risk and fossil fuel price risk. Long-term arrangements can provide visibility and mitigate risks for investors and keep financing costs low. These arrangements are likely to remain technology specific.
- For example, providing support in the form of a market premium that is modulated while strengthening carbon pricing would contribute to integrating low-carbon investments into the market while mitigating market price risk.
- Auctions can introduce competitive forces to determine the level of support needed, on top of market revenues. Auctioning can help reduce asymmetry of information on costs and market forecasts.

Attracting capital to build low-carbon power plants entails investors "making a bet" on the policies and technologies of the low-carbon energy transition. Some investments in low-carbon power plants are made purely on a cost-competitive basis, independent of emission-constraining goals. But when cost-competitiveness is an insufficient basis for investment, policy makers and regulators currently employ two approaches to incentivise low-carbon investment and underwrite this "bet". The first approach imposes carbon emission controls on fossil-fuel generation, including but not limited to carbon pricing. The second approach offers various types of preferential arrangements, including long-term power purchase agreements (PPAs).

In the early years of policies to decarbonise the power sector, many policy makers focused on introducing a strong carbon price to ensure decarbonisation. This pricing approach corresponds to a textbook vision of correcting the climate externality, and could be in the form of a carbon tax or a cap-and-trade emissions market, such as those that have been introduced in the European Union in 2003, in parts of the United States and Canada since 2008, in South Korea since January 2015, and in China, where a national carbon market will be introduced as part of the 13th Five-Year Plan for 2016-20 (World Bank, 2015). In theory, strengthening and expanding the reach of carbon price signals could be sufficient to incentivise future low-carbon generation investment and achieve 2050 objectives. However, this could lead to very high and volatile wholesale electricity prices, and carbon prices have proven politically challenging in many countries.

In practice, alongside carbon pricing, recent investments in new low-carbon technologies (such as wind, solar and biomass) have benefited from additional measures to facilitate their deployment, often in the form of preferential long-term PPAs, usually at a fixed uptake price. The vast majority of such long-term contracts have, in one way or another, been backed by governments. Such contracts often not only subsidise low-carbon investments, but also insulate investors from electricity market risks. By 2014, non-hydro renewables amounted to 6.3% of electricity production in OECD countries.

This model is quite different from that of the older low-carbon fleet of nuclear and hydropower generators, which was largely financed and built directly by vertically integrated utility monopolies. The eventual retirement of this existing fleet – in particular in countries that choose not to replace ageing nuclear facilities or accelerate their retirement schedule – will create additional demand for new low-carbon installations. In recent years, nuclear and hydropower have accounted for roughly one-third of total electricity production in OECD countries (31.9% in 2014), and nuclear power alone accounted for around 18.4% in 2014 (down from 22.4% in 2005), three times the production of non-hydro renewables.

In addition, experience in Europe during the past decade has highlighted a number of unpredicted outcomes that stem from the interactions between energy policy and the wider economy, which have compounded the uncertainties surrounding climate policy. In Europe, the chief executives of traditional power utilities have argued that the power sector has become "uninvestible" (Magritte Group, 2015), and that the current policy framework is unlikely to provide sufficient incentive for new low-carbon investment on the scale envisaged under EU roadmap scenarios.

Consequently, a perceived tension has emerged between the carbon pricing approach and the long-term PPA approach. In the view of some stakeholders, relying on attracting low-carbon investment primarily by means of preferential long-term PPAs could lead to a split market: a regulated market for low-carbon generation and a competitive one for conventional power. Others argue that a shift to a "single driver" carbon price policy would be inadequate, both because existing policies also seek to achieve non-climate objectives and because of the political feasibility of strong carbon pricing.

Within this context, this chapter provides a short discussion on market design in relation to the investment needs of decarbonised power generation. It initially considers the end goal – what might be the different technology configurations of a decarbonised power sector in the year 2040 or 2050, and what might each of these mean for the functioning of the market? The next section discusses the various market and regulatory failures that might reduce investment in low-carbon generation during the transition period. Finally, concluding that interventions are necessary if ambitious decarbonisation is to be achieved at a steady "walking pace" rather than through an inefficient "crawl, then sprint" model, the chapter briefly examines the range of instruments that are available and what trade-offs they typically involve.

2.1. Aligning electricity market design and low-carbon electricity: What does it take?

Designing power markets that can facilitate investment in decarbonisation is a multifaceted challenge. The central objectives include pricing in externalities, enabling returns on capital-intensive technologies, overcoming the lock-in of existing high-carbon generation, and in addition, the need to ensure operational efficiency while also securing sufficient investment in flexible resources.

Two elements of these challenges need to be distinguished when mapping an electricity market design that can deliver in a decarbonised electricity system:

- First is the question of what the generation mix of a future fully decarbonised system may look like. This revolves around which technologies will be available in low-carbon power systems – the relative potential and cost of renewables, nuclear, carbon capture and storage (CCS), power storage and demand response, which are all still under development.
- And, second, linked to this, is the question of the extent to which the current paradigm of electricity market design remains applicable under these circumstances.

An often-mentioned problem in this regard is that prices based on short-term marginal costs cannot cover the costs of technologies requiring high levels of capital expenditure, which are characteristic of any deep decarbonisation scenario. Is this a true challenge or a fallacy?

Low-carbon market design and technology options

Electricity is unlike other commodities. It presents three distinctive features that have shaped the design of existing systems. First, electricity demand is highly volatile and remains inelastic to prices in the short run. Second, it can only be transported via a grid where supply and demand are balanced in real time to maintain the technical integrity of the system (i.e. keeping the system frequency at its target level of 50 or 60 Hertz). Third, electrical energy cannot currently be stored at reasonable cost in very large quantities.[3] Consequently, wholesale electricity market prices are highly volatile and trade is constrained by the physical requirements of the grid.

By 2050, technological progress may have overcome a number of these constraints. First, electricity storage technologies are progressing both for short-term applications, in intervals of a few seconds to a few hours, and for periods of day-into-night (crucial to solar photovoltaic [PV] deployment). Technologies that can provide longer-term storage spanning multi-day weather patterns and seasonal cycles are also being researched. Significant investment in short-term battery storage solutions is being driven by investments in electric vehicles and synergies with PV and mobile information technology (IT) applications. While this will not eliminate the need to balance electricity supply and demand in real time, storage has the potential to radically reduce the need for peaking and mid-merit capacity, and to ensure better use of other capital-intensive power plants. In a more disruptive scenario, storage could greatly affect the plant mix, allowing for a much higher penetration of variable resources.

Second, demand response technologies are already a mature option, bringing a degree of flexibility to large industrial consumption. Tapping into the much larger but highly fragmented potential of the commercial and residential sectors is becoming possible thanks to the exploding capabilities of

[3] Pumped hydro storage is an exception, but in order to store very large amounts of energy, reservoirs need to be very large both for storing and pumping water when electricity is available.

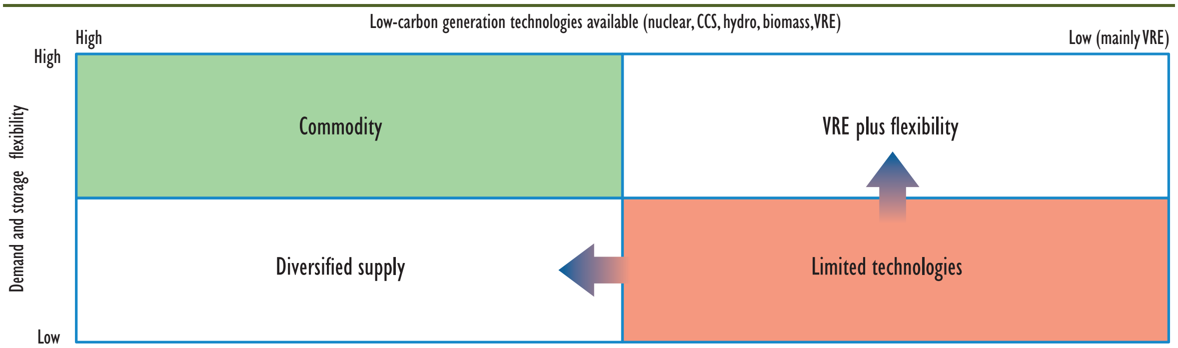
IT, at ever-decreasing cost. As for storage, demand response helps reduce the need for peaking and mid-merit plants and can mitigate generation surpluses by increasing demand when low-cost power is available. Both technologies have been the "holy grail" of the power sector for decades, and technological progress now offers the potential to make them a reality.

On the supply side, a range of low-carbon generation technologies is available. The course of the low-carbon transition and current technological development strongly indicates a future in which large volumes of generation will come from variable renewable energy (VRE) sources. But the relative roles of nuclear and CCS in complementing VRE generation are harder to forecast; all of these options are subject to technological development uncertainties – at least to ensure cost discovery – and also to doubts about public acceptance and/or consumer uptake.

Restricting the number of available technologies rapidly creates complications in making reasonable assumptions in the market design question. If nuclear is not acceptable, and if no reservoir is available to store carbon dioxide (CO_2) by implementing CCS, or if the potential to use biomass sustainably is limited, then wind and solar power may emerge as the solutions by exclusion of alternatives. While several studies conclude that 100% renewable power is technically feasible, these scenarios do not envisage VRE without demand response and storage. Under a scenario in which low marginal cost generation without the option to defer output (e.g. VRE without supply-side storage) constitute the only participants bidding onto a short-term spot market, it would be difficult to imagine sufficient revenues to cover costs in competitive market environment.

To a large extent, the long-term market design discussion is overwhelmed by the uncertainty surrounding the future availability and acceptability of technologies. The various possible futures can be represented in a matrix of four potential technological outcomes, as presented in Figure 2.1, differentiated according to two axes. The horizontal axis corresponds to the number of low-carbon generation technologies available: hydro, wind and solar, but also biomass, nuclear and CCS. The vertical axis corresponds to the degree to which technologies enable the development of demand and storage flexibility. They are currently relatively low and are expected to improve in the coming decades.

Figure 2.1 • Market-technology matrix



The matrix defines four possible futures for the design of the market:

- Diversified supply corresponds to situations in which the available low-carbon technologies have different cost structures. While capital intensity will broadly be higher than in today's systems, one may hope that existing market design could, in principle, work in combination with a high and robust carbon price. This might entail keeping some non-abated gas technology, which would lead to very low but not zero CO_2 emissions.
- Limited technologies envisages a situation in which wind and solar (VRE) power are largely dominant, certain countries have decided to phase out nuclear, CCS is not ready at commercial scale, while demand and storage flexibility remain limited. Decarbonisation would

then imply installed capacities much higher than peak demand, and low marginal cost generation setting the price almost all the time. Under this scenario, marginal cost pricing of energy is unlikely to ensure the recovery of upfront investment costs. However, this case could expect to result in pressure to innovate and invest in other dispatchable and base-load renewable sources – such as hydropower, geothermal and bioenergy.

- VRE plus flexibility represents the combination of VRE and highly flexible storage and responsive demand. This enables electricity to be stored for relatively long periods. Under this scenario, high flexibility smooths out electricity prices, but hourly price differences would still be important in driving investment in storage and demand response.
- Commodity refers to a future in which many technologies are available, power can be stored and demand is driven by power availability. Variable or base-load generation technologies can run at maximum full-load hours. The electricity market would then become even more comparable to other commodity markets, such as gas, in which prices show a lower volatility within days or months.

Among these possible futures, the advantage in increasing the portfolio of available technologies is clear, by moving from the "Limited technologies" case to the "VRE plus flexibility" case or the "Diversified supply" case. The choice crucially depends on the potential to increase storage and demand response. With storage costs significantly lower and demand more responsive to prices, the "Commodity" scenario would then become possible. If it is economic to store the output of wind and solar power, it should also be easier to store the output of a plant running around the clock – such as from CCS or nuclear. In the "Commodity" case, the cost of meeting variations in electricity demand and supply should be sufficiently low to smooth out the volatility of electricity prices.

Power systems dominated by hydro today provide an interesting insight into how such a system might perform. In Brazil, for example, hydropower has enabled electricity prices to remain very stable from one week to the next. When reservoirs contain insufficient water, electricity has to be generated from gas-fired power plants, which sets the opportunity cost of hydro in reservoirs. Short-term volatility does not disappear but is significantly reduced. This model, however, has recently comes under strain in Brazil due to enduring drought conditions, which are likely to become more frequent in certain parts of the world as a result of climate change, even under a 2°C scenario. Other examples of hydro-dominated power systems can be found in Canada and Norway, all of them highlighting the importance of seasonal and inter-annual supply-side variability. Inter-annual variability can be higher for water than wind and solar resources.

To sum up, low-carbon electricity market design for 2050 depends on the availability of key technologies. As more VRE sources are integrated into the electricity system, the capability of markets based on marginal cost pricing to recover the cost of new investment will be increasingly challenged, and will depend on the technological progress of demand response and storage.

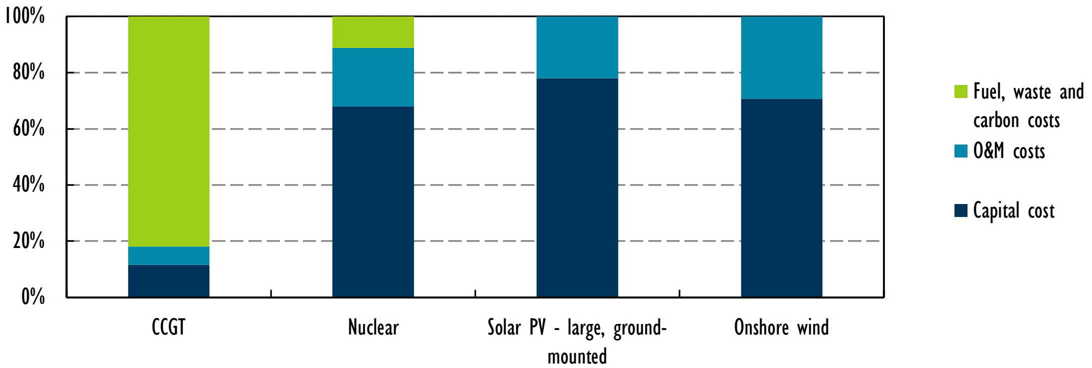
Capital-intensive investments: The infrastructure financing puzzle

One feature of today's low-carbon generation technologies is that the up-front investment cost per kilowatt (kW) of installed capacity is usually two to five times higher per unit of electricity generated than for gas-fired power plants (Figure 2.2). Over time, the lower fuel costs of wind, solar PV and nuclear partly compensate for the higher capital cost, but VRE technologies generate only when the wind is blowing or the sun is shining. Thus, solar and wind capacity utilisation factors (or load factors) are relatively low, in the range of 9% to 30% for solar and 20% to 50% for wind. As discussed later, high up-front investment costs have important consequences for the risk profile of low-carbon investments under different market arrangements.

Low-carbon investment projects differ in size and maturity. Large hydropower and nuclear power plants are large-scale projects well understood by many electric utilities. Solar PV and onshore wind are more recent technologies, but are now mature in a growing number of markets. Offshore wind and CCS are less mature, large-scale projects that still present technological risks. Despite massive cost reductions, most low-carbon technologies are still relatively expensive, and some might remain so during the transition to low-carbon electricity systems (see IEA, 2015).

Consequently, policy makers are generally concerned with facilitating the financing of low-carbon investment projects.

Figure 2.2 • Breakdown of the levelised cost of various technologies by cost component, United Kingdom by 2020 (7% discount rate)



Notes: CCGT = combined-cycle gas turbine; O&M = operation and maintenance.
Source: IEA/NEA, 2015.

The primary finance structure for renewable energy investment has been and continues to be corporate finance via the balance sheets of electric utilities and project developers. Such projects are financed on the basis of the strength of the developer's balance sheet, and are therefore dependent on investors' willingness to purchase the developer's debt and equity, and thus on the credit worthiness and balance sheet health of the developer. In particular, well-capitalised state-owned enterprises are often well positioned to finance projects on their balance sheets. However, if developers' balance sheets are constrained, project finance will be the mechanism of choice.

It is possible to finance capital-intensive investment at low cost, but this requires revenue certainty. Onshore wind, for example, is already reported to have been tendered at a price of around 50 USD/MWh in Brazil and the United States, and 41 USD/MWh in Egypt. Solar PV projects have been tendered at 58.4 USD/MWh in Dubai, and 63 USD/MWh in South Africa (IEA, 2015).

Financing capital-intensive investment is also possible with revenue uncertainty, but this comes at high cost of capital. Seeking finance for capital-intensive projects at the lowest possible cost of capital in a context of long-term uncertainty is a puzzling problem. Managing the energy transition will require new solutions.

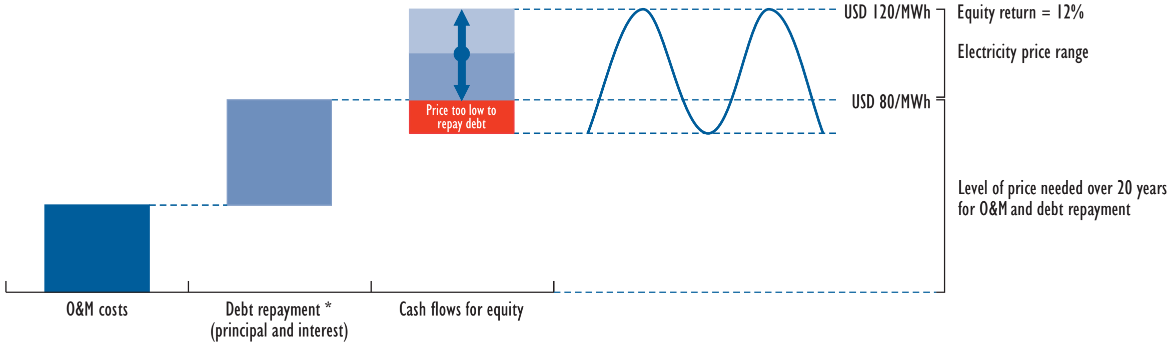
Note that the existence of high fixed costs is not *per se* a market failure and should not therefore constitute a justification for a regulatory intervention. It is often said that "an industry with high fixed costs cannot cover its costs with a market design based on marginal cost pricing". It is true that a wind turbine, for example, has proportionately high fixed capital costs and a low marginal cost of production. But this does not mean that renewables are natural monopolies. Most low-carbon generation plants have a modular structure, so that wind and solar power plants can be built as long as the revenues they can generate on the market are sufficient to recover the investment cost of a new unit, limiting investors' exposure to risk.

The infrastructure financing puzzle results largely from the risks associated with the cash flows that are needed to recover investment costs, including a return on capital invested. The cash flows of a typical project are illustrated in Figure 2.3 below.

Note that:

- The first 20 USD/MWh would recover the O&M costs.
- Assuming that a price of 80 USD/MWh is certain enough, this would provide resources to repay the principal and interest of a project debt representing 75% of investment costs. A lender would probably refuse to finance more debt than the amount the guaranteed revenues can reimburse each quarter.
- The remainder would have to be financed with equity, and the equity investor would have to be confident that the price could reach around 100 USD/MWh sufficiently often to secure a reasonable return on investment. If electricity prices turned out to remain at 80 USD/MWh, the equity investor would never be able to get its money back. If electricity prices turned out to fall below 80 USD/MWh, the cash flows would be insufficient to reimburse the debt. This would lead to a default in the case of project finance.

Figure 2.3 • Cash flows of a wind project (USD/MWh)



* Average over 20 years, assuming a project financed with 75% of debt and a cost of debt of 5%

In industries with high market price risk, such as oil, investors usually have the option of choosing projects with high profitability. For example, new oil projects were typically developed with a reference break-even price of USD 40-50 per barrel even in a market environment with prices at USD 80-100 per barrel. In the case of low-carbon investments, market prices are not yet high enough to recover their costs (even with a low cost of capital) and ensure remuneration commensurate with the market price risk.

In theory, adding a carbon price on top of market prices solves the problem. But even assuming perfect certainty about the level of the carbon price (which is unlikely with a cap-and-trade system unless it has a price corridor [price cap and price floor]), its level would have to be high enough to compensate for the uncertain electricity prices associated with fossil-fuel commodity prices. The level of carbon price necessary to attract and remunerate risky investments is likely to be much higher than has been seen in any carbon market to date. One key success factor for decarbonisation is to keep the cost of decarbonisation as low as possible. Solving this puzzle involves attracting low-cost capital to finance risky investments.

Distributed generation and retail pricing design

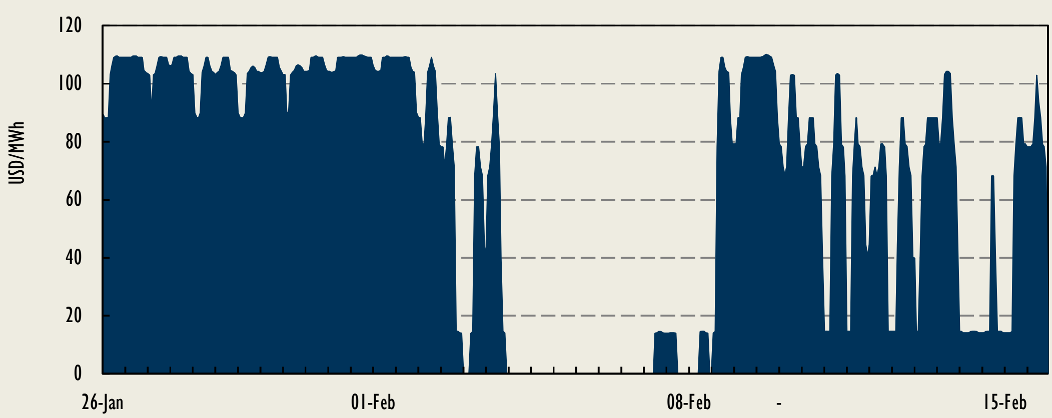
By and large, renewable generation tends to be connected to the distribution network, rather than the transmission grid. Even if wind, solar PV or biomass power plants change the way distribution networks are operated, the traditional regulatory framework remains fairly robust; the generated electricity is injected into the public network and supplied to other customers. It can be metered, billed and, importantly, taxed.

Box 2.1 • What can be learned from modelling electricity prices in 2050?

A specific illustrative example of a simplified dispatching model is used to explore the question of prices, using data derived from the International Energy Agency (IEA) *Energy Technology Perspectives* (*ETP*) model to provide a first order estimate for the winter month of February in Europe.[1] The scenario corresponds to the Diversified supply field of the 2050 electricity matrix presented above, with a high diversity of generation technologies and relatively little demand response and storage capacity.

In the IEA 2°C Scenario (2DS) presented in *ETP 2014*, electricity is almost fully decarbonised by 2050. Conventional fossil-fired generation has almost disappeared in Europe. More than 60% of electricity is generated by a combination of nuclear (21.5%), wind (31%) and solar PV (11%), all technologies with low marginal cost. Given the variability of wind and solar output, Figure 2.4 shows that hourly prices look like a "canyon landscape", with deep ravines between hours of high prices. Prices are high during hours when gas or coal, most often with CCS (10%), is needed. Prices are nil or very low during hours with high wind and solar PV output. A degree of routine curtailment is applied when supply exceeds demand, as is security of supply curtailment when grid stability constraints become important.

Figure 2.4 • Electricity prices, three weeks in 2050



In the IEA 2°C Scenario (2DS) presented in *ETP 2014*, electricity is almost fully decarbonised by 2050.

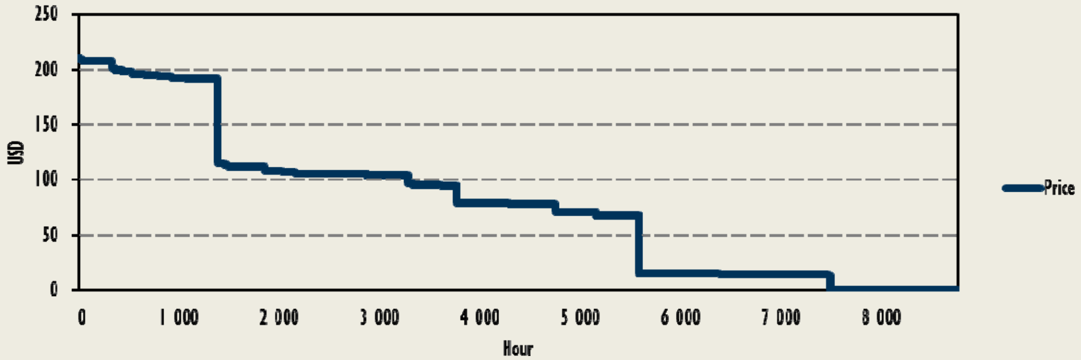
Due to the fact that wind and solar output are weather-dependent, all plants of each type tend to produce at the same time (this is sometimes referred to as auto-correlation or the cannibalisation effect). Depending on the level of system flexibility, this leads to situations of low or zero prices when the output of wind and solar power is high.

The hourly prices can also be represented using price-duration curves, where prices are ranked by declining value over one year with 8 760 hours, as in Figure 2.5. Assuming that 20% of conventional synchronous capacity must run in order to maintain system stability, wholesale prices in 2050 are equal to zero for 1 000 hours per year, with very low prices between zero and 20 USD/MWh for around 3 000 hours per year, or one-third of the time, for a share of wind and solar PV output representing 43% of total electricity generated. Beyond this point, the number of hours with zero prices increases further.

This "price suppression" effect reduces the market revenues that generators can expect. By 2050, it is calculated that the average wholesale price remains at a relatively high level of 78 USD/MWh. The increasing number of hours at zero marginal price are compensated by high CO_2 prices at USD 100 per tonne of CO_2 (tCO_2), which push prices above 100 USD/MWh for more than 3 500 hours. Despite significant volatility, prices remain at a sustained level on average. While average prices decrease with increasing variable renewables, most conventional technologies and nuclear can expect significant revenues even in close to fully decarbonised power systems with a Diversified supply scenario.

Box 2.1 • What can be learned from modelling electricity prices in 2050? (continued)

Figure 2.5 • Modelled price duration curve in *ETP* scenarios, 2050



As a result of zero and lower prices when there is a lot of wind or sun, the market revenues of VRE also decline rapidly below the average market price.

Wind generates 31% of electricity by 2050, and its market revenues represent 40 USD/MWh, or 50% of the average market price. These market revenues still represent two-thirds of the levelised cost of electricity (LCOE) of onshore wind, which is assumed to be around 60 USD/MWh by 2050. Note that the model relies on the scaling-up of historical wind infeed time series, which tends to yield lower market value at high shares than new turbine models.

Solar generates only 11% of electricity by 2050 and its market revenues represent 70 USD/MWh, or 90% of the average market price. Solar PV market revenues are close to the LCOE of utility-scale solar PV by 2050.

In light of these indications, it can be inferred that revenues from electricity markets could deliver a relatively high fraction of the revenues of low-carbon generators in 2050. While the model results are easy to understand, it relies on a number of simplifying assumptions that can play in different directions. On the one hand, the model does not fully capture the technical constraints of conventional power plants, causing high ramp-up and start-up costs, and it considers only one typical year. On the other hand, the model does not capture extreme demand events, which could lead to higher prices, while more optimistic assumptions regarding storage or demand shifting capacity would also reduce the number of hours with zero prices, leading to higher average electricity prices. All in all, the model is almost certainly wrong, but uncertainties point in different directions.

More crucially, the model uses an exogenous capacity mix. Taking a scenario with very high shares of wind and solar power could lead to lower average electricity prices and even lower market revenues for wind and solar power generators. In contrast, imposing a requirement that all low-carbon technologies cover their costs with revenues from the energy market would imply a different generation mix by 2050. For instance, Lion Hirth (2013) calculates the optimal share of wind under different assumptions. The variability of wind significantly affects the modelling results and many details have to be factored in.

Note: The key assumptions of the model are described in annexes that can be found at www.iea.org/media/topics/electricity/repoweringmarkets/annexes.pdf. This simplified model assumes that a minimum level of conventional capacity must operate to ensure network stability and be able to ramp up quickly, but does not take into account the plant-specific start-up costs and ramp rates that also play a role in price formation.

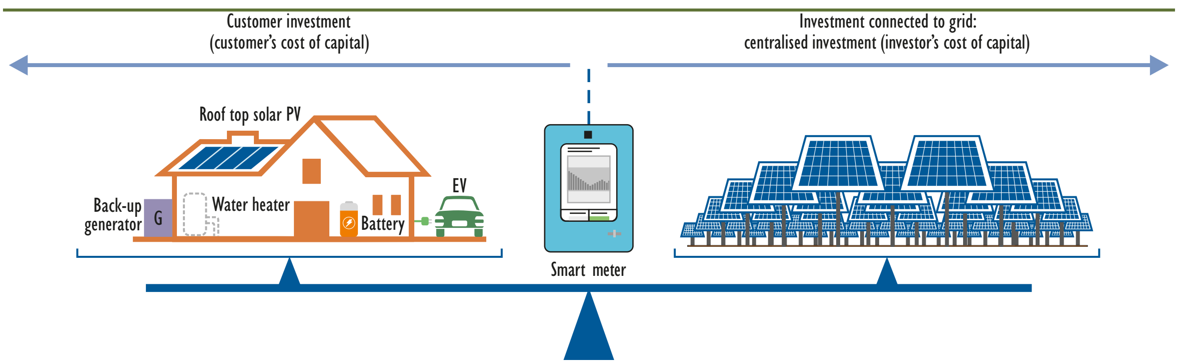
The rise of smaller-scale distributed generation, by contrast, has the potential to be more disruptive for the electricity sector. In particular, rooftop solar PV and micro-generators can be installed behind the meter, i.e. not connected to the public distribution grid directly (Figure 2.6). Associated with local battery storage, electric vehicles (EVs), smart water heaters and other local equipments, distributed generation makes it possible for households to generate their own electricity. Some consumers are already willing to invest to become more “self-sufficient”. Looking further into the future, this could lead to a fragmentation of electricity systems, with a large amount of behind-the-meter generation and many micro-grids, micro-utilities or small co-operatives.

This trend has often been connected to the concept of "grid parity", which is used in a number of different ways. Sometimes it implies that competitiveness of solar PV (or other technologies) is reached as against alternative options. Other uses relate to retail prices. To avoid any possible confusion, the IEA has used the term "socket parity" to describe the point at which the LCOE of a given technology (e.g. solar PV) falls to, or below, the per-kilowatt hour retail price of electricity obtained from the grid, i.e. the variable part of a consumer's electricity bill. Socket parity has been reached in a number of markets; however, the possible mismatch in time between solar PV supply and customer demand effectively limits the amount of electricity can be directly self-consumed (IEA, 2014a).

The development of energy service companies is expected to further facilitate the deployment of innovative financing solutions, in particular for solar PV. Solar City in the United States, for example, raised several hundred million dollars of capital in the form of "solar bonds" with an interest rate of between 2% and 4% to finance the upfront cost of rooftop solar PV (Greentech Media, 2014). An additional aspect that makes the economic analysis of behind-the-meter generation more challenging is the uncertain economic value that plant owners attach to producing "their own" electricity.

Unlike generators connected to the transmission or distribution grid, the development of behind-the-meter generation depends on retail prices, not wholesale electricity prices. From this perspective, it is the design of retail prices that matters most, including 1) the possibility to net electricity generated and consumed (net metering), 2) the tariff structure (fixed, capacity charge, variable charge), 3) the surcharges to cover energy policy costs and 4) the taxation of electricity. A key element is that price signals between retail and wholesale markets need to be co-ordinated (real-time price, dynamic pricing) in order to balance the contribution of centralised and decentralised resources (Chapters 8 and 9).

Figure 2.6 • Finding the right electricity price and structure to optimise distributed resources



In order to ensure an efficient decarbonisation pathway, incentives to deploy behind-the-meter generation and storage must be aligned with the costs and benefits for the electricity system (Figure 2.6). The degree to which the system-wide value of PV outweighs its cost depends on a number of factors, including avoided fuel costs, increased or decreased transmission and distribution costs, capacity value (i.e. extent to which PV deployment reduces the need for building other generation capacity), reductions in transmission and distribution losses, and the pricing of externalities (CO_2 and other emissions etc.). Electricity prices should send the right investment signals at the meter.

Conclusion: Can an energy-only market with a strong carbon price signal enable investment in a low-carbon power system by 2050?

This section discusses whether the current paradigm of electricity market design, based on an energy-only market plus a carbon price, might work in decarbonised capital-intensive systems. Our analysis concludes that there is a clear indication that markets should be designed according to the technologies available for decarbonisation and the associated energy mix.

- Under long-term IEA assumptions, with a high carbon price and a generation mix including renewables, nuclear, CCS and gas power plants, as well as a demand response and storage, revenues from electricity markets can represent a significant share of the revenues needed to recover the fixed costs of low-carbon power sources. This remains the case under a 2DS by 2050. The energy-only market with a sufficiently high carbon price can plausibly work in attracting low-carbon investments under these conditions.
- Under other scenarios, it is not possible to rely for transition solely on energy-only markets with a carbon price in decarbonised power systems. The primary issue becomes one of ensuring investment cost recovery in technology-constrained scenarios and/or sending the right investment signals at the retail level.
  - If nuclear, CCS and sustainable biomass are highly constrained, decarbonising will have to take place mainly with wind and solar power. In the absence of significant reductions in the cost of energy storage, the load would have to adapt to the generation available. Low-carbon power would then require very large installed capacities, with low marginal cost generation very often setting prices. Under this scenario, marginal cost pricing of energy is unlikely to ensure the recovery of upfront investment costs. Such a scenario would challenge current market design in a more fundamental way.
  - If behind-the-meter generation experiences rapid deployment, which remains to be seen, then the question is no longer only about the design of wholesale markets. Instead, the more general question of design of retail tariffs and their link to wholesale markets matters more (see Chapter 9).

The above analysis has demonstrated that discussion of long-term market design is overwhelmed by the uncertainty over future availability and acceptability of technologies. Different end point technological scenarios have markedly different implications for market design. At one end of the spectrum of possible outcomes in 2050, current markets plus a robust carbon constraint could deliver low-carbon investments in the long term. At the other, the current paradigm would no longer be applicable, calling for a more fundamental redesign of investment frameworks.

Given this uncertainty, a more appropriate approach to market design is to treat it as an evolutionary process. At this stage, there is a broad consensus that energy-only markets and carbon prices are required, if only for short-term efficiency. However, there is also broad consensus that this will be insufficient in the absence of additional low-carbon support policies.

What is more, delivering decarbonisation means that market design rules have to attract new low-carbon generation investment during the transition period, over the next 15 years, and under current market realities. Consequently, before any conclusion can be drawn on how markets need to be adjusted, the current challenges to low-carbon investment need to be investigated.

2.2. Market and carbon price uncertainty can hinder low-carbon investment

Investors considering market-based investment in low-carbon projects are confronted with four key issues: current wholesale market prices are too low; the prospect of a high carbon price is unclear; fossil fuel prices are uncertain; and capital markets may not be prepared to take on and diversify investment risk.

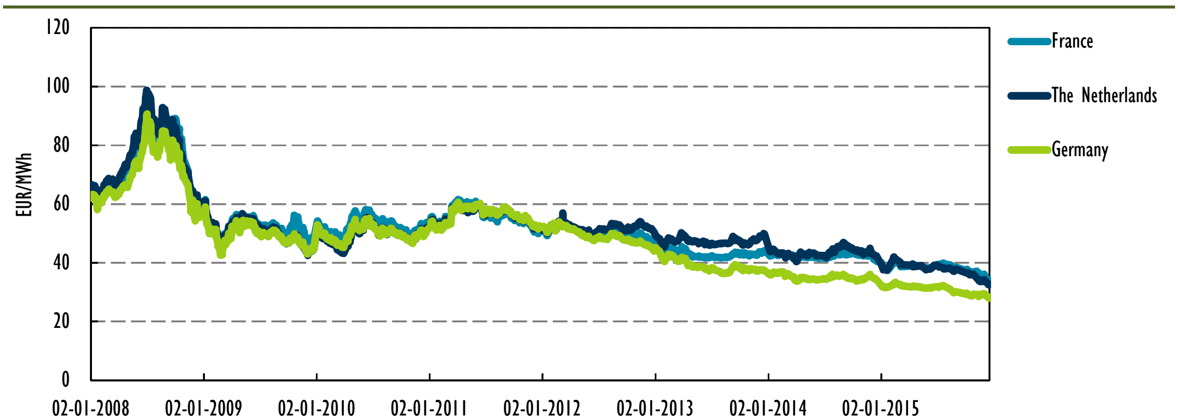
Wholesale electricity prices could remain low during the transition

Power prices at the electricity exchange are lowering due to an increased supply of low marginal-cost power, most often VRE (the merit order effect). Once built, many low-carbon technologies provide low-cost generation and thus consistently take first place in the merit order. Consequently, wholesale market prices decline and all capacities earn less revenue for the hours they operate.

In liberalised markets, the revenues of all generators will be negatively affected if the rate at which new low-carbon generation is added outpaces the need for new investments to meet growing demand or replace ageing infrastructure. The result is that incumbent assets with comparably higher fuel costs experience reductions in their operating hours compared to what they might have anticipated in a situation with higher demand growth and less new capacity in the market for generation. In addition, overcapacity also mutes scarcity prices on the wholesale market (see chapter 4).

This accounts for the economic challenges observed in markets where incentives have prompted rapid growth in wind and solar PV, despite demand growth being sluggish. This situation diminishes the value both of existing assets and new generation, creating a generally poor climate for investment, including low-carbon and conventional generation. Consequently, increasing numbers of generators, particularly gas turbines, are facing a financial challenge to remain in operation and the risk of being mothballed, which can present issues for reliability and adequacy (Chapter 4).

Figure 2.7 • Year-ahead forward market prices for Germany, France and the Netherlands, 2008-15



Source: Bloomberg.

Current electricity prices in Europe (Figure 2.7) and North America are decreasing, mainly driven by low gas, coal and CO_2 prices and excess capacity. The previous analysis suggests that wholesale prices will stay low, and that the situation of excess capacity will last until ageing generators are slowly retired (Green and Léautier, 2015). In pursuing enhancement of market design to cater for low-carbon generation, it is important resolve the reality of transitional overcapacity so as to attract investment in efficient low-carbon generation.

Carbon pricing credibility

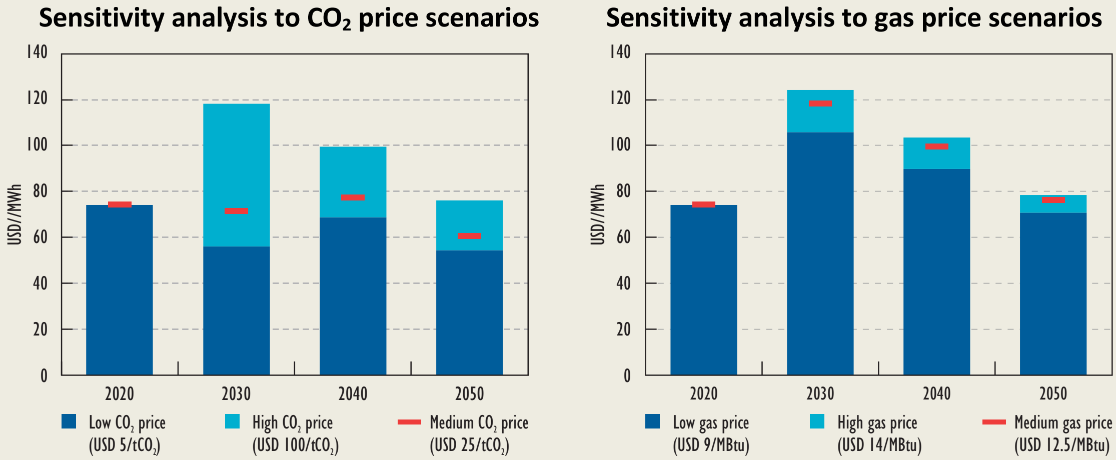
The 195 governments which adopted the UN Paris Agreement have committed to keeping the increase in global average temperature to well below 2°C above pre-industrial levels. This multilateral framework gives a newly clear long-term signal to investors. Nonetheless to translate the commitment into measures and actions, one critical issue for governments to tackle is the political risk associated with carbon pricing.

In Europe, the EU Emissions Trading System (EU ETS) introduced an emissions cap in 2003, which has so far proven to be too loose to create scarcity. In addition, the EU ETS supply mechanism has to date operated independently of economic conditions and demand, resulting in a carbon price too low to affect either operational or investment decisions. The introduction of a new Market Stability Reserve aims to strengthen the scheme, but a yet prices remain below 10 EUR/tCO_2 range. EU leaders have also agreed headline targets for 2030 within which the annual linear reduction factor of the EU ETS cap is significantly tightened to give long-term visibility to investors to 2030 and beyond.

Box 2.2 • Sensitivity analysis of electricity prices to CO_2 and gas prices

In order to quantify the electricity price uncertainty associated with carbon and gas prices, the simplified dispatching model presented in Box 2.1 above has been tested for a range of price assumptions for the years 2020, 2030, 2040 and 2050.

Figure 2.8 • Sensitivity analysis of electricity prices to CO_2 and gas price assumptions, 2020-50



Note: MBtu = million British thermal units.

By 2030, it is calculated that the corresponding average wholesale price can range from 55 USD/MWh to 117 USD/MWh depending on CO_2 price. By 2050, the impact of the carbon price is somewhat reduced as more generation is low carbon, but it continues to significantly influence wholesale electricity prices. This exposes generators to a very high policy risk during the energy transition, on top of the usual market risks.

Based on the long-term gas price assumptions, the difference in electricity prices between low gas price and high gas price scenarious could reach 20 USD/MWh by 2030. This range declines after 2030 as gas becomes less important in the mix.

In the United Kingdom and Australia, carbon pricing has also been a source of uncertainty for investors. In the United Kingdom, in 2011 the government introduced a carbon price floor as a top-up tax on the ETS, which was supposed to reach 70 GBP/tCO_2 by 2030. One year after its introduction in 2013, the government decided to postpone to 2020 implementation of the floor price. In Australia, an emissions trading system with an initial fixed price of 23 AUD/tCO_2 was introduced in 2012, with the objective of coupling this price to the EU ETS. Subsequently, the new government elected in 2013 cancelled the policy.

In the United States, in 2010 an attempt was made to introduce a carbon price at the federal level, but the Waxman-Markey Bill failed to pass in the Senate. At a state level, several regional initiatives, such as the Regional Greenhouse Gas Initiative (RGGI) and the California ETS, have been developed, but the resulting carbon prices remain fairly low, in the range of 5-15 USD/tCO_2.

After more than a decade of carbon pricing, it is fair to say that, to date, this approach has failed to deliver a level consistent with a 2°C transition. The failure of governments to correct the climate externality with a predictable and high carbon price has deep root causes. Carbon pricing creates winners (existing low-carbon producers) and potential losers (high-emitting industries that do not develop lower-carbon processes). The interests at stake are such that intense lobbying activity has repeatedly derailed informed policy-making.

High carbon prices have potentially large distributive and competitive impacts. With a carbon price, say, above 50 USD/tCO_2, average wholesale electricity prices would increase by around 20-40 USD/MWh. This would create significant rents for existing low-carbon or lower-carbon generators such as nuclear and hydro. For example, assuming that a carbon price of 50 USD/tCO_2 increases electricity prices by 20 USD/MWh, the revenues of one single nuclear reactor would increase by around USD 150 million per year.

To date, governments have failed to ensure that the social cost of CO_2 emissions is properly internalised into investment decisions. A track record of stop-and-go policies has also somewhat damaged the credibility of carbon pricing. All in all, it remains difficult to envisage a high and robust carbon price as a single or key driver for either high-carbon retirement or low-carbon investment in the foreseeable future in many jurisdictions. Under current circumstances, more than just a carbon price will be needed to incentivise low-carbon investments.

That said, carbon prices should nevertheless play an important role, and governments should continue their efforts to introduce and strengthen them. Not least, the potential exists for global linking of carbon markets to support an international climate agreement, and the revenues they generate from auctioning of emissions permits could provide a source of low-carbon funding, either for domestic use or international climate finance.

Confronted with the difficulties associated with carbon pricing, some governments have taken alternative measures to constrain carbon emissions using direct regulation. While this can be instrumental in catalysing the transition to a 2°C pathway, it may not ensure sufficient revenues for low-carbon generators and may create regulatory risk for investors in power markets.

As one example of direct regulation, in the United States the Obama administration has turned to the Environmental Protection Agency (EPA) to implement regulations restricting power-sector CO_2 emissions through the Clean Power Plan (CPP). The CPP creates national standards for carbon pollution from power plants, with EPA setting emissions performance rates at the state level, and states developing and implementing individual plans to achieve reductions. States may implement a wide variety of mechanisms, both from the technology perspective, including retrofits, operating limits, energy efficiency or renewable investments, and the market perspective, including pricing carbon emissions, carbon taxes and single or multi-state emissions trading programmes.

Fossil fuel price uncertainty

Besides carbon prices, fossil fuel prices are the other major source of long-term uncertainty in the electricity price. In particular, natural gas is expected to be the marginal fuel and set power market prices for many hours in the move towards a fully decarbonised scenario. Future gas prices and CCS costs therefore largely determine the extent to which wholesale markets can recover the fixed investment costs of low-carbon generation (Newberry, 2012). Gas price uncertainty differs somewhat by region.

In the United States, the downward trend in natural gas prices since 2000 has already hit the profitability of power generation investments, including those in the natural gas industry itself. After reaching a peak of 12.7 USD/MBtu in 2008, the natural gas price dropped to 3 USD/MBtu in 2013. It is now expected that the abundance of shale gas will maintain natural gas prices at relatively low level of 3-7.5 USD/MBtu for the coming decades, but the possibility of higher or lower prices cannot be ruled out over the next 30 years.

In Asia, liquefied natural gas (LNG) prices remain indexed to the oil price through long-term contracts, and therefore follow oil price swings. After reaching around 12 USD/MBtu during the years of USD 100 per barrel oil prices, natural gas prices were cut by half in 2015. The creation of new trading hubs in Asia (IEA, 2012) might play a role in reducing the use of oil-indexed formulas, but such initiatives are not expected to reduce the long-term natural gas price uncertainty.

In Europe, gas prices are also partly linked to oil prices in the case of long-term contracts with major supplying countries (Russia, the Netherlands, Norway and Algeria). Natural gas prices tend to be higher than in the United States and also more stable over time. Needless to say, other issues, such as availability of LNG in the global market and dependence on Russian natural gas pipeline imports, also matter when it comes to the role of gas power generation in the European electricity mix.

Long-term gas price assumptions usually seek to be consistent with CO_2 emission scenarios. According to most energy market modellers, the lower CO_2 emissions fall, the lower gas usage will be, and therefore the lower gas prices should also fall. This correlation is captured in the IEA *World Energy Outlook* scenarios, where the gas price assumption is lower in the 450 ppm NPS by 3 USD/MBtu compared to the CPS. Such assumed negative correlation is often important to reach the conclusion that decarbonisation scenarios are no more expensive than business-as-usual scenarios. During the transition to a low-carbon economy, however, natural gas use may increase, depending on the policy. For example, in the United States, the CPP is likely to drive demand for natural gas in the medium term.

Climate policy scenarios are, however, only one dimension of the uncertainties for gas, alongside the gas cost curve (as seen in the shale gas revolution), the investment cycle over 15 years, the persistence of oil-indexation formulas and the geopolitical situation of gas-exporting countries such as Algeria, Qatar and Russia.

Against this background, an intense debate is taking place about the opportunity to expose low-carbon investments to natural gas price risk. Investors are in principle exposed to fossil fuel price risk. Indeed, many capital investments are made in other industries with long-term price uncertainty, for instance in the oil and gas industry itself, the mining industry or in the telecommunications industry, and these investments are usually expected to be profitable. If the expected returns are high enough, it is possible to find investors willing to take on the fossil fuel price risk. The problem here comes from the high level of carbon price that would be needed to incentivise investment in low-carbon projects exposed to gas price volatility.

Capital market constraints

In the past, financial institutions looked upon investment in large diversified utilities as low risk, and large power utilities could borrow as much money they needed, fuelling large investments. In traditionally regulated markets, strong utility balance sheets have been a standard way to finance power sector growth. For such traditionally regulated utilities, investment costs for new generation can be passed through to bill payers by means of tariff increases. Tariff increases are generally allowed such that the rate of return is competitive, encouraging financial investment. Additionally, utilities can diversify their risks across a portfolio of projects and geographies, and can therefore absorb project-specific risks. However, this model is currently under strain.

On the one hand, deregulation in many markets has changed utilities' investment strategies, and increased the desirability of using project finance rather than financing on balance sheets. On the other hand, low-carbon generation investments will have to take place in a context of financial regulatory changes (most notably Basel III with its tighter capital adequacy requirements), that leads to credit rationing and creates a challenging environment to finance the most risky investments. Any new investment added to the balance sheet is now carefully examined by lenders, because it can affect the credit rating of large utilities, even the largest ones. Consequently, a new power plant will tend to be assessed according to its own merits as if it were project finance.

Capital markets present a challenge because of imperfect information. The lender usually has less information about a project's risks and the functioning of electricity markets than the borrower. In addition, most institutional investors are naturally risk-averse. While this is not specific to the power sector, lenders usually prefer low-risk and easy-to-understand projects, such as contracts with a PPA and little or no exposure to market risks.

Apart from participation through equity markets, to date, institutional investors have not been very active in financing energy infrastructure investments. Yet they represent a large source of potential finance with a long-term investment horizon that fits well with the financing of energy infrastructure. The assets held by institutional investors and looking for investment opportunities represented USD 926 trillion in 2013 (OECD, 2015b), more than ten times the investment needed for sustainable energy in the 450 ppm scenario, which is USD 88 trillion, (IEA, 2014b). Pension funds and sovereign wealth funds, however, have to meet long-term liabilities and usually seek to hedge the risks. They invest in infrastructure on the basis that they benefit from a long-term contract with a guaranteed price.

This trend is reinforced by the fact that interest rates on government bonds are currently very low in OECD countries. In Europe, real interest rates are even negative in certain countries, pushing institutional investors needing to meet their obligations into diversifying their portfolio of investments in order to achieve a higher return and meet their liabilities.

Equity investors are looking for higher remuneration on equity stakes. Equity is remunerated only after the debt has been reimbursed (principal and interest) which concentrates most of the project cash flow risks onto equity investors. For a typical project with 30-40% equity, equity investors expect a return that is in the range of 6-8% above the risk-free rate, or higher. In addition, when an investor does not fully understand the complexity of a project, it is a common and convenient practice to add a risk premium, a mark-up on top of the cost of capital. This leads to underinvestment in projects characterised by larger risk exposure. Ultimately, high risk is not only a question of the cost of capital for low-carbon investments, but a question of the availability of capital. If projects or technologies are judged too risky by investors, it is impossible to attract financing sources.

It is of the utmost importance to increase understanding of potential sources of funding and finance when designing low-carbon policies. The role of the financial sector is to ensure financial intermediation between sources of finance and investment needs, in a way that diversifies risk across projects and across the economy. Under current circumstances, the financial sector is less able to play this role. Attracting low-carbon investment requires a careful assessment of the different sources of funding and the associated cost of capital.

Conclusion: Policies to incentivise low-carbon generation investment must address several key challenges

Several key issues must be addressed to encourage investment in low-carbon generation. First, wholesale electricity prices are currently low and could remain low during the transition. Acknowledging this challenge is necessary when considering how to effectively drive investment in low-carbon generation. Second, although many governments have introduced carbon pricing, the credibility of carbon prices high enough to drive the transition to low-carbon electricity is low. Third, fossil fuel prices are variable and hard to predict, and low-carbon generation investors may be exposed to fossil fuel price risk. Fourth, the landscape for investment in low-carbon generation is in flux, and policy makers must consider how policies can help attract different sources of investment.

2.3. The regulatory transition: Low-carbon investment support instruments

Attracting low-carbon investment while keeping financing costs low in a context of uncertainty (i.e. the infrastructure financing puzzle described previously) will require continued policy intervention during the energy transition, as explained in the previous section. Wholesale electricity prices are currently too low and are expected to remain low if further capacity is added during the energy transition period. Long-term gas prices are expected to increase, but this prospect remains uncertain and capital markets have a limited ability to take on such risk. Many governments remain committed to introducing carbon pricing or strengthening existing schemes. But they also have to recognise that restoring the credibility of carbon pricing will take time. This introduces a regulatory risk from the perspective of investors. In summary, the factors holding back low-carbon investment are leading to a situation in which progress is falling short of what is required to meet the 2°C target. As such, additional measures are required.

In fact, governments already intervene to attract low-carbon investment. Current low-carbon support policies, mainly applied for renewable energy deployment, work by increasing low-carbon revenues or mandating certain shares of clean energy. In many cases, these support instruments take the form of long-term contracts that provide visibility to investors, or they create the conditions in which an investor will find an appropriate counterpart to sign a PPA, thus shifting risk away from low-carbon generators.

Existing experience from renewable energy support policies can provide valuable insights into how to supplement electricity market design with additional instruments during the transition. However, these instruments need to be developed further, reflecting their new role: away from bridging a large cost gap for non-mature technologies, towards providing revenue predictability and visibility during the energy transition.

Existing types of low-carbon investment support instruments

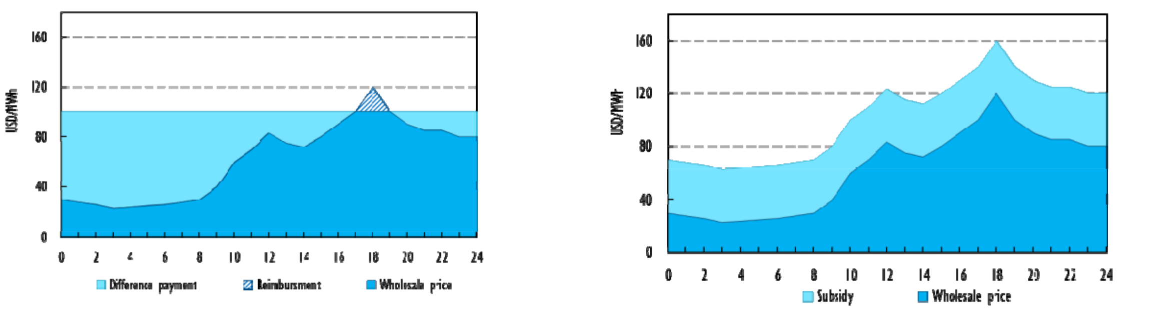
Low-carbon investment support instruments can be designed in many ways, and these various designs result in a varying division of market risk among governments, investors and low-carbon generators. The most relevant risks categories for electricity generation come from markets:

price, volume and imbalance risks. At one end of the spectrum, a policy may entirely shield low-carbon generators from all of these risks, while at the other end, generators carry all three (Figure 2.9).

Instruments that provide revenue certainty offer investors full predictability of future prices, guaranteeing sales volumes and socialising imbalance costs. Standard feed-in tariffs (FITs) fall under this category (see below for a more detailed discussion).

Subsidies on top of market revenues ("market plus subsidies") are instruments that either increase deployment by reducing the cost of low-carbon energy projects for investors or increasing revenues on top of the market price revenues. Tax incentives such as the US Production Tax Credit (PTC) are market plus subsidy instruments. These instruments increase revenues, but can expose investors to the full set of market risks.[4]

Figure 2.9 •FIT (left) and market price plus subsidy (right) (illustrative)



Between these two poles lie a variety of intermediate approaches that divide risk between actors. As policies, particularly decarbonisation policies, evolve, the allocation of risk under these instruments will also vary, resulting in the need to revisit support policies and reformulate them to allocate risk as appropriate.

Existing policy examples can provide a basis for developing the next generation of low-carbon support instruments and are therefore discussed in further detail next.

Low-carbon generators face a variety of risks that may affect revenues, including energy price, production levels (volumes) and imbalance risk. The following list introduces a variety of low-carbon financial support instruments, beginning with policies that shield generators from risk, continuing through shared-risk instruments, and concluding with systems that shift all risk to generators.

- Fixed-price instruments shielding generators from market risk:

  FITs function in a similar way to a standardised, long-term PPA, usually signed with a utility or a network company and backed by government, although the stability and consistency of the FIT depends on the durability of its supporting legislation.[5] The cost of support is usually born by consumers in the form of surcharges on electricity prices. When combined with priority dispatch and curtailment compensation, it removes virtually all market risk from investors. Construction risk and technology risk still rest with investors.

  FITs have been an effective measure to rapidly install new renewable capacity. From 2009 to 2013, OECD countries installed 75 GW of wind capacity and 91 GW of solar capacity,

[4] Strictly speaking, a fixed subsidy on top of market revenues mitigates volume risk, since generators will have an incentive to bid below true short-run cost and thus have a reduced volume risk. However, provisions can be put in place that preclude payment of premiums during negative market price events.

[5] Where the long-term stability of legislation is an issue, FITs are frequently implemented as contracts between government-backed companies that are subject to the law of contracts, as this gives greater power to investors to enforce the contract in the long term.

representing 36.4% of the generation capacity of OECD countries. FITs frequently unlock deployment at lower costs than instruments based on trading green certificates. However, as illustrated by the recent experience of solar PV, safeguards need to be put in place to avoid uncontrolled rapid deployment. The advantage of FIT systems is that even small project developers are able to finance projects with a high level of debt, driving down financing costs. For example, if the cost of capital decreases from 8% to 6% in real terms, the LCOE of wind declines from 100 USD/MWh to 90 USD/MWh. Shifting wholesale power market risk away from investors greatly helps to keep the budget cost of subsidies as low as possible.

- Shared-risk instruments:

  Variable premium systems, including the United Kingdom's Contracts for Difference (CfDs) and the variable renewable premium in Germany (Box 2.3), are similar to FITs in that they provide a standardised, long-term PPA for renewable energy – and most recently also nuclear energy – projects. One difference as against FITs is that under a variable premium, generated electricity is sold directly by generators on the market, thus subject to imbalance risk. If market revenues fall short of a predetermined price (strike price), investors receive additional compensation such that market revenue and support payments equal the strike price. Conversely, if market revenues exceed the strike price, investors reimburse what is surplus to the strike price.

Box 2.3 • Variable renewable premium in Germany

Establishing a market premium system through which generators can sell their electricity directly onto the market, and thus shoulder the balancing risk, can be a first move towards increased market integration. The premium level is calculated such that the average generator of the particular technology (e.g. wind) would receive a payment (market revenue plus market premium) that matches the FIT (Figure 2.9, left) plus a management payment for covering the cost of organising sales to the market.

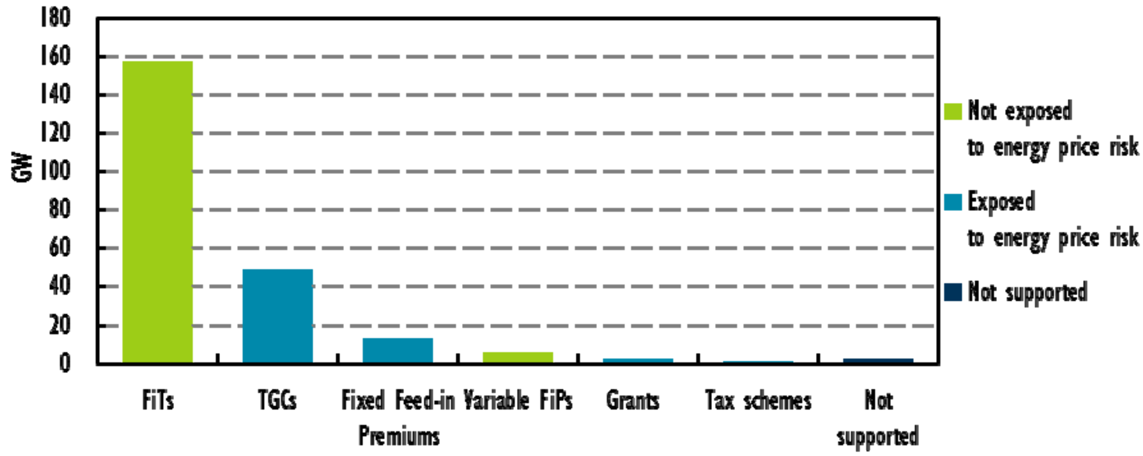
An important property of the German premium system should be noted. For a given technology, if a generator is able to produce an output of higher value (market price) than the average, it is possible to make an additional profit, because the per-MWh premium level is calculated for the average generator. This induces competition within each generator category to secure sites and build power plants that will generate when prices are particularly high, i.e. generate electricity when it is most valuable.

- Subsidies on top of market revenues:
  - Direct cash grants, rebates and tax incentives or credits can be used to reduce investment costs and so improve returns for investors. Under cash grant schemes, renewable energy project developers recoup a percentage of the investment cost in cash. This can similarly be done through a reduction in tax liability. Tax incentives or credits are often used to reduce the cost of renewable energy projects from an investor perspective. Mechanisms include reduced tax rates or waiving certain taxes for equipment or revenues from energy sales. Tax incentives may also take the form of accelerated depreciation of renewable energy assets. The timing of the tax liability reduction affects the allocation of risk.
  - Fixed market premiums are intended to complement revenues generated on the standard electricity market by paying investors a fixed premium according to the amount of electricity they generate, which supplements market revenues. This can be done through direct remuneration or through a reduction in tax liability, as in the United States' PTC. When implemented through the tax code, a number of financial arrangements may exist to allow an entity other than the generator to take advantage of the reduction in tax liability (Box 2.5).

- Low-carbon generation quotas with TGCs work by setting a specific amount of electricity that needs to be covered by generation from low-carbon sources. This obligation is usually imposed on electricity suppliers. In order to allow for meeting this obligation more efficiently, a market is established for certificates that are issued for each unit of green electricity that is generated towards meeting the quota, and thus the owner of the low-carbon energy benefits from an additional revenue stream. The certificate market is an additional market based on the idea of separating the actual power and its "greenness". The electricity component is remunerated in the same way as non-renewable electricity, for example via the wholesale power market. TGC schemes usually include a fine that the entities under the obligation have to pay if they fail to buy enough certificates. In most cases, this penalty rate determines an upper bound for the value of certificates (Box 2.5).

Not all low carbon investments are supported. In the absence of a financial support policy, but when low-carbon energy is cost-competitive, low-carbon generators may be entirely exposed to market price risk through merchant power sales, i.e. selling directly onto the spot market without a long-term contract in place. In Europe, only 2.4 GW of capacity has been installed without support over the period 2005-14 (Figure 2.10).

Figure 2.10 •Renewable capacity built by support instruments, OECD Europe 2005-14



Notes: GW = gigawatt; TGCs = tradable green certificates.

Existing low-carbon support schemes are either largely de-risking investments or fully exposing them to market risks (market plus subsidy). Shared-risk instruments in Europe are just starting to expose low-carbon generators to imbalance risks, but not to the long-term evolution of the generation mix, or carbon and gas price uncertainty. The next section discusses intermediate support instruments consisting of the partial pass-through of market risks.

Towards the partial pass-through of market risk: Modulated premiums

An issue with fixed-price instruments is that low-carbon generators are not incentivised to maximise the value of their project to the system. This mutes any incentive to factor in the value of the generated electricity when making investment decisions. Conversely, market plus subsidy schemes fully expose investors to market price risk, including carbon price risk. If governments manage to strengthen the carbon price, this can create windfall profits for low-carbon generators, and in any case, such support schemes expose investors to carbon price risk, a regulatory risk that unduly increases financing costs.

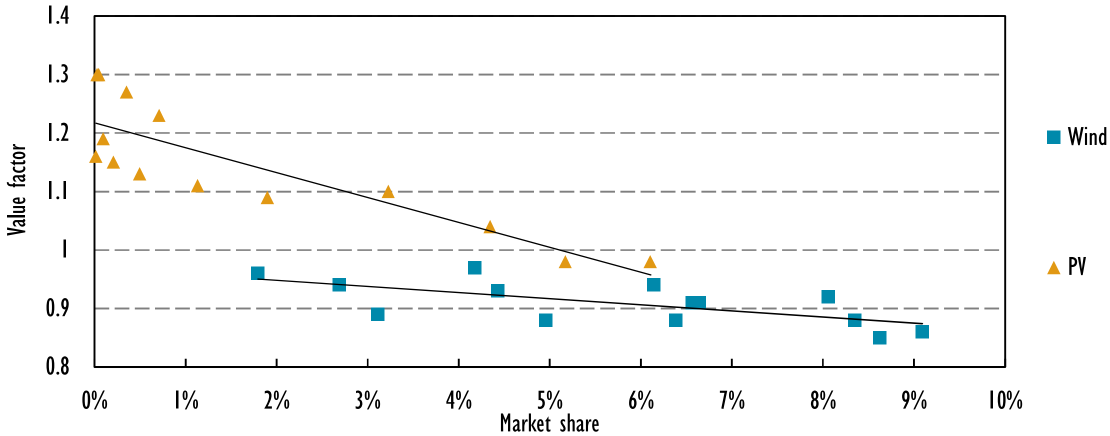
Keeping financing costs low means in particular that low-carbon investments should not be exposed to unnecessary policy risk. In many jurisdictions, moving away from FITs directly to fully market-based arrangements could be a significant step, and might prove unsuccessful in delivering continuing low-carbon investment.

At the same time, market prices provide valuable feedback on the value of different assets to the power system, and are thus vital for steering effort in the right direction. For example, the market value of VRE can experience a significant decline at growing penetrations (Figure 2.11) in the absence of actions to increase the flexibility of the power system (IEA, 2014a).

Between instruments that entirely shield low-carbon generators from all market risks and instruments that fully exposed them to long-term price uncertainties, it is possible to identify instruments that expose low-carbon generators to some, but not all, of energy market price risk. Advancing these instruments entails finding a balance between providing certainty for capital-intensive investment while maintaining market feedback.

One option is to combine the properties of instruments. Such a combination can be termed modulated premium. This type of instrument integrates low-carbon projects into electricity markets while increasing their deployment by mitigating the market price risk from an investor's perspective. Modulated premium is an intermediate category of instrument between fixed-price instruments and market plus subsidy instruments (Figure 2.10). Such modulated premium systems should be thought of as a family of different instruments that allow fine-tuning of the degree of risk that is passed onto investors. They are thus a promising option for incentivising investments while not muting market signals altogether.

Figure 2.11 • Market value factor of wind and solar PV as a function of their market share in Germany



Source: Hirth, 2015.

In Europe, small steps are already being taken in the direction of better market integration, in the form of gradual changes to FITs or the introduction of market premium systems (see also Box 2.4). For example, the non-payment of the FIT in case of over-generation leading to negative prices is a step in that direction; low-carbon generators are partially taking a market risk on the volume of electricity sold (in addition to the volume risk resulting from weather conditions). Notably, the move in certain European countries towards sophisticated market premium systems provides incentives to locate technologies in the best places, to perform better than other low-carbon technologies and to schedule maintenance when the market value of electricity is low.

Yet, long-term prices on the electricity market are the greatest source of risk for investors. The degree to which low prices are passed on to investors thus directly influences the risk levels they need to bear. Regulators therefore have to strike a balance between imposing a higher level of market risk that generators can realistically be exposed to and allowing sufficient certainty to keep financing costs low.

Box 2.4 • Multiple policy layers in the United States

The United States has employed various overlapping policies to drive renewables installation, which in concert have created a large and growing market. An examination of US policies reveals how employing multiple drivers and incentives can lead to substantial growth in renewable energy.

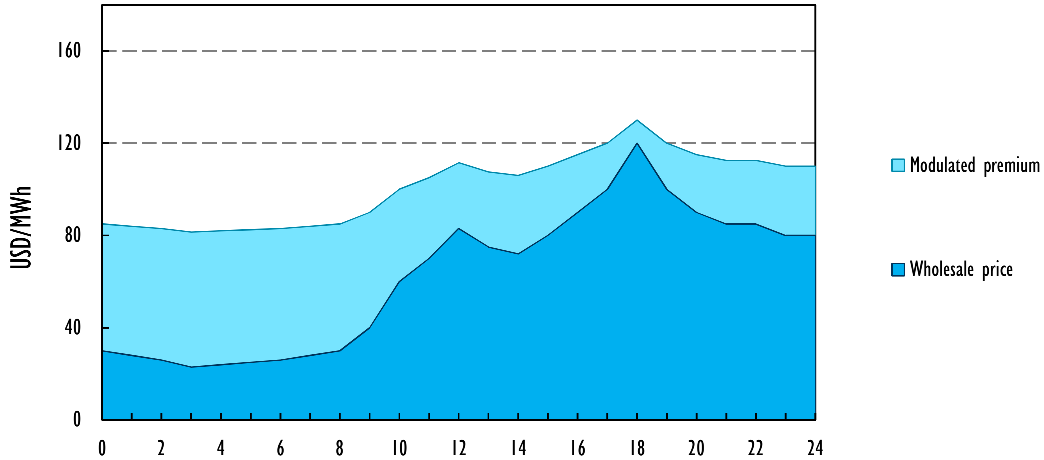
The primary policy drivers of renewable energy installation in the United States are 30 state-level Renewable Portfolio Standards (RPSs), which require a certain amount of load to be provided by renewable energy. Roughly half of the states with an RPS are traditionally regulated, while half are restructured (i.e. partially deregulated). RPS compliance mechanisms vary substantially between the two market types. In traditionally regulated states, utilities tend to own projects directly or sign long-term PPAs with renewable energy providers. In restructured markets, where load-serving entities (LSEs), which may or may not be utilities, have less ability to forecast their future demand load, LSEs tend to comply with RPS requirements through the acquisition of renewable energy certificates (RECs). RPS policies are often structured with an alternative compliance payment (ACP), which acts as a *de facto* cap on REC prices. The ACP may commonly be in the USD 50 to USD 100 range, with separate ACPs for higher-cost technologies possible in states with, for example, a “solar carve-out” requirement. In the case of restructured markets, the electricity from the project and the REC may be sold separately, potentially exposing the renewables project to energy price risk. However, projects in restructured states also generally sign a PPA with an offtaker, such as a utility or power merchant, which is not necessarily the purchaser of the RECs.

PPAs with creditworthy offtakers are often critical to financing renewables projects, but in some cases equity investors are willing to finance projects that do not sign long-term PPAs and instead take on the energy price risk by selling through short-term contracts or on the wholesale spot market. In these cases, the project may be able to profit from energy price variability. This primarily occurs in the wind sector, and in 2014 almost all such capacity was in Texas, where wind energy prices can be competitive with wholesale market prices (even before the PTC is applied) and where various other market factors make risk-hedging more feasible. These so-called “merchant” projects generally include a 10- to 12-year price risk hedge to reduce risk to the project owner.

The vast majority of renewable energy installed in the United States has benefited from tax schemes to support the industry, including the PTC, which provides a tax credit based on the kilowatt hours produced, the Investment Tax Credit (ITC), which provides a tax credit based solely on upfront investment, and accelerated depreciation schedules for renewables investment. These policies have served to lower the costs of renewable energy, ensure compliance through new renewable generation rather than ACPs, and de-risk investment. Some projects, particularly in the solar sector, have also benefited from federal loan guarantees from the Department of Energy, increasing access to capital. These various incentives serve to reduce the energy price risk associated with renewable energy investment for whoever bears it. Combined with the state RPSs, these policies have helped to spur huge investment into the US renewable energy sector. In recent years, however, many investments in renewable energy have been voluntary and not RPS-driven, motivated by positive economics, sustainability missions of utilities, or voluntary corporate procurement.

Modulated premium mechanisms that expose generators to some – but not all – price signals coming from wholesale power markets (Figure 2.12) are a way to strike this balance. A modulated premium can depend on realised market price. As illustrated in Figure 2.11, a high market price translates into a lower premium (the modulated premium is 60 USD/MWh if the market price is 20 USD/MWh, the premium is 40 USD/MWh if the market price is 60 USD/MWh and the premium is 20 USD/MWh if the price is 100 USD/MWh). Such a support instrument aims to supplement market revenues to partially compensate for the gap between required revenues and actual market revenues.

Figure 2.12 • Modulating premiums and level of support as a function of wholesale market price (illustration)



Rather than modulation as a function of electricity market price, another option is to modulate the premium depending on realised carbon price. A high carbon price would translate into a lower premium (say the modulated premium is 60 USD/MWh if the CO_2 price is zero, the premium is 60 USD/MWh if the CO_2 price is 25 USD/tCO_2 and the premium is zero if the CO_2 price is above 50 USD/tCO_2). Modulation according to a carbon price index would, in practice, largely shift the carbon price risk away from investors onto customers or taxpayers, better aligning the regulatory risk with those most able to mitigate it. This would reduce the risk premium demanded by investors to compensate for the CO_2 price risk, and would contribute to keeping the cost of financing low.

Modulated premiums can partially pass through market price risk to low-carbon investors. As illustrated in Box 2.4, the risk profile of a project under a modulated support scheme can be intermediate, in between a very risky project and one benefiting from a FIT and thus not exposed to any market risk. Modulated support schemes would improve the following incentives:

- For project developers, providing the incentive to choose system-friendly equipment (such as low-speed wind turbines) and locate the plants where good wind resources are found.
- For investors, designing modulated payments in such a way that they reduce the risk of extreme losses. For example, introducing a floor or stop loss in the sharing agreement can prevent extreme tail risks.
- For governments, using risk-sharing agreements to establish more predictable renewable deployment policies and facilitate an increase in CO_2 prices. Such policies can reduce the cost of modulated support schemes for consumers (the higher the carbon price, the lower the support).

Modulated premium instruments have their own implementation challenges, such as determining the relationship between market price or carbon price and premium level. A careful analysis of the market context in each specific case will be needed to determine the appropriate set of instruments to optimise low-carbon deployment. Implementation challenges include the determination of premium levels and deployment quantities, technology-neutral vs. technology-specific instruments, and the determination of the modulation formula that allocates risks. However, the general principle of exposing investors to some but not all market risk should prove applicable across a wide range of circumstances.

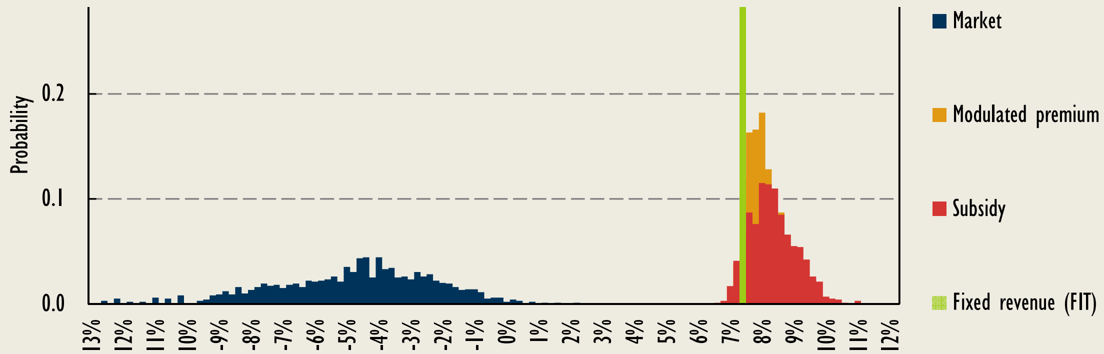
Box 2.5 • Modelling investors' risks under different support instruments using Monte Carlo simulation

Quantifying risks associated with investments can be done using a number of techniques, including assessment of risk premiums by the financial community to calculate the risk-adjusted cost of capital. In economics, the notion of risk aversion is also commonly used but more difficult to quantify. Another widely used technique is to perform stochastic Monte Carlo simulations.

The analysis presented in this box is a stochastic analysis that uses Monte Carlo simulations to calculate the internal rate of return (IRR) of an onshore wind project under a pure market framework and under different support mechanisms.

In addition to the hourly dispatch model already presented in the previous boxes, two other model components have been added to calculate the business plan of a power plant starting operation in 2020, and a module to perform Monte Carlo analysis using the software Crystal Ball. Four different stochastic variables are considered in the simulation. The gas price can vary between 9 USD/MBtu and 14 USD/MBtu, with a uniform distribution; the CO_2 price follows a log normal distribution with a median of 25 USD/tCO_2 and a standard deviation of 13 (leading to a range of approximately 5-100 USD/tCO_2). In addition to fuel price uncertainty, the pace of deployment of renewables is also uncertain. For instance, in a system of 56 GW of peak demand, the installed capacity varies in the range of 27-35 GW for wind and 16-18 GW for solar PV by 2030. These different variables are not correlated. The investment cost of onshore wind is 1 479 USD/kW, the lifetime is 25 years and the load factor of the wind turbine is assumed to be 23%. These figures are taken from the study, *Projected Costs of Generating Electricity 2015* (IEA/NEA, 2015).

Figure 2.13 • Probability distribution of the IRR for onshore wind with and without support



Under this set of assumptions, the IRR of a wind power plant earning only the wholesale price is almost always highly negative (Figure 2.13). When the IRR is slightly positive, corresponding to high gas price and high CO_2 price scenarios, it is far below the cost of capital that most investors use to finance their projects. In this example, no market-based investment takes place.

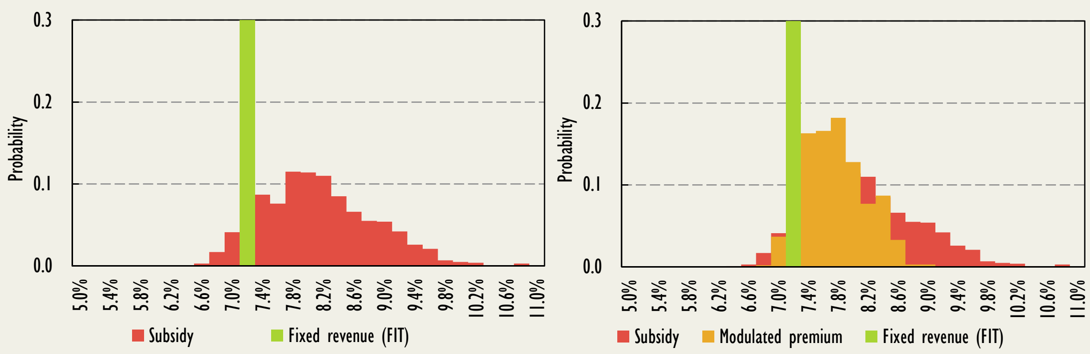
In order to attract investment, the revenues have to be increased thanks to a support mechanism, to ensure the Internal rate of return (IRR) is positive and at a level superior to the cost of capital most of the time.

Fixed revenue instruments, such as FITs, can be set at a level that sets the IRR precisely at the cost of capital (construction and operational risks are not considered here) in our simulation, a FIT set at 104 USD/MWh would lead to an IRR of 7%. This solution, however, insulates investors from any market risk, as illustrated in Figure 2.14 by a vertical bar (the graph is cut at 0.3, but the probability is 1).

Box 2.5 • Modelling investors' risks under different support instruments using Monte Carlo simulation (continued)

A subsidy such as a fixed premium a (an investment premium would have the same effect) can also increase the IRR but fully expose the investment to the market price risk more. The subsidy has to be sufficiently high to ensure a return most of the time. In this calculation, a fixed premium of 58 USD/MWh ensures that the IRR is above 7% with a probability higher than 90%. In some cases the return is lower, but the probability distribution of IRR in red shows that the expected return is around 8.5%, higher than 7%, which ensures the remuneration of risk.

Figure 2.14 • Probability distribution of the IRR for onshore wind under different support schemes



However, subsidies can lead to remuneration of around 10% in high CO_2 and gas price scenarios and low VRE deployment. These levels of return carry the risk of being considered too high for activities that are subsidised. This scenario could occur if governments succeed in strengthening carbon pricing.

A modulated premium depends on wholesale electricity prices. In this calculation, the premium is set at 55 USD/MWh if the wholesale price is 65 USD/MWh, so the total revenue is 120 USD/MWh. The market risk is shared on a 50/50 basis with investors. So if the wholesale price falls to 25 USD/MWh, for example, the premium is increased by (65-25)/2=20 USD/MWh to 35 USD/MWh and total revenues are only 25+35=60 USD/MWh. Therefore, investors are exposed to market price risk. As the premium is reduced when wholesale market prices are high, the probability distribution of the IRR is narrower. Consequently, modulated premiums contribute to keep financing costs low, even if it is not a risk-free rate, while at the same time integrating low-carbon power into electricity markets.

A description of the model and assumptions used for the Monte Carlo simulation can be found on the IEA website (www.iea.org/media/topics/electricity/repoweringmarkets/annexes.pdf).

Designing support instruments

When implementing a low-carbon investment support mechanism, several design factors will influence short-term market prices and the build-out of various technologies.

As discussed in Section 2.2, the introduction of low-carbon generators affects the merit order and the impact of low-carbon policies on existing generators. In addition, the design and level of premiums also changes the bidding behaviour of low-carbon generators, further influencing short-term market prices. Several markets have experienced negative prices caused by generators bidding below their actual marginal cost in order to secure the premium payment, exacerbated by the inflexibility of conventional generators. Empirical evidence from the German market indicates that negative price events remain rare at a combined share of roughly 15% wind and solar PV, with 64 hours of negative prices per year in 2013 and 2014, i.e. 0.7% of total hours.

Another important issue for governments is controlling the pace of deployment of low-carbon technologies. In a number of countries, solar PV deployment has reached much higher than expected levels at a time when costs remained very high. This has translated into a legacy of high costs for consumers that will endure for 20 years, leading to abrupt policy changes and policy discontinuation in a number of countries. In Germany, over 7 GW per year were installed between 2010 and 2012 when the government's target was roughly half that. Also, despite frequent revision of FITs, governments have sometimes failed to keep pace with the decline in the cost of solar PV, which has led to overshooting of targets and the installation of more solar PV capacity than anticipated and a higher than needed remuneration of capital invested. This has been the case in particular in Spain, Italy and Japan.

To reduce the asymmetry of information about the costs of wind and solar power, US utilities and European governments are increasingly introducing competitive auctioning processes for the construction of new installations. In France, auctions have been used for biomass power plants since 2004, and the first offshore wind projects were auctioned at a price of EUR 150/MWh in 2012. In April 2015, Germany ran its first auction for ground-based PV plants at capacities between 100 kW and 10 megawatts. These auctions resulted in remuneration levels between 84.8 EUR/MWh and 94.3 EUR/MWh (96-106 USD/MWh). Auctions enable competitive pressure to select the least cost projects at the investment stage, and thus enable better control of the quantities and pace of deployment of supported renewables and their associated costs.

The above discussion, however, assumes a technology-specific approach, which promotes technologies with no or low carbon emissions while specifying which technologies should be supported. Policy makers may find setting technology-specific policies desirable for reasons such as driving cost reductions through learning, the potential to reduce final consumer prices, or the value of certain technologies to the overall system.

In this context, determining the optimal mix of technologies is challenging and requires an assessment of the value of each technology to the overall system. From the perspective of economic efficiency, technology-neutral policies encourage the selection of lowest-cost technologies and thus lowest-cost compliance, and avoid the political challenge of choosing technologies.[6]

Implementing modulated premiums in practice requires determination of the premium that different projects seek for their investment, by unit of output generated. Consequently, if a project has a low market value, its wholesale revenues are lower and the premium asked will be higher. This paves the way for the introduction of competition between different low-carbon technologies that are not based solely on their costs. Competition can also reveal the value of different technologies and projects to the electricity system. For example, a technology can be less expensive than wind on a LCOE basis. But if this technology has a low market revenues it will have to bid a higher premium than other and this is an indication that this project should not be selected.

The example of solar PV in a summer peaking system illustrates this effect. Assuming that solar PV receives an average market remuneration of 80 USD/MWh while it comes at a cost of 100 USD/MWh, it would require a premium of 20 USD/MWh. If land-based wind generation can make 60 USD/MWh on the market while costing 85 USD/MWh, it would require a premium of 25 USD/MWh. In this example, a FIT would be lower for wind, but wind has less value to the system than solar. Competition between technologies for the level of premium needed would reveal that solar PV needs a lower level of support. While wind has a lower cost than solar PV, it requires a higher premium to cover the gap between costs and market revenues. These considerations are particularly relevant because the market value of wind and solar PV generation

[6] However, a number of arguments can be made as to why a purely technology-neutral approach may not bring a sufficiently broad set of technologies to market maturity on time to achieve decarbonisation (for example, see: Heptonstall P, Gross R, Greenacre P, Cockerill T. et al., 2012).

is very system-specific and drops with increasing penetration (Figure 2.12, Hirth, 2013; Mills and Wiser, 2012; NEA, 2012; IEA, 2014a).

An example of an auctioning system that seeks to factor in the system value (albeit not the market value) of different technologies is currently being implemented in Mexico. Under this system, generators bid a certain base price, which is then adjusted according to the expected system value of the generation for the next 15 years. Technologies that produce in locations and during hours where electricity is most valuable receive a premium on top of the base price they bid. This allows the comparison of different technologies in the same auction. However, the problem of calculating the system value of technologies is then shifted to determining the exact value of the premium that is granted on top of the base price for different locations and times of generation.

Allocation of risk

Risk is not a fixed quantity. How risk is allocated can reduce the overall level of risk by incentivising the stakeholder that bears it to take risk-mitigating actions. In general, it is most efficient to allocate the risks associated with a project to the stakeholder that can take actions to mitigate each risk (Table 2.1). For example, project construction risk should be borne by the project developer, which will incentivise it to select the right equipment manufacturers and manage the project and its construction effectively.

The CO_2 price is particularly relevant in this context. Governments rather than investors are better placed to address the regulatory risk stemming from the carbon price. From this perspective, the support system that modulates the level of support according to the CO_2 price should allocate this risk to governments and/or consumers. In this case, governments would be expected to take into account these impacts when deciding to change carbon-pricing rules. Investors exposed to such a regulatory risk would demand a high risk premium that would unnecessarily increase the cost of capital. By contrast, modulating the level of premium according to the carbon price would contribute to allocating this risk to governments, its natural owner.

In practice, however, risk analysis is a complex task and different risks cannot easily be isolated, separated and allocated individually. For instance, CO_2 price risks, renewable deployment, fossil fuel and demand risks all have an impact on the outcome of wholesale market price risk, and it is usually not possible to assess these risks individually in a practical and non-controversial manner. Practically, for example, it might be difficult to isolate the effect of the price of CO_2 on electricity market prices, leaving in its place a simple risk-sharing mechanism.

Conclusion: Support policies beyond a carbon price are needed, with many possible options available

Although carbon pricing is, in theory, an efficient driver for low-carbon energy deployment, in practice additional long-term support arrangements are required to drive the substantial amount of new low-carbon generating capacity required to meet the 2°C target. Many types of support policies have been developed and implemented, including fixed-price instruments, shared-risk instruments, and subsidies on top of market revenues. Different support schemes expose low-carbon energy investors to varying levels of risk.

In an environment characterised by a high degree of uncertainty about future electricity prices (e.g. depressed prices due to transitional overcapacity, uncertainty around CO_2 pricing, lack of clarity about how to achieve system flexibility), transferring all associated risks to low-carbon generators may inhibit investment or lead to very high support levels to unlock investment by compensating for higher costs of capital. Unless and until the outlook for purely market-based revenues becomes more certain, supplementary mechanisms such as modulated premium systems could be an appropriate intermediate step to compensate for transitory risk factors and successfully move low-carbon investments into the mainstream.

Table 2.1• Major risks of a power plant project, allocation and possible mitigation actions

Risk	Description	Magnitude	Allocation	Allocation instruments	Owner	Risk owner's mitigation action
Operating risk	If the plant is not performing as planned or is not available when needed, then revenues and profitability are lower	Medium	Investor Gov.	- Electricity price	Investor	- Selection of plant site - Technology choice - Maintenance plans
Construction risk	If there are cost overruns or delays, then the power plant profitability will be low or negative	Low (wind and solar) High (nuclear offshore)	Investor Gov.	- FiT or PPA - Support contingent on CO_2 price	Investor	- Better project management - Technology choice - Engineering, Procurement and Construction (EPC) contractor
Electricity demand	If the electricity demand is low, then market prices are depressed and the power plant will not recoup its investment costs, including a fair return	Medium	Investor Gov.	- Integration of low-carbon investments into electricity markets	Investor	- Improve resources dedicated to the forecast of electricity demand growth (net and growth)
Gas, coal prices	If the gas and coal prices are low, then the power plant will not recoup its investment costs, including a fair return	High	Investor Gov.	- Risk-sharing agreement - Support contingent on gas price	Investor	- Forecast of fuel prices -Vertical upstream integration
Wind and solar deployment	If wind and solar deployment are higher than expected, then the power plant will have a lower load factor and lower revenues (volume and price risks)	Medium	Investor Gov.	- Risk sharing agreement - Support contingent on renewables policy objectives	Investor (technology) Gov. (policies)	- Industry R&D investment and competitive intelligence - Use of renewables quotas
Political acceptability	If the licensing delays or stops the project, then the cost of the power plant will increase	Depends on technology	Investor Gov.	- Standby support in case of project delay due to licensing delay or onsite protest	Gov.	- Governments to clarify policy objectives ex ante and ensure that the project can be built
Carbon price	If the carbon price is low, then the power plant will not recoup its investment costs, including a fair return	High	Investor Gov.	- FiT or PPA - Support contingent on CO_2 price	Gov.	- Governments can increase carbon price level and certainty

References

Greentech Media (2014), "SolarCity starts selling 'solar bonds' online to public investors", www.greentechmedia.com/articles/read/SolarCity-Innovates-Again-With-a-Public-Offering-of-Solar-Bonds.

Green, R. and T-O. Léautier (2015), "Do costs fall faster than revenues? Dynamics of renewables entry into electricity markets," Working paper, Toulouse School of Economics, July 2015.

Hirth, L. (2015), "The market value of variable renewables: The effect of solar and wind power variability on their relative price", personnal communication.

Heptonstall, P. et al. (2012), "The cost of offshore wind: Understanding the past and projecting the future", *Energy Policy*, Vol. 41, pp. 815-821.

IEA, (2015), *Medium-Term Renewable Energy Market Report*, OECD/IEA, Paris.

IEA (2014a), "The power of transformation: Wind, sun and the economics of flexible power systems", http://dx.doi.org/10.1787/9789264208032-en, OECD/IEA, Paris.

IEA (2014b), "Attracting finance for low-carbon generation", in *Energy Technology Perspectives 2014*, OECD/IEA, Paris.

IEA (2012), "Developing a natural gas trading hub in Asia", www.iea.org/media/freepublications/AsianGasHub_WEB.pdf, OECD/IEA, Paris.

IEA/NEA (Nuclear Energy Agency) (2015), *Projected Costs of Generating Electricity, 2015 Edition*, OECD/IEA, Paris.

Magritte Group (2015), "Magritte group call for more convergence and integration in European Energy Policy", www.engie.com/wp-content/uploads/2015/04/magritte-group-calls-for-more-convergence-and-integration-in-european-energy-policy.pdf.

Mills A. and R. Wiser (2012), "Changes in the economic value of variable generation at high penetration levels: A pilot case study of California", Ernest Orlando Lawrence Berkeley National Laboratory, https://emp.lbl.gov/sites/all/files/lbnl-5445e.pdf.

NEA (Nuclear Energy Agency) (2012), *Nuclear Energy and Renewables: System Effects in Low-carbon Electricity Systems*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264188617-en.

Newbery, D.M. (2012), "Reforming competitive electricity markets to meet environmental targets", *Economics of Energy and Environmental Policy*, Vol. 1, No. 1, pp. 69-82.

OECD (Organisation for Economic Co-operation and Development) (2015a), *Aligning Policies for the Transition to a Low-Carbon Economy*, OECD Publishing, Paris.

OECD (2015b), "Mapping channels to mobilise institutional investment in sustainable energy", OECD report to G20 finance ministers and central bank governors, Paris, https://g20.org/wp-content/uploads/2015/04/Mapping-Channels-to-Mobilize-Institutional-Investment-in-Sustainable-Energy.pdf.

The World Bank (2015), "carbon pricing is expanding: Initiatives now valued at nearly $50 billion" www.worldbank.org/en/news/feature/2015/05/26/carbon-pricing-initiatives-nearly-50-billion.

Chapter 3 • Short-term markets

HIGHLIGHTS

- Wind and solar energy sources create new operational requirements for electricity system operators. They do not contribute to meeting demand when there is no wind or sun, but can potentially lead to over-generation when they are abundant. Their variations and forecast errors need to be managed.
- Existing market designs fall into two categories:
  - High-resolution markets, such as locational marginal pricing (LMP) in the United States, provide detailed temporal and geographical price information about transmission network constraints and marginal generation costs.
  - Low-resolution markets, such as those developed in Europe, which are decentralised and have been successful in promoting cross-border trade in electricity.
- As decarbonised electricity systems become more volatile, system operators need to take action to ensure prices correspond to actual marginal generation costs. Market design needs to provide a high resolution of the physical reality of the network.
- Short-term markets must be upgraded. Intraday and real-time markets could be improved according to the following principles: have high geographical resolution to more accurately price congestion; use uniform prices for all real-time energy used for balancing, in order to reflect the marginal costs; and be updated continuously during the last few hours before operations, reflecting improved forecasts.
- Upgrading prices in short-term markets would reveal to distributed resources how to contribute to system needs and also assist co-ordination between complex and large power systems. These upgrades could build upon existing markets, extending them into the intraday timeframe and making information more transparent.

This chapter looks at the design of electricity markets necessary to ensure the secure and efficient operation of decarbonised electricity systems. Reliable electricity supply has never been as important for the functioning of modern technology-driven economies. With new technologies and decarbonisation policies, however, electricity markets are entering unchartered territories.

While some markets have already implemented sophisticated arrangements to reflect the physical reality of system operations at the transmission level, other market designs are based on an oversimplified representation of the physical reality, reflecting a low degree of congestion at the time of market restructuring.

It is increasingly clear, however, that all existing markets will, to some degree, need to adapt to accommodate rising shares of distributed and weather-dependent generation, for the following reasons.

First, low-carbon electricity systems tend to be decentralised. Onshore wind and solar photovoltaics (PV) are connected at a medium- or even low-voltage level, possibly behind the meter in some cases. System operators should be able to control these resources, either directly or indirectly, especially during tight conditions.

Second, wind and solar power generation is weather-dependent. This limits the degree to which they can respond to system needs. Their output varies widely depending on irradiation levels (solar PV) and weather patterns, and may be subject to relatively rapid swings. This leads to errors in forecasting the state of the power system and electricity flows. Variability and reduced predictability make it more challenging for system operators to ensure system security.

Third, low-carbon electricity is deployed in large regional power systems that are already interconnected and therefore interdependent in respect of system security and reliability. On the one hand, larger integrated market areas contribute to smoothing out the variability of renewables and reducing aggregated forecast errors. On the other, large areas lead to volatile power flows across balancing areas or countries, making the secure operation of networks more challenging.

While these issues have been analysed from a technical and cost perspective in a previous International Energy Agency (IEA) publication, *The Power of Transformation* (IEA, 2014), this chapter focuses on the design of short-term markets to provide the necessary operational reliability and flexibility under the new energy paradigm.

This chapter is structured as follows: the first section describes new operational requirements for power systems with large shares of variable and distributed resources. The second section describes in greater detail the design of short-term markets and discusses their ability to ensure secure system operations with high shares of variable renewable energy (VRE).

3.1. New operational requirements for large shares of variable and distributed resources: Increasing flexibility

Many technical studies have analysed in great detail the flexibility needed to accommodate variable renewables in different contexts, including NREL (2015), PJM (2014) and IEA (2014). They conclude that accommodation of high shares of wind and solar power is technically possible, which is also already confirmed by experience in several electricity systems. This section reviews the key requirements and operational challenges for the efficient design of markets.

Electricity systems were developed according to the technical characteristics of centralised generation. Reaping the benefits of economies of scale has long been the mantra of electricity providers. The deployment of distributed resources, most notably onshore wind and solar PV, is a paradigm shift for the sector.

Co-ordinated transformation of the entire capital stock, hand-in-hand with the deployment of new resources, would keep down the cost of decarbonisation (IEA, 2014). However, given long asset lifetimes, infrastructure is slow to adapt and may lag behind the rapid deployment of low-carbon power. While several gigawatts of solar PV or onshore wind can be installed within one year, this usually takes place in a grid that has not been designed to accommodate such sources. System transformation cannot keep pace with low-carbon installation and this raises technical challenges for the operation of power systems.

Most existing generation plants were built 20 to 40 years ago, and around half of the capacity has been designed to run around the clock, rather than to follow demand or wind and solar fluctuations. In member countries of the Organisation for Economic Co-operation and Development (OECD), 55% of existing capacity will still be operating in 2030. However, in the longer run, by 2050, most of the ageing infrastructure will have to be replaced. This creates a window of opportunity to effectively transform the industry.

The design of short-term markets has to address several operational challenges during the transition to low-carbon generation: controllability of distributed generation, short-term adequacy, over-generation, ramp requirements, forecast errors and network congestion. Examining these factors leads to the conclusion that the short-term market must allow for greater adjustment in the hours before operation in order to function efficiently and securely.

Box 3.1 • Technical-economic features of power systems, the law of physics and long-lived assets

Increasing deployment of renewable energy is taking place in a context where certain electricity system capabilities are evolving, while others are not.

Electricity systems will always have to manage two physical features. Since the adoption of alternating current, load and generation need to be balanced every second to maintain frequency at its target level, usually 50 or 60 Hertz (Hz). In addition, electricity flows around networks according to the laws of physics (Kirchhoff's laws), and can only be controlled within the boundaries set by physics. In the absence of modern control equipment, electricity flows cannot be controlled effectively. The design of electricity markets has to reflect this physical reality.

Two other characteristics have long shaped constrained electricity markets but are currently in flux. First, electricity demand has traditionally been highly inelastic to prices and will continue to be so in the short term, but information technology-enabled demand response has the potential to increase demand elasticity in the future. Second, electricity storage continues to be very costly but new technologies are driving down costs, in particular for batteries. Unlocking demand response and storage potential could significantly change the design of electricity markets; electricity could become similar to gas or other commodities (see Chapters 2 and 6).

Due to these four features (generation must equal load in real time, Kirchhoff's laws, inelastic demand, costly storage), operational challenges are not new to power systems. Demand forecast errors, outages of power plants or the loss of a transmission line have always been handled by system operations. Similarly, the output of run-of-river hydropower is also variable. The variability and unpredictability of very high shares of wind and solar power, however, are expected to be greater than those of hydropower and come in addition to traditional operational challenges.

The distribution grid has been designed according to load patterns, and not with the purpose of hosting distributed generation in mind. Distribution investments are very costly and there is a need to make the most of existing cables and wires.

For transmission lines, building a new line takes at best five years, usually ten years and up to twenty years for projects facing severe local opposition. Consequently, the grid will tend to fall short of the adaptation needed: congestion and voltage issues are expected to become more frequent. A rapid transition to low-carbon system with a high share of distributed and variable generation is likely to face many operating constraints.

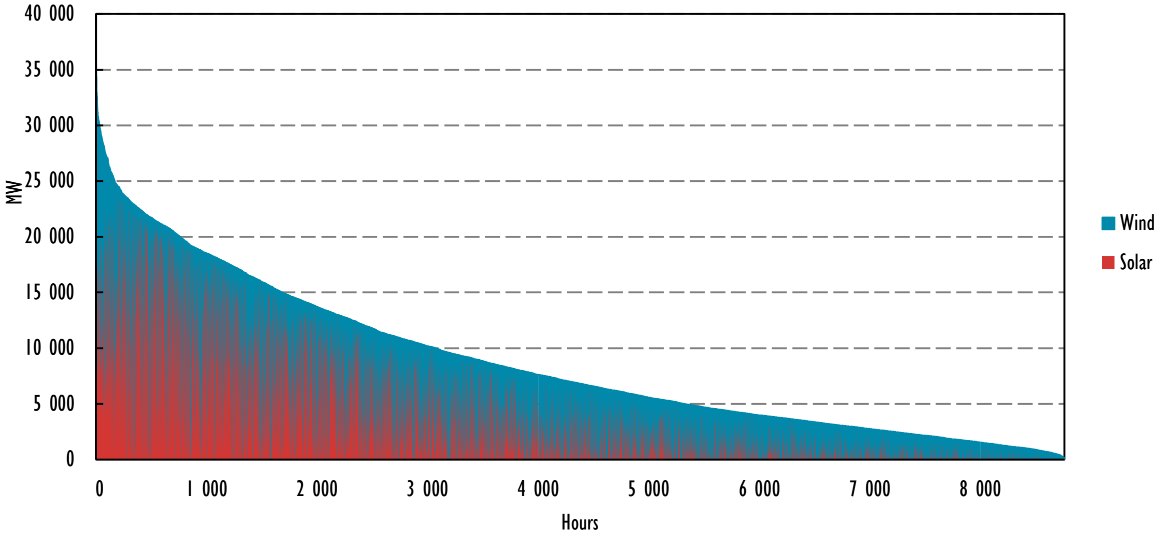
Forecasting and controllability of distributed generation

The foremost operational requirements of distributed variable generation are real-time visibility, forecasting and controllability. At low shares, distributed variable generation can be accommodated by the electricity system at all times without significant operational constraints; wind and solar generation only reduce the energy withdrawn from the transmission network. Apart from having accurate forecasts (which require decent real-time measurement) there is little requirement to change procedures. Priority dispatch can be implemented and the output of distributed resources need not be actively controlled.

As the share of wind and solar power increases, network constraints can occur due to the fact that wind and solar outputs are peaky (Figure 3.1). Locally, the concurrence of solar and wind generation creates a hot spot (voltage or thermal constraint) either on the distribution or transmission grid. Being able to control the output of distributed power plants is therefore important for system security.

As smart technologies already enable the remote control of loads of one kilowatt or less, the control of generators producing a few kilowatts is easily conceivable. It is therefore important that a sufficient proportion of distributed generation devices be equipped with two-way communication systems and be remotely controllable.

Figure 3.1 • Load duration curve of wind and solar generation in Germany, 2013



Note: MW = megawatt.

Limited curtailment may be more cost effective than upgrading grid infrastructure. Curtailment of distributed generation (or "DG shedding") has the potential to considerably increase the connection capacity and therefore accelerate the deployment of wind and solar power. According to a study from the German distribution company, EWE Netz, the dynamic curtailment of 5% of the energy generated from solar PV increases the grid connection capacity by around 225% without new grid investment (EWE Netz, 2015). While this might sound surprising for project developers, curtailment can lower the overall cost and accelerate the deployment of wind and solar PV.

Availability of dispatchable generation

Sufficient generation capacity is required to balance supply and demand irrespective of the availability of wind and sun. This capacity needs to be available and perform when it is needed, possibly at short notice and with the capability to ramp up production rapidly, possibly from a low initial output level.

Power plant cycling and start-ups are also expected to increase. For example, in systems with a high share of solar PV, mid-merit power plants may need to start twice rather than once per day; they will need to operate during morning peak demand, then stop operating when the sun is shining, resume operation for the evening peak and stop again at night. Such operation will tend to entail increased start-up costs and may reduce the technical lifetime of power plants.

The specific operational requirements depend on the characteristics of the given electricity system. In summer peaking systems, solar power can contribute to system needs during daylight hours, when peak demand is driven by air conditioning. Conversely, in winter peaking systems, peak demand typically occurs at night, when there is no sun. At growing penetration levels, however, additional wind and solar capacity will contribute less and less to adequacy.

Consequently, conventional dispatchable capacity remains needed but will see its average load factor or utilisation factor decline. Lower load factors at conventional plants during the transition raise issues both for existing capacity and new investments. There is a perception that certain power plants, in particular mid-merit plants (e.g. gas-fired generation in Europe), could be pushed out of the market during the transition to low-carbon power. While this perception is largely the result of excess capacity and declining electricity demand in many countries, the ambitious climate policies analysed here would undoubtedly lower the load factor of the conventional generation fleet even after the resorption of excess capacity. Lower load factors can lead to unrecoverable sunk costs for plants initially expected to run as baseload.

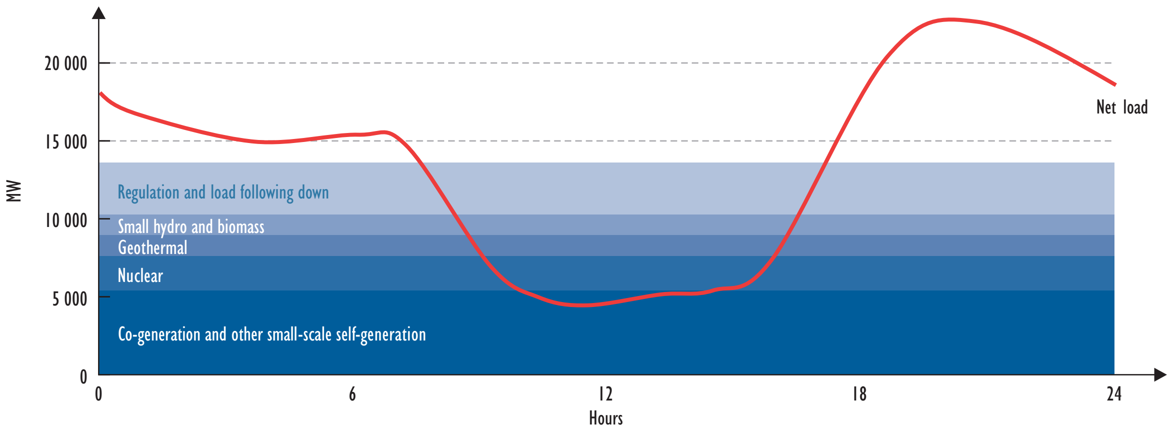
Where excess capacity occurs, older plants will be retired and dismantled while newer plants can be mothballed. Some utilities have considered the possibility of relocating new plants, but such attempts are expected to remain uneconomic and therefore extremely rare. Retirement of older, more polluting plants is a likely outcome for decarbonisation. Mothballing for a few years can reduce excess capacity, and these plants can be brought back on line rapidly in case of need.

The economics of plant mothballing are complex. Labour skills have to be maintained to bring the plant back to life. Notwithstanding, assuming typical annual operation and maintenance costs of USD 30 per kilowatt per year (IEA/NEA, 2015), existing plants can remain operational with relatively low electricity prices. As discussed in Chapter 4 on adequacy, incentivising new investment in peak or mid-merit capacity is another issue.

Over-generation

Over-generation of electricity can occur during hours that combine low consumption (for example, on public holidays or during the summer period of cold-climate countries) and high wind and solar power output. Figure 3.2 shows such a situation modelled by the California independent system operator (ISO) for one spring day in 2024. Solar PV generation during the daytime reduces the "net load" to 5 000 MW, which is lower than the minimum output from co-generation,[7] nuclear, geothermal and small renewable plants in this figure. Over-generation means that the electricity available at very low or zero cost exceeds demand.

Figure 3.2 • California ISO Long-Term Procurement Proceeding Scenario 24 March 2024



Source: CAISO, 2015.

During hours of over-generation, some of these low-cost generation sources have to be curtailed. In practice, system operators might in some cases invoke system security as the reason to curtail wind and solar power, particularly when markets do not reflect such operating constraints. In principle, market prices will become very low or possibly negative during such events, and generators may choose to shut down, i.e. curtail, output. In addition to purely technical requirements, many contractual arrangements also introduce rigidity, for example priority dispatch rules or output-based payments for co-generation plants. Curtailing wind and solar power that are the last resources added may be a practical technical solution, but not always the least-cost one.

[7] Co-generation refers to the combined production of heat and power.

Efficient system operation requires balancing supply and demand during over-generation events, while taking into account not only fuel costs, but also start-up costs, ramp rates, minimum output of plants and other technical rigidities. Optimisation calls for the system operator to consider a large number of distributed plants using many technologies, comparing the prices of each. As the number of distributed plant and storage device increases, system operators will no longer have the capability to centralise all the information. Electricity prices have a key role to play in ensuring decentralised co-ordination.

Ramp requirements

The effects of variability in demand are compounded by the variability of wind and solar power generation, which increases the volatility of the power system. Demand can vary by 10% within an hour in certain countries, for instance during the morning when people wake up, switch on the lights or electric heating and start work. In France, for example, demand increased from 52 gigawatts (GW) to 64 GW between 5:15 and 7:45 in the morning of 22 March 2012. If such a change occurs at the same time that wind generation decreases, conventional generation capacity would have to compensate for even larger variations than those caused simply by jumps in demand.

The variability of the electricity system results in ramp rates of the net load (demand minus wind and solar output), expressed in MW per hour. Figure 3.2 shows that California expects a steep ramp in the evening, when the sun sets and solar generation declines while people are still working and switch on the lights. One fundamental issue is that conventional generation in this case needs to ramp up from a very low base – the system operator faces a variation in residual demand of 400%, from 5 GW to more than 20 GW. The new operating challenge is therefore to ramp up capacity from low generation levels.

From a technical perspective, different solutions are available to address ramp requirements, including flexible generation capacity, smoothing out the variability across large geographic areas, demand response by customers or storage, and adjusting the output of renewables generators themselves. In particular, it is also possible to control and reduce the output of wind or solar PV before the sun sets or the wind stops. Such limits on the ramps of renewables have proven to be helpful in maintaining reliable operations (Smith et al., 2010).

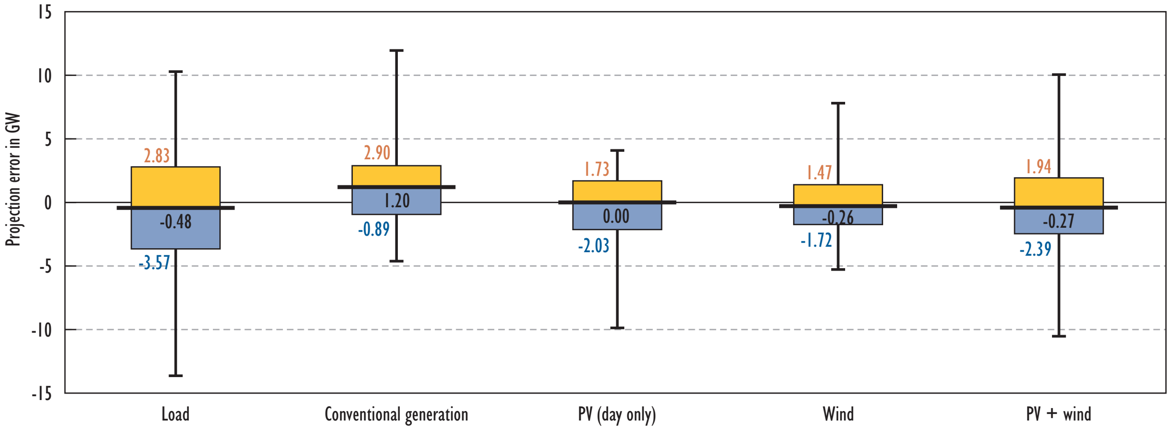
Predictability and forecast errors

Uncertainty of wind and solar generation tends to be higher than, or at least equivalent to, that of load (Figure 3.3). Progress has been made on the accuracy of forecasts. Commercial providers of forecasting solutions claim that they can attain a forecast error rate of 5%, 24 hours before real time. Centralisation of weather forecast information also reduces the uncertainty of aggregated forecasts over large geographic areas. At the distribution level, however, forecast errors remain higher.

From a reliability perspective, the potential for very rare but extreme events should be analysed. The probability of extreme forecast errors determines the size of the margins and reserves that need to be maintained in the electricity system at the different stages of unit commitments and operations. In addition, instances of automatic wind farm disconnection, where the wind blows too hard and the wind farm needs to come off line, present the risk of a significant impact on system security. In Spain, Red Eléctrica de España (REE) estimated that 500 MW of additional reserves are needed for 10 GW of wind. Similarly, the German Energy Agency, DENA (2010), forecast that 3 GW of additional reserves are needed for 36 GW of wind.

The new technical requirements placed on system operations by wind and solar power have important consequences for market design. In general, greater volatility of electricity systems merits more frequent adjustments to production schedules within a few hours before plant operation. Variations can also be smoothed out when considered over large geographic areas.

Figure 3.3 • Forecast errors of different system components 24 hours before real time, Germany 2014



Note: 90% of the errors fall within the range of the bars, the end of the lines mark the maximum over- and under-forecasts. For example, actual load is lower than forecasted load by more than 2.83 GW in only 5% of the hours. Source: based on data from EEX-transparency and ENTSO-E.

Network congestion

Rapid decarbonisation based on renewables means that the grid, at both the transmission and distribution levels, are likely to be increasingly congested. System operators will have to manage increasingly volatile and less predictable power flows, while ensuring system and operational security. Failure to use existing infrastructure efficiently would dramatically increase the cost of decarbonisation.

Transmission and interconnections

There is little doubt that network congestion will increase during the energy transition, at both the distribution and transmission levels, despite the fact that low-carbon plants are decentralised. The reasons for this may lie in a possible mismatch between the existing historic network and the location of newly decarbonised plants. In addition, strong opposition to a new network of overhead lines and the high cost of underground power lines constrain their construction.

New low-carbon generating capacity is not usually constructed at the same location as a retired higher-carbon plant. Historically, vertically integrated regulated monopolies have planned the construction of large centralised plants and transmission lines in a well-co-ordinated fashion, resulting in low levels of congestion within their service area.

Attaining such co-ordination is more difficult in restructured electricity systems. Unbundling generation and networks implies that new investments can be located anywhere, since connection charges and network prices usually do not deliver efficient locational signals to influence the location of new plants. While new gas plants can be located at existing sites or closer to consumption centres, wind and solar generation frequently need to be transmitted over long distances.

Large or regional balancing areas enable efficient integration of VRE sources and smooth the variability of demand. In addition, according to the law of large numbers, large systems reduce aggregated forecast errors. Depending on output variations, regional integration results in large swings in power flow from one control zone to the other, from one day to another, within days or even considerable changes within an hour. Even where transmission infrastructure can be built easily and at low cost, it would be too costly to prevent any congestion at all times, particularly for extreme situations that last only a few hours per year. Least-cost solutions indeed imply a degree of network congestion (see Chapter 7).

Moreover, while it takes from one year for solar PV and from two years for onshore wind to be built (or longer in case of permitting hurdles), the lead time for construction of new transmission lines can often exceed 10 years. Even if the planning process is smooth, examples can be found of wind farms being completed before network reinforcements, resulting in curtailment during the first few years of operation.

Last, local opposition has been, is and will remain an issue for any new large overhead transmission lines. Local opposition argues that new lines have an impact on health and the natural landscape. While underground direct current transmission lines are technically feasible, they are four to five times more expensive than overhead lines, and cannot be built extensively to relieve congestion.

Distribution

The distribution grid will also become increasingly congested. Contrary to what is often said, distribution networks will not disappear with the development of distributed generation such as solar PV and wind. Rather, above a certain level, large amounts of distributed solar PV may imply two-way power flows on electricity networks and voltage or thermal constraints on medium-voltage lines.

Distribution costs are high due to the mileage of lines and the associated civil works, in particular when lines are located underground. Distribution costs represent 40% of total investment in the power sector (see Chapter 8). It would be very costly to reinforce the grid every time a new rooftop solar PV was installed. It is therefore important to adapt planning standards to future needs, accounting for distributed generation. Moreover, evacuating “the last kWh” on sunny summer days during hours of over-generation, when the electricity has little or no value, would be uneconomic. As such, the electricity distribution network will have to manage congestion by using a growing number of distributed resources and relying on curtailment of solar PV and wind plants, back-up gas or oil generators, demand reduction and distributed storage.

Other technical challenges

Other technical aspects of deep decarbonisation scenarios need to be carefully analysed. They include inertia, voltage control, common mode failures and black-start capability to re-start power plants and the entire system after a power failure (O’Malley, 2015). These important technical challenges should not be underestimated, as they raise non-negligible risks for operational security. It is not clear, however, to what extent market-based solutions can be developed to address these more technical issues. Alternatively, technical regulation or connection contract arrangements will be required to ensure system security.

Result: More adjustments in the last few hours before operation

In the context of increasing network congestion caused by generators dependent on weather conditions, the last few hours before real time will become critical to ensuring efficient system operations and security. This creates the need to improve many aspects of existing markets. Existing electricity systems are largely based on the day-ahead forecast, with management of small deviations during the following 24 hours. In future electricity systems with variable renewables, more adjustments will have to be made in the last few hours, i.e. 3 to 12 hours before operation. A more dynamic electricity system will require more frequent ramping of power plants, steeper ramps, short-notice start-up and ramp-up to mitigate the variability and relieve congestion. In this context, it is necessary to define what decentralised market arrangements can achieve without putting system security at risk. Future market arrangements need to reflect this technical reality.

3.2. Resolution of market design

This section describes existing market designs to identify best practices for low-carbon power systems with high shares of VRE. Looking ahead, electricity systems in OECD countries can expect to face similar physical challenges, as their energy mix is expected to decarbonise. While it is clear that one size does not fit all, similar issues should create the opportunity for many different markets to adopt the most efficient set of common market rules.

High-resolution vs. low-resolution market design

Operating an electricity system efficiently requires the use of the lowest cost generators available to meet load, taking into account grid constraints. In more technical terms, the frequency of the system has to be kept at 50 or 60 Hz. In addition, electricity flows on the grid must be kept within certain acceptable limits to ensure system security in case a transmission line is suddenly cut off. In unbundled electricity systems, these tasks are the responsibility of system operators.

The cost of providing electricity usually varies by time (from one minute to the next) and by location. To understand this, consider for instance a consumer turning on the air conditioning system at noon. If this consumer is connected to a medium-voltage line where many other consumers have installed solar PV systems, the cost of this instantaneous increase in electricity consumption by the air conditioning system is very low, perhaps nil or even negative if too much electricity is being generated.

Conversely, if no solar PV systems are present and the day is calm but very hot, the action of this consumer turning on the air conditioning can be very expensive. It might require the operation of costly oil power plant on the generation side with a high marginal cost of USD 300 per megawatt hour (USD/MWh). And if no further generation capacity is available, it is possible that the system operator will be compelled to reduce consumption by industrial consumers and compensate them at a price of 1 000 USD/MWh or even more, in order to accommodate the consumer's requirement.

Although the operation of electricity systems ultimately obeys the laws of physics, the design of electricity markets can differ significantly to deal with this type of operational challenge.

Existing markets were designed to address the most salient issues at the time of their introduction. In the United States, the primary objective of regional transmission organisations (RTOs), such as PJM in the Northeast of the country, was to ensure the co-ordination of small balancing areas that were poorly interconnected. In Europe, the primary objective of market design has been to enable trading of electricity across borders, between large national balancing areas. The significant differences in design are thus not surprising.

Certain markets have opted for a simple design with "low resolution", i.e. they capture few of the underlying physical properties of the system, which they leave to system operators to handle. Others adopted a market design with "high resolution", to factor the physical reality of power systems into the process of price formation on the market itself. The resolution refers to the geographical resolution (nodal pricing vs. large bidding zones), as well as the temporal resolution (five-minute real-time prices are the highest resolution found in existing markets). The notion of resolution also includes the quality of market information in the intraday time frame.

High-resolution market design has been implemented in about one-half of US states. As illustrated in Table 3.1, this approach provides a still-simplified but much more detailed representation of electricity systems. This market design is more demanding and complex for market participants.

Table 3.1 • Technical resolution of market operations

	High resolution	Low resolution	Intermediate (high temporal resolution with low geographic resolution)
Example of market	PJM	Germany	Australian National Electricity Market
Power market platform	System operator	Power exchange	Power exchange
Bidding information	Unit/plant, complex bids	Portfolio, aggregated bid	Unit/plant
Geographic resolution	Nodal	Single national price	Zonal
Primary market	Real-time	Day-ahead	Real-time
Real-time balancing prices	Single marginal price	Asymmetric prices	Single marginal price
Dispatch internal	5 minutes	15 minutes or longer	5 minutes
Operating reserves	Co-optimised with energy	Separate markets	Separate markets

Low-resolution market design can be found in Europe. Electricity prices provide a rough economic representation of actual system conditions. The main advantage of a low-resolution system lies in its simplicity, which was sufficient in the 1990s to open up the electricity system to competition and facilitate cross-border trade of electricity on the day-ahead timeframe.

High-resolution market design

High-resolution electricity markets seek to provide an accurate economic representation of the operation of power systems in practice. To that end, system operators directly manage the market platform where bids are collected using software called a market management system (MMS). For example, PJM uses a software program called e-terramarket developed by Alstom Grid. Each generation unit submits complex bids, including an energy price and a fixed component corresponding to start-up and minimum running time. In addition, system operators take into account start-up duration and the ramping capability of specific units.

The system operator uses the MMS software to calculate the least-cost security-constrained dispatch (LCSCD) of power plants, which results in different real-time electricity prices. In PJM, for example, more than 10 000 price nodes are regrouped into 12 bidding hubs. A centralised algorithm calculates the price at each node every 5 minutes, corresponding to a vast amount of information.

The primary market in high-resolution market designs is the real-time market. PJM, for example, calculates locational marginal prices (LMPs) for a given five-minute period of time based on actual system conditions. The LMP values posted to the PJM website are available to participants within 10 minutes of their calculation. Transactions between buyers and sellers are settled hourly; invoices are issued to market participants weekly.

> "The [real-time] price tells PJM market participants the cost to serve the next megawatt of load at a specific location. The calculations factor in all the available generating sources to come up with the mix that creates the lowest production cost, while observing all limits on

the transmission system. The use of actual operating conditions and energy flows in determining LMPs encourages the efficient use of the electric grid and enhances reliability." (PJM, 2015).

Box 3.2 • Day-ahead markets in nodal pricing systems – US case study

In parts of the United States where competitive wholesale markets prevail, the ISOs and RTOs operate both a day-ahead market and a real-time market. According to the Federal Energy Regulatory Commission (FERC), 95% of transactions are agreed upon in the day-ahead market, leaving only 5% to be scheduled in the real-time market. Real-time markets run hourly as well as in five-minute intervals.

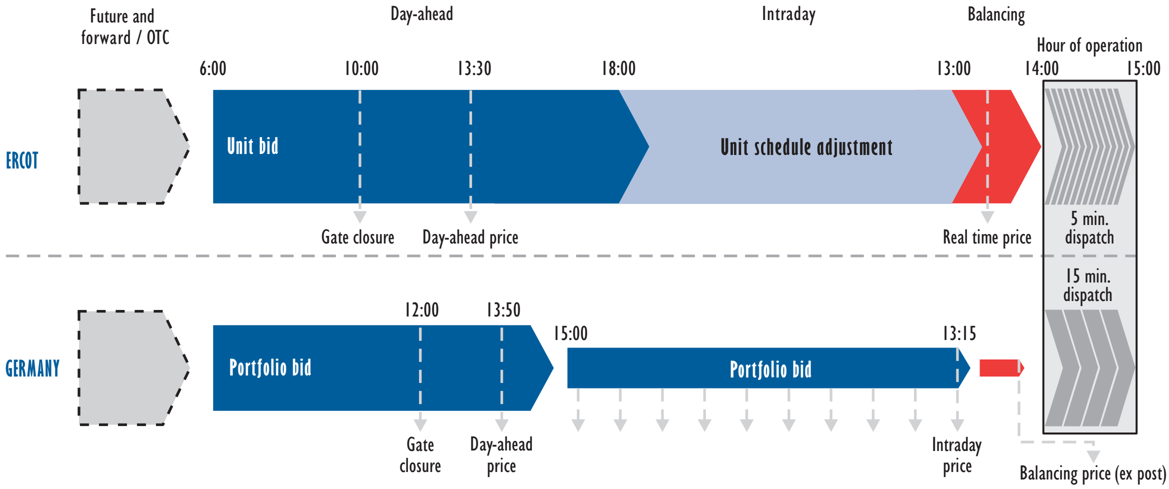
A variety of physical (i.e. forward) and financial (i.e. future) products are available in the market. In addition to physical transactions through the ISOs, trading can occur as bilateral transactions via direct interaction but often occur through an exchange such as the New York Mercantile Exchange (NYMEX) and the Intercontinental Exchange (ICE), which offer longer-dated products. The ISOs/RTOs also offer virtual bidding, which allows traders to financially participate in the day-ahead market even if they do not have physical generation assets.

As a specific example, PJM includes day-ahead and real-time markets for energy, as well as markets for ancillary services, capacity, and financial transmission rights (FTRs). FTRs are contracts to hedge against transmission congestion costs. They pay the holder of the FTR for transmission congestion costs over a specific grid path, and are typically distributed by auction. In PJM, load-serving entities (LSEs) are issued auction revenue rights (ARRs), which entitle them to either a share of funds from FTR auctions or the right to convert ARRs to FTRs. FTRs are also exchanged minimally through a secondary bilateral market. An FTR market is a critical market design component as PJM's energy pricing relies on market-clearing nodal prices, i.e. LMPs.

PJM's system includes over 10 000 price nodes across 20 transmission control zones, with trading available at nodes, at aggregates of several nodes, at 12 hubs consisting of hundreds of nodes each, and at 17 import and export external interfaces. Trading in the PJM system and other systems across the United States is liquid in the day-ahead markets and drives price convergence between day-ahead and real-time LMPs. In particular, in the virtual bidding market, traders can offer increment bids (INC), decrement bids (DEC), and up-to-congestion transactions (UTCs) in the day-ahead markets. INCs simulate generation offers, DECs simulate load buy bids, and these are factored into PJM's market clearing. UTCs are bids to purchase congestion and losses between two points in the system. These bids affect day-ahead scheduling, including dispatch, resource commitments and pricing. This nodal pricing system facilitates adjustments to dispatch in the real-time market, efficient use of variable resources and demand-side response, and limits to market power by individual generators.

PJM's day-ahead market is a forward market in which hourly LMPs are calculated for the next day, based on the amount of energy generators have offered to produce, the amount of energy needed by consumers, and scheduled transactions between buyers and sellers of energy. It is important to understand that day-ahead prices are forward prices. Figure 3.4 represents the natural link that exists between day-ahead and real-time markets. In this simplified representation of the sequence of short-term markets, it is clear that the real-time is equal to the day-ahead price if demand forecast errors and generation forecast errors are limited. In reality, in addition to net-load forecast errors, it is also necessary to take into account the fact that fewer generation units are available to balance generation and demand, and these units tend to have higher costs. The cost of last-minute changes can be slightly higher, but day-ahead prices represent the best estimate of real-time prices and are therefore intrinsically correlated.

Figure 3.4 • Timeline of centralised and decentralised markets



Notes: ERCOT = Electric Reliability Council of Texas; OTC = over the counter.

System operators must also be able to balance the power system in case of unplanned generator outages or unexpected deviations in energy demand. They contract operating reserves as part of the ancillary services needed to operate the system. Different categories of operating reserves are available, such as synchronised reserves or regulation reserves, and different markets have different definitions of the products needed to balance system deviations.

From a market design perspective, these operating reserves take the form of generating capacity that remains available instead of actively generating electricity. In some markets, the provision of reserves is co-optimised with the calculation of real-time energy prices. Such a calculation algorithm is centralised and takes into account the trade-off between energy provision and the supply of capacity reserve, with the objective of finding the least-cost solution.

High-resolution markets involve a high degree of centralisation. To be able to calculate prices for thousands of nodes every five minutes on a real-time basis, while observing transmission system limits, the MMS needs to centralise all the information about the bids of all generators and the state of the transmission system. The sophistication of the algorithm reflects the complexity of the market clearing process. In addition, the system operator instructs power plants to operate directly and, in practice, has direct control over the generation assets at all times, not the generating company.

Due to this high degree of centralisation, the participation of distributed resources might require adaptations. Small generators and demand response resources do not usually attain the minimum size to participate directly in centralised markets, because a) transaction costs are high for bidding in and dispatch by such complex markets, and b) there are technical limitations to the number of market players and bids that the security constraint dispatch algorithm can handle in real time. In effect, distributed resources cause centralised systems to no longer be as centralised, in the sense that bids are not unit-based for distributed resources and system operators do not have direct control over them. Instead, aggregators of demand response or renewables can bid, proving to be an effective way of accommodating distributed resources in centralised systems.

In some instances, system operators have activated (emergency) demand response a few hours before real-time prices. These activations have had an important price suppression effect on real-time markets. As renewables are deployed, the generation schedule will increasingly need to be adjusted to compensate for wind and solar forecast errors in the intraday time frame, that is, between the day-ahead prices and real-time prices. In the absence of intraday price signals, it might be difficult for renewable aggregators and other market participants to reschedule their generation programme efficiently to contribute to system balancing. Indeed, renewables themselves can provide flexibility and offer bids below which they will not produce, but these bids have to be activated as soon as better information is available on system conditions, most notably around six to two hours before real time. Certain high-resolution centralised power markets lack price signals during the intraday time frame.

In addition, this centralised model has not been implemented in many markets where such centralisation is considered excessive or would involve a loss of control or sovereignty that local governments or regulators are not ready to accept. Around two-thirds of Americans are served by an ISO or RTO, but many areas have not implemented the standard market design approach that FERC proposed in 2001-03. The rest of the electricity system in the United States remains very fragmented, with around 130 balancing areas that are not well co-ordinated and are less prepared to accommodate high shares of wind and solar power.

Another perceived drawback of high-resolution market design is its high complexity for systems with low levels of network congestion. For example, in European countries, most national markets have experienced few internal congestion problems until recently. In France, two regions suffer from structural congestion, Bretagne and Provence-Alpes Côte d'Azur, but the cost implications remain limited. Similarly, Germany did not experience internal congestion problems until 2011, with the development of wind power in the north of the country. Other European countries, such as Belgium and Spain, still do not experience significant internal congestion. Despite high shares of renewables, Spain has no significant internal congestion, thanks to important network investments during the 2000s and a decline in electricity demand since the economic crisis of 2007. LMPs have not been implemented.

Low-resolution market design

Europe has adopted a simplified decentralised market design – called here a low-resolution electricity market design. In Europe, the first objective was to enable trade of electricity across borders. The balancing areas are much larger than those that existed in the United States before the introduction of RTOs and ISOs. Incumbents have not been horizontally unbundled into several generating companies and competition between generators has been introduced mainly by cross-border competition.

Within a given price zone, power exchanges rather than system operators calculate European power prices, as if congestion and network constraints did not exist (actual congestion is relieved by redispatching more expensive power plants after the day-ahead market, but this has no impact on prices). Congestion between price zones has historically been taken into account in a very simplified way. An often-mentioned objective of this market design is ultimately to have just one price for the entire area. For example, while France and Germany have a size comparable to PJM's in terms of consumption and peak demand, each country has only one price, while PJM has thousands of prices.

These price zones have been refined in a certain number of countries. Sweden, for example, introduced five price zones in 2010. A process is under way in Europe to introduce more bidding zones leading to a higher resolution in the geographic representation of the electricity system used by electricity markets. Given that network utilisation and congestion are likely to become more unpredictable with the deployment of renewables, it will become necessary to define such price zones in a more dynamic way.

The primary market in low-resolution designs is the day-ahead market. Most transactions take place bilaterally over the counter (OTC) and directly via supply contracts. Consequently, liquidity on the day-ahead market is not usually high, with notable exceptions. In Germany and Austria

50% of consumption was traded on EPEX Spot in 2014, as was 15% in France (EPEX Spot, 2015). In Spain, 78% was traded on the OMIE spot market in 2013, while in Nordic countries 85% was traded on Nordpool Spot (ACER/CEER, 2012). Nevertheless, market participants usually consider power exchanges as the reference.

Day-ahead market coupling in Europe is a major achievement of the internal energy market, and now links the Nordpool area, Great Britain, central Western Europe, the Iberian Peninsula and Italy. The same algorithm is used to clear markets simultaneously, to ensure that electricity always flows from the lower price zone to the higher price zone. Flow-based market coupling, introduced in Western European countries in 2015, uses a more sophisticated presentation of the meshed transmission network than previously, helping to mitigate issues such as loop flows at the borders. Interestingly, the European intraday and balancing markets are considered to be independent and relatively small residual markets, contrary to high-resolution markets. Their depth is limited to a few gigawatts of power needed to balance the system. Market participants have to bid on these different markets and their outcome is difficult to predict.

On the intraday market, the target European model is to have continuous trading of power. Trades are bilateral and executed based on the best price. In 2014, continuous EPEX Spot intraday volumes represented 30.7 terawatt hours across Germany, Austria, Switzerland and France (EPEX Spot, 2015). In practice, the only liquid intraday market is in Germany because of the marketing of renewables there. Trades to adjust wind forecast errors are the main source of liquidity on the intraday markets.

Gate closure on intraday markets has been gradually moving closer to the time of operation, and currently stands at 30 minutes for local trades and 60 minutes for cross-border intraday trades. Market players may submit bids until gate closure. Gate closure close to real time enables market players to reduce their own imbalances and therefore system imbalances, but provides system operators with less time to react and ensure system security.

Consequently, numerous balancing markets have been designed to reduce imbalances that system operators will have limited time to manage. The “responsible balancing parties” (in practice mainly large suppliers) are incentivised to balance their own portfolios of generation and load. The incentive comes from an asymmetric price system and the choice of a pay-as-bid or an average price rather than uniform marginal pricing system to settle or cash out imbalances. In the Netherlands, for example, upward balancing energy is on average more expensive than the day-ahead price, while downward balancing energy is less expensive. In terms of market design, balancing prices do not reflect the marginal cost of the marginal unit needed to balance the system.

In practice, however, a system operator cannot wait until the very last minute to redispatch generation to relieve network constraints securely. Many system operators have to instruct certain power plants to generate and others to reduce output before market gate closure and sometimes already in the day-ahead timeframe. Such redispatching actions clearly take place out of short-term market platforms and are not reflected in market prices. There is a significant discrepancy here between the design of markets and the needs of system operators. In some cases, redispatching actions are less automated than market dispatch. If the network becomes more congested, as is likely to be the case with more renewables, this could even lead to greater risk for system security.

Operating reserves are traded on separate markets. Traders and generators (not the system operator) have to decide whether they want to bid and in which market. Given the lack of correlation and predictability of different markets, trading decisions are more complex than with a centralised approach, where a computer algorithm co-optimises the provision of operating reserve and real-time energy. In addition, the decisions are taken for a smaller portfolio of resources, and these products cannot be traded close to real time as this favours large players and reduces overall efficiency.

Low-resolution electricity markets may have a low degree of centralisation. This has proven to be an advantage where political circumstances do not facilitate the introduction of a single market design. They have been popular because they are not mandatory, are relatively inexpensive to implement and allow trade in electricity over large geographical areas.

The drawback of this low-resolution approach is that the markets cannot manage network congestion, which then has to be dealt with separately. System operators take redispatching actions within price zones before gate closure, and this interacts with market prices. Consequently, low-resolution market design is likely to lead to increasing inefficiencies as more renewables are introduced into electricity systems and increase congestion.

Ultimately, the disconnection between system operations and market representation will increase system security risk with high shares of renewables. European system operators are increasingly confronted with markets creating production schedules that do not respect the grid's capabilities. If market participants do not take into account the physical reality of the grid, the operation of the system becomes more difficult and requires more interventions. This could become a concern for security of supply in particular, as the gate closes only one hour before real time, allowing very little time for system operators to adjust production plans.

In practice, system operators continuously perform security analysis in close co-operation with power exchanges and use two market platforms, the intraday market operated by power exchanges and brokers, and the balancing market, operated by system operators and used to correct the imbalance of the system during the intraday timeframe. In Germany an additional platform enables system operators to redispatch. Markets with low resolution rely on a complex ensemble of parallel markets and constantly introduce new market products to meet system flexibility needs; this leads to a high complexity of market design and poses interaction and co-ordination problems between platforms.

Europe is already experiencing these difficulties at the German border. In countries where flow-based market coupling is yet to be implemented, unplanned electricity flows can occur. For instance, interconnection capacity at the Germany-Austria border is relatively low, but the two countries are part of the same price zone. This has resulted in high electricity flows through Polish-German interconnections, raising concerns for system security and decreasing trading opportunities between Poland and Germany. Similar issues arose in 2012-15 at the border between Germany and the Czech Republic, leading to the installation of a phase shifter to better control electricity flows at the border.

In summary, low-resolution market design does not represent best practice for short-term markets, as the transition to a low-carbon system with high shares of VRE is likely to increase transmission congestion and forecast errors in the day-ahead timeframe. At best, this leads to out-of-market operations that are not priced in. If these out-of-market operations become too frequent, system operators increase network reserve margins and reduce the transmission capacity available to the electricity market, leading to inefficient use of existing assets on the grounds of system security.

Such a poor utilisation of network infrastructure would make the transition more costly and therefore more difficult to implement. Finally, too large a disparity between market resolution and physical reality increases the probability of security-of-supply events.

Market power and geographical resolution of short-term markets

Electricity markets are vulnerable to market power, which is enjoyed by specific power plants that are frequently dispatched in short-term markets to relieve congestion. Many economists argue that high-resolution markets can make the market less vulnerable to market power (Holmberg and Lazarczyk, 2012; Green, 2007). However, the interactions between market power and electricity markets are complex to analyse and depend on the geographical distribution of the power plants of each generation owner. For example, market power can increase the profit of one unit while reducing the revenues of other units of the same generation owner. Ultimately, market power issues have to be mitigated using plant-by-plant regulatory measures in both high-resolution and low-resolution electricity market designs.

In the United States' ISO/RTO markets, energy offers are normally capped at 1 000 USD/MWh, and market power mitigation provisions are employed if a generator is determined to have market power based on its bid. Generally, in this case, the generator's offer will be reduced to a reference level determined for each generator and based on current fuel prices. Local market power also exists in low-resolution markets that have uniform energy prices across larger areas. During the California electricity crisis of 2000-01, the so-called DEC game[8] revealed that the zonal, i.e. low-resolution, pricing system is vulnerable to gaming because intra-zonal congestion is not addressed in the day-ahead market but only by redispatching. Alaywan et al (2004) suggest that market power issues have been reduced by the transition from zonal pricing to nodal pricing in California.

In a European context, Germany is implementing highly detailed regulation of local market power: generators have to bid their marginal cost on the redispatching platform. Similarly in France, the existence of market power in the Provence-Alpes Côte d'Azur region used to be addressed with long-term contracts with one or two power plants, defining in advance the bids under the supervision of the regulator.

Moving from a low-resolution to a high-resolution market design does not by itself eliminate local market power issues; the way to address this is with plant-by-plant regulatory interventions, such as those being pursued in the United States and Germany. In high-resolution market design, generators should have to bid prices in advance (see Figure 3.4) and have less ability to exploit transmission constraints close to real time. Intuitively, as high-resolution markets are expected to ensure better use of existing transmission infrastructure, they should in principle reduce the frequency of network congestion, rather than increasing it, further mitigating the vulnerability of the system to market power. Only new transmission investment can structurally reduce market power (Wolak, 2014).

Against this background of existing market design, the next section identifies best practices that could be implemented in both a European and North American context, as well as in other competitive electricity markets with high shares of low-carbon VRE. While these best practices do not constitute a "blueprint" for all situations, they seek to highlight useful lessons learnt.

3.3. Designing short-term markets fit for a high share of renewables: A strawman proposal

This section provides a roadmap for policy makers endeavouring to improve the design of the market for electrical energy. Features of existing market design described in the previous section are not totally fit for purpose for the five challenges described in the first section of this chapter – adequacy, over-generation, ramp rate flexibility, predictability and congestion. While there is no one-size-fits-all approach to market design, identifying best practices can help identify the market rules best adapted to local conditions.

The issues encountered in North America differ as compared with Europe. In North America, existing RTOs have already implemented high-resolution market designs and gathered best

[8] The DEC game consists of a company submitting low bids for power plants located behind a constraint, in order to be scheduled to run on the day-ahead market and earn the uniform zonal price, knowing that they will probably be redispatched down (DEC'ed) because of the constraint.

practices. Compared to Europe, they could be improved by introducing intraday price signals that would aid the participation of distributed resources. For regulated markets in North America, however, better integration of neighbouring balancing areas into larger regional markets is needed, and the experience of decentralised electricity markets in Europe could be useful to them.

In Europe, the market design has a low resolution and it is suggested that it should evolve, particularly in the intraday and balancing timeframes.

Despite their apparent differences, similarities also exist between North American and European markets. The laws of physics are the same in all electricity systems, and this has led system operators to hand over the control of the system in a centralised manner 45-60 minutes before the time of operation. At this stage, system operators have direct visibility and remote control of the dispatch of all large units, information is always available unit by unit for redispatching reasons, and substations can be controlled or curtailed. One hour before the time of operation, the market information available to system operators is largely the same. It is possible to build on these similarities and existing market arrangements to enhance energy markets.

Principles for designing markets during the adjustment period (the last few hours before operation)

The adjustment period comprises the last few hours before time of operation. During this adjustment period, renewable and demand forecasts improve considerably. In addition, system operators have to make sure that network constraints are respected. The adjustment period is critical for the integration of renewables over large geographic areas and for system security.

Adapting short-term markets to low-carbon power systems with high shares of wind and solar power consists mainly of improving the design of markets over the adjustment period. To this end, five high-level principles that markets should meet are proposed: 1) locational pricing, 2) uniform pricing, 3) cost-reflective bids, 4) administrative reliability pricing, and 5) intraday transparency.

1) Locational pricing

Differentiate electricity prices by local geographic area, in order to reflect the differences in electricity generation costs due to the limitation of network capacity.

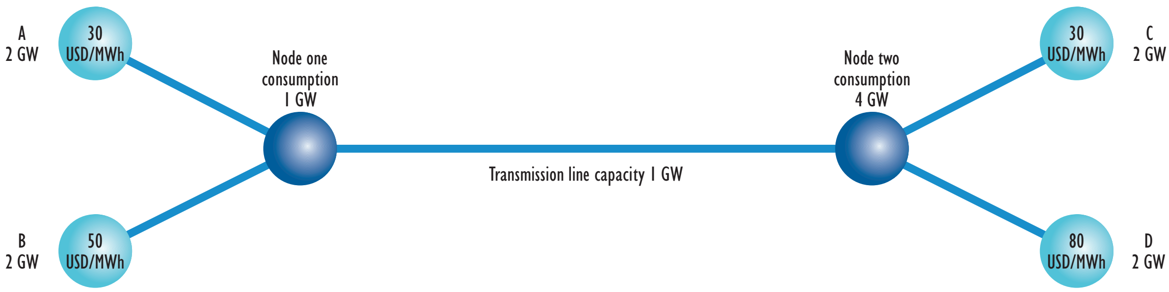
All electricity networks are already operated in such a way that the marginal generation costs differ by location at the end of the adjustment period, when grid capacity becomes congested. This is a technical requirement to ensure security of supply. Implementing locational prices implies that the prices are published and the associated financial settlement sufficiently reflects the reality of system operations.

The design of existing markets reflects a trade-off between simplicity and accuracy of locational pricing. Even where locational marginal pricing is implemented, it is common to trade electricity at trading hubs (for instance, the PJM Western Hub), with high liquidity and prices often equal for a large number of neighbouring nodes (Box 3.2). Where network constraints are always binding for the same lines and for radial networks, such as in Australia, zonal pricing might be sufficient to reflect the physical reality of networks.

Failure to implement locational marginal prices with a proper geographic resolution, however, can result in inefficiencies. Consider, for example, one line with four generators A, B, C and D, each with a capacity of 2 GW at 30, 50, 30 and 80 USD/MWh, respectively, as shown in Figure 3.5. The load is 1 GW in node one and 4 GW in node two, and a transmission line with a capacity of 1 GW connects them. If there is a single price, the market clears at a price of 50. This market outcome would violate the transmission capacity limit. The system operator has to redispatch generators, having scaled back production at B from 1 GW to 0 GW to take into account the grid constraint and ask for extra production from D. The system operator asks generator D to produce at a price of 80 rather than generator B. After redispatching, the market price (50) does not reflect the marginal cost of the system (80). Generator A receives 50 while it should be paid 30. Generator B receives 50 but does not generate. While C and D are in the same location, they receive different prices. Consumers are not exposed to the costs, consume too much in node 2 and leave ample capacity in node 1.

As wind and solar power are expected to increase the volatility of electricity flows and lead to congestion, efficient locational pricing will be needed even in power systems yet to face major congestion. Given that these instances of congestion will be revealed during the adjustment period, locational prices are needed several hours before the start of operations, not just one hour before.

Figure 3.5 • Illustrative locational pricing



In addition, calculating efficient locational marginal prices is the only solution to ensure that system operators do not take undue security margins on the transmission capacity, particularly across borders. In ERCOT, for example, the implementation of LMP in 2010 has increased the number of hours with the same price across the balancing area. In other words, this could seem paradoxical, but locational pricing algorithms are needed to make sure that the existing physical infrastructure is fully utilised and triggers the convergence of nodal prices as often as technically possible.

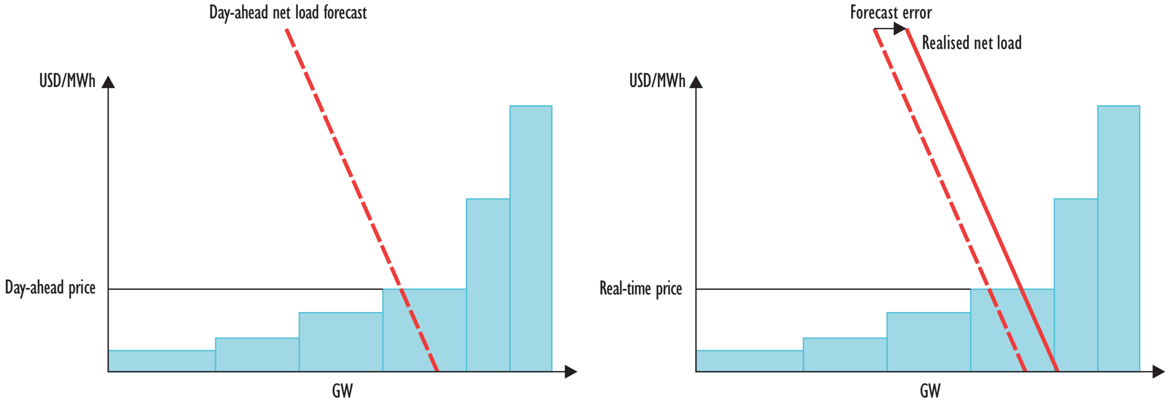
2) Uniform pricing

Apply uniform prices to all real-time energy used for balancing, to reflect the marginal cost of the marginal resource used to balance the system at each location.

Uniform prices should be used during the entire adjustment period. Such uniform prices should reflect the marginal cost of the marginal resource used to balance the system at each location; this would send the right signals for resources to adjust their schedules. This is not the case at the moment in many European markets, where balancing prices are equal to average costs instead of marginal prices and balancing prices are often 50% lower or higher than day-ahead prices. In principle, uniform prices should lead to well-correlated prices in the intraday and balancing timeframes.

Consider a simple merit order. In the day-ahead market, the market clearing price reflects the marginal cost of production, given the best forecast of the load, net of wind and solar output. Due to forecast errors, however, net load ends up being higher in real time, but in this example the marginal cost does not change. The intraday and the real-time price should be equal in this simple example.

Figure 3.6 • Evolution of prices during the adjustment period



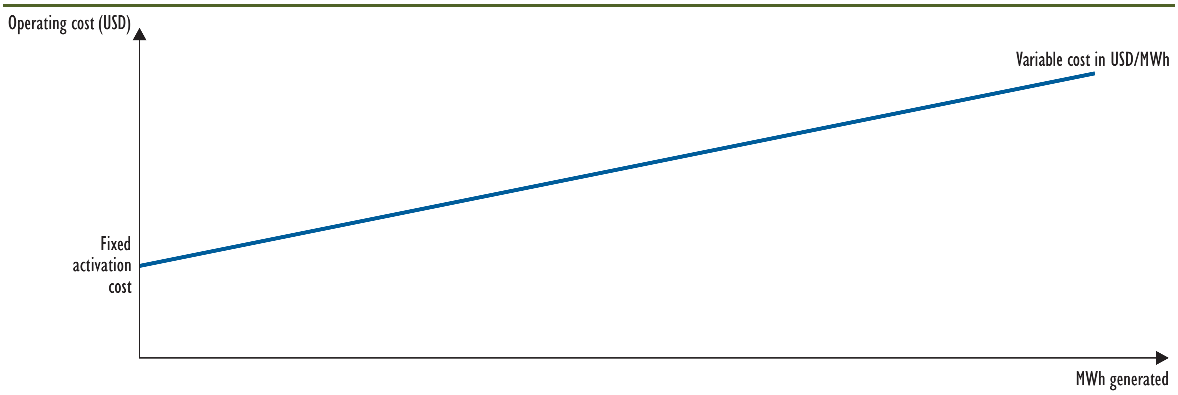
In practice, however, due to the unit commitment problem and rigidity of the power plant mix, the marginal costs are rarely exactly the same in day-ahead and real-time markets. Certain power plants have long ramps and high start-up costs and are not available to compensate for forecast errors. These constraints increase the overall cost of operating the system. As illustrated in Figure 3.6 below, the marginal cost of power systems depends on the marginal cost of the last unit, which can vary but does not change radically in the case of forecast errors. The day-ahead, intraday and real-time prices should be very close.

3) Cost-reflective bids

Use bids that reflect the marginal costs of different resources in different locations during the entire adjustment period.

Marginal cost pricing in competitive electricity markets requires that bids from market participants reflect their costs. In high-resolution market design, the corollary of this proposition is that bids have to be location-specific. They have to be unit-based for centralised generation and aggregated by location for portfolio bids for distributed resources and virtual power plants. Ideally, these bids should also reflect the variable component of marginal costs and the fixed component of operating costs, such as start-up costs or the activation costs of aggregated demand response. These start-up costs are reflected in the energy prices when the power plant is dispatched (Figure 3.7).

Figure 3.7 • Cost-reflective, complex bids



The market outcome can result in negative prices. This reflects the fact that it would be more costly for some power plants to reduce output for, say, only one hour rather than receive a negative market price. System operators must also take into account other technical constraints, such as ramp rates.

Regulators sometimes have to regulate the bids of generators that enjoy market power at a specific location. The existence of locational market power, however, is independent of the implemented market design, as locational market power exists, and can be exercised, under locational pricing as well as under a single price with redispatching (see discussion on market power in Section 3.2).

4) Administrative scarcity pricing

Regulate energy prices during capacity shortage conditions, i.e. when there is insufficient capacity to meet, in addition to the load, the reserve requirements needed for reliability.

Administrative scarcity pricing pricing is a form of government intervention in markets for electrical energy. Despite all the attention that governments devote to security of electricity supply, rare instances of capacity shortage will nonetheless occur. If this were not the case, it would be an indication of the existence of excess capacity. The occurrence of a shortage of capacity brings an increased risk of involuntary load curtailment (Figure 3.8).

Some form of regulation of scarcity prices is necessary to ensure accurate price formation during scarcity hours (see chapter 4). Allowing market participants to bid extremely high prices that would lead to peak prices raises a number of issues. First, large market participants do not wish to expose themselves to a potential *ex post* competitive behaviour investigation, and in practice only a small number of small traders typically bid such prices, leading to scarcity price formation that is not robust. Second, as these capacity shortage situations rarely materialise, most traders do not devote significant resource to anticipating them, which can lead to high prices even when the system is not stressed or conversely to too low prices when there is actually a shortage.

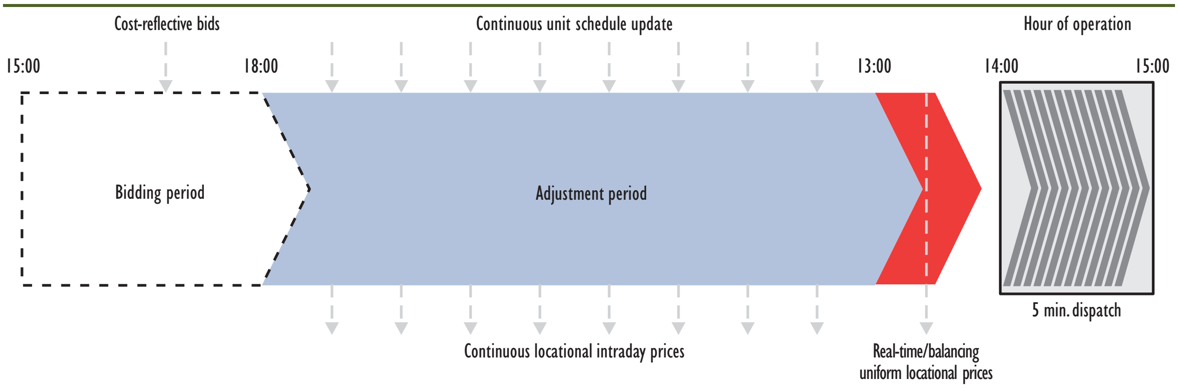
The design details of administrative scarcity pricing can vary depending on the pre-existing market design. If prices are well correlated, the regulation of operating reserve scarcity pricing might suffice and will find its way into the real-time price, the real-time price will find its way into the intraday price, which will find its way into the day-ahead price.

A more detailed discussion of experiences of reliability pricing is provided in Chapter 4, including a detailed discussion of the rationale for administrative pricing, indications for the construction of the regulated pricing curve, and experiences of market power mitigation rules.

5) Intraday transparency

Transparent intraday prices are necessary to inform all market participants about the cost of serving the next megawatt.

Figure 3.8 • Illustrative timeline during the adjustment period



This evolution is critical in all markets, in Europe and in North America. Thanks to transparent intraday price information, aggregators of distributed resources, such as demand response and virtual power plants, can adjust their schedule in a decentralised fashion to complement intraday variations caused by increasing shares of renewables.

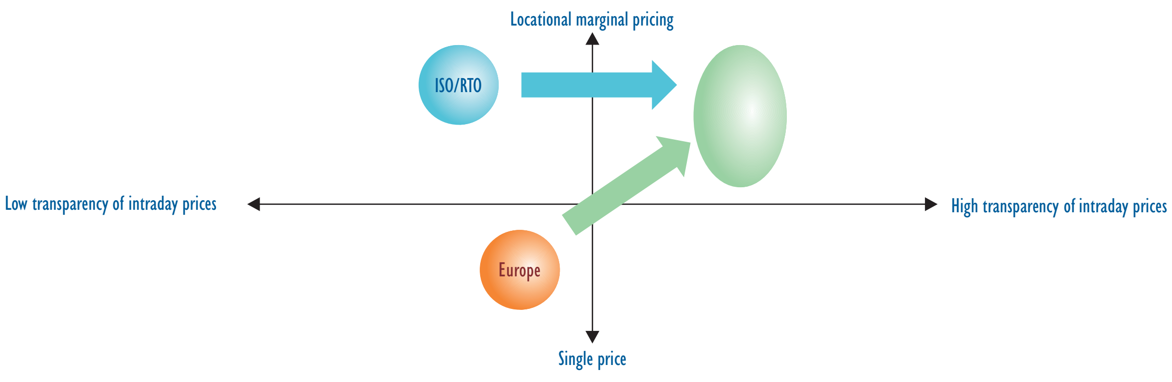
Because the cost of last-minute generators available at short notice is high, it is important to continuously update the schedule of other resources during the adjustment period. The least-cost resources have to be activated as soon as system operators receive better information about actual demand and variable outputs. In a system with abundant distributed generation, the continuous update of intraday prices can ensure the price co-ordination of many decentralised resources.

To sum up, the design of markets becomes critical for low-carbon power systems during the load and generation forecast adjustment period. As illustrated in Figure 3.9, the experience of IEA countries leads to the identification of the following best practice for the sequence of market participant and system operator activities, (represented for a start of operation at, say, 14:00).

- Immediately after the day-ahead market, participants submit location-specific bids for all the available resources, reflecting their marginal operating costs. The same bids are used by the market management system throughout the adjustment period, which prevents gaming in instances of tight system conditions.
- After the bidding period, market participants may not change their bids, but can continuously update their unit schedule to take into account generator availability and wind and solar forecast errors at each location, as well as the evolution of intraday prices.
- Market operators use all the information available to continuously calculate the least-cost dispatch, while respecting the technical limits of all the resources in respect of lead time, start-up duration, costs and ramp rates.
- Market operators publish continuously updated location-specific intraday prices.
- The gate for schedule adjustment closes around one hour or less before real time. After this closure, market operators calculate the real-time balancing price at each location.
- After gate closure, the system operator relies on operating reserves to balance generation and load.

This section provides a very detailed and specific example of how a market could be designed for efficiency, building upon experiences in both Europe and North America. Notwithstanding this detail, the proposed market framework is not a complete revamp of existing market design, but actually builds upon existing structures. The necessary changes to existing market design may in fact be minimal.

Figure 3.9 • Overview of the evolution of the design of power markets



Minimal changes required to existing market designs

Based on the compilation of best practices in existing power markets, the following market design recommendations are made for the operation of the power system (Figure 3.10).

North America

In North America, markets with RTO/ISOs implement most of the best practices identified in this report. The design could be further improved by providing greater transparency of the evolution of locational prices during the intraday timeframe. This would contribute to helping aggregators of demand response and virtual power plants better respond to system needs and reschedule their resources, which are decentralised.

In regulated markets in North America, the application of best practices during the adjustment period could be considered. This would help to reap the benefits of integrating balancing areas, on a voluntary basis, while keeping a high degree of autonomy and decentralisation. A specific balancing area, for instance, could join a platform with portfolio bids rather than unit-specific commitment. Flexibility in market participation is needed.

European markets

With regard to the balancing markets, most designs already rely on unit-specific bids in order to relieve congestion. In Germany, a specific platform for redispatching power plants has been put in place. The market information and bidding interfaces already exist in all markets. The only changes needed for balancing would be the introduction of uniform prices and their publication location by location. At the moment, redispatching prices are not published. Greater transparency is needed on balancing price information.

In Europe, the intraday markets would see the most important changes. Given that these markets are relatively small and recent, they already need to be modified because of network codes, and therefore changes could be implemented.

The principles advocated in this chapter would imply that intraday markets should be an extension of the balancing markets. Like balancing, economic dispatch could be done by the system operator based on detailed information about the transmission networks and technical constraints, using the same unit-specific bids for all resources, and publishing the intraday prices continuously for each location.

References

ACER/CEER (Agency for Cooperation of Energy Regulators/Council of European Energy Regulators) (2012), *Annual Report on the Results of Monitoring the Internal Electricity and Natural Gas Markets in 2011*, ACER, Ljubljana.

Alaywan, Z., Wu, T. and Papalexopoulos, A.D. (2004), “Transitioning the California market from a zonal to a nodal framework: An operational perspective”, Power Systems and Exposition 2, pp. 862-867.

CAISO (California ISO) (2015), “CAISO Energy Imbalance Market”, presentation given by CAISO at the IEA Electricity Security Advisory Panel (ESAP) and CEER Workshop on Regional Resource Adequacy, www.iea.org/media/workshops/2015/esapworkshopv/Rahman.pdf.

DENA (Deutsche Energie-Agentur) (2010), *Integration of Renewable Energy Sources in the German Power Supply System from 2015-2020 with an Outlook to 2025*, DENA, www.dena.de/fileadmin/user_upload/Projekte/Erneuerbare/Dokumente/dena_Grid_Study_II_-_final_report.pdf.

EPEX Spot (2015), *Annual report 2014*, www.epexspot.com/document/33486/EPEX%20SPOT_Annual%20report_2014_SD.pdf.

EWE Netz (2015), "EWE's approach to double grid connection capacity by using DER flexibility", presentation to co-operation workshop on energy on 19 March 2015, https://webhotel2.tut.fi/units/set/ide4l/EWE_%2020150319_Aachen_Merkel.pdf.

Green R. (2007), "Nodal pricing of electricity: how much does it cost to get it wrong?" *Journal of Regulatory Economics*, Vol. 31, pp. 125-149.

Holmberg, P. and E. Lazarczyk (2012), "Congestion management in electricity networks: Nodal, zonal and discriminatory pricing", Cambridge Working Papers in Economics 1219, Faculty of Economics, University of Cambridge, www.econ.cam.ac.uk/research/repec/cam/pdf/cwpe1219.pdf.

IEA (International Energy Agency) (2014), *The Power of Transformation*, IEA/OECD, Paris, www.iea.org/bookshop/465-The_Power_of_Transformation.

IEA (2013), *Electricity Networks: Infrastructure and Operations*, IEA Insights Series, www.iea.org/publications/insights/insightpublications/ElectricityNetworks2013_FINAL.pdf.

IEA/NEA (2015), *Projected Costs of Generating Electricity*, OECD Publishing, Paris.

NREL (National Renewable Energy Laboratory) (2015), "Western wind and solar integration study", www.nrel.gov/electricity/transmission/western_wind.html.

O'Malley (2015), "Integration of variable renewables: What the market can do to help or hinder", Presentation at the conference, Elements of a New Target Model for European Electricity Markets, www.ceem-dauphine.org/assets/dropbox/Mark_OMalley_Target_Model_Final.pdf.

PJM (2015), *Locational Marginal Pricing*, fact sheet, www.pjm.com/~/media/about-pjm/newsroom/fact-sheets/locational-marginal-pricing-fact-sheet.ashx.

PJM (2014), *Renewable Integration Study Reports*, www.pjm.com/committees-and-groups/subcommittees/irs/pris.aspx.

Smith J.C. et al. (2010), *Impact of Variable Renewable Energy on US Electricity Markets, Power and Energy Society General Meeting, Conference publication, 2010 IEEE*, http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5589715.

Wolak, F, A. (2014*), Measuring the Competitiveness Benefits of a Transmission Investment Policy: The Case of the Alberta Electricity Market*, http://web.stanford.edu/group/fwolak/cgi-bin/sites/default/files/alberta_transmission_benefits-energy-policy-wolak.pdf.

Chapter 4 • Reliability, adequacy and scarcity pricing

HIGHLIGHTS

- Current levels of electricity supply reliability are very high in member countries of the Organisation for Economic Co-operation and Development (OECD); in practice, most small power interruptions are caused by incidents at the distribution level.
- Reliability standards will have to evolve with the development of new technologies and in particular demand response. For the time being, reliability standards are still needed and probabilistic methods are more suited to variable renewables (for example, the one-in-ten event rule in the United States or the three hours of loss of load expectation in France, Great Britain and Belgium).
- When defining reliability standards, which is a complex task, governments tend to adopt conservative values. While this can lead to overinvestment, the additional costs remain much lower than other policies and their impact on electricity bills is relatively limited.
- Scarcity prices play a key role in ensuring reliability. On the demand side, high prices are useful for reducing demand. On the generation side, scarcity prices incentivise plants to be available when most needed and can remunerate the fixed investment costs of peak capacity.
- To improve scarcity price formation while addressing political and market power concerns, regulators should develop administrative scarcity pricing rules, putting a price on reliability and addressing market power for when there is capacity shortage.
- Scarcity pricing can ensure adequacy if there is sufficient demand response to meet the reliability standards on average. However, meeting high reliability standards at all times might require additional measures, such as capacity mechanisms (see Chapter 5).

Given the importance of electricity to the day-to-day functioning of modern economies, electricity disruptions can have huge consequences for industry, the service sector and the population at large. Not surprisingly, governments of OECD countries make electricity security a top priority.

Security of electricity supply is a broad notion comprised of three building blocks:

- security of fuel (i.e. availability of gas/coal/nuclear/hydro to generate electricity)
- security of system operations (avoiding blackouts)
- resource adequacy (avoiding load curtailment in case of capacity shortage).

The resilience of electricity systems is also increasingly important when setting technical standards of reliability regulation, as more frequent extreme heat waves or cold snaps can affect the availability of power stations, the thermal limits of networks, and the incidence of extreme load events.

Today's electricity security performance and regulatory arrangements are largely a legacy of investments made in the 1960s and 1970s. But in the coming years, ensuring electricity security is likely to prove challenging. Ageing capacity will have to be replaced, and moreover, replacement should ideally take place within a competitive market framework while also decarbonising the electricity sector. The conditions are present to create a perfect storm for electricity security.

Electricity security is by no means a new factor in market design. The transition of electricity systems raises questions about whether markets will be able to deliver the new investment needed to ensure security of supply – characterised by an intense academic debate which opposes purists of energy-only markets against supporters of capacity markets.

This chapter focuses on the question of resource adequacy. The discussion about the security of system operation largely pertains to the previous chapter on short-term markets. While fuel security and resilience are also important for electricity security, they are less directly related to the design of electricity markets and are not discussed in detail in this publication (see IEA, 2013; IEA, 2014a).

4.1. Regulation of reliability

Regulatory frameworks for reliability broadly involve setting *ex ante* standards for the bulk power system and reporting on the performance of the electricity sector *ex post*.[9] When defining a loss of load expectation (LOLE) or unserved energy, regulators administratively define a standard to override the market. In brief, LOLE can be defined as the expected number of hours in a specified period during which the daily peak load is higher than the available generating capacity.

Smart technologies will empower customers to choose their level of "reliability". Meanwhile, regulation of reliability today assumes that involuntary curtailments are impossible to avoid.

In many other industries, reliability is set by the market. In the car manufacturing industry, for instance, there is no regulator requiring that a car will start with a probability of 99.998%. Purchasers are free to choose a car that is less reliable – and with that a risk that it will not start one morning. Likewise in the telecommunications industry, a telephone call has a probability of failing to connect, due to an unavailable or overloaded network; for example, the network is often congested on 31 December at midnight.

Why regulate reliability?

Since the beginning of the electricity industry, a lack of real-time metering technologies has prohibited real-time billing of consumers. As a result of demand inflexibility, and because storage is costly, there has always been a risk of resorting to non-price restriction of demand in the form of load curtailment or rolling blackouts. Regulation of reliability developed due to these technical constraints.

Smart technologies could be a game changer. In principle, if consumers are enabled to respond to prices and reduce hourly demand when wholesale prices are high, then the market could potentially balance supply and demand at all times. If supply were scarce, prices would rise until there was enough voluntary load reduction to balance available capacity. Consumers would never suffer involuntary rationing and reliability would always be ensured (Cramton, Ockenfels and Stoft, 2013).

In the long run, it is possible to envisage a situation in which different consumers can express different preferences for the quality of their electricity supply. Some consumers might be willing to pay a high price for electricity in order never to reduce their consumption. Other consumers might accept reducing their consumption from time to time in order to pay a lower price. We might consider this second group as consumers with a low preference for reliability, because they accept voluntarily curtailment of some of their electricity usage. Estimates of the value of lost load for different categories of consumers provide a useful indication of the value that consumers place on reliability (Box 4.1)

---

[9] According to the North American Electric Reliability Corporation (NERC), reliability standards are the planning and operating rules that support and maintain a reliable electricity system (NERC, 2014). NERC's traditional definition of reliability rests on two different concepts: adequacy and operating reliability. Adequacy refers to "the ability of the electric system to supply the aggregate electric power and energy requirements of the electricity consumers at all times, taking into account scheduled and reasonably expected unscheduled outages of system components." Operating reliability is defined as "the ability of the electric system to withstand sudden disturbances such as electric short circuits or unanticipated loss of system components."

Reliability and adequacy notably refer to different time scales. System adequacy is analysed in terms of transmission and generation adequacy, the latter being defined as the ability of the generation to match the consumption on the power system. This is now commonly referred to as "resource adequacy", acknowledging the increasing role of demand-side resources.

Box 4.1 • Value of lost load (VoLL)

When setting a reliability standard, regulators must ensure a trade-off between the value that consumers place on supply reliability and the overall cost of the power system. A higher reliability standard reduces the costs associated with supply losses, but increases the price consumers must pay. The standard should seek to deliver reliability at minimum cost to consumers.

Value of lost load (VoLL) is a useful and important measure in electricity markets. It represents the price that an average customer would be willing to pay to avoid an involuntary interruption of electricity supply. In electricity markets, VoLL is usually measured in USD per megawatt hour (USD/MWh). It is used mainly in two ways, both on the planning side of the market and on the operational side. In planning, VoLL is used in the cost-benefit analysis of investment in generation, transmission and distribution in relation to customers' willingness to pay. On the operational side, VoLL can be used to calibrate resource adequacy rules and scarcity pricing algorithms.

As electricity cannot be delivered during an involuntary interruption, there is no transaction information on which to calculate VoLL. Instead, VoLL assessments must rely on econometric analysis. Four key methodologies are used for estimating VoLL in the field of economics: revealed preference survey; stated choice survey; macroeconomic analysis; and case study analysis.

Irrespective of methodology chosen, VoLL is highly variable depending on: 1) the sector or customer type; 2) the timing of outage; 3) the duration of outage; and 4) the time of advanced notification of outage and preparation (Ofgem, 2012). VoLL has been found to vary significantly, for example ranging from 713 GBP/MWh to around 59 000 GBP/MWh (London Economics, 2013) depending on which methodology is used, the duration and timing of the interruptions and different consumer categories. Typical figures used in many countries range around 10 000 USD/MWh, with some countries using higher values up to 20 000 USD/MWh.

In preparation for the introduction of a capacity mechanism in Great Britain, London Economics estimated the VoLL for electricity consumers there (2013). They used a variety of methods, but focused on choice experiments in which domestic and small and medium-sized enterprise (SME) consumers could state their willingness to accept an electricity outage by choosing between two scenarios. The results indicated a peak winter workday VoLL of 10 289 GBP/MWh for domestic users and GBP 35 488 for SME users based on their willingness-to-accept. It can be inferred that the higher value for SMEs results from their higher time-value of output and fewer possibilities to substitute other non-electricity-using activities during peak times, as compared to households. For industrial and commercial customers, a variety of value-at-risk approaches suggested an average VoLL of about 1 400 GBP/MWh.

At the time of a capacity shortage, system operators may direct distribution network operators (DNOs) to reduce the voltage in order to prevent demand disconnection. In Great Britain, around 500 megawatts (MW) of demand reduction may be achieved through voltage reduction. The results of analysis of the potential costs of voltage reductions indicate that, given the statutory range of voltages and the maximum 6% reduction, this is unlikely to cause significant costs to household and SME consumers.

London Economics (2013) suggests that a weighted-average winter peak workday VoLL is the most appropriate single number for the purposes of security of supply calculations, given that customers who experience an outage cannot in general be identified or ordered in terms of preference. Furthermore, given that large industrial and commercial customers may now, or in the future, have the option of demand-side response, self-supply and other types of protection, only the VoLLs across domestic and SME customers were used as the basis for London Economics' estimates. These calculations yield a weighted-average VoLL figure of 16 940 GBP/MWh for peak winter workdays in Great Britain.

Rationale for reliability regulation

Despite significant technological progress in metering, reading and billing, institutional barriers stand in the way of transforming the vision of price-responsive consumption into reality.

First, the electronic communication of electricity consumption data raises data privacy issues, and several countries are unwilling to let distribution or supply companies track the time at which (domestic) consumers make use of home appliances.

Second, retail consumers may have a smart meter and be billed based on real-time electricity prices, and yet choose not to respond to those prices, the reason being that the benefits of adjusting consumption are low compared to the time spent and effort needed to respond. Although this is eased by the automation of consumer demand reduction, the development of demand response is slow and, from a reliability perspective, the result continues to be an absence of demand response.

Third, electricity in general has a dimension of public service that limits the role of market-based solutions. Introducing real-time prices and more interruptible contracts could be considered a reduction in the quality of service. The result can be public and political resistance to the introduction of innovative electricity prices, even though this would be more efficient. Evolution of retail tariff structures is a slow process (Chapter 9).

There is a further reason, however, for reliability regulation. This is the risk of a large-scale blackout or the collapse of an entire network. Network collapse also implies that there is no longer a functioning market: if electricity cannot be delivered, no transaction takes place and there therefore cannot be a market-based electricity price (Joskow and Tirole, 2007). Such a market failure would justify continued regulatory intervention.

For the time being, electricity reliability is likely to remain an issue for regulators rather than for markets alone. There will continue to be periods when generation capacity and demand a response resources will be insufficient to clear the market and determine the price. In the absence of market clearing prices, markets will not be able to determine the optimal installed capacity and regulators will have to intervene to regulate reliability.

Does reliability regulation always result in excess capacity?

In practice, many electricity systems enjoy higher capacity than is needed to meet the strict application of their reliability standards, with several reasons to explain why reliability in OECD countries is so high. First, system operators and policy makers tend to be conservative and prefer to be on the safe side by having high capacity margins. The CEO of a system operator can be ousted if there is a shortage of capacity, but few people will notice if there is excess capacity.

Second, the perceived risk of major blackout in case of lack of capacity means that system operators prefer to have comfortable generation capacity. Even if system operators have shown that they know how to prevent major blackouts caused by lack of generation capacity, they face considerable risk aversion *vis-à-vis* such large-impact low-probability events.

The third reason why actual reliability exceeds the standard is that forecasts tend to overestimate demand, while the pace of deployment of new capacity can be faster than anticipated, resulting in the installation of excess capacity. Electricity demand growth plays an important role in adequacy forecasting, given the lead-time needed to build new power plants (from two years for an open cycle gas turbine [OCGT], to eight to ten years for a nuclear plant). Faced with uncertainty over electricity demand growth, conservative policy makers and system operators are likely to prefer to size the electricity system based on familiarly optimistic estimates.

Finally, and this is perhaps the most important reason, excess reliability is not very expensive. According to several quantitative analyses, "even a several percentage points increase in the target reserve margin would only slightly increase the average annual costs, but substantially reduce the likelihood of experiencing very high-cost events" (Brattle Group and Astrape consulting, 2013). According to International Energy Agency (IEA) calculations, even if the last megawatt hour costs

60 000 USD/MWh to generate, this increases the average cost of electricity by only a few cents. Compared with the amount of money spent on other electricity policy objectives, such as renewables support schemes, excess reliability is clearly a second-order issue.

It should not be surprising that governments do not let the market determine the level of electricity security of supply and tend to err on the safe side of reliability. Despite the interest in the theoretical question of adequacy, governments need practical and simple solutions to make sure that there is sufficient capacity.

Reliability with increasing shares of wind and solar generation

With the rapid deployment of wind and solar power, many gas and coal power plants are running fewer hours and several generators are losing money. In Spain, for instance, the capacity utilisation factor of conventional capacity declined to less than 15% for gas and less than 40% for coal in 2014. This trend has led the industry to argue that an energy-only market with renewables cannot provide the incentives to invest in new conventional generation capacity and therefore cannot ensure the reliability of the power system.

The first reason why variable renewables raise concerns for security of supply is due to the variable nature of wind and solar power (known as variable renewable energy [VRE]). As wind and solar power outputs are variable and their capacity is not available around the clock, their contribution to meeting peak demand is limited. This is particularly the case for solar power in Europe, where the electricity system peaks on winter evenings when domestic lights and electric heating are switched on.

As a result, investment in VRE makes little contribution to ensuring reliability. At low deployment levels, system operators usually calculate the capacity credit of renewables by looking at the expected wind and solar production during peak demand. The contribution of wind capacity to peak demand is usually in the range of 8% to 12% of installed wind capacity.

Box 4.2 • Effective load-carrying capability (ELCC)

The objective of ELCC is to calculate a capacity value which corresponds to the contribution that a given generator makes to overall system adequacy. While the notion of ELCC is applicable to all types of resource, it is especially relevant in the case of VRE generation, where capacity credit is difficult to estimate using methods based on a plant's availability.

ELCC can be calculated either as an increase in load or the equivalent in generating capacity. It is generally based on LOLE, but other suitable reliability metrics – such as expected unserved energy (EUE) – can also be used. For example the calculation of ELCC based on LOLE for wind generation (NERC, 2011) is done in several steps:

- calibration of the power system without the wind plants, so that it meets the desired reliability target (e.g. 0.1 day/year)
- subtraction of wind production time series from the load time series
- load addition to the system until the reliability target is met again.

This additional load is the ELCC of the wind plants. In this way, the system operator has calculated that the contribution of wind turbines to overall reliability amounted to 8% of their nameplate capacity. The capacity credit of renewables decreases when more wind is added into the system.

This capacity credit metric, however, is not accurate at higher levels of wind and solar power deployment. Periods of system stress do not necessarily occur during peak demand conditions, but may happen during periods of relatively high load combined with low wind and solar output – potentially 1 000 to 2 000 hours in the year. Analysis of the contribution of renewables to reliability therefore calls for a more sophisticated stochastic approach, such as the effective load-carrying capability (ELCC) (see Box 4.2), which assesses the additional peak load that can be added to the system with renewables.

The second reason why variable renewables raise concerns for security of supply is that they increase the need to invest in mid-merit and peak power plants. Renewables reduce the capacity utilisation of conventional capacity and increase volatility of the energy generated and the price of this electricity, making investment in gas power plants a more risky value proposition for investors. There is little track record of market-based investment in mid-merit and peaking units (rather, recent market-based investment has mainly focused on gas and coal plants expected to run as baseload plants). This makes price formation during scarcity hours particularly important for future investment in order to recover the capacity costs.

The third reason why renewables can affect reliability is not related to their variable nature, but to the fact that their pace of deployment has been relatively uncertain. Different interest group pressures and different policy viewpoints create high uncertainty over the timing, the location and the nature of the renewable technologies that will be deployed. While there is little doubt that renewables play a role in explaining the poor economics of conventional power in Europe, the low load factor of existing gas power plants is also explained by a low carbon price, declining electricity demand and erroneous investment decisions. Meanwhile, investors face increasing regulatory uncertainty with regard to future market opportunities for investment in conventional power, the future mix of generation and future prices. These circumstances support the argument that it may become useful to co-ordinate investment decisions or at least create a safety net, for instance by means of a capacity market.

A further dimension of VRE sources is that they expose the electricity system to weather conditions. This problem is not new, as networks have always been exposed to damage by storms or flooding and weather-related variations in demand. However, many wind farms can automatically switch off at the same time when the wind speed is too high, putting stress on the system. Solar power can also drop abruptly in case of snow. The solar eclipse that took place in Europe on 20 March 2015 created an extreme ramping event and was a real-life test of power system flexibility in Germany. In addition, the efficiency of photovoltaics (PV) decreases at high temperatures. For example, in 2015 Spain experienced its hottest month on record and demand increased substantially, but PV production decreased because of the heat. This creates new challenges for system security.

High shares of wind and solar power change the way that system operators manage power systems. Deployment of these renewables can lead to increases in the frequency and scale of start-up, ramp-up and ramp-down of conventional generation, and can result in power plants operating at their minimum output more often. It can also require maintaining more operating reserves in order to tackle forecast errors. Most of these phenomena have been already discussed in Chapter 3.

Finally, VRE presents technical challenges related to maintaining local voltage and frequency levels within prescribed boundaries. In existing distribution systems, VRE generation may cause voltages to rise above permitted levels. This issue can be mitigated by enhancing voltage control capabilities to adapt to the operation of VRE, for example through additional controllers at solar PV inverters or transformers with online tap changers (IEA, 2014a). Frequency deviation can occur when the grid control system is not fast enough to compensate for short-term fluctuations. When system frequency drops, the control system (such as supervisory control and data acquisition – SCADA) needs to pick up the imbalance by increasing the dispatch. Curtailment to avoid stability problems during critical periods can be limited where VRE sources do not have the technical capability to provide fast frequency response.

Reliability regulation and market footprint

The mandate of institutions in charge of reliability is usually limited to country or state borders. This is clearly inadequate in light of market developments during recent over the past few decades; the institutional framework for reliability regulation needs to evolve regionally in parallel with the market footprint. Generation adequacy forecasting should occur at the relevant geographic level in order to better align the market and reliability regulation footprints.

The European Union's electricity market integration process has focused on the integration of markets without paying much attention to reliability regulations, other than by restating that security of supply falls under the subsidiarity principle (Directive 2005/89/EC on electricity security of supply). As a result, the regulatory framework in Europe is extremely fragmented. It should therefore not be surprising that many European countries are now introducing national capacity markets with little co-ordination. This pattern could increase the overall cost of ensuring reliability.

In the United States the regional co-ordination of reliability could also be further improved. NERC regions, which were defined almost 50 years ago, do not match the footprint of the regional transmission organisations (RTOs) created 10 years ago. Robust overall assessments involve a regional resource adequacy assessment over the footprint of several RTOs and balancing areas, irrespective of state or country borders.

Survey of reliability and its regulation

In the 1960s and 70s, during periods of high electricity demand growth, governments usually relied on vertically integrated, regulated monopolies to make adequate and timely investment. Relatively light regulation by ministries was sufficient, and only a few countries introduced specific legislation regarding security of supply and reliability standards.

As competitive and unbundled electricity systems developed, however, responsibility for security of supply had to be clarified. This section provides a survey of the reliability metrics used in IEA member countries and discusses the notion of optimal reliability and different approaches to setting reliability standards.

Reliability can take different meanings in different contexts. NERC's definition encompasses two dimensions of electricity security (system operation security and adequacy). The notion of reliability standards usually refers to a metric to characterise the risk of involuntary load curtailment.

Reliability standards

In the majority of IEA member countries, reliability standards are explicitly set by governments: of 30 electricity systems surveyed by the IEA, 22 had reliability standards. This analysis does not include Japan, Korea or New Zealand. The methods used to set these standards can be described as either deterministic or probabilistic.

The reliability standards applied in OECD countries to the bulk power system translate into expected annual outages for end-use customers that are below a few minutes per year on a system-wide basis (The Brattle Group and Astrape consulting, 2013). For instance, the reliability standard in Australia (0.002%) means that, on average, consumers might be cut off for ten minutes per year.

The most commonly employed deterministic metrics are planning reserve margins, which measure available capacity over and above the capacity needed to meet peak demand levels under normal weather conditions (NERC, 2013a). NERC independently assesses reliability and uses a default reserve margin target of 15% for predominately thermal systems and 10% for predominately hydro systems.[10] This means that a thermal system with a normal peak demand of 100 gigawatts (GW) should have of at least 115 GW of installed capacity.

In Europe, the European Network of Transmission System Operators for Electricity (ENTSO-E) also uses a deterministic definition of reserve margin. The clear advantage of this approach is its simplicity for policy makers, given the ease of understanding the concept of a deterministic margin. This deterministic metric, however, is not well suited to taking into account the capacity of VRE sources, because their contribution to peak demand depends on weather conditions.

The other common approach is to use probabilistic methods. Here the metric is the result of a stochastic model that predicts the likelihood that demand will be served (NERC, 2012). The most common probabilistic metrics are:

- Loss of load expectation (LOLE): the expected number of firm load shed events an electricity system expects in a given year (Astrape Consulting, 2013).
- Loss of load probability (LOLP): either the probability of firm load shed events, typically expressed as a percentage of total hours in a year (Astrape Consulting, 2013), or the probability that the load will exceed the available generation at a given time (NERC, 2013b).
- Loss of load hours (LOLH): the expected number of hours of firm load shed events a system expects in a given year.
- Expected unserved energy (EUE): the expected energy in MWh that is shed, taking into account the magnitude of the outage.

Deterministic and probabilistic methods are interrelated: a target planning reserve margin may be derived from a probabilistic study and can lead to the same outcome. From a regulatory perspective, it is also important to express reliability in a way that is easy to understand for policy makers. For this reason, regulation usually favours simple metrics.

Europe

Reliability regulation is not uniform across Europe. Certain European countries have no reliability standards; in others, reliability standards exist but are not binding. Table 4.1 summarises reliability standard provision in selected European countries.

Europe's various reliability standards tend to be probabilistic and expressed in terms of LOLH. Several countries, including Belgium, Great Britain, the Netherlands and France, have similar standards of three to four hours per year.

Despite the creation of the European Union Internal Energy Market and the progressive harmonisation of many market and technical rules across Europe, security of supply remains a national competence in the European Union.

Within European countries, transmission system operators (TSOs) are responsible for monitoring and reporting on generation adequacy (CEER, 2014). At a European level, ENTSO-E also publishes a European generation adequacy outlook (ENTSO-E, 2015), which assesses adequacy at three different levels – for individual ENTSO-E member countries, regional blocks and the whole ENTSO-E area. These results are not binding on member countries.

Certain European countries have taken the initiative to better define reliability as a regional group. The Pentalateral Energy Forum is working on a common methodology for assessing the security

[10] NERC does not have authority to set reliability standards for resource adequacy (e.g. reserve margin criteria) or to order the construction of resources or transmission.

of supply at regional level.[11] In addition, a political declaration for regional co-operation on security of electricity supply in the framework of the European internal market was signed by Germany, Denmark, Poland, the Czech Republic, Austria, France, Luxembourg, Belgium, the Netherlands and Sweden, as well as the neighbouring countries Switzerland and Norway. It sets political commitments to improve co-ordination of national energy policies, including those on security of supply.

Table 4.1 • **Reliability standards and metrics in Europe**

European countries	No RS	Deterministic RS	Probabilistic RS	Binding RS	Non-binding RS
Austria	✓				
Belgium			3 hr/yr		✓
Czech Republic	✓				
Estonia	✓				
Finland			✓		✓
France			3 hr/yr	✓	
Germany	✓				
Great Britain			3 hr/yr	✓	
Hungary			✓		✓
Ireland			8 hr/yr	✓	
Lithuania	✓				
Malta	✓				
The Netherlands			4 hr/yr		✓
Norway	✓				
Romania	✓				
Spain		✓		✓	
Sweden		✓		✓	

Notes: hr/yr = hours per year; RS = reliability standards.
Source: CEER (2014).

North America

Explicit reliability standards and criteria are typical in North America. Most markets use the deterministic reserve margin approach, but it is usually derived from, or is benchmarked against, a probabilistic criterion (such as the 1 in 10 standard). Some regions use an economic approach, for example setting reliability targets at a level which aims to minimise customers' costs. Other markets that are highly dependent on hydro generation also have an energy criterion in order to manage the occurrence of low water inflows. Table 4.2 shows a sample of North American regions. The standard of 1 in 10 years is widely used, but may be interpreted either as one event in ten years (0.1 LOLE) or one day in ten years (2.4 LOLH).

NERC provides co-ordination of reliability regulation across North America. NERC was set up after the 1965 blackout event as a not-for-profit international regulatory authority whose mission is to ensure the reliability of the bulk power system in North America. It has direct access to detailed plant-level data and annually assesses seasonal and long-term reliability according to different geographic areas.

[11] The Pentalateral Forum is the framework for regional co-operation in central Western Europe. It was created in 2005 by energy ministers from the Benelux countries, Austria, Germany and France (with Switzerland as a permanent observer) in order to promote collaboration on cross-border exchange of electricity (http://europa.eu/rapid/press-release_IP-15-5142_en.htm).

Table 4.2 • Reliability standards and metrics used in selected areas of North America

Regions	Target reserve margin	1-in-10 standard (0.1 LOLE/yr)	1-in-10 standard (2.4 LOLH/yr)	Economic assessment	Energy criterion	Other probabilistic criteria
MISO (Midcontinent Independent System Operator)	✓	✓				
PJM	✓	✓				
NYISO (New York ISO)	✓	✓				
ISO-NE (ISO-New England)	✓	✓				
SPP (Southwest Power Pool)	✓ *		✓			
Maritimes	20%	✓				
Québec	✓	✓			✓	
Saskatchewan	Based on EUE					EUE
Manitoba	12%				✓	
SERC/SoCo				✓		
SERC/Duke Energies Carolinas	✓			✓		
ERCOT (Electric Reliability Council of Texas)	✓	✓				
CAISO (California ISO)	15%, benchmarked with LOLE studies					

* 12% for steam-based RTO members and 9% for hydro based, benchmarked with LOLH studies.

Source: The Brattle Group and Astrape Consulting, 2013.

Australia

Australia's reliability regulation is overseen by the reliability committee of the Australian Energy Market Commission (AEMC). The reliability standard is defined in terms of the maximum EUE, or the maximum amount of electricity expected to be at risk of not being supplied to consumers, per year (Henderson, 2014; AEMC, 2014). Essentially, this reflects a trade-off between the value consumers place on supply reliability and the overall power system costs associated with achieving a certain reliability level. The EUE is measured in gigawatt hours (GWh) and is expressed as a percentage of the annual energy consumption for the associated region or regions. Currently, the reliability standard is set at 0.002%, which means that out of 100 000 MWh of demand, no more than 2 MWh of outages would be allowed.

Actual reliability

Bulk power system

Power system performance is usually measured in terms of continuity of supply. System security indicators focus on the frequency, duration and impact of interruptions. They provide a system-wide and outcome-oriented perspective on power system performance, and are generally easy to interpret and apply in a high-level context. These indicators are collected by system operators and regulators.

A discrepancy exists, however, between the regulation of reliability and actual reliability. The reliability standards discussed in the previous section (e.g. 3-4 hours per year and 1 in 10) apply to the bulk power system and are primarily used to determine the adequacy of generation capacity. In practice, few OECD countries have experienced significant adequacy issues during recent decades. On November 2006, seven European countries experienced a blackout resulting from the switching of a transmission line. Japan (in 2011) and Korea (in 2013) are two other recent examples, resulting from the closure of nuclear plants in the aftermath of the Fukushima accident.

Table 4.3 • Large-scale system blackouts involving several power system areas

Date	Region	Population affected (indicative)	Affected power system areas	Cause
1965, 9 November	US Northeast	30 million	5 (St Lawrence-Oswego, Upstate New York, New England, Maine)	Relay with faulty trips, setting off power line overloads
2003, 14 August	US Northeast, central Canada	50 million	5 (Ontario, MISO, PJM, NYISO, ISO-NE)	Plant outage, line failure led to a chain reaction
2003, 28 September	Italy	56 million	3 (France, Switzerland, Italy)	Failure of a transmission line in Switzerland; lack of communication
2006, 4 November	Western Europe	15 million	7 (France, Germany, the Netherlands, Belgium, Italy, Spain, Portugal) – the entire continental European system was affected	Human error in a substation

Source: IEA, 2013.

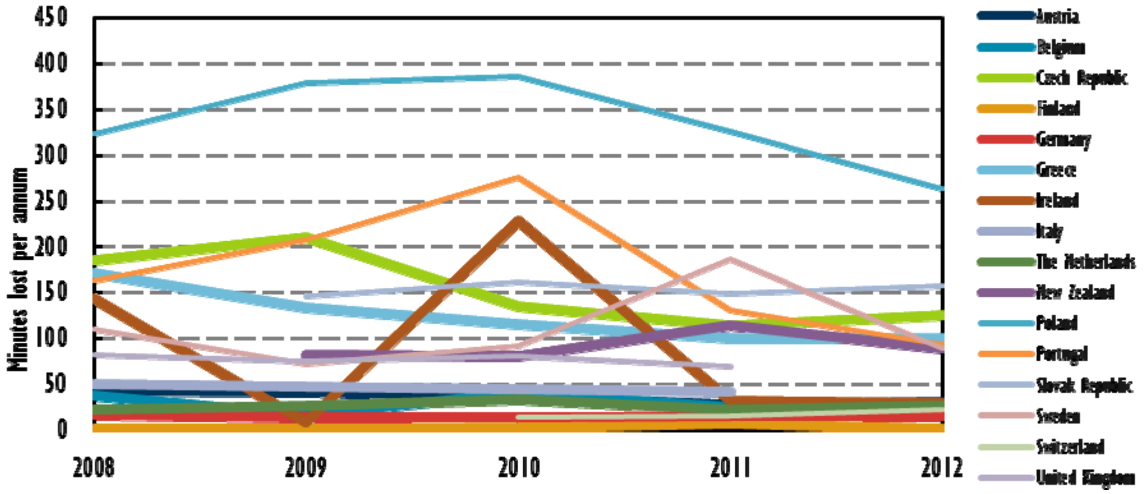
Lack of co-ordination among system operators is at the root of almost all recent major blackouts occurring in the systems of IEA member countries (Table 4.3). For instance, the Italian blackout in 2003 involved co-ordination problems between Italy and Switzerland. The Great Northeast blackout in 1965 led to the creation of NERC in 1968.

Distribution

In practice, the total duration of customer interruption reaches several hundred minutes per year on average, mainly due to distribution system outages. Interruption at distribution level is usually a local problem. Such interruptions are often the result of local weather events, such as storms or snow. They have a relatively limited effect and distribution companies are able to rapidly install emergency generators and repair power lines. For these reasons, such outages usually do not make the headlines of national newspapers.

In 2013, the IEA organised a survey to provide information on the frequency, duration and impact of interruption episodes. Figure 4.1 presents one commonly used indicator to measure trends in the duration of power system interruption. The duration of interruptions varied considerably, ranging from fewer than 20 minutes per year in Finland, Germany and Switzerland, to nearly 400 minutes during two years in Poland. Such large variations can reflect a combination of factors, including the impact of exceptional events, which tend to have a substantial effect on interruption duration indices, and differences in the nature of each power system. Power system differences include the size and topology of networks, the distribution of users, and differences in the data collection and calculation methodologies used to create indices.

Figure 4.1 • Duration of unplanned power interruptions (2008-12)



Notes: Austrian, German and Greek data exclude exceptional events; Belgian and Italian data exclude distribution; Belgian data are weighted by volume of consumption and number of consumers; complete data for Australia were not available; data for New Zealand and the Slovak Republic were not available for 2008; data for Switzerland were not available for 2008 and 2009; data for Italy were not available for 2012.

Source: CEER, 2012; Ofgem 2012; EnelDistribuzione, 2013.

Action to regulate the reliability of the distribution network is often referred to as regulation of quality rather than of reliability. This mainly requires new investment; for example, meshing the medium-voltage grid and building underground cables contribute to improving the resilience of the grid to extreme weather events. Such measures have a cost that regulators must consider when approving the investment plans of regulated distribution system operators.

4.2. Market design implications: Scarcity pricing

One of the primary objectives of market design is to ensure an adequate supply of resources to meet system needs over the long run. This section discusses the role that energy prices can play in ensuring reliability and adequacy.

From this perspective, the first approach consists of implementing efficient scarcity or shortage pricing. This approach is often referred to as the "energy-only market". Based on the experience of ERCOT in the United States and the National Electricity Market (NEM) in Australia, it is shown that, in order to work properly, scarcity pricing should be implemented with:

- high price caps, often above existing ones, consistent with reliability standards
- *ex ante* market power mitigation
- some form of regulation of scarcity price formation during system stress.

This section also discusses whether efficient scarcity pricing is sufficient to meet existing reliability standards, given the fact that price spikes remain rare and given the investment cycle experiences in the power sector. We conclude that, in addition to efficient scarcity pricing, capacity mechanisms may be necessary to create a safety net during the transition to low-carbon power system (Chapter 5).

Market failure and regulatory failure regarding scarcity pricing

Much has been written about whether or not energy-only markets can incentivise sufficient investment to maintain system reliability. In general, two energy market "flaws" are highlighted.

First, as energy markets do not allow for sufficient demand response (at least at present), the price cannot always clear the market. In particular, most consumers are not exposed to real-time electricity prices because there is neither the physical nor market infrastructure in place. It is possible that electricity markets will not clear even in the absence of wholesale market price caps. Given that in any particular moment the supply of generation is fixed, the only alternative to a catastrophic blackout for a system operator in such a situation is controlled load-shedding. In addition to that, in the event of a system outage or blackout, generators receive no remuneration at all, and therefore markets cannot optimise the risk of large-scale blackout (see Box 4.3).

Box 4.3 • Can markets optimise blackouts?

One important insight from the economic literature is that electricity markets cannot optimise the duration of blackouts and involuntary load curtailment. The reason is that the duration of blackouts depends on the generation capacity built to avoid them, and the incentive to build generation to avoid blackouts depends on the price being paid during blackouts. Yet there exists no competitive market price during blackouts; the price paid to generators during blackouts must be set by administrative rules.

The failure of markets to optimise blackouts goes beyond the case of rolling blackouts. For instance, when capacity becomes scarce, the probability of a network collapse increases (Joskow and Tirole, 2007; Joskow, 2008). But a network collapse implies a market collapse, because, as electricity cannot be delivered during a system collapse, consumers are unwilling to pay. As a result, market mechanisms cannot capture the cost of catastrophic blackouts and thus cannot optimise their occurrence.

The literature on peak-load and scarcity pricing, and investment incentives in electricity markets, begun with Boiteux (1949). Scarcity pricing relies on market clearing prices. The basic idea is that, where all available generation capacity is fully utilised, there may be excess demand at a spot price that is equal to the marginal production cost of the last unit provided by the physically available generating capacity. Because supply cannot meet demand in such a scarcity event, the demand side is then required to bid prices up until the market clears. At the resulting "scarcity prices", all generators that are supplying energy earn scarcity rents, which in turn are needed to cover their fixed capital costs. This mechanism is essential to providing an incentive to invest in all energy markets (Grimm and Zöttl, G, 2013). But it cannot help in optimising blackouts or in finding efficient prices when there is a possibility that no market-clearing price exists due to demand-side flaws. The adequacy problem is ultimately the result of demand-side market failures and not the result of regulatory price suppression (Cramton, Ockenfels and Stoft, 2013).

Second, energy market prices are capped and cannot incentivise investment in sufficient generation to avoid load shedding in the first place. Investing in generation that may only run for a few hours a year means that the investors must earn back all of their investment costs over a relatively short period of time. Because these generators are unable to recover their fixed costs via infra-marginal rents (as most generators do), during these few hours prices must be allowed to rise above their marginal costs. In other words, these generators must be allowed to exercise a certain degree of market power.

Such high prices have been considered, for the most part, politically untenable, and so, lacking a natural mechanism for keeping such market power abuse in check, regulators have often applied some cap on wholesale market prices. Limiting scarcity prices, however, potentially dis-incentivises investment in the peaking generation required for resource adequacy. It further leads to the so-called "missing money" problem, where resources are unable to recover their full investment costs through the wholesale market alone.

The existence of price caps has been the most popular explanation for the introduction of capacity markets. More recently, several studies have shed new light on this debate (Hogan, 2013; The Brattle Group, 2013; Cramton, Ockenfels and Stoft, 2013; FERC, 2014; RTE, 2014).

From a regulatory perspective, both an energy-only market and a capacity market involve a high degree of intervention from regulators or system operators; purely decentralised market solutions alone are unlikely to provide the accurate scarcity prices needed to meet reliability standards.

Increasing price cap to a value consistent with the reliability standard

To improve scarcity price formation during tight system conditions, regulators may need to intervene in an energy-only market to control market power and introduce administrative scarcity pricing curves.

In energy-only markets, generators make revenues only when they generate electricity. (Energy-only markets can also strictly speaking, include part of the revenues for operating reserves based on a capacity term.) Markets based only on energy revenues were introduced in the 1990s in North America and Europe, and are still working in eastern Australia, the US state of Texas and several European countries.

In such markets, power plants that are needed to ensure reliability might only run for a few hours per year. Compared to the average electricity wholesale price, usually in the range of 30-50 USD/MWh depending on fuel and carbon prices, in principle the spot prices needed to cover the costs of such plants have to jump up to 10 000 USD/MWh or higher. This means that prices have to rise well above the marginal cost of even the most expensive generators.

Not surprisingly, the energy-only market design has raised a number of concerns associated with price spikes:

- First, prices above the marginal cost of the marginal unit result from the exercise of market power by generators, which is in principle prohibited under competition law. In addition, extreme price events tend to make the headlines of newspapers, triggering political intervention.
- Second, price spikes have been less frequent and lower than expected – which raises the question of price formation during peak hours.
- Third, even if price formation is accurate, it is not clear to what extent investors can actually build new capacity based on revenues that depend on these spike prices.

Market power should be regulated ex ante

During periods when all the available generation capacity is needed to meet demand, every generator can enjoy market power and can bid a price above its marginal cost. Although this is efficient in theory, it raises practical difficulties.

Price spikes are possible because electricity demand is not price-responsive. Generators can offer their output at 1 000, 10 000 or even 100 000 USD/MWh, which can lead to high costs for consumers. For instance, when prices reach 10 000 USD/MWh, turning off an electric heater of 1 kW for one hour would save USD 10. However, except for large industrial users, consumers are not usually in a position to alter consumption because they are not exposed directly to real-time prices.

A further concern is that frequent peak prices could lead to extremely high profits for generators. Even if price spikes are needed in theory to ensure coverage of fixed costs for peak plants, the associated exercise of market power is usually prohibited by law. Market participants refrain from bidding prices above marginal costs in order to avoid being legally prosecuted *ex post*. Another concern is that extreme peak prices usually make headlines in newspapers, triggering policy makers to take measures to prevent them from happening again.

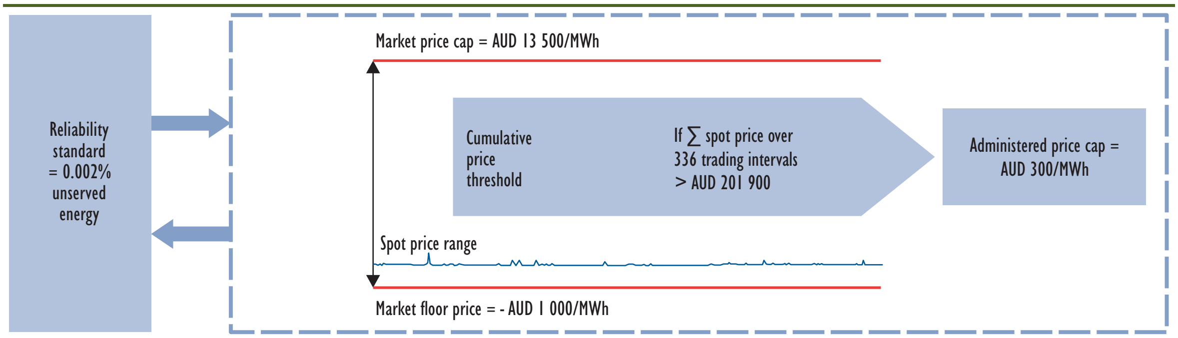
In order to mitigate the power of market actors, some regulators have introduced caps on prices or on bids. For example, setting caps at the marginal cost of the most expensive power plant would lead to a maximum price of around 300 USD/MWh. In this situation, a peak plant would see its variable costs reimbursed, but would never be able to cover its fixed costs.

In practice, most existing price caps have been set at around 2 000 to 3 000 USD/MWh. While this increases revenues, such price levels would result in revenues of USD 6 000 to USD 9 000 per MW for a plant running on average only three hours a year – too low to cover annual fixed investment costs in the range of USD 60 000 to USD 90 000 per MW per year.

To be consistent with existing reliability standards, the right level of price cap should be set at the VoLL. This solution has been adopted by the AEMC in Australia, which has calculated that a reliability standard of 0.002% EUE translates into a price cap of 13 500 AUD/MWh. In a situation of equilibrium in the market, this should in principle cover the fixed costs of optimal capacity, including peak plants.

Recognising that markets are never in equilibrium and that insufficient capacity could result in frequent peak prices and high profits for existing generators, regulators in Australia and Texas have set a limit on the revenues that generators can make during such periods. In Australia regulators have introduced a cumulative price threshold: if the sum of spot prices over 336 trading intervals exceeds AUD 201 900, the administered price cap is lowered to 300 AUD/MWh (Figure 4.4).

Figure 4.2 • Scarcity pricing in Australia (NEM)



Source: Henderson, 2014.

In Texas, the public utility decided in 2014 to progressively increase the price cap in the ERCOT region from 3 000 USD/MWh in 2011 to 9 000 USD/MWh from June 2015 (Potomac Economics, 2015). Along with other technical measures discussed later in this chapter, this increase in the price cap is expected to trigger new investments that would restore reserve margins.

In summary, price spikes and market power are inseparable and remain an issue in wholesale electricity markets.

Scarcity pricing might have to be regulated

System security is ultimately the responsibility of system operators. They have been and will remain in charge of avoiding blackouts. For this reason, all market designs hand control of the electricity system to its system operators, rather than relying on purely decentralised operations. And these system operators have developed operating protocols to prevent large-scale incidents.

System operators then take preventive action to avoid load shedding or large-scale blackouts. Such actions include activating emergency demand response contracts or interruptible contracts, or using part of the operating reserves to produce energy instead of shedding load. Ultimately, system operators can also rely on temporary voltage reduction, with the same effect as demand reduction.

From a market perspective, there is a risk that all these actions are undertaken "out of the market". Indeed, even where certain operating decisions are formalised, "grey areas" are likely to exist because tight system conditions remain rare and are often due to unexpected situations. In PJM, for example, emergency demand response capability activation can reduce demand by several gigawatts and suppress scarcity prices on the real-time market. Other out-of-market operating decisions, such as operating reserve depletion or voltage reduction, also have the effect of suppressing prices. All system operations that have the effect of lowering real-time market prices are therefore not monetised.

A further difficulty is that traders usually lack information about the exact state of the electricity system. For example, the highest recent peak prices in France 2 000 EUR/MWh occurred on 9 February 2012 between 10:00 and 11:00, while the actual demand peak occurred the day before at 19:00. Trading desks do not usually maintain the resources necessary to properly price rare events; this is particularly an issue in decentralised markets.

Another example of poor price formation can be found during the "polar vortex" in the United States in 2014. While energy prices reached only 800 USD/MWh for a few hours, PJM had to take out-of-market actions to ensure reliable operations. The resulting costs reached USD 438 million for the period 21-30 January, which could not be allocated to specific market participants and were assigned as "uplift costs" uniformly paid by all users. The polar vortex recalls the inherent difficulty in accurately pricing energy during tight system conditions, even in one of the most sophisticated RTO markets.

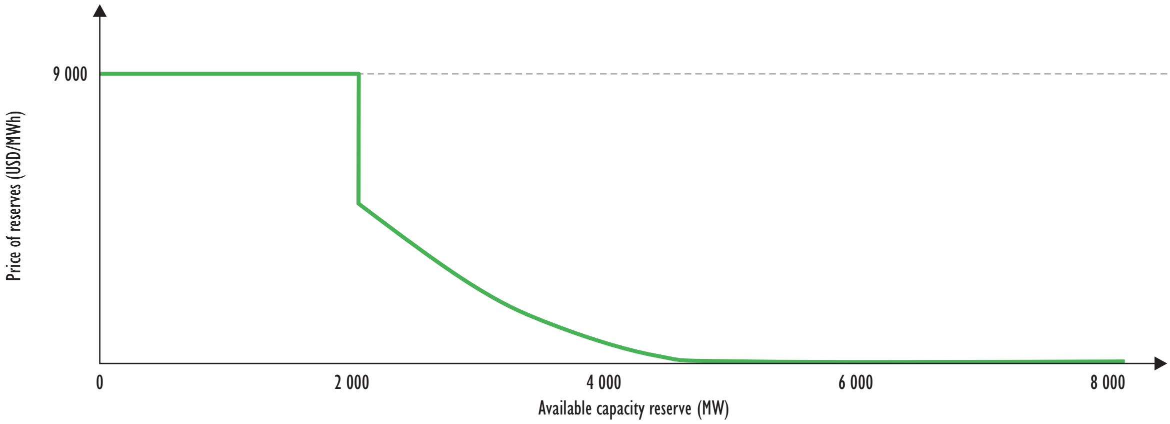
Following the polar vortex, the Federal Energy Regulatory Commission (FERC) initiated a series of technical meetings and issued technical papers to address these issues (FERC, 2014).

As FERC explains:

> "When the system operator is unable to meet system needs, it applies administrative pricing rules to ensure that costs, including the costs associated with the failure to meet minimum operating reserve requirements, are reflected in market prices. Ideally, these prices would reflect the valuation consumers place on avoiding an involuntary load curtailment. Under such conditions, prices should rise, inducing performance of existing supply resources and encouraging load to reduce consumption so that the system operator would not need to administratively curtail load to maintain reliability. A failure to properly reflect in market prices the value of reliability to consumers and operators' actions taken to ensure reliability can lead to inefficient prices in the energy and ancillary services markets leading to inefficient system utilization, and muted investment signals. Reducing such inefficiencies may lead to more reliable and more economic electric services."

Regulators can intervene to improve scarcity price formation. For example in Texas, the Public Utility Commission of Texas decided to introduce an operating reserve demand curve, starting on 1 June 2014, which is a form of administrative determination of prices during scarcity conditions (Figure 4.3) (see Hogan, 2013; Pfeifenberger, 2014).

Figure 4.3 • Proposed operating reserve demand curve in the ERCOT region



Sources: Potomac Economics, 2015.

The implementation of administrative scarcity pricing rules requires regulators to set prices. This is a form of regulatory intervention in the energy market, as opposed to pure market-based pricing. In markets where system operators calculate the real-time price of energy and operating reserves using co-optimisation techniques, high administrative scarcity prices for operating reserves translate into higher energy prices. High operating reserve prices find their way into real-time prices, day-ahead prices and forward prices.

Note that with such a scarcity price curve, scarcity revenues increase when there is a depletion of operating reserves, even without an actual load curtailment event. The number of hours with price spikes is therefore higher than the number of loss-of-load hours, thereby increasing scarcity rents for generators. Assuming that scarcity revenues were USD 30 000 per MW per year, and that revenues from operating reserve shortage were USD 20 000, then total revenues and would be USD 50 000 per MW per year. The administrative scarcity pricing curve can be defined in such a way that it increases revenues.

From a regulatory perspective, administrative scarcity pricing constitutes a form of *ex ante* regulation of prices. It is important to mention that regulators implementing administrative scarcity pricing commit to accepting high price spikes and such a regulation has to be stable over time. In turn, this should contribute to reassuring potential investors that policy makers or regulators will not intervene when such high prices materialise.

Note that administrative scarcity pricing can also be implemented in markets with a capacity market. Accurate price formation during tight system conditions should, in principle, increase the revenues that generators can attract on the energy market, and consequently reduce the bids of generators on the capacity market. All in all, better scarcity price formation leads to lower capacity prices and reduces the relative importance of capacity markets.

New Zealand has adopted an approach to scarcity pricing that reflects its unusual generation mix, dominated by hydro and constrained in energy due to limited storage and reliance on rainfall. The scarcity price is set at 10 000 NZD/MWh when emergency load shedding occurs. In addition, if the risk of non-supply exceeds 10%, the system operator can call for a public conservation campaign. If this occurs, retailers are required to compensate consumers for being asked to conserve energy, up to NZD 10.50 per week, a measure intended to be an incentive for retailers to contract for sufficient generation.

Whether the administrative pricing of energy alone is sufficient to trigger new investment in demand response and adequate generating capacity (which would drive down the price of capacity to zero where capacity markets exist) deserves additional consideration.

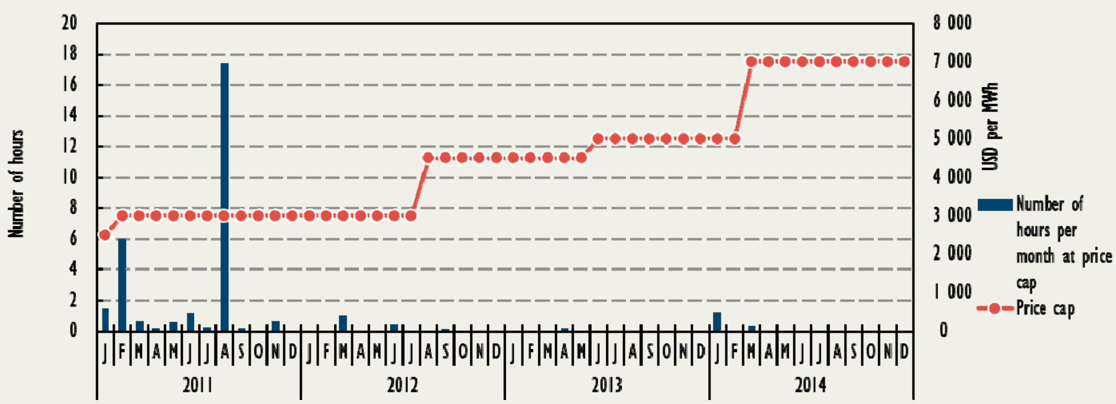
Box 4.4 • Scarcity pricing in the ERCOT region

ERCOT operates an energy-only wholesale market for electricity. That is, it does not include an explicit capacity remuneration mechanism (see Chapter 5). ERCOT also has a target (i.e. non-mandatory) reserve margin of 13.75% based on a 1-in-10 reliability standard. In ERCOT's case, 1-in-10 is defined as one load shedding event in 10 years, a relatively stringent standard. Demand response competes with generation by bidding directly into the day-ahead wholesale electricity market (Pfeifenberger, 2014).

In the absence of a centralised capacity market, generators must be able to earn sufficient revenue from the wholesale market to recover their fixed and variable costs. To this end, ERCOT has introduced the scarcity pricing mechanism (SPM), where wholesale prices are allowed to rise just enough to ensure a hypothetical peaking generation unit can earn sufficient revenues. Wholesale electricity offers in ERCOT's region are subject to a system-wide offer cap, which in effect functions as a cap on wholesale electricity prices, as any offer submitted above the cap is rejected. The offer cap has been raised on a roughly annual basis, from 3 000 USD/MWh in 2011 to 7 000 USD/MWh in 2014. On 1 June 2015, the cap was raised again to 9 000 USD/MWh (PUCT, 2012). This cap is automatically lowered when the calculated net margin for a peaking power plant reaches a total of USD 300 000 per MW per year. Figure 4.4 shows the number of hours where prices have reached the system-wide offer cap from January 2011 to December 2014.

Since 2011, relatively few shortage events have been experienced, and wholesale prices in general have been relatively low. In 2014, wholesale prices rose above 300 USD/MWh in only 34 hours. As recently as 2013, there was concern that the wholesale market was not providing sufficient revenues to incentivise a level of investment in the power sector to meet the target reserve margin, with ERCOT projecting reserve margins falling below the target by 2017. Recently, however, ERCOT has revised its demand forecast to reflect the fact that, contrary to projections, peak demand has been declining (Potomac Economics, 2015). The most recent forecast suggests that the reserve margin will remain above the target until at least 2020.

Figure 4.4 • Number of hours per month with prices at system-wide offer cap



Source: Potomac Economics, 2015.

Frequency of price spikes and incentives to invest

The low frequency of price spikes is a key concern for investors in conventional generation technologies. Peak prices are difficult to predict; even with accurate scarcity price formation, the 1-in-10 or 3 hours per year criteria imply that high prices only occur very rarely. In practice, high prices are unlikely to occur every year, but perhaps on average once every few years, depending on weather conditions. For example, France experienced cold weather for several weeks in 1956 and 1963, and then again in 1985 and 1987. France reached its peak consumption of 102 GW during a cold spell in 2012, which was less extreme and shorter. Based on a criterion of 3 hours per year, the LOLE is likely to be 30 hours every 10 years.

For investors, the low probability of revenues from price spikes is not an attractive investment proposition. Consider a merchant project undertaken by an independent power producer that needs to service its debt on a quarterly basis. Peak power plant enters into operation in a given year, but may only earn its first revenues after five to ten years. It is highly unlikely that such a plant can be financed purely on the basis of such unpredictable cash flow.

The issue of extreme weather events is not specific to the electricity sector. Insurance companies provide financial products to manage such risks, essentially transforming an infrequent cash flow (in case of accident) into a stable recurrent payment. The insurance and reinsurance sectors, in this way, spread risk across the economy and generally make it possible to buy a hedge against unpredictable events.

From this perspective, a peak power plant may be seen as an option to call up energy in case of system stress (Pöyry, 2015). If we assume accurate formation of scarcity prices, financial companies would offer hedging products that could create a more stable revenue stream for potential investors and thereby provide incentives to invest.

To date, however, this purely financial approach has not been developed. Several barriers stand in the way of developing such products, including a lack of information on which the finance industry can assess the probability distribution of electricity price spikes, and the risk of political intervention. In practice, the electricity industry itself has better information, and solutions might therefore have to be found within the industry, involving regulators and system operators.

This brings us back to the initial question of investment incentives in the context of low frequency of price spikes. In an energy-only market, even with accurate administrative scarcity pricing, investors still have to take investment decisions and this is what will determine the installed capacity. Given the financial risk, investors might require a high risk premium. For instance, using a cost of capital of 12% in real terms translates into a cost of new entry of USD 120 000/MW per year. Assuming a value of lost load of USD 20 000/MWh, this would correspond to an LOLE of around 6 hours per year, which is higher than most of reliability standards. In short, an energy-only market is unlikely to deliver the reliability standard set by regulators.

Meeting reliability standards during investment cycles

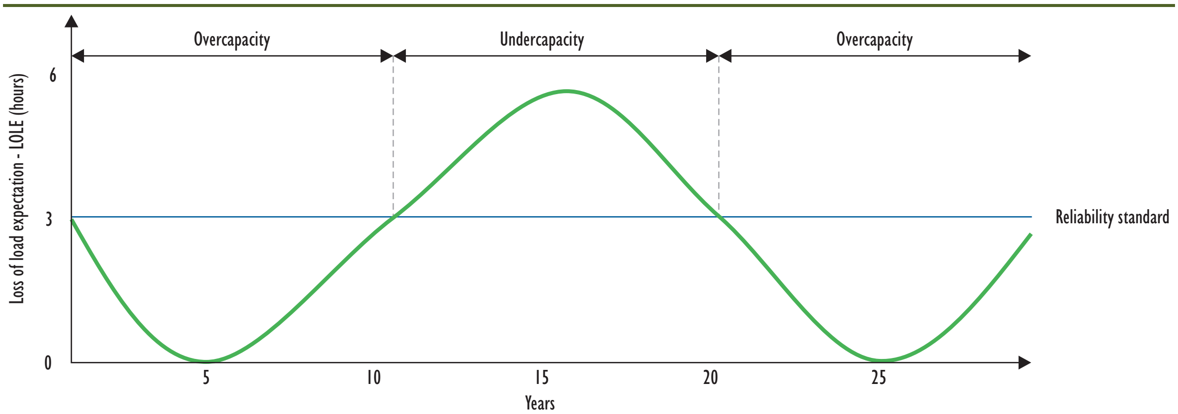
Although reliability regulation usually assumes that adequate capacity, in practice, means just enough to meet the standard, existing electricity systems are rarely in equilibrium and the dynamics of generation investment need to be managed (RTE, 2014). Many IEA member country markets have been in a situation of excess capacity for more than 20 years. This fact largely explains why actual reliability is higher than the standards.

Excess capacity can occur even in liberalised markets. In the United States, the dash for gas in the 2000s created a position of excess capacity that endures 15 years later. In Europe, investment created a wave of combined cycle gas turbine (CCGT) plants that came online in the 2010s, once again creating a situation of excess capacity after the economic crisis of 2008. With power plant lifetimes at 25 to 30 years, overinvestment can create overcapacity for long periods. This is especially the case in countries with sluggish or falling demand. In this situation, excess capacity is not reabsorbed by demand and can last for the technical lifetime of the asset.[12]

As a result of excess capacity, the frequency of price spikes is low and does not signal investment needs. When there is no concern about adequacy, governments have little reason to intervene to restore economic signals.

Conversely, when an adequacy assessment concludes that security of supply is at risk, governments are very likely to intervene in the market. Assume an LOLE close to zero during five years of excess capacity, followed by five years of LOLE of six hours per year; on average over ten years, the loss of load would not exceed three hours per year, meeting the reliability standard (Figure 4.7). But a government experiencing an LOLE of six hours for a period of five years is likely to intervene. In practice, governments tend to implement reliability standards as floors that need to be met every year, rather than average targets over long periods of time.

Figure 4.5 • Investment cycles and reliability standards



Accordingly, an energy-only market is unlikely to ensure that a reliability standard floor is met at all times. Scarcity prices and rents will be insufficient during periods of excess capacity. During periods of tight system conditions, governments are likely to intervene to ensure that the reliability floor is met. On average, the consequence of these interventions is that, even with accurate scarcity prices, their frequency is too low to ensure the revenues needed to cover fixed costs.

Conclusion

Market design has to take into account how governments regulate reliability. Although demand response has the potential to reduce or replace involuntary curtailment, for the time being and for the foreseeable future, most consumers are not exposed to real-time electricity prices because there is neither the physical nor market infrastructure in place. Consequently, reliability standards are implemented.

Scarcity price formation has to be administered and also needs to address market power issues with a revenue cap. Under these conditions, an energy-only market with accurate scarcity prices can ensure that the market provides capacity. But if governments wish to maintain a higher level of reliability or to make sure that reliability never falls below a floor, despite investment cycles in the power industry and

[12] In competitive electricity markets, excess capacity should, in principle, be reduced as utilities and merchant investors close down excess capacity. In practice however, investors do not immediately shut down excess capacity for several reasons, including governmental regulation prohibiting closure or because they hope to cover at least fixed operating costs, once the capital costs are sunk.

even if scarcity price formation is efficient, it may be necessary to introduce a capacity market in order to create a safety net.

Box 4.5 • Modelling the economics of resource adequacy with high shares of renewables

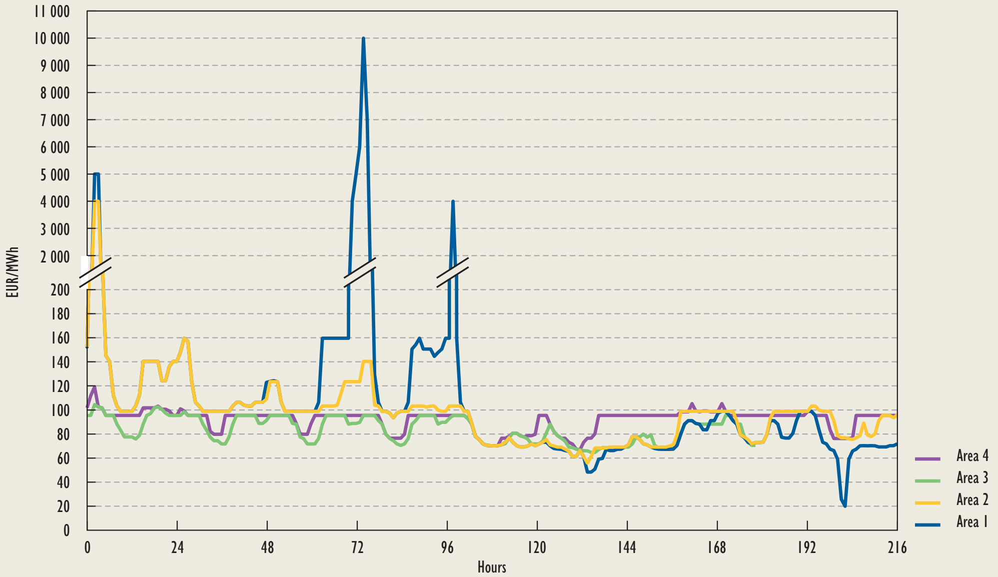
This box presents the key results of a modelling exercise undertaken to analyse the costs and benefits of different solutions to ensuring adequacy. The analysis looks at the costs of involuntary load curtailments and of actions to reduce these events using the approach developped by The Brattle Group and Astrape Consulting (2013). It reveals a trade-off between increased generation costs and the reduction of involuntary curtailment costs. The model shows how the design of energy markets can ensure a reliable electricity system. A detailed description of the model can be found at www.iea.org/media/topics/electricity/Repoweringmarkets/annexes.pdf.

The model assumes a high penetration of wind and solar power in a hypothetical interconnected electricity system. The modelled electricity system consists of four interconnected regions, each with a different capacity mix and with a certain degree of interconnection. The model calculates the least-cost dispatch of all the regions simultaneously, taking into account available network transfer capacity between regions.

System stress situations are manifested in peak electricity prices. To that end, the model assumes an administrative scarcity-pricing curve. Figure 4.2 presents the results for one week in February (for the 2007 weather year). On 18 February, areas 1 and 4 violate operating reserve requirements and use operating reserve capacity to generate electricity rather than shedding load, which, according to the modelling assumptions, pushes prices up to 5 000 EUR/MWh. On 21 February, load has to be curtailed in area 1, which pushes up the price to the VoLL. Import capacity is used to its maximum and this results in price increases in other areas. On 27 February, high renewables generation depresses prices to less than 20 EUR/MWh in area 1.

In this initial scenario, load has to be curtailed which means an adequacy issue arises. The exact reason does not matter for the analysis and could include demand growth being higher than expected, unexpected retirement of nuclear capacity or a lower-than-expected deployment of renewables. The modelled LOLE of 17 hours per year is far above a reasonable reliability standard.

Figure 4.6 • Modelled prices in four areas during one week in February



Box 4.5 • Modelling the economics of resource adequacy with high shares of renewables (continued)

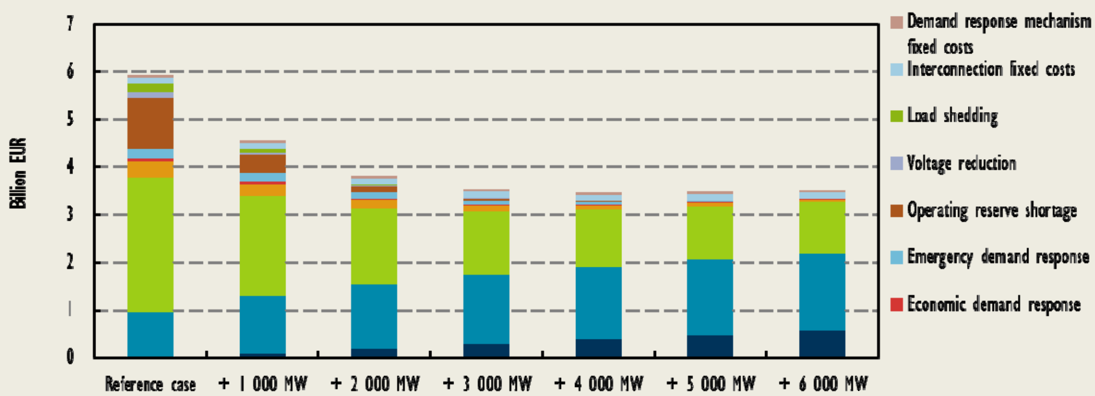
Moreover, the number of hours with tight system conditions (reserve margins lower than the target) in the electricity system exceeds 150 hours per year during most years – and reaches 500 hours in the extreme weather of year 2011. Consequently, market prices are expected to be very high and should provide an incentive to invest in new capacity.

Different scenarios are then simulated in order to restore an adequacy situation and meet the reliability standard: building new plants; increasing demand response; and creating interconnections.

Building more capacity is the first option to restore adequate capacity. From the perspective of area 1, adding gas-fired capacity (in a system with an installed capacity of 104 GW) would minimise the "reliability cost" (see Figure 4.7), defined as the sum of the investment costs of new marginal units, the production costs of marginal units, the cost of demand response, operating reserve depletion and voltage reduction, in addition to the VoLL in case of curtailment. In these calculations, it is found that the LOLE that minimises the reliability cost is very low (i.e. 11 minutes).

The model also considers demand response as an alternative to generation capacity. Potential demand response is assumed to be 5% of peak demand for industrial consumers, with a low cost of demand response of 9-12 EUR/kW per year. The demand response is assumed to be 8.5% of peak demand for residential and small business consumers. With these assumptions, it is always less costly to develop demand response rather than rely on involuntary load curtailment.

Figure 4.7 • Total reliability costs for different gas-fired capacity additions



Finally, new interconnections are another alternative to generation investments. The model suggests that new interconnectors offer an efficient solution to reduce the LOLE; where there is no interconnection, increasing interconnections by only a few percentage points of installed capacity considerably increases reliability; when interconnections already represent 5%, the marginal benefits of interconnections decrease. Nevertheless, interconnections are less expensive than new capacity.

The intuition behind this result is that reliability depends on the adequacy position on the other side of the interconnection. The contribution of capacity imports to generation adequacy in one country must take into account the availability of conventional units, wind and solar generators, and demand levels in the entire electricity system. As a result, interconnection capacities cannot usually be attributed their maximum capacity, but a lower level instead.

Box 4.5 • Modelling the economics of resource adequacy with high shares of renewables (continued)

Defining the optimal level of capacity that a system requires to ensure adequacy is a data-intensive task. The model presented here is purely illustrative, and is not intended to provide recommendations or guidance on the right level of capacity needed in a system with renewables. With the set of assumptions used, nevertheless, it indicates that increasing demand response and interconnections have the potential to be less costly than adding generation capacity to ensure reliability. If demand response costs are sufficiently low, involuntary curtailments could even be avoided altogether.

The model also enables quantification of the costs associated with overinvestment in reliability. Excess capacity, if limited to a few percentage points (2% or 3%) above the optimal reliability criteria, has a modest impact on the average bill. However, this can have important consequences for the functioning of electricity markets during tight system conditions, as discussed above.

References

AEMC (2014), *Fact sheet: the NEM reliability standard*, AEMC, www.aemc.gov.au/Media/docs/Fact-sheet---reliability-standard-d451eb88-58bf-4efd-8b18-f3bc38b24fbb-0.pdf.

Astrape Consulting (2013), *The Economic Ramifications of Resource Adequacy White Paper,* www.naruc.org/grants/Documents/Economics%20of%20Resource%20Adequacy%20WhitePaper_Astrape_Final.pdf.

Boiteux, M. (1949), « La tarification des demandes en pointe: Application de la théorie de la vente au cout marginal » [Pricing for advanced applications: Application of the theory of the sale at the marginal cost] *Revue Générale d'Electricité*, Vol. 58, pp. 321–40.

Brattle Group and Astrape Consulting (2013), *Resource Adequacy Requirements: Reliability and Economic Implications*, report prepared for FERC, www.ferc.gov/legal/staff-reports/2014/02-07-14-consultant-report.pdf.

CEER (Council of European Energy Regulators) (2014), *Assessment of electricity generation adequacy in European countries,* www.ceer.eu/portal/page/portal/EER_HOME/EER_PUBLICATIONS/CEER_PAPERS/Electricity/Tab3/C13-ESS-32-03_Generation%20Adequacy%20Assessment%20Elec_10-Dec-2013.pdf.

CEER (2012), *5th CEER Benchmarking Report on the Quality of Electricity Supply 2011*, Council of European Energy Regulators, April 2012, Brussels.

Cramton, P., A. Ockenfels and S. Stoft, (2013), "Capacity Market Fundamentals", *Economics of Energy and Environmental Policy* 09/2013; Vol. 2, No. 2, ftp://www.cramton.umd.edu/papers2010-2014/cramton-ockenfels-stoft-capacity-market-fundamentals.pdf.

Enel Distribuzione (2013), "Electric Network Operation Key Challenges", presentation to the *IEA 2013 Emergency Response Review of Italy*, 31 October 2013, Rome.

ENTSO-E (European Network of Transmission System Operators for Electricity) (2015), *2015 Scenario Outlook and Adequacy Forecast*, www.entsoe.eu/Documents/SDC%20documents/SOAF/150630_SOAF_2015_publication_wcover.pdf.

FERC (Federal Energy Regulatory Commission) (2014), *Price Formation in Organized Wholesale Electricity Markets, Staff Analysis of Shortage Pricing in RTO and ISO Markets*, FERC staff report, www.ferc.gov/legal/staff-reports/2014/AD14-14-pricing-rto-iso-markets.pdf.

Grimm, V. and G. Zöttl, (2013), "Investment incentives and electricity spot market competition", *Journal of Economics and Management Strategy 2013*, Vol. 22, No. 4, p. 832-851.

Henderson N. (2014) "Role of scarcity pricing in Australia", AEMC, presentation to the IEA Electricity Security Advisory Panel (ESAP) workshop on scarcity pricing, 2 July 2014, Paris, www.iea.org/media/workshops/2014/esapworskhopi/Neville_Henderson.pdf.

Hogan, W, (2013) "Electricity Scarcity Pricing through Operating Reserves", *Economics of Energy and Environmental Policy*, Vol. 2, Issue 2, pp. 65-86.

IEA (International Energy Agency) (2014a), *The Power of Transformation: Wind, Sun and the Economics of Flexible Power Systems*, OECD/IEA, Paris, http://dx.doi.org/10.1787/ 9789264208032-en.

IEA (2013), "Resilience of the energy sector to climate change", webpage, www.iea.org/topics/ climatechange/subtopics/resilience/.

Pfeifenberger, J.P. (2014), "Resource adequacy requirements, scarcity pricing, and electricity market design implications", IEA Electricity Security Advisory Panel (ESAP) presentation, 2 July 2014, www.iea.org/media/workshops/2014/esapworskhopi/Hannes_Pfeifenberger.pdf.

Joskow, P. (2008), "Capacity payments in imperfect electricity markets: Need and design, *Utilities Policy,* Vol. 16, pp. 159-170.

Joskow, P. and J. Tirole (2007), "Reliability and Competitive Electricity Markets" *Rand Journal of Economics*, Vol. 38, No. 1, pp. 60-84.

London Economics (2013), "The value of lost load (VoLL) for electricity in Great Britain", final report for OFGEM and DECC, www.gov.uk/government/uploads/system/uploads/ attachment_data/file/224028/value_lost_load_electricty_gb.pdf.

NERC (2013a), *2013 Long-Term Reliability Assessment,* www.nerc.com/pa/RAPA/ra/ Reliability%20Assessments%20DL/2013_LTRA_FINAL.pdf.

NERC (2013b), *NERC Probabilistic Assessment, Addendum to the 2012 Long-Term Reliability Assessment*, www.nerc.com/pa/RAPA/ra/Reliability%20Assessments%20DL/NERC_2012_ Probabilistic_Assessment_Final.pdf.

NERC (2012), *Reliability Assessment Guidebook,* NERC, www.nerc.com/pa/RAPA/rg/ ReliabilityGuidelines/Reliability_Assess_Guidebook_3_1Final.pdf

NERC (2011), *Methods to Model and Calculate Capacity Contributions of Variable Generation for Resource Adequacy Planning*, NERC, www.nerc.com/files/ivgtf1-2.pdf.

Ofgem (Office of Gas and Electricity Markets) (2012), *Electricity Distribution Annual Report for 2010-11*, Ref: 46/12, Ofgem, 30 March 2012, London.

Pfeifenberger P. J. (2014), *Resource Adequacy Requirements, Scarcity Pricing, and Electricity Market Design Implications*, IEA Electricity Security Advisory Panel (ESAP), Paris.

Potomac Economics (2015), *2014 ERCOT State of the Market Report*, Potomac Economics LTD, July 2015, www.potomaceconomics.com/uploads/ercot_documents/2014_ERCOT_State_of_the_ Market_Report.pdf.

Pöyry, (2015), *Decentralised reliability options*, Pöyry Management Consulting, www.poyry.co.uk/ sites/www.poyry.co.uk/files/decentralisedreliabilityoptionsfullreport_v300.pdf.

PUCT (Public Utility Commission of Texas) (2012), *Texas Administrative Code*, *Electric Substantive Rules*, Chapter 25, Subchapter S, 11 November 2012, www.puc.texas.gov/agency/ rulesnlaws/subrules/electric/25.505/25.505.pdf.

RTE (2014), "French capacity market – report accompanying the draft rules", RET, Paris, www.rte-france.com/sites/default/files/2014_04_09_french_capacity_market.pdf.

Chapter 5 • Designing capacity markets

HIGHLIGHTS

- Capacity mechanisms, or capacity markets, have been introduced in several power systems with the objective of ensuring reliability needs are met. They need to be carefully designed to prevent market distortions.
- Capacity mechanisms should not be considered a replacement for ensuring wholesale market price signals are right in the first place, but rather as a safety net to meet policy-driven reliability goals.
- Targeted volume-based capacity mechanisms, such as strategic reserves, are quick to implement and can address short-term electricity security issues or ensure a high level of reliability. But they do not ensure that the energy market delivers adequate investment in the long run.
- With that in mind, market-wide capacity mechanisms should be technology neutral, should include both supply- and demand-side resources, and should be forward looking. Sound penalties can ensure the availability of contracted capacity.
- In order to allow cross-border participation, clear and transparent rules for contracting of neighbouring generation and short-term cross-border flows are essential – in particular rules that reflect the reliability standards in the respective markets.

In response to concerns over medium- and long-term electricity security, several jurisdictions with liberalised electricity markets have implemented, or are considering implementing, some form of capacity mechanism. A capacity mechanism seeks to incentivise sufficient investment in, or to prevent the economic retirement of, capacity in order to ensure resource adequacy. Capacity mechanisms take many forms, from targeted reserve requirements that focus only on the marginal generation needed to maintain reliability, to market-wide mechanisms that involve all participants.

This chapter examines the design elements of both targeted and market-wide capacity mechanisms. While many jurisdictions around the world have implemented capacity mechanisms in one form or another, this chapter focuses mainly on experiences in the United States, where capacity markets have been in place for more than a decade, and more recent developments in Europe.

5.1. Capacity mechanisms are increasingly used

Energy-only market vs. capacity mechanism

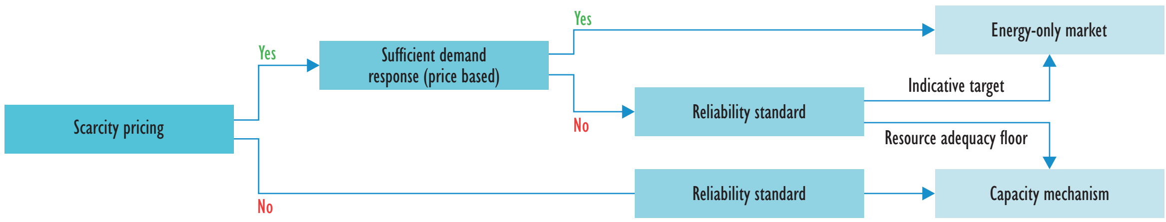
The main argument for the implementation of capacity markets is that energy-only markets are not able to incentivise sufficient investment in generation (and alternatives to generation, such as demand response) to ensure resource adequacy (see Chapter 4 for a longer discussion).

In a nutshell, capacity markets are needed if scarcity prices are capped at too low a level and if demand response is insufficient to meet the set reliability standard at all times. Figure 5.1 presents a simplified decision tree for policy makers: in the absence of scarcity pricing, or where price caps are too low, some kind of capacity mechanism will be necessary to ensure that generating resources are able to recover their fixed costs.

In theory, an energy-only market design with sufficient demand response can clear at all times. However, even with scarcity pricing in place and a certain level of demand response, a capacity market might be necessary. This is highly dependent on both the level and the nature of the reliability standard. If the reliability standard is just an indicative target and policy makers can accept high prices and lower reliability over limited periods of time (for example, a couple of years), then an energy-only market with scarcity pricing is likely to be sufficient. But if the standard is defined as a resource adequacy floor that must be met at all times, then a capacity mechanism will be necessary.

This section discusses this decision tree in further detail and draws implications for the design of capacity markets.

Figure 5.1 • Simplified decision tree, energy-only market versus capacity mechanism



Scarcity prices can remunerate capacity

When price caps are imposed on wholesale energy prices, it is often at a level too low to incentivise investment in sufficient generation to avoid load shedding. Generation that may only run for a few hours a year must earn back all of its investment costs over a relatively short period – which means that prices during these short periods must be allowed to rise above the generator's marginal cost, as the generator has no other opportunity to earn infra-marginal rents.

Allowing entirely unrestricted prices has been considered for the most part politically untenable, in part because there is the potential for these marginal generators to increase prices well above the level required to recover their investment costs. Lacking a natural mechanism for keeping such market power abuse in check, regulators have often applied some cap on wholesale market prices. It is difficult, however, to set the price cap at an appropriate level that accurately reflects the value of reliability to the consumer.

To understand why, assume for instance that the value of lost load (VoLL) is USD 20 000 per megawatt hour (USD/MWh). A typical price cap may be closer to 3 000 USD/MWh. At a reliability standard of three hours per year – that is, assuming three hours of scarcity prices per year – the average revenue for the marginal peaking plant would be 3 000 USD/MWh for 3 hours per year, totalling 9 000 USD/MWh per year. In that case, the "missing money" (the gap between revenues earned and the VoLL) is 51 000 USD/MWh (20 000 less 3 000, multiplied by 3).[1]

Recognising this situation, scarcity prices should be regulated *ex ante* in order to make price spikes politically tenable. Price formation should be improved by setting the price cap at a sufficient level, consistent with reliability standards. The price cap needs to be set at the VoLL or an operating reserve demand curve must be introduced. Scarcity pricing might have to be defined by an administrative price curve that sets prices on behalf of consumers in situations of capacity shortage (see Chapters 3 and 4).

[1] VoLL represents the price that an average customer would be willing to pay to avoid an involuntary interruption of electricity supply. The actual value may vary depending on the customer or the jurisdiction. For a more in-depth discussion of VoLL, see Box 4.1 in Chapter 4.

Some jurisdictions have chosen to use these or similar methods to avoid the need to introduce a capacity mechanism at all (see Box 4.1 in Chapter 4 on scarcity pricing). Moreover, it should be emphasised that the design and implementation of capacity mechanisms is, to say the least, controversial. Where they have been introduced, forward capacity mechanisms can lead to a higher quantity of capacity than that strictly necessary to meet the reliability standard – a reflection of the fact that the introduction of capacity mechanisms is driven by a general preference on the part of policy makers for higher levels of resource adequacy, so as to minimise the potential for capacity shortages.[2]

Improving scarcity pricing reduces the magnitude of the missing money problem, although it cannot eliminate other challenges that derive from relying on scarcity pricing, for example the uncertainty of cash flows. The following dimensions have to be considered before opting for scarcity pricing: demand response and reliability standards.

Demand response

Even in markets with scarcity pricing, sufficient price-based demand response is needed. To date, however, the challenge facing energy-only markets is an insufficiency of demand response during a scarcity event. It is therefore possible that there is no price high enough for the market to clear, even in markets with no price cap. Although demand response has future potential to reduce or possibly even replace involuntary curtailment to ensure that supply will meet demand, for the time being most consumers are not exposed to real-time electricity prices because there is neither the physical nor market infrastructure in place.

Reliability standards

This lack of demand response means that policy makers usually continue to define and set a reliability standard on behalf of consumers (see Chapter 4). Where the reliability standard is an indicative target, an administrative pricing curve can be designed to reach the reliability standard on average, with periods of higher reliability and lower reliability depending on investment needs.

Where the reliability standard is defined as a minimum, administrative scarcity pricing is not sufficient to meet a pre-defined reliability standard at all times. The power industry is subject to investment cycles. Before new investment decisions are taken, it is possible for a market to experience several years of lower reliability and higher scarcity prices. Under these circumstances, capacity mechanisms can ensure adequacy in a context of increasing uncertainty over demand, plant retirement and capacity additions during the transition to low-carbon power.

More generally, capacity markets might be needed to create a safety net, especially in times of system transformation and politically set goals, and where governments wish to maintain a higher level of reliability or to ensure that reliability never falls below a pre-determined floor.

State of play

Capacity mechanisms exist or are being introduced in one form or another in a number of liberalised markets around the world. Table 5.1 summarises the experiences in selected markets in the United States and the European Union.

[2] A number of alternative mechanisms for ensuring resource adequacy and remunerating capacity have been proposed, including decentralised reliability options (Pöyry, 2015) or the adoption of an operating reserve demand curve (Hogan, 2013). Addressing all possible capacity remuneration mechanisms and their various alternatives is beyond the scope of this publication, and it would be presumptuous to choose a single design as the ideal for all markets. This chapter will therefore limit its focus to the experiences of capacity markets as they have been implemented.

Table 5.1 • Capacity mechanism experiences in selected markets

Region/market	Capacity mechanism	Comments
United States		
PJM	Market-wide	Oldest and largest capacity mechanism in the United States
NYISO (New York ISO)	Market-wide	Notable for being a monthly spot market
ISO-NE (ISO-New England)	Market-wide	Uses a vertical demand curve
CAISO (California ISO)	Capacity auction	Currently considering alternative capacity mechanisms, with the aim of meeting reliability and flexibility needs
MISO (Midcontinent ISO)	Capacity auction	
ERCOT (Electric Reliability Council of Texas)	No explicit capacity mechanism	Generators earn additional revenues via adders to balancing and energy market; auctions for demand response have been organised
European Union		
Great Britain	Market-wide	Option not to participate in the capacity auction
France	Market-wide	Decentralised
Italy	Market-wide	
Germany	Targeted volume-based mechanism	
Sweden	Targeted volume-based mechanism	Strategic reserve
Spain	Targeted mechanism	Capacity payments
Belgium	Targeted volume-based mechanism	Strategic reserve

Note: ISO = independent system operator.

United States

The United States has a mixture of markets, with fully regulated, vertically integrated utilities, and markets that have been either partially or completely restructured. Much of the United States is organised into various regional transmission organisations (RTOs) and independent system operators (ISOs) – namely, ISO New England (ISO-NE), the New York ISO (NYISO), the PJM ISO (PJM), the Midcontinent ISO (MISO), the California ISO (CAISO), the Southwest Power Pool (SPP), and the Electric Reliability Council of Texas (ERCOT). Of these, ISO-NE, NYISO and PJM have fully functioning capacity markets; MISO has a limited capacity mechanism in the form of capacity auctions; CAISO places a capacity requirement on load-serving entities (LSEs) and has a standardised capacity procurement mechanism but, at present, has no formal capacity market; and ERCOT remains an entirely energy-only market.[3]

This chapter primarily describes capacity markets in two regions: PJM and NYISO.[4] As PJM's capacity market is the oldest and most mature, most of the theoretical discussion around capacity markets can be illustrated with this example. NYISO is presented because the design of its capacity mechanisms is somewhat unique, in that it is a near-term spot market, as opposed to the forward-looking mechanisms employed in PJM and ISO-NE.

[3] SPP is made up entirely of vertically integrated utilities, which meet their own reliability requirements. While it does have a formal reserve margin requirement, there is no enforcement mechanism in place to ensure that requirement is met.

[4] Examples will also be drawn from the capacity market in ISO-NE, when relevant to the discussion.

European Union

Capacity mechanisms have been implemented in various forms within individual European countries for some time. The United Kingdom, Italy, Ireland and Spain have used capacity payment schemes, while Sweden and Finland implemented a strategic reserve in 2003. Over time certain mechanisms have been retired, while others have remained in place (Süssenbacher, 2011).

In recent years the question of sufficient security of energy supply has re-emerged. Great Britain, France and Italy decided to implement market-wide capacity mechanisms, while Belgium opted for a targeted volume-based mechanism – the strategic reserve. Other countries – for example Ireland, Poland and Denmark – have yet to decide in which direction to move or have yet to implement their planned mechanism.

At the European level, broad guidelines for capacity mechanisms currently exist in the form of the guidelines on state aid (EC, 2014a). These are mainly related to ensuring a common playing field for technologies that can contribute to the security of energy supply and the consideration of capacity in neighbouring countries. They also contain pre-requirements for the introduction of a capacity mechanism in EU member states, including *inter alia* measures to foster energy efficiency and demand flexibility. The guidelines require a thorough assessment of the causes of the generation adequacy problem and a demonstration of the reasons why the market is not expected to deliver adequate capacity in the absence of intervention.

Configuration of capacity mechanisms

Capacity mechanisms can take different forms, but in each case the goal is the same: to ensure sufficient capacity to meet resource adequacy needs. The European Commission has identified two broad categories of capacity mechanism: 1) targeted mechanisms, or mechanisms that provide out-of-market remuneration to the resources needed to meet the reliability target; and 2) market-wide mechanisms, which remunerate all resources in the market.[5]

This chapter focuses on two specific types of capacity mechanism:

- a targeted volume-based reserve (category 1)
- a market-wide, volume-based, central buyer model (category 2).

To varying extents these mechanisms are already in place in certain jurisdictions in the United States and Europe, and are representative of the mechanisms that are likely to be implemented in other countries.

A targeted volume-based capacity mechanism is mostly used as an instrument to contract generation and demand response for use solely in scarcity situations. For this reason, it is mainly called strategic reserve or capacity reserve.

Under such a mechanism, generation that would most likely be decommissioned or mothballed in the near future is kept available in case of scarcity events. It is also possible, however, for new generation to be built under such a mechanism. Availability is usually only required in months with higher probability of scarcity events. A targeted volume-based mechanism does not provide additional revenue for generation or demand response but is the only source of revenue for the contracted generation, as participation in the energy market is not allowed in order to avoid distortions.

A market-wide capacity mechanism can be broadly defined as a regulatory instrument designed to create revenues for all capacity – whether it be in the form of generation, demand response or

[5] This categorisation omits methods for meeting reliability needs that do not provide an explicit capacity payment – for example, energy-only markets with scarcity pricing or options for reliability (which allow for the smoothing out over time of energy market revenues).

some other technology ‒ available during a specified period (generally when system operations are tight). These capacity mechanisms are meant to complement revenues from the sale of electricity in restructured electricity systems, in order to ensure that sufficient capacity is available to meet peak demand. This is especially the case for the recovery of fixed costs of peaking capacity, which is rarely used and generally faces capped wholesale electricity prices. Market-wide arrangements remunerate all capacity in order to avoid the negative bias that could appear with targeted mechanisms. No market participant is favoured by the mechanism; this leads, in theory, to an efficient – that is, least cost – portfolio of technologies.

One way to think of a market-wide capacity mechanism is as a tool for procuring sufficient *reliability options* to ensure resource adequacy. A reliability option is the right (but not the obligation) to call a resource into service in order to meet reliability needs. The resource receives a payment in exchange for agreeing to come into service as needed, according to a pre-established set of criteria.

Overall, there are three fundamental components to a capacity mechanism, which need to be considered regardless of the form the mechanism takes.

- First, the level of demand for capacity must be determined. Unlike a typical wholesale market, where total demand is the result of many individual decisions aggregated together, in capacity markets the level of demand must be determined administratively. This is because there is currently no mechanism through which load can collectively express its preference for reliability. For that reason the level of demand for capacity is usually defined as the reserve margin required to meet some specific reliability standard – for example, the 1-in-10 standard.
- As part of determining demand, the entity responsible for bearing the reliability costs must also be established. In the United States, for example, the reliability need itself is determined by the system operator, and the obligation to meet that need is placed on the LSE, which must therefore bear the cost of the capacity market payments. These costs are often passed on to consumers in the form of higher tariffs.
- Second, and related, the administrator must develop a mechanism for price discovery, ideally in the form of an auction. For market-wide mechanisms this leads to the development of a demand curve, which as noted above must be administratively determined, and which depends on the resource adequacy target and the expected level of compensation required to incentivise new entry into the market. For a targeted volume-based mechanism such as the strategic reserve, the administrator needs to determine which costs should to be factored into the bidding process.
- Third, there must be a defined capacity product. Capacity is essentially an option to deliver electricity, and therefore the amount of capacity a resource can provide may differ from the amount of electricity actually delivered under typical market conditions. A plant's capacity is equivalent to the amount of reserve margin that it can dependably meet. For example, a natural gas turbine may have a high availability, because it is reliably dispatchable, even if under actual operating conditions it only runs a fraction of the time. A wind turbine, on the other hand, may be seen as having low availability, even if it has a relatively high load factor, if only a small fraction of its production can be relied upon to provide power when needed (i.e. during peak load or during a scarcity event). The capacity product can also take into account the performance characteristics of the technology in question.

Regardless of design, the intent of a capacity mechanism is to act as a complement to wholesale markets in order to ensure resource adequacy. Crucially, capacity mechanisms should not be seen as a substitute for getting wholesale market design right in the first place. In particular, capacity mechanisms should not be implemented merely as a way to ensure generator profitability. In cases where there is an oversupply of capacity relative to demand (perhaps because of a legacy of overbuilding when markets were regulated, or because out-of-market mechanisms are incentivising too much investment in certain types of generation), wholesale prices may not be sufficient to allow all generators to recover their costs. It is tempting in such situations to see capacity markets as a way to allow otherwise unprofitable generators to remain in the market. For capacity markets to function properly, however, they must be focused on the singular goal of ensuring sufficient resource adequacy in order to meet reliability goals.

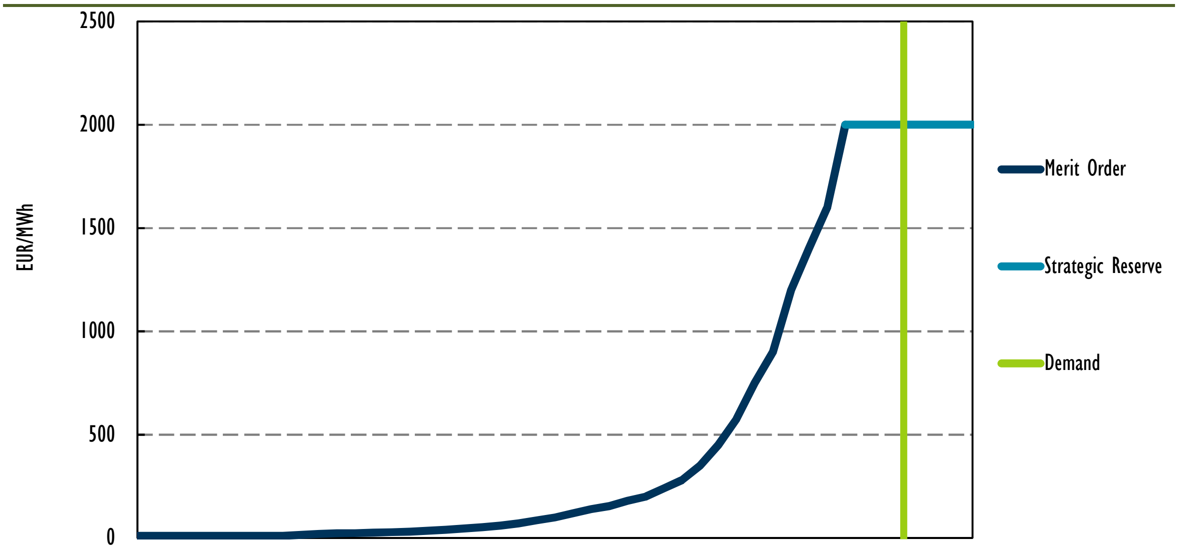
As the nature of power markets changes – in particular, with the introduction of large volumes of variable renewable generation – the definition of "reliability" may change. Significant penetrations of zero-marginal cost variable renewable power can exacerbate the missing money problem by lowering wholesale prices, and can also increase the need for flexibility services – both on the supply and the demand side. Certain jurisdictions are looking to implement capacity mechanisms with the explicit purpose of ensuring system reliability as the penetration of variable renewable power increases.

5.2. Targeted volume-based capacity mechanisms

General principle of strategic reserves

As a targeted volume-based capacity mechanism, a strategic reserve is a useful instrument to ensure short-term security of supply by contracting mainly old generation that would otherwise leave the market. It is quick to implement and has low implementation and transaction costs. Nevertheless, it does not reduce the long-term risk of new generation investments, which derives equally from uncertainty as to future energy policy and the electricity market itself. Therefore, over time, some countries that have implemented a strategic reserve might be forced to increase the amount of generation in the strategic reserve, or even invest in new generation if investors continue to see market risks as too high and demand response is not sufficiently developed (Cramton, Ockenfels and Stoft, 2013). This could lead to a further reduction in the number of participants in the energy-only market and an increase in long-term contracts for new generation in the strategic reserve. Therefore, a strategic reserve should be seen as a suitable mechanism for addressing short-term reliability needs. Other capacity mechanisms may be more suitable to ensure long-term electricity security, if the aim is to avoid falling short of a minimum level of security of supply.

Figure 5.2 • Wholesale market supply curve with strategic reserve (illustrative)



Strategic reserves are essentially generating units, storage or demand response that are kept exclusively available for use during a reliability event – that is, during times when the market is not able to provide sufficient supply to meet demand (Figure 5.2). Strategic reserve units are called into service by an independent body, such as the transmission system operator (TSO). The specification of the amount and type of capacity and demand resource is usually based on a specific analysis of the system's reliability needs. While the amount of contracted capacity is usually rather limited, it does, however, amount to a meaningful intervention in the wholesale electricity market.

As the strategic reserve is intended to operate only when the market does not provide sufficient capacity, reserved units should be dispatched at a price above a reference level signalling scarcity. These units can be utilised either in the day-ahead, intraday or balancing markets. The rules that determine when exactly strategic reserves are to be dispatched also directly determine the impact on market prices.

The activation of the reserve is usually linked to a predetermined threshold price. This can, in effect, act as a market price cap on the market, shielding energy consumers from scarcity pricings (so long as the threshold price is below the VoLL). As a result, there is the potential for a reduced incentive for investors to build new generation. Therefore, effectively assessing the VoLL – often a difficult proposition – is an important factor in the design of a strategic reserve.

Capacity within a strategic reserve is usually procured through a tendering procedure for a specified quantity (in megawatts [MW]), for example on a year-to-year basis. Before launching such a tender, a thorough assessment is necessary to determine whether there is sufficient capacity available to have a competitive bidding process, and the auction must be designed accordingly so as to limit the potential for market power abuse.

In markets with tight capacity margins, the potential exists for gaming and market power abuse in the wholesale market. For example, generators may withhold supply in order to increase prices. The presence of a strategic reserve is unlikely to reduce the potential for abuse, as these resources are only dispatched as a measure of last resort. There remains, therefore, the potential for significant price spikes in the energy market, and measures should be put in place to reduce the abuse of market power in scarcity situations. Dispatching a strategic reserve at a lower price would reduce gaming opportunities but would lead to a much larger reserve.

The strategic reserve can consist of existing generation, or new generation built for the purpose of reserve capacity, and it may include demand response. The latter comprises users who are normally obliged to reduce electricity consumption sufficiently rapidly and to a specified level when called upon. Whether or not to allow demand response in the strategic reserve depends on the circumstances of the country in question. If there is an oversupply of old generation capacity available, it is unlikely that demand response would be the least-cost method of ensuring sufficient capacity. Including demand response may therefore be a pure policy choice. To ensure its participation, for example, Sweden implemented a requirement that an increasing share of the strategic reserve be derived from demand response.

The compensation schemes for the providers of strategic reserves are specified in the tendering documents and may vary from case to case. These schemes may involve direct payments, payments in the form of an option or mixed forms of payment. Strategic reserve contracts may also contain provisions for notice periods, duration of activation, etc. The more diverse these contracts are, the more complex the strategic reserve becomes, making it more difficult to assess whether the contracted capacity fully meets the reliability goals.

The costs of strategic reserve schemes are typically recovered through system charges included in the transmission tariff or balancing charges. Hence they are effectively passed on to consumers.

Delivering the desired level of reliability, at efficient cost, is a key consideration for any capacity mechanism. A strategic reserve is likely to deliver the desired level of security of supply by keeping additional capacity in case of shortages. However, employing a strategic reserve means accepting the risk that plants not selected for the reserve could instead choose to close down. This can lead to the "slippery slope" effect, whereby the reserve must grow larger and larger in order to ensure sufficient capacity in the system to meet reliability goals.

An energy-only market leads to plants being deployed according to their place in the merit order, i.e. their short-run marginal cost level. With the implementation of a strategic reserve, however, this mechanism is distorted as some plants are held outside of the market. As a result, it is possible that electricity may not be generated by the most cost-efficient plants available. However, any market inefficiency introduced by a strategic reserve is likely to be small as long as the reserve is small in volume and is only dispatched under exceptional circumstances.

Furthermore, where the reserve is no longer required, exit costs should be limited as long as the amount of capacity procured is small and the reserve itself does not take the form of long-term contracts.

Sweden

Sweden introduced a strategic reserve in 2003 due to a high percentage of its electricity being generated from hydro-power resources, which have highly variable, weather-dependent capacities (with weather affecting both reservoir levels and peak demand). Initially envisaged to last for only a year, it has since been extended until 2020.

The Swedish TSO, Svenska kraftnät (SvK), is legally responsible for ensuring that there is sufficient capacity available in case of scarcity. SvK runs a yearly tender for the upcoming winter (defined as 16 November to 15 March) the size of which is stated in the regulation concerning the strategic reserve. Both generation and demand response may participate. In the period 2011-13, Swedish law required the strategic reserve to be at most 1 750 MW. Current plans are for it to decrease to 750 MW between 2017 and 2020, and after that for it to be reduced to zero (Figure 5.3), though there are discussions about extending the reserve beyond 2020. The legal maximum for the most recent tender was 1 500 MW.

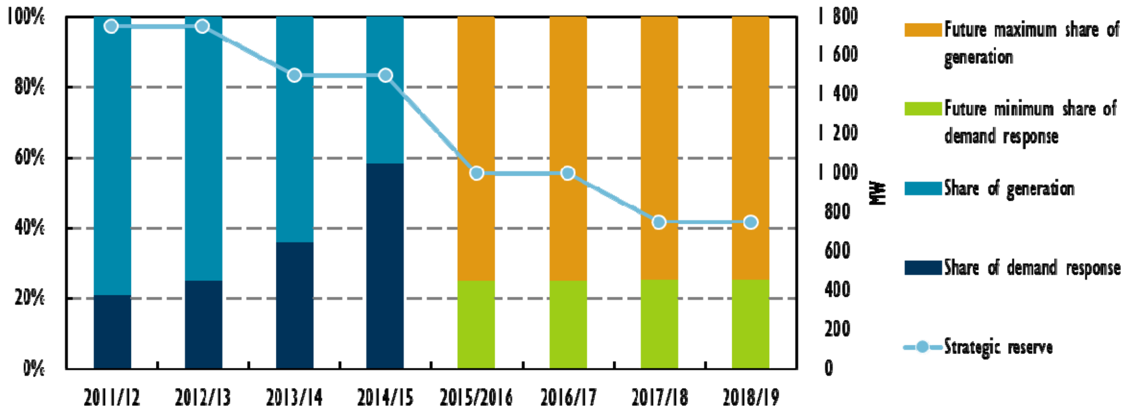
For 2014/15, SvK elected to procure 1 346 MW, a lower amount than the legal maximum. The strategic reserve amounts to 5.7% of Swedish peak demand of 26 gigawatts (GW).

The requirements for demand-side participation in the strategic reserve are that the participant must:

- be connected to the Swedish grid
- consume of at least 5 MW in a specific electricity area
- continuously offer one or more consumption bids of at least 5 MW on the balancing market, either as a balancing responsible party or through a balance responsible agent
- have an activation time of less than 30 minutes
- offer a reduction of at least 2 hours' duration
- be able to restart consumption units within 24 hours.

The reserve can consist of a group of different consumption units as long as the units are in the same bidding zone. The demand-side resources bid their administrative costs on a per-megawatt basis. Additionally, they make regulating bids in the balancing market for every hour the consumption reserve is available and contracted. The price of the bids is the variable cost plus a mark-up. The owners must report to the TSO on an ongoing basis the state of the resource and inform the TSO promptly if a resource becomes unavailable. The demand-side participant is allowed to bid the resource into the day-ahead market but is not paid by SvK for those occasions. This allows users to decrease consumption in scarcity situations and decrease the risk of curtailment. If the demand bid has not been activated in the spot market, the resource owner is required to bid the resource into the regulating market. In the regulating market the demand reduction is used when all other available resources have been used, and priced at the highest commercial bid, or the *ex ante* agreed bid, whichever is higher.

Figure 5.3 • Development of Swedish strategic reserve



Source: Svenska kraftnät (2015), (2012).

The share of demand-side resources in the reserve was initially, to be increased to 100% from 2017. This requirement was softened by amendments to the regulation, effective June 2014, but the overall amount and the schedule for phasing out the strategic reserve remains. This change increased the procurement options for SvK, but for 2014/15 it procured a split of almost 50:50 supply and demand resources. SvK appealed for the changes in the demand reduction on the grounds that it is important to keep a mix of highly responsive reserves as well as reserves able to tackle longer-lasting problems. The objective of using the strategic reserve as a tool to increase demand participation in the short-term energy market has not really been realised. Additionally, the long planning horizon and the requirement to be available at all times make demand reduction less suitable.

The electricity price of the strategic reserve is based on the highest commercial bid in Elspot, which is the day-ahead power auction on the Nordic power exchange, Nord Pool. If unused strategic reserve capacity remains after Elspot is closed, it will become available for the balancing market, where it will only be used after all market bids have been activated.

The principles for activating the strategic reserve are restrictive. The main criterion is that the buying bids exceed available sales bids. The available options are either curtailment of demand bids or use of the strategic reserve. The strategic reserve is then bid into the spot market with a price that is 0.1 EUR/MWh higher than the highest available commercial bid for increased selling or decreased buying of electricity.

The generation portion of the strategic reserve is outside the market as used today, because it cannot bid into the market. However, generation in the strategic reserve may be called into service below the market cap, and so may impact wholesale prices. SvK also notes that the strategic reserve should not address the need for investment in baseload power plants. The reserve has been used for spot market purposes eight hours between its introduction in 2003 and May 2014: 17 December 2009, hours 17 and 18 at approximately 1 400 EUR/MWh; 8 January 2010, hours 8, 9 and 10 at approximately 1 000 EUR/MWh; and 22 February 2010, hours 9, 10 and 11 at approximately 1 400 EUR/MWh (Elforsk, 2014).

Belgium

In 2014, Belgium introduced a strategic reserve into the Electricity Act, which concerns the organisation of the electricity market. The intent of the reserve was to ensure security of supply at a time of short-term problems with certain nuclear power plants, as well as the envisaged phase-out of nuclear by 2025.[6] A strategic reserve of 850 MW was contracted for winter 2014/15, comprising 750 MW of generation and 100 MW of demand-side response. For winter 2015/16, the Minister for Energy decided to increase the amount of capacity in the strategic reserve to 3 500 MW (Elia, 2015a).

Elia, Belgium's electricity transmission system operator, runs the strategic reserve tender process. Any aggregator, Elia grid user or access responsible parties (such as an electricity producer, major consumer, electricity supplier or trader) is authorised to participate in calls for tender for the strategic reserve. Generation capacity that successfully bids into the tender must be available for five winter months each year and with a notification time of 5.5 hours. Generation units in the Belgian control area that have already shut down can participate in the tender. Contracted generation units will be considered to be operating off-market for the share of capacity contracted by Elia.

Load, whether individual or aggregated, is allowed to participate in the tender as demand response. Contracted demand response receives a one-year-contract. Two different contracts were envisaged for demand response: a contract for a four-hour duration with a gap between activations of four hours, and a maximum of 40 activations per year; and one with a duration of 12 hours with a gap between activations of 12 hours, and a maximum of 20 activations per year.

Each year, the Federal Minister for Energy may instruct Elia to establish a strategic reserve, following the advice of the authorities (the Directorate-General for Energy) and a statistical analysis of security of supply conducted by Elia. In the decision, the Minister sets the required strategic reserve volume in megawatts, with a specific volume per year. The strategic reserve changes from year to year depending on requirements, following the same procedure. In other words, the Minister decides on the required volume, while the market sets the price of the strategic reserve by bidding into the tendering process.

The strategic reserve is activated once a risk of an energy shortage on the electricity market has been detected. If the results at the Belgian electricity exchange indicate a shortage in the total volume of energy on offer *vis-à-vis* the demand for energy, on day D-1 or in the intraday, the exchange launches a process for allocating additional energy from the strategic reserve. These exchanges of energy are made at the maximum price that applies on the Belpex DAM (currently EUR 3 000/MWh) (Elia, 2015b).

5.3. Market-wide capacity mechanisms

General principles (based on PJM's Reliability Pricing Model)

While there are many designs for market-wide capacity mechanisms, the PJM mechanism provides a wealth of experience and interesting lessons learned. This section uses the example of PJM to discuss the general principles needed for the design of market-wide capacity mechanisms.

The PJM Reliability Pricing Model (RPM) is the largest – and most complex – capacity mechanism in the United States. Given its complexity, it may not necessarily serve as a model for capacity market design in countries or regions seeking to serve a narrower purpose (for example, to meet

[6] Nuclear amounted to 35% of electricity production in 2013.

temporary reliability needs). However, the RPM is worth examining in detail for a number of reasons. First, it is the most evolved example of an initiative to incentivise investment in capacity using market forces. Second, it has been successful in meeting its fundamental goal of ensuring resource adequacy in a very large and complex region. Third, examining its composition offers insights that are relevant to all capacity mechanisms, regardless of their form.

Formed in 1999, PJM is the now the largest RTO in the United States. It includes 13 states plus the District of Columbia, and in June 2014 had a peak load of 141 673 MW. Installed capacity in 2014 was 183 724 MW.

The push to create a capacity market began soon after the formation of PJM, in response to market restructuring – in particular, the introduction of retail competition. New retailers were required to meet the same capacity obligation as the original LSE, which meant either entering into bilateral contracts or becoming a full-fledged utility and building generation. These newcomers found themselves at a disadvantage, as their load obligation was often smaller than a single generating unit. In addition, the annual capacity product was not flexible enough to allow an LSE to efficiently meet its capacity obligation given how quickly it could gain or lose customers.

Recognising that wholesale energy market revenues were insufficient to maintain reliability in the long run, and that a single, market-wide capacity price would fail to incentivise new capacity where it was needed most, in 2007 PJM introduced the RPM. Under the RPM, capacity was defined as an annual product and the obligation to meet the reserve requirement continued to fall on the LSEs. In addition, the RPM introduced a three-year forward market, a locational component to ensure prices would reflect system constraints, and *ex ante* rules to mitigate the potential for market power abuse.

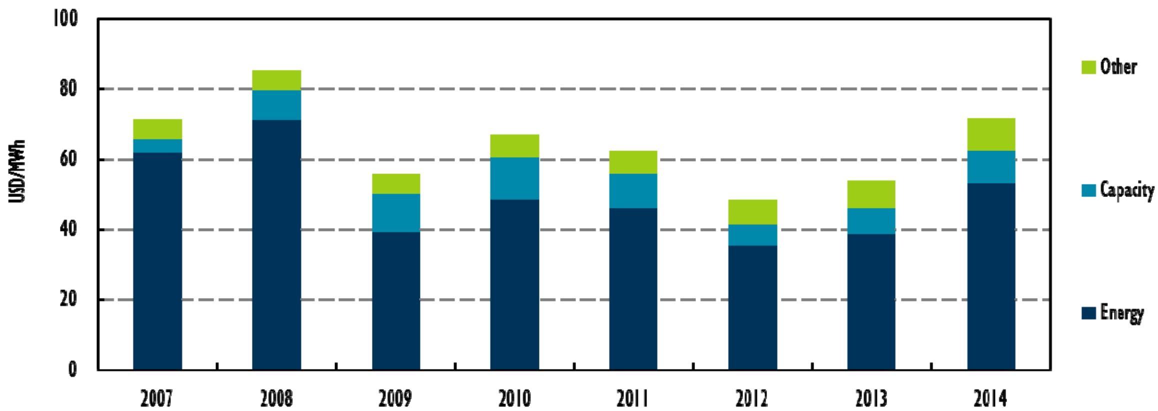
The capacity product

At the core of any capacity market is the definition of the capacity product. Defined most generally, a capacity product is a resource that is available to meet reliability needs. It is important to highlight, however, that meeting reliability needs is not the same as delivering energy. The capacity product in PJM is a physical product, and it therefore requires the actual delivery of energy when called upon. It is not possible, for example, to substitute the capacity obligation with a financial product.

A general capacity definition, however, does allow for a wide variety of resources to participate in the capacity market, including dispatchable generators, variable renewable generators, demand-side resources such as demand response or energy efficiency, and transmission investments. From a system operator's perspective, all that matters is that it contributes to the reliability requirement – that is, it can either produce electricity, reduce demand, or reduce overall reliability needs. For energy resources, this means that generators that participate in the capacity market must also offer all of their committed capacity into the energy market on a daily basis. This both gives the system operator an assurance that the resource is in fact available, if needed, and reinforces the fact that the capacity market is meant as a complement to, and not a replacement for, the wholesale energy market (Bowring, 2013). In PJM, capacity market costs make up a significant proportion of the wholesale price, although it is still a relatively small component compared to the energy cost (Figure 5.4).

Variable resources also contribute to system reliability, but as the system operator is unable to dispatch them on an as-needed basis, determining their actual capacity credit is more challenging. The first challenge comes from a general lack of experience with the performance of variable generation under actual operations. The second challenge is that the operational performance of a variable generator is highly location specific. Understanding how a wind turbine in a western part of PJM's area will perform under differing conditions says little about how that same wind turbine would perform in an eastern part. Therefore, performance metrics must take into account the location of the resource, in addition to the resource type. Location alone, however, is not sufficient, as the operator must also know whether the resource is in a transmission-constrained region, and therefore only able to serve power locally, or if it is in an unconstrained region, and is therefore able to contribute to the adequacy needs of the larger system.

Figure 5.4 • Components of PJM wholesale price



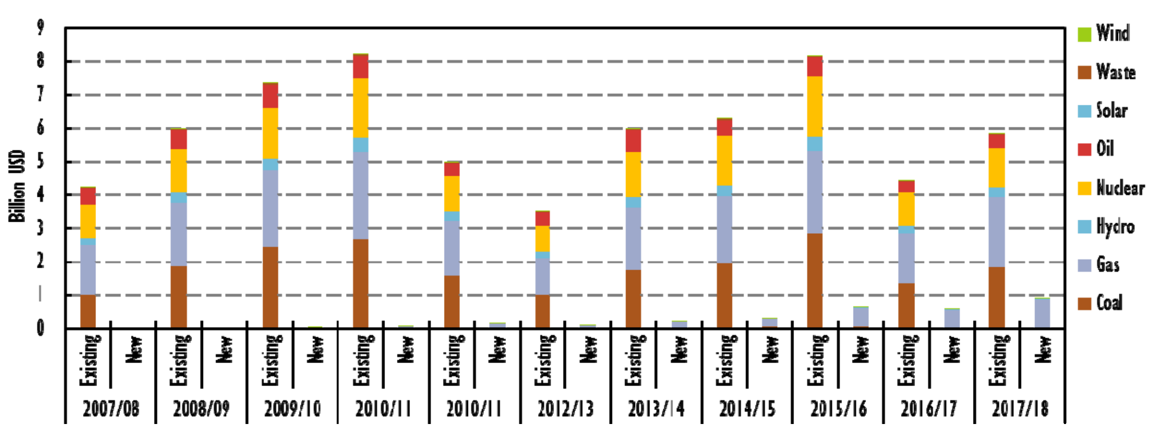
Source: Monitoring Analytics, 2015.

For new variable resources, PJM will apply a default derating factor. For a wind resource this is 13% of nameplate capacity, while for a solar photovoltaic (PV) installation the default rate is 38% (PJM, 2014a). Note that the default value for solar PV installations is higher than the typical annual load factor for a solar PV resource. This is because solar PV production and peak load are relatively coincident, and both are heavily influenced by the weather – peak load in the PJM area occurs in the summer, typically on a hot, sunny day, when a solar resource is more likely to be producing.

Once the renewable generator begins to deliver power, the default rate can be adjusted based on its actual performance. In PJM, this is based on the three-year rolling average of the resource's actual production, compared to actual peak demand.

One source of controversy over capacity markets is which resources, exactly, should be allowed to participate. In particular the question revolves around whether existing resources should be allowed to receive capacity revenues, or if only new resources should receive revenues.

Figure 5.5 • PJM capacity market revenues by resource type



Source: Monitoring Analytics, 2015.

As Figure 5.5 shows, in PJM existing resources receive the vast majority of capacity market revenues. While new resources in the market are growing, in the 2017/18 delivery year they will still receive less than USD 1 billion in revenues, compared to nearly USD 6 billion received by existing resources.

Some question why resources that are already earning sufficient revenues from the wholesale markets should receive an additional benefit in the form of a capacity payment. One answer is to recognise that the provision of capacity is a service, just as the provision of energy is. By participating in the capacity market, the resource in question is committing to be available for reliability needs, regardless of what happens in the wholesale market. It is therefore reasonable that these resources should also receive capacity revenues.

Another way to look at it is to ask what would happen without a capacity market and with no cap on wholesale market prices. In that case, during times of scarcity, the wholesale market price would rise well above the marginal cost of all generators in the market, and all generators would receive these same infra-marginal rents. Capacity markets replace or offset infra-marginal rents earned during scarcity events with a steady revenue stream. Just as there is no discrimination against specific resources that participate in the wholesale markets, there should be no discrimination in the capacity market.

The demand curve

No mechanism exists within PJM through which load can express an explicit demand for capacity. Therefore, the demand for capacity must be determined through an administrative process. This means determining both the quantity of capacity to be procured, and the price that should be paid for that capacity.

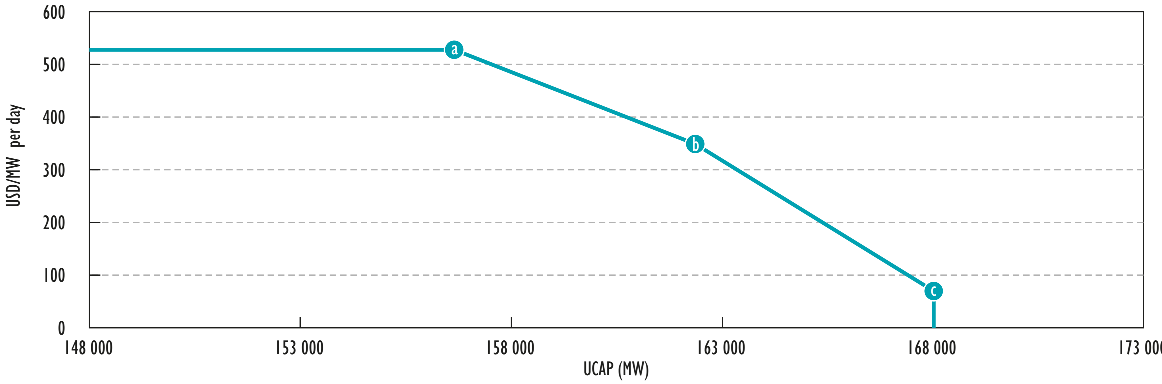
As the intent of a capacity market is to incentivise sufficient investment to ensure resource adequacy, the demand for capacity is taken to be the resource adequacy target – usually, peak load plus the installed reserve margin (IRM). In its simplest form, the demand curve is one where the demand is fixed at the resource adequacy target for all prices – that is, the demand curve is vertical.

The actual price paid for capacity is determined by the intersection of the supply of capacity and demand, which is the resource requirement curve (Figure 5.6). However, as demand is fixed (and, furthermore, known in advance), suppliers have the opportunity to withhold capacity and raise prices far above what would be required to recover their fixed costs. In addition, capacity markets suffer from a problem of monopsony – that is, there are multiple sellers, but only a single buyer. Therefore there is the potential for buyer-side market power abuse. Market power mitigation will be discussed in greater detail later in the chapter. From the perspective of the demand curve, the mitigating response to market power abuse has been to impose caps on capacity prices, and, in some cases, price floors. As the cap is meant to represent the maximum amount that customers would collectively pay for reliability, it is often set at the VoLL.

PJM's initial capacity market design used a vertical demand curve. However, it soon became apparent that vertical demand curves result in significant price volatility, reaching the price cap when capacity was too scarce, and then quickly reaching the price floor when new capacity was added. The reason for this is that actual capacity needs may often be quite small – of the order of a few tens of megawatts – while a new generator may be an order of magnitude larger. In such a situation, markets that are just shy of meeting their reserve margin target may find themselves in the awkward situation of paying high capacity prices, but being unable to incentivise new entry, as any new entrants would not be able to take advantage of the high capacity price for a sufficient duration.

The alternative to a vertical demand curve is a downward sloping demand curve,[7] according to which the quantity of capacity procured varies with the price. PJM uses a concave demand curve, named the variable resource requirement (VRR) curve, as shown in Figure 5.6.

Figure 5.6 • PJM variable resource requirement curve for 2017/18 delivery period



Note: UCAP = unforced capacity.
Source: PJM, 2014b.

The PJM curve is defined by three points:

- Point (a) is the intersection of the price cap, which in this case is either 150% of the net cost of new entry (CONE) of a combined cycle gas turbine (CCGT) or 100% of the unadjusted, gross CONE (whichever is larger), and the target installed reserve margin (IRM) minus 3%.[8]
- Point (b) is the target reserve margin plus 1%, with the price at 100% of net CONE.
- Point (c) is the cap on capacity IRM plus 5% and where the price is 20% of net CONE. The reserve margin is calculated based on a 1-in-10 loss of load expectation (LOLE) requirement.

Using a downward sloping curve reduces the potential for price volatility by allowing the market to clear somewhere between the price cap and the price floor even if the supply of capacity is below or above the desired resource adequacy target. Downward sloping curves also reduce the potential for market power abuse by limiting the impact on prices should a supplier withhold capacity.

As illustrated, however, it is clear that with downward sloping curves it is possible to procure either more or less capacity than is actually desired. Therefore, the slope of the demand curve must be carefully determined. In PJM, the demand curve is concave – that is steeper when capacity exceeds the reserve margin target than when it is below the target, indicating that excess capacity has relatively less value. This is to disincentivise overinvestment in generation. However, from a reliability perspective, while the optimal investment level is one that meets the reserve margin requirement exactly, between sub-optimal outcomes, overinvestment may be preferable to underinvestment. PJM addresses this by explicitly targeting an amount of capacity 1% above the optimal reserve margin requirement. Some markets, though, may prefer to implement a convex demand curve. Among other things, this will help to encourage investment when supply is short of the reliability requirements by a relatively small amount compared to the size of capacity it is economical to build, by reducing the volatility of capacity prices when supply exceeds demand.

[7] ISO-NE continues to use a vertical demand curve in its capacity market, with prices set by descending clock auction.

[8] To put these costs into perspective, gross CONE for PJM as a whole is estimated to be slightly more than USD 143 000 per MW-year (or USD 392 per MW-day), while net CONE is estimated to be approximately USD 121 000 per MW-year (or USD 332 per MWh-day). Both gross and net CONE can also be measured on a locational basis, to take into account differences both in costs and revenues.

The cost of new entry

A key component of the demand curve is the estimated CONE. This is because, regardless of the amount of capacity to be procured, the price of that capacity must be sufficient to incentivise new investment into the market. For PJM, the slope of the demand curve is determined by CONE. Therefore, before the demand curve can be established, the administrator must estimate the appropriate CONE for the market.

CONE is based on the estimated levelised cost of a reference technology for the delivery year – that is, the year for which the auction is procuring capacity. In PJM, this means a three-year forward basis.

CONE can either be calculated in gross terms (the total cost of new entry) or in net terms (the cost less revenues that a hypothetical plant would receive from the wholesale and ancillary services markets). Ideally, net CONE should be calculated based on expected (forecasted) revenues, and not historical revenues, so as to avoid the potential for temporary market conditions to bias net CONE analysis.

Box 5.1 • The reference technology for estimating CONE

Determining the reference technology is a key – and controversial – component of the CONE estimate. It is certainly market specific, needing to reflect what would likely be built in that market, under ideal financial circumstances, in the timeframe under consideration. Generally, however, the reference technology has been taken to be a gas-fired combustion turbine (CT) or a CCGT (FERC 2013a).

As the intent of a capacity market is to ensure sufficient capacity to meet peak-load requirements, a CT may seem the more obvious choice, as it is a peaking technology and therefore the one most likely to be deployed to meet peak requirements. However, in wholesale markets there is relatively less experience of building CT plants than there is of CCGT plants, and therefore relatively less cost data available for that technology. As a result, it is more difficult to estimate the actual going-forward cost of building a new CT in a given market, relative to estimating the cost of a newly built CCGT.

One alternative, suggested by the Brattle Group, is to estimate CONE based on the average costs of CCGT and CT (Pfeifenberger et al., 2014). This would reduce the influence of market fluctuations or estimation errors. A second alternative would be to choose a different technology altogether. In particular, an argument could be made in favour of using demand response, which can be thought of as a proxy for the consumer's willingness to curtail consumption. There is, however, a tremendous diversity of technologies that can qualify as demand response. Estimating a single CONE value to represent the entire range of potential new entrants (and, for PJM, doing so on a three-year forward basis) is extremely difficult.

Estimating CONE requires that the administrator make a certain set of assumptions, including the future price of electricity and fuel, operating and maintenance costs, and the cost of construction. CONE may also vary by location – for example, it is likely to be more costly to build a new CCGT plant in a densely populated area than in a rural area. For that reason, a particular market may need multiple separate CONEs, which ideally should align with specific locational capacity market auctions.

Regardless of the technology choice, it should be recognised that any decision will come with certain inherent biases. While over the long term, assuming well-functioning wholesale and capacity markets, investment decisions should lead to an optimal generating mix, in the short-term choosing a particular technology or technology mix can lead to over- or underinvestment, if CONE is calculated under incorrect assumptions – for example, by assuming that short-term fluctuations in the values of key parameters (say, fuel costs) are actually long-term.

As the underlying market conditions change over time, the assumptions behind CONE must also be updated on a regular basis. In PJM, CONE is reviewed on a triennial basis, although it is also updated annually based on an inflation index. The process for establishing CONE should be as open and transparent as possible, with a wide range of stakeholders involved. Nevertheless, even in an ideal process, the fact that CONE is based on a specific technology means that, in the long run, that technology (or technologies of similar or lower cost) may be the dominant new entrant (see Box 5.1).

In markets with downward-sloping demand curves, net CONE marks the point where the demand curve meets the reserve margin target. In the case where the market has too little capacity, the clearing price will be above net CONE, which in theory should incentivise new entry into the market. When there is too much capacity, the capacity price should be insufficient to incentivise new entry, and ideally low enough that some existing capacity will choose to exit the market.

Over the long term, capacity prices should converge with net CONE. Non-convergence may suggest that net CONE estimates are incorrect, or that some other factor is affecting supply (for example, additional policy or regulatory interventions).

Forward auctions and commitment periods

As capacity markets are developed, two time dimensions must be kept in mind. First, the goal is to procure sufficient capacity to meet projected reliability needs. Therefore, capacity markets must establish how far in advance capacity will be required to be online (the forward period). Second, capacity markets must provide a steady payment for some period of time in exchange for a generator's agreeing to remain available (the commitment period). Determining how long to provide a payment means striking a balancing between the risk borne by the investor (who faces more price uncertainty when capacity prices are only guaranteed for short periods of time) and the consumer (who bears the cost of long-term capacity payment obligations).

PJM's capacity market has a three-year forward period in order to allow new, unbuilt capacity to compete. New generation can commit to providing a certain level of capacity at a future date, and in return it receives a degree of price certainty. Forward auctions also provide value to existing assets, who attain greater certainty as to their own future revenues and can therefore decide whether to commit to staying online or to retire. From an investor perspective, therefore, longer forward periods may be more desirable. They are also beneficial to system planners, who need to understand long-term supply availability as they develop their transmission plans.

The challenge with long forward periods, however, is the general uncertainty that comes with any forecast. Not only must investors commit to building generation within a particular timeframe and to a particular specification, but planners must be sufficiently confident in their view of the reliability needs at the end of that forward period to determine the system's reliability requirement. Overestimation of reliability needs (for example, by incorrectly assuming higher load growth) can lead to overinvestment, which is not necessarily a problem in the long term, as overcapacity should lead to retirements. More worrying from a system operator's perspective is the possibility of underestimating reliability needs, which means underinvestment and therefore the potential need for additional investment in a timeframe where such investment is not possible.

The forward period in PJM is determined from the assumed construction time of a new CCGT. As with the reference technology for determining CONE, the choice of technology as the basis for the forward period brings with it an inherent bias. Investors in generation that requires longer construction times (for example, nuclear) may not see the capacity market as providing sufficient price certainty. Technologies that require relatively short construction periods – such as solar PV, or demand response – may also be disadvantaged by long forward periods, as they have to choose whether to come online well in advance of receiving a capacity payment or to delay the investment until closer to the actual commitment period.

One way in which PJM addresses this problem is by using reconfiguration auctions. These allow the system operator to revise the reliability requirement in advance of the commitment period to reflect updated market conditions, while also allowing investors to update their own bids to reflect any changes on their side. For example, a generator that is facing construction delays can choose to remove itself from the capacity auction in order to avoid the risk of paying a penalty for being unable to deliver capacity as promised.

The default commitment period in PJM is one year, although, in constrained areas (where the addition of a single generator can have a significant impact on capacity prices) participants can choose to fix the capacity price they receive for up to three years.

The length of the commitment period has significant implications on who bears which risks. Long commitment periods are conservative from a reliability perspective and have the potential to reduce the cost of the procured capacity, but they also place a long-term cost burden on the ratepayer. They do, however, place one risk on investors, namely that they must continue to be available regardless of market conditions. If the capacity price proves to be insufficient in the long term for a particular generator, it may need to choose to either stay online at a loss, or to retire and face a penalty for non-performance.

Short commitment periods, on the other hand, reduce the risk of overpayment on the part of bill payers, but may act as a disincentive for investment in more capital-intensive technologies. If investors do not see capacity markets as offering sufficient price certainty, they may choose to invest in whichever technology offers the most natural hedge against electricity price risk at the lowest upfront cost. Alternatively, the capacity market may not attract sufficient new investment to meet reliability needs, leading to higher capacity prices.

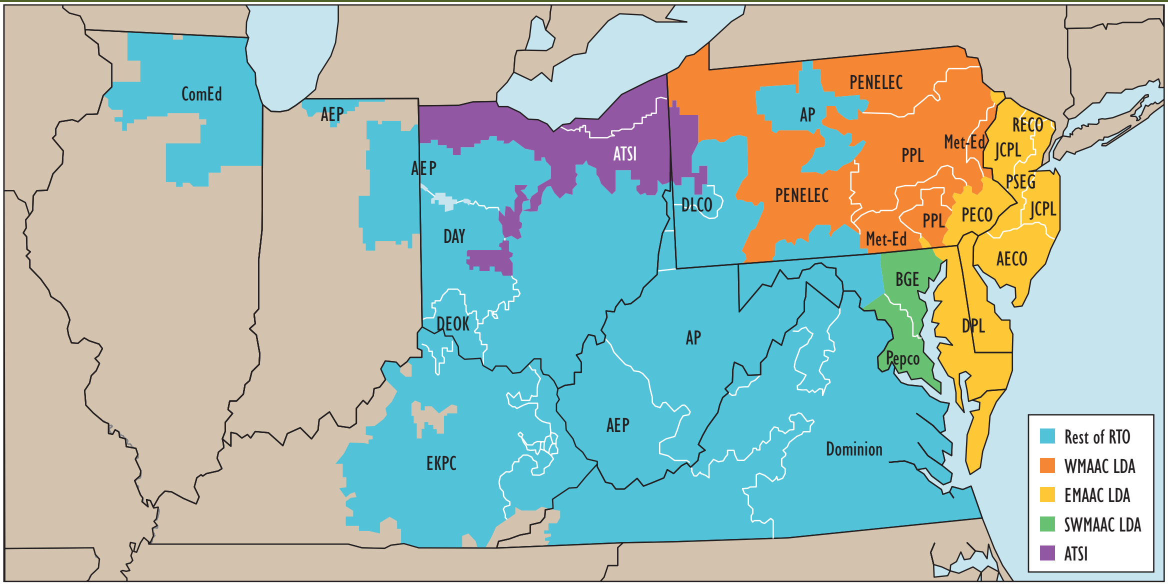
While there is no single answer to how long the commitment period should be, as long as other markets are functioning properly and regulatory uncertainty is limited, shorter periods should offer a more appropriate balance of risks than longer periods. This does not negate the need for long-term price signals. Rather, it is to say that capacity mechanisms (which focus exclusively on reliability needs) may not be the appropriate way to provide such long-term signals. While investors may prefer capacity payments that provide greater long-term price certainty, it should be recognised that long-term commitments shift risk away from the investor and onto the consumer. Some markets have sought to compromise by only offering the possibility of longer commitment periods to new entrants into the market.

Locational capacity prices

The existence of transmission constraints within the PJM system mean that while it is possible for PJM as a whole to meet its reliability goal, pockets of load within PJM are in fact underserved. To address this, the PJM capacity market includes a locational component, in the form of locational deliverability areas (LDAs). PJM has 27 LDAs in total, although this includes regional zones and sub-zones within regional zones. The capacity market process works the same at the LDA level as it does at the RTO level, although each LDA has its own reliability requirement to reflect local conditions.

The RTO capacity price sets the floor price for each LDA. If transmission constraints do not allow the lowest cost capacity to meet the locational reliability requirement, then higher cost capacity within the LDA will set the capacity price for that zone (or sub-zone), and the LDA capacity price will clear above the RTO capacity price. In practice, this has happened in nearly all of the PJM capacity auctions (Figure 5.7).

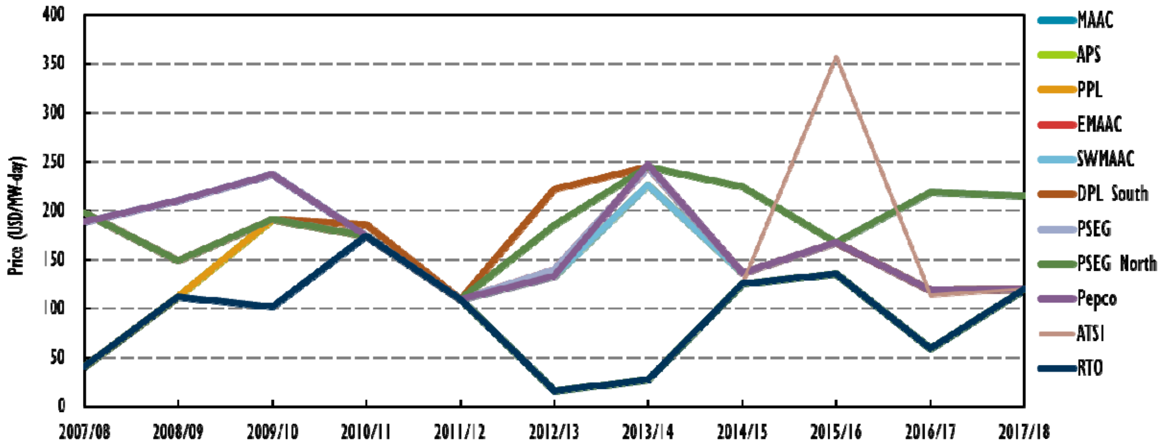
Map 5.1 • PJM locational delivery areas



This map is without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

Source: Monitoring Analytics, 2015.

Figure 5.7 • RPM clearing prices by LDA



Source: Monitoring Analytics, 2015.

Resources outside PJM

Starting with the 2017/18 delivery year, PJM will include external transmission import limits in its assessment of LDA resource needs, so that capacity resources outside PJM can potentially contribute to local reliability requirements. External capacity resources must be deliverable into the LDA in question, and so PJM has limited participation to five external source zones where sufficient transmission capacity has been demonstrated to be available.

PJM defines the capacity import limit for each external zone, which determines the maximum amount of capacity that can be imported into PJM. An individual capacity resource, however, can apply for an exemption to that limit, if it can demonstrate that it has long-term transmission rights into PJM, that it will follow the same must-offer requirements of resources within PJM, and that the capacity resource is "pseudo-tied" – i.e. that it is subject to the same re-dispatch and locational pricing rules as generation physically located within the PJM service area.

Performance requirements

Regardless of the resource type, it is vital that the system operator has as accurate a picture of each available resource as possible. If the assumed capacity credit for a given resource or a resource type is too small, the capacity market will procure too much capacity, whereas if it is too large, the system operator will operate under the false assumption that the reserve margin has been met.

From the perspective of the capacity resource, however, the incentive is to offer as much capacity as possible. Left unchecked, some resources may seek to influence their capacity credit by falsely representing their actual expected availability. Over time, false reports will be uncovered through actual operations, but system operators can and should do all they can to remove this incentive in the first place. For variable resources, the use of default capacity values removes the possibility of manipulation completely, although it does not remove the potential for under- or over-procurement of capacity. In addition, the system operator can apply a significant charge for non-performance. If a resource is called into service under its capacity obligation, and is unable to deliver energy or delivers less than its obligation, then the resource owner must pay a deficiency charge (or deficiency penalty). For forward markets, a penalty can also be applied if the resource does not come online within the required commitment period.

In PJM, capacity resources that fail to perform receive a penalty in the form of a lower capacity payment. In the first year of non-performance, the capacity payment is reduced by 50%. In the second year, the capacity payment is reduced to 25% of the capacity price, and in the third year the resource receives no payment at all. These non-performance penalties are not applied to hydroelectric power or to variable resources such wind and solar.

Performance requirements are of particular importance during an emergency. During the 2014 "Polar Vortex" in the United States, peak demand in PJM was 25% above what was typical for that time of year (Paulos, 2014). At the same time, 22% of PJM's capacity was out of service, due to mechanical failures and limited natural gas supplies. In response, PJM has created a new "capacity performance" product, where capacity that meets a certain performance requirement is paid a premium above the capacity price.[9] In effect, the capacity performance product creates two capacity categories: a "base" category for capacity that meets minimum capacity market requirements, and a "dependable" category for capacity that takes additional steps to ensure its availability. For example, a natural gas plant could respond by adding dual-fuel capabilities, or by entering into a firm fuel delivery contract, or a wind turbine could install some form of storage. Complementary resources can also be "coupled" together – for example, by combining an inflexible nuclear plant with demand response. Demand response resources themselves would also be affected, by putting more emphasis on their winter performance, instead of focusing mainly on their summer performance.

Market power mitigation

Capacity markets generally suffer from the potential for market power abuse, as the level of supply of capacity is commensurate with demand for capacity. The PJM market does have an issue of structural market power, having failed the three pivotal supplier (TPS) test in nearly all

[9] PJM was not the first to introduce performance payments into its capacity mechanisms. In May 2014, the Federal Energy Regulatory Commission (FERC) approved ISO-NE's Pay-for-Performance Initiative (PI). Under PI, capacity payments are divided into two portions. The capacity resource first earns a base payment as determined by the capacity auction. Then, during an actual scarcity event, capacity resources may earn an additional payment, or be charged a penalty, depending on how much capacity they are able to deliver compared to their obligation.

auctions, both at a system level at the LDA level.[10] Despite this, the PJM capacity market is considered competitive in terms of its actual performance (Monitoring Analytics, 2015).

A primary reason for PJM's competitive performance is the existence of *ex ante* market power mitigation rules. Under PJM's rules, any supplier that fails the TPS test must offer its capacity into the capacity market at its marginal cost. In this case, the marginal cost of capacity is defined as the resource's annual variable costs, net of any other PJM market revenues, plus any fixed costs required to keep the generation online in order to participate in the capacity market (Bowring, 2013). These costs can be measured on a resource-specific basis, or the resource owner can elect to use default costs determined by PJM for each technology (although marginal cost is still calculated using the specific unit's revenues from other markets).

PJM's capacity is tight both on the supply and the demand side, and so rules are in place to mitigate the potential for buyer-side market power as well. The Minimum Offer Price Rule (MOPR) states that new, natural gas-fired resources in the market must, in the first year they participate in the auction, offer their capacity at a minimum price. This is in order to prevent the submission of artificially low capacity bids that have the effect of lowering the capacity price.

The minimum offer price is determined as the net CONE of the generating resource type, as measured by PJM. Notably, this rule only applies to gas-fired generation – coal, nuclear and renewable generation are all excluded. These resources are excluded because, either due to lack of experience building these generation types, or because of significant cost variation within a particular technology category, it is relatively difficult to determine a standard reference price. In addition, nuclear is excluded because the long lead time required to build nuclear generation makes it less likely to be built with the intent of lowering capacity prices.[11] Similarly, it is unlikely that an investor would build a renewable generator with the intent of lowering capacity prices, as the capacity credit of such technologies is relatively low.

Demand-side resources

Typically, two types of demand-side resources participate in capacity auctions: demand response and energy efficiency. Demand response resources can be called upon by the system operator when needed (generally during a scarcity event) to reduce load, thereby reducing the amount of generating resources required to maintain system reliability. With demand response, total load is not necessarily reduced, as it can often come in the form of load shifting – for example, a manufacturer moving operations from peak to off-peak hours. Energy efficiency resources are permanent reductions in energy use, in particular during peak hours. Demand-side resources are included in all of the major capacity markets to varying degrees, but PJM has been the most successful – at least in quantity terms – with more than 12 GW clearing in the 2017/18 delivery year auction (11 GW of demand response and 1.4 GW of energy efficiency).

PJM separates demand response resources into three separate categories, depending on their actual availability: annual, extended summer, and limited. The existence of these categories is a reflection of the fact that demand response involves an active choice to reduce economic activity, and is therefore potentially not as available a resource as, for example, a gas-fired generator. Annual demand response resources are those that, as the name suggests, are available to the system operator year round. Extended summer demand response resources are only available during particular months – roughly, summer and the seasonal shoulder months. Finally, limited demand response resources are only available for particular days of the week – in the

[10] The TPS test measures whether the capacity market is able to clear without the contribution of the supplier being tested plus the two largest suppliers in the market.

[11] Answer of PJM Interconnection, L.L.C. to Complaint and Request for Clarification, Docket No. EL11-20-000. A similar point could be made for coal generation, which also tends to have long lead times and is a relatively capital-intensive technology.

case of PJM, weekdays except holidays. In each case, the demand response resources are also limited both in terms of the duration the resource is required to reduce load, and the times of the day during which it may be called.

This separation of demand response resources into different product categories marks a distinct division between these and supply-side resources. Generators that participate in the capacity market have historically not been distinguished according to their operating characteristics, but only by the amount of capacity they are able to provide to the system – that is, their capacity credit. The fact that demand response resources are distinguished based on the time in which they can be called into service points to a challenge in the way they are integrated into the capacity market.

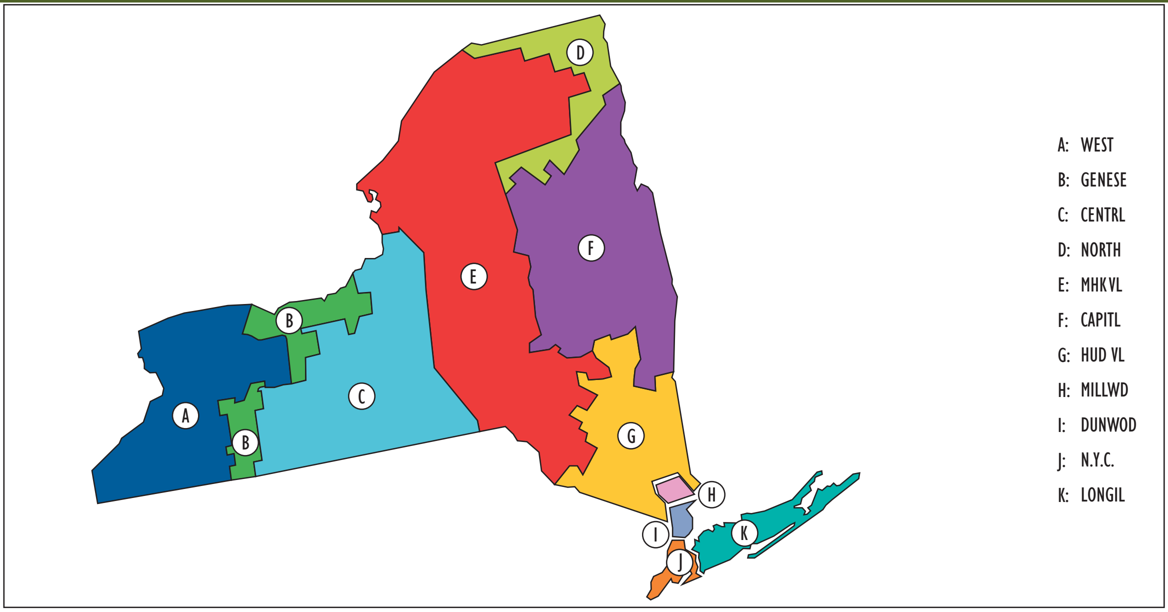
Whether and how demand-side resources should participate in capacity markets remains a source of controversy. It is certainly clear that, in the United States, capacity markets have been a major driver of investment in demand response (IEA, 2013). Capacity markets provide a steady, forward payment stream that aggregators and other demand response providers can use to pre-pay programme participants. However, as demand response resources also participate in retail markets, they derive an additional benefit in the form of avoided retail electricity costs (see Chapter 6). One solution is to continue to allow demand response resources to participate on the supply-side, but to require that they continue to pay the equivalent retail rate – or reduce their capacity payment by the equivalent retail rate – when they are called into service.

One alternative to the inclusion of demand response resources in the capacity market as a supply resource is to simply provide them with a separate payment for service and include them as part of the demand side – reducing the demand for capacity by the extent to which they contribute to reducing peak load.

NYISO

NYISO runs an installed capacity (ICAP) market that offers an interesting contrast to PJM's RPM. While it shares certain common elements with RPM, its main defining characteristic is that it is a short-term (spot) market. While the intent of the ICAP market is to ensure long-term resource adequacy, it does not offer long-term price signals or the ability to lock in specific capacity prices for multiple years.

Map 5.2 • NYISO load zones



This map is without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

Source: NYISO, 2015.

NYISO was formed in December 1999, and is one of a few ISOs in the United States to confine their service territories to a single state. In 2014, NYISO had an installed capacity of 41 297 MW and an estimated peak load of 33 666 GW. The capacity market was launched in 2003.

NYISO is a nodal pricing market, although the ISO is also broken down into 11 load zones labelled A to K (Map 5.2). As with PJM, NYISO's ICAP market has a locational component. In addition to the system-wide New York Control Area (NYCA), NYISO's capacity market allows for the possibility of locational capacity prices for New York City (Zone J) and Long Island (Zone K). In April 2014, NYISO added a new Lower Hudson Valley capacity zone that includes Zones G, H and J (and which therefore includes the NYC zone as a sub-zone).

The capacity auction and capacity product

The NYISO ICAP market has two six-month capability periods: winter and summer. Capacity is secured through three auctions: a strip auction, which procures capacity for the upcoming capability period; a monthly forward auction, which covers all months remaining in the current capability period; and a spot auction, which procures capacity only for the month in question. Only participation in the spot auction is mandatory.

Unlike the RPM, by only offering a short-term capacity price the ICAP market does not provide signals for investment. Instead, the ICAP market functions as a way for LSEs to meet their reliability requirements at times when they are unable to do so through long-term contracts or through self-supply (Kirsch and Morey, 2012). The capacity market, however, does standardise the capacity product, so that regardless of how the capacity is procured, the system operator can be assured that the reliability requirement has been met.

To qualify as a capacity resource, each generator submits to winter and summer Dependable Maximum Net Capability (DMNC) testing. This applies equally to new resources, meaning resources cannot qualify to participate in the capacity market before they are fully commissioned.

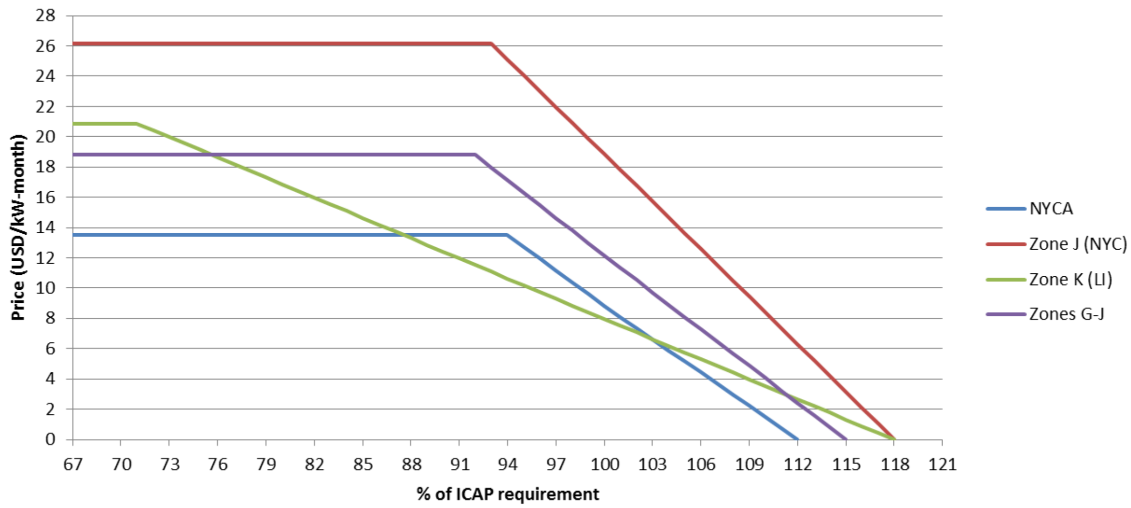
The DMNC test is technology specific, so as to account for differing operating characteristics, and must be based on actual operating data. Baseload fossil and nuclear plants must demonstrate their sustained maximum output over a four hour period, whereas a combustion turbine must do so only over a one hour period. The DMNC for variable resources such as wind, solar PV and run-of-river hydro is simply the net combined nameplate value for all generating units. Capacity resources must also report, on a regular basis, operating data and their maintenance schedule.

The demand curve

Like RPM, the ICAP market uses downward-sloping demand curves. These curves, however, are somewhat simpler than those used in RPM, as they have no inflection point (that is, they are linear, not concave). Each of the four ICAP zones has an associated demand curve, reflecting the specific locational reserve margin requirement (Figure 5.8). The reserve margin requirement is determined on an annual basis by the New York State Reliability Council (NYSRC), based on a 1-in-10 LOLE.

Resources outside NYCA may also participate in the ICAP market, as long as they can demonstrate that they are fully deliverable (that is, that there is sufficient transmission capacity, and that the resource will not be curtailed by its own control area at the expense of NYISO).

Figure 5.8 • NYISO ICAP demand curves for 2014/15

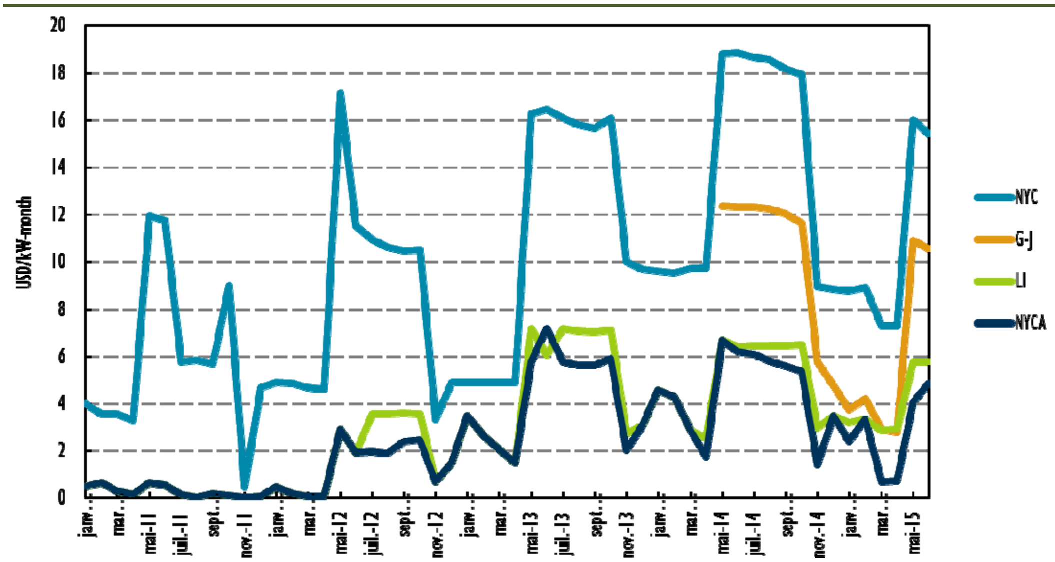


Note: kW-month = kilowatt-month.
Source: NYISO, 2015.

Capacity prices and market power mitigation

The ICAP market's short forward period means greater price uncertainty, and therefore the potential for greater price volatility (Figure 5.9). From an investment perspective, therefore, the NYISO capacity market does not act as a tool for mitigating long-term price risk. Rather, it serves as a price signal indicating a need for investment when available capacity may not be sufficient to meet near-term reliability needs. Investors who would prefer to invest with some sort of price guarantee can choose to enter into long-term power purchase agreements (PPAs) with an LSE – which can in turn count that capacity toward its own capacity obligation (Nelson, 2014).

Figure 5.9 • Historical NYISO ICAP spot prices



Source: NYISO, 2015.

Short forward periods, however, mean that only existing generation competes in the capacity market. This limits the total potential pool of supply, which may limit market liquidity. This raises additional concerns over market power, in particular because, like PJM, the New York market has regions with limited transmission capacity and relatively high demand requirements. This is particularly the case for New York City, which has both limited import capacity and limited land on which to build new generating capacity. For that reason, the NYISO ICAP market includes a set of *ex ante* market mitigation rules, to limit the potential for market power abuse.

United Kingdom

A centralised capacity mechanism has been implemented in the United Kingdom as part of electricity market reform. The first auction was run in 2014 for delivery of capacity in Great Britain for a one year period beginning the winter of 2018/19. The decision whether to run the capacity auctions will be taken annually and will be informed by an electricity capacity assessment carried out by the National Grid, the system operator for Great Britain. National Grid will assess the likely evolution of future capacity margins for the next 15 years, taking into account the contribution of interconnected capacity and demand-side response, and recommend the amount of capacity needed to deliver the enduring reliability standard. The government will then assess whether a capacity auction is needed (EC, 2014b).

The capacity market was introduced with the aim of ensuring the availability of sufficient electricity generation capacity at all times to meet projected levels of demand, as the government recognises that the market may not make this capacity available without some form of incentive. Capacity market participation is not mandatory, and both generation and demand-side response can participate.

A pay-as-cleared auction takes place four years ahead of the relevant delivery year, with plants able to opt out either on the grounds that they will remain open without the capacity payments or that they intend to close before the delivery year. A second year-ahead pay-as-clear auction is held in advance of the delivery year to enable fine adjustment of capacity positions and provide room for demand-side response participation, which is better suited to a short lead time.

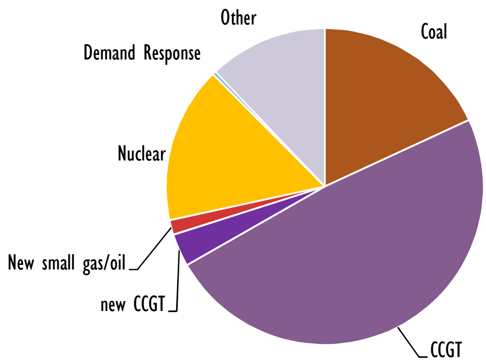
Existing plants have access to a one-year capacity agreement. Existing plants requiring major refurbishment are allowed access to agreements with a term of up to three years, and longer agreements of up to 15 years are available for new plants. The cost of the capacity payments will be recovered from licensed electricity suppliers according to a forecast of each supplier's demand at the time of the system's peak total annual demand, reconciled against the supplier's actual demand when meter data are available. Those that also provide "relevant balancing services" will be able to participate in the capacity market. Providers of relevant balancing services will be deemed to be delivering energy if they comply with National Grid's instructions in a period of system stress. However, if they fail to respond to a dispatch instruction from National Grid for the relevant balancing service, then they will be exposed to penalties under both the capacity market and the relevant balancing service contract.

A secondary market will also be established, with participants able to hedge their positions through secondary trading between the auction and delivery in the delivery year.

The capacity market requires "delivered energy", meaning that capacity providers are obliged to deliver energy whenever needed to ensure security of supply, and they face penalties if they fail to do so. The model also includes additional physical testing of capacity. Failure to demonstrate capacity to the required level on the requisite number of occasions will result in capacity payments being forfeited until successfully demonstrated.

Units which perform below the expected level of performance will be penalised, while those that exceed the expected level will receive over-delivery payments. The penalty consists of three main elements: a monthly liability cap of 200% of a provider's monthly capacity revenues, which, given the weighting of monthly payments according to system demand, may expose providers to a penalty liability of up to 20% of their annual revenue in any one month; an overarching annual cap of 100% of annual capacity revenues; and a penalty rate set at 1/24th of a provider's annual capacity payments.

Figure 5.10 • Successful bidders in the UK capacity market 2014



Source: Ofgem (2015)

Parties who have agreed to generate or reduce demand at times of system stress receive four hours' notice from National Grid, and will have to generate electricity or reduce their demand before the notice elapses – otherwise, they will have to pay a penalty linked to the VoLL.

The first capacity market auction results for 2018/19 saw 49.3 GW of capacity procured at a clearing price of 19.4 GBP/kW. Existing capacity volume, 54.9 GW, exceeded the procured volume in the auction by 5.6 GW. New capacity represented 2.8 GW, while 8.4 GW of older existing coal and CCGT plants failed to secure a capacity agreement, leaving these plant owners in a potentially precarious position (DECC, 2015).

France

The market rules for the capacity mechanism in France were approved in February 2015. In this decentralised mechanism, each power supplier has to be able to guarantee that it can provide sufficient electricity for all its customers even during peak consumption periods, starting on 1 January 2017. Power suppliers will be able to buy capacity certificates from power producers or demand-response operators.

Because of France's heavy reliance on electric heating, its power system is highly sensitive to temperature. A 1°C fall in temperature leads to an extra 2.4 GW of power demand, equivalent to more than two nuclear power reactors. During the cold winter of 2012, France experienced peak power consumption of 102 GW, while in 2014 peak consumption was just 82.5 GW. Related to this, RTE, the national transmission system operator, forecasts a continuous reduction in security of supply margins, amounting to a possible deficit of 900 MW for winter 2015/16 and 2 000 MW for winter 2016/17.

Under the capacity mechanism, RTE issues capacity certificates to power producers for keeping sufficient generation capacity available, and to demand-response aggregators for reducing power demand. The system rewards all generation assets, depending on their availability, which is about 80% of the time for nuclear, 85% for gas plants and 20-25% for wind power.

Certified capacity has the obligation to commit to its forecasted availability during defined peak periods, starting in 2017. At the same time, suppliers have the obligation to own capacity certificates corresponding to the consumption of their own customers during those peak periods. Certificates will be tradable on the EPEX power market, but also over the counter between parties.

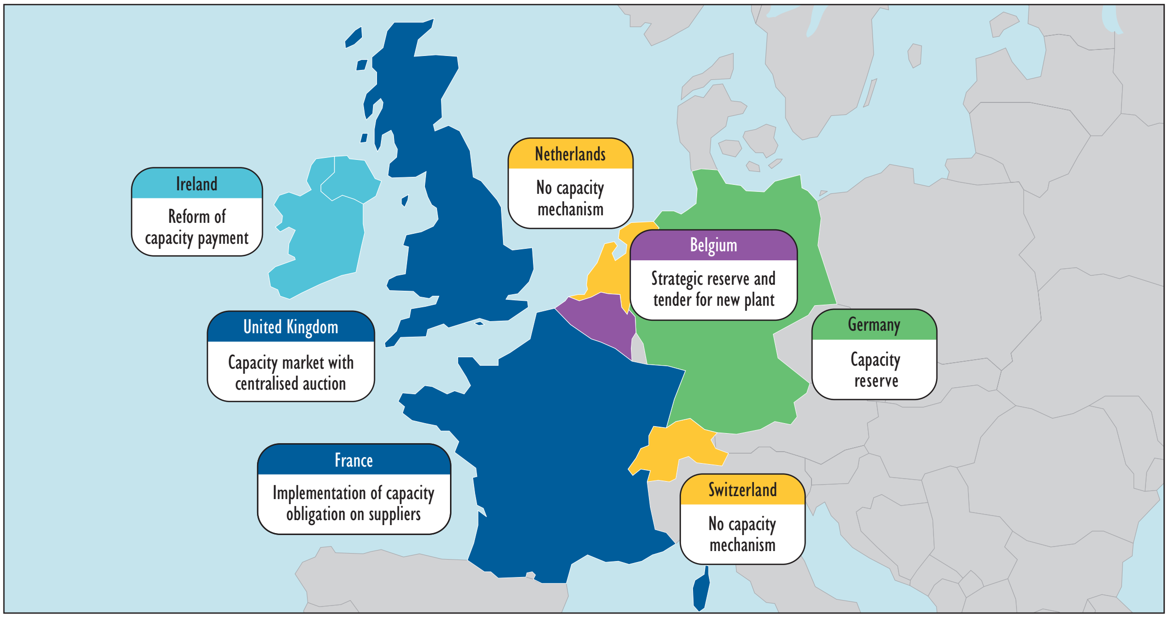
The definition of so-called "peak periods" is an important component of the system's design. They are expected to amount to 100-250 hours per year, corresponding to periods when security of supply is at risk (which, in the case of France, is during winter). The peak period days are notified one day ahead (D-1) and are triggered by a demand criterion (days when demand is expected to be highest).

5.4. Regional markets and capacity trading

In Europe and the United States, as well as other parts of the world, electricity markets are becoming increasingly interconnected. Under such circumstances, meeting reliability needs can no longer be considered purely from a national or jurisdictional perspective. Despite this, both reliability standards and capacity markets continue to be developed according to political boundaries.

Within Europe, a discussion has evolved as to whether the implementation of national capacity mechanisms is distorting the European energy market. This discussion has primarily emerged in reaction to the increasing number of countries in the centre of the European Union (for example, France and Germany) that have begun to discuss the necessity of implementing some kind of mechanism to ensure security of supply. Countries such as Italy, the United Kingdom and Ireland have also either been discussing the implementation of such a mechanism or have already implemented one (Map 5.3).

Map 5.3 • Neighbouring capacity markets in selected countries of Western Europe



This map is without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

While capacity mechanisms are not a new phenomenon, greater co-ordination between countries on cross-border flows through day-ahead market coupling raises questions as to whether country-specific capacity mechanisms should interact. In particular, does the existence of national or regional capacity markets undermine the functioning of inter-regional energy markets? And, how can external resources participate in capacity markets?

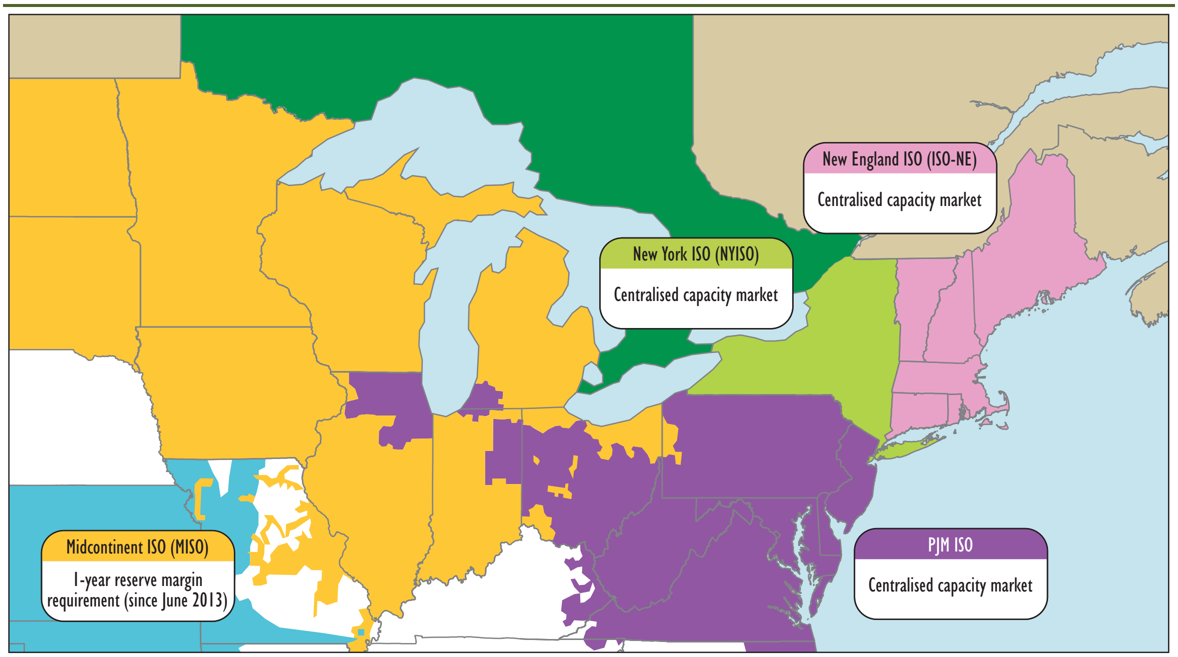
Does the existence of national or regional capacity markets undermine the functioning of inter-regional energy markets?

A capacity market should ideally have the same physical boundaries as the wholesale market within which it operates. One possible impact of a capacity market is to reduce the frequency of scarcity prices. Replacing scarcity rents with capacity revenues reduces revenue risk, and so an investor choosing between two markets – country A that is an energy-only market and country B with a capacity market – may decide that the one with a capacity market is more attractive. The increase in capacity in country B will lower wholesale prices. But because country A is interconnected with country B, it will benefit from those lower wholesale prices as well. Therefore, a capacity market that exists within a larger wholesale market may reduce wholesale prices for all participants, while only allowing participants within the territory of the capacity market to recover the lost revenues.

In the case where there is a regional energy market but multiple sub-regional capacity markets, capacity markets can potentially create distortions if they have inconsistent designs. In particular, inconsistent designs can lead to four types of distortion: 1) changes in utilisation patterns of installed capacity; 2) underestimates of the wind and solar contribution to resource adequacy; 3) potential market power abuse within the capacity markets themselves; and 4) influence over the location of new investments (IEA, 2014).

If capacity mechanisms allow external capacity resources to participate on an equal footing with local resources, usage patterns should in fact become more optimal over time. For example, surplus capacity in one market could participate in markets with a capacity shortfall. Allowing external capacity resources to participate, however, requires the markets in question to agree on how to handle so-called "seam" issues, to ensure that market rules to do not prevent or limit the full transfer of capacity between the markets in question. This includes issues such as transmission allocation, the capacity product definition, and co-ordination of transmission and generation outages (FERC, 2013b; MISO, 2012). Capacity markets in the United States have extensive experience in allowing external capacity to participate (Map 5.4), although the regulatory frameworks within these markets limit the relevance of this experience to other jurisdictions.

Map 5.4 • Neighbouring capacity markets in Northeastern United States



This map is without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

Studies looking at the interaction of different capacity markets and cross-border participation in capacity markets conclude that cross-border participation is attractive, but raises a number of issues as to who should participate and as to design. How to quantify the benefits – contribution to security of supply, economic efficiency and competition – is not entirely clear (FTI Compass Lexecon, 2015).

How can external capacity resources participate in capacity markets?

A number of different options are available to account for external capacity resources in the design of capacity markets, including the simplest option: not taking it into account at all by restricting participation to internal resources. In France, the contribution of interconnections is assessed unilaterally in a statistical fashion. In the United Kingdom, the first capacity auction held in 2014 did not allow for the participation of external capacity resources, although interconnectors have the possibility to bid starting in 2015. Under pressure from the European Commission, both markets are enabling the explicit participation of external capacity resources in their national capacity markets (State Aid Guidelines 2014). Another option is to design a capacity market that covers several regions (FTI Compass Lexecon, 2015).

In PJM, external capacity resources have an additional constraint that by necessity differentiates them from local resources – namely, the need for firm transmission capacity in order to ensure deliverability. A system operator can only control local resources, and must therefore assume that external capacity resources will be available when called upon. Often this will involve some form of network analysis. For example both MISO and PJM – neighbouring markets in the United States – use point-to-point (PtP) analysis, which estimates the amount of capacity that is fully deliverable given locational constraints. However, as each jurisdiction performs this analysis separately, co-ordination is needed at least in terms of methodology. Otherwise there is a risk that one market may view a resource as having more deliverable capacity than it actually does, raising the possibility that the exporting market will curtail a generating resource that the importing resource assumes will remain firmly available.

Another concern is the distinction between transmission capacity allocation for energy and capacity usage. Both energy and capacity suppliers can bid for the right to firm transmission capacity. However, during a scarcity event, only the capacity supplier is obligated to serve energy. In the United States, the competition between energy-only and capacity markets potentially limits the supply of capacity available for export. There is also a temporal component to this issue, as the allocation of transmission rights may not match capacity obligations. If the external capacity resource wishes to participate in a three-year forward capacity market with a one-year commitment period, it must obtain one year of firm transmission capacity three years in advance. Therefore, if the external capacity resource is to participate, the export market must have transmission allocation rules that match the importing market's requirements.

It should be noted that, from the perspective of system operators, the degree of certainty as to the availability of external capacity resource is lower than for internal capacity. Under scarcity conditions, curtailment procedures usually give priority to local consumers, and interconnections could be cut. In a European context, it is difficult to anticipate how a neighbour would react if load has to be shed in their country to allow for the export of power. From this perspective, a clear definition of curtailment procedures across borders is a prerequisite to ensure trust among system operators and create the conditions for cross-border trade of capacity.

In the event that both regions have capacity markets, the potential exists for resources on both sides of the seam to sell capacity in opposite directions, using the same transmission line. This means that transmission capacity must be co-ordinated simultaneously and bi-directionally, on a forward and firm basis, in order to prevent the possibility of over-allocation of the transmission line.

Principles for efficient co-ordination of capacity markets

For the optimal participation of external capacity resources in regional markets, a common method is needed to determine deliverability between all interconnected markets.

In Europe, Eurelectric (2015) has developed a reference model for European electricity markets. The trade association defines capacity as availability and considers that cross-border participation in capacity markets should be seen as a stepping-stone towards regional capacity markets. The discussion about capacity market co-ordination is only beginning and is likely to receive increasing attention in the coming years.

Short of complete market integration, several measures could be taken to lay the groundwork for integrating capacity markets:

- Regional resource adequacy forecasting should be used to determine capacity needs and have to be consistent with the energy market footprint. Such regional adequacy forecasts can be used to calculate the contribution of interconnections to different local markets and the level of capacity delivery between interconnected areas.
- Avoid conflicting capacity product definitions to enable cross-border trade of capacity. The product definition includes the availability period (which hour), the lead time of the product and the penalty regime. Deliverability of imported capacity products is also an important consideration.
- Avoid interference with the energy trade across borders. Capacity markets should not distort price formation on wholesale energy markets, in particular cross-border capacity should not be reserved for capacity in order to avoid distortions of the forward, day-ahead, intra-day and balancing markets, which determine the actual direction of the energy flow.

Conclusion

In liberalised markets, capacity mechanisms can play an important role in ensuring sufficient resource adequacy. They can also allow alternative resources to participate in a market that has traditionally focused mainly on supply-side resources. Properly designed, capacity markets can help to resolve the "missing money" problem without distorting energy markets.

Capacity markets should not, however, be seen as a tool for resolving problems in wholesale electricity markets. For capacity markets to function properly, it is important for the design of wholesale markets to be right. Capacity markets can fill a revenue gap for energy resources, but should not be seen as a tool to ensure profitability. Instead, capacity markets can be seen as a safety net and can complement energy market scarcity rents.

While many different types of capacity mechanism can be found, this chapter has mainly focused on the strategic reserve (the targeted volume-based model), as applied in parts of Europe, and a model that is most common in the liberalised portions of the United States (the market-wide central buyer model). Regardless of the form, a properly designed capacity market has three key components: a pre-determined level of demand, based on the system operator's assessment of resource adequacy needs; a mechanism for price discovery, preferably in the form of an auction; and a well-defined capacity product, which takes into account the contribution of the capacity resource to meet adequacy needs, but is, to the greatest extent possible, technology neutral.

Capacity mechanisms should also capitalise on regional diversity in the resource base by allowing external capacity resources to participate. This requires regional collaboration to ensure that capacity is truly deliverable across borders, but does not require that capacity market mechanisms be completely harmonised.

The problem of ensuring adequacy will become more pronounced as high penetrations of zero-marginal cost variable generation enter the grid. As capacity markets exist in many competitive markets, it is important to better understand how to implement capacity markets in such a way as to provide the most reliability, at least cost, and in as market-friendly a way as possible.

References

Bowering, J. (2013), “Capacity Markets in PJM”, *Economics of Energy and Environmental Policy*, Vol. 2, No. 2, www.iaee.org/en/publications/eeeparticle.aspx?id=47.

Cramton P., A. Ockenfels and S. Stoft (2013), “Capacity market fundamentals” www.pserc.cornell.edu/empire/2_2_a02.pdf.

DECC (Department of Energy and Climate Change) (2015), “Cross-border participation in gb capacity market”, www.iea.org/media/workshops/2015/esapworkshopv/OConnell.pdf.

Elia (2015a), “Adequacy calculations for Belgium – Strategic reserve volume”, http://publications.elia.be/upload/UG_upload/BG3KPBOHX4.pdf.

Elia (2015b), “The strategic reserve – A mechanism to cover structural shortages in generation”, www.elia.be/~/media/files/Elia/Products-and-services/ProductSheets/E-Evenwicht/E9_E_Strategic_Reserve.pdf.

Elforsk (2014), “Demand response in the strategic reserve - The case of Sweden”, www.elforsk.se/Documents/Market%20Design/projects/ER_14_29.pdf.

Eurelectric (2015), “A Reference Model for European Capacity Markets”, www.eurelectric.org/media/169068/a_reference_model_for_european_capacity_markets-2015-030-0145-01-e.pdf.

EC (European Commission) (2014a), “Guidelines on State aid for environmental protection and energy 2014-2020”, http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52014XC0628(01)&from=EN.

EC (2014b), “State aid: United Kingdom electricity market reform – Capacity market”, http://ec.europa.eu/competition/state_aid/cases/253240/253240_1579271_165_2.pdf.

FERC (Federal Energy Regulatory Commission) (2013a), “Centralized capacity market design elements”, www.ferc.gov/CalendarFiles/20130826142258-Staff%20Paper.pdf.

FERC (2013b), “Order on initiatives to address SEAMS issues”, Docket Nos. AD12-16-000, www.ferc.gov/whats-new/comm-meet/2013/121913/E-3.pdf.

FTI Compass Lexecon (2015), “Towards a European approach to capacity mechanisms”, www.ceem-dauphine.org/assets/dropbox/Charles_Verhaeghe-_FTI-Presentation_-_Dauphine_-_Target_Model_-_08072015.pdf.

Hogan, W., (2013) “Electricity Scarcity Pricing through Operating Reserves”, Economics of Energy and Environmental Policy, Vol. 2, No. 2, pp. 65-86.

IEA (International Energy Agency) (2013), “Energy provider-delivered energy efficiency: A global stock taking based on case studies”, www.iea.org/publications/insights/EnergyProviderDeliveredEnergyEfficiency_WEB.pdf.

IEA (2014), *Seamless Power Markets: Regional Integration of Electricity Markets in IEA Member Countries*, Insights Paper, OECD/IEA, Paris, www.iea.org/publications/freepublications/publication/SEAMLESSPOWERMARKETS.pdf.

Kirsch, L. and Morey, M. (2012), "New York state capacity market review", Christensen Associates Energy Consulting, 19 September 2012, www.publicpower.org/files/PDFs/CAEnergy_NY%20Capacity%20Market%20Study_120919_Final.pdf.

MISO (Midcontinent Independent Transmission System Operator) (2012), "Comments of the Midwest Independent Transmission System Operator", Docket No. AD12-16-000, www.misoenergy.org/Library/Repository/Tariff/FERC%20Filings/2012-08-10%20Docket%20No.%20AD12-16-000.pdf.

Ministry of the Environment and Energy Sweden (2015), "The strategic reserve - why and how?", http://ec.europa.eu/competition/sectors/energy/strategic_reserve_en.pdf.

Monitoring Analytics (2015), "PJM state of the market – 2014", www.monitoringanalytics.com/reports/PJM_State_of_the_Market/2014.shtml.

Nelson, E. (2014), "Written statement of Emilie Nelson, Vice President – Market Operations, on behalf of the New York Independent System Operator, Inc.", FERC Docket No. AD14-18-000.

NYISO (2015), "Installed capacity manual", www.nyiso.com/public/webdocs/markets_operations/documents/Manuals_and_Guides/Manuals/Operations/icap_mnl.pdf.

Ofgem (2015), Annual report on the operation of the capacity market, https://www.ofgem.gov.uk/sites/default/files/docs/2015/06/annual_report_on_the_operation_of_the_cm_final_0.pdf.

Paulos, B. (2014), "PJM tightens capacity market rules to improve reliability", POWER, www.powermag.com/pjm-tightens-capacity-market-rules-to-improve-reliability/.

PJM (2014a), "Resource Obligations in RPM", www.pjm.com/~/media/committees-groups/committees/pc/20140610-energy/20140610-rpm-obligations-overview.ashx.

PJM (2014b), "Triennial review of VRR curve shape", www.pjm.com/~/media/committees-groups/task-forces/cstf/20140630/20140630-item-04c-vrr-curve-background.ashx.

Pöyry (2015), "Decentralised reliability options", Pöyry Management Consulting, www.poyry.co.uk/sites/www.poyry.co.uk/files/decentralisedreliabilityoptionsfullreport_v300.pdf.

Süssenbacher, W. (2011), "Kapazitätsmärkte- und Mechanismen im internationalen Kontext" [Capacity markets and mechanisms in the international context], www.eeg.tuwien.ac.at/eeg.tuwien.ac.at_pages/events/iewt/iewt2011/uploads/plenarysessions_iewt2011/P_Suessenbacher.pdf.

Svenska kraftnät (2015), *Annual Report 2014*, www.svk.se/siteassets/om-oss/organisation/finansiell-information/annual_report_-2014.pdf.

Svenska kraftnät (2012), *Effektreservsupphandlingen 2012/2013* [*Power Reserve Contract],* www.svk.se/Energimarknaden/El/Effektreserv/Effektreserven-20112012; Förordning (2010:2004).

Chapter 6 • Demand response

HIGHLIGHTS

- Demand response, can play a role in the decarbonisation of the power system by decreasing demand when the system is tight, and also by adjusting the timing of power consumption to when supply from low-carbon resources is more abundant.
- Large consumers already respond to prices by participating directly in wholesale electricity markets. They buy their expected consumption in advance and respond to price variations by re-selling on the short-term markets.
- In addition, smart meters and progress in automation technologies increasingly enable smaller consumers to be price responsive. Dynamic pricing options, such as critical peak pricing (CPP), are a straightforward way to tap into this potential.
- To date, however, revenues from participation in the wholesale energy markets are rarely sufficiently abundant or predictable to cover the (fixed) cost of investing in the equipment needed to develop demand response.
- Another option is to treat demand response as generation, and “dispatch it” on wholesale electricity markets. Direct participation of demand response aggregators in capacity markets has been effective in kick-starting demand response in several markets, such as the US regional transmission organisation, PJM.
- But treating demand response as generation requires complex market rules, with the need to define a baseline of consumption against which demand response can be assessed. Defining the correct level of remuneration is difficult and can be controversial.
- Lastly, the protection of data to safeguard consumer confidence is an additional and important prerequisite for the significant deployment of demand response.

A major challenge for regulators in the successful transformation of the electricity sector is the integration of new technologies into the power system. This is not only about electricity generation, but also about new technologies that change the way we consume electricity.

Demand response programmes offer the opportunity for electricity consumers to intentionally shift or reduce their load either in response to price signals or in exchange for an incentive. To date, demand response has mainly been the preserve of large industrial users. However, new smart appliances and technologies are now empowering smaller consumers (or energy service providers on behalf of consumers) to manage their own electricity demand.

While many smart technologies already exist, four principal challenges remain: the need to build consumer engagement; the lack of a supportive regulatory framework in many markets; privacy and cyber security issues that can be a major constraint unless factored in the design of demand response arrangements; and the large number of fragmented stakeholders involved in restructured electricity markets, which introduces added complexity.

This chapter begins with a discussion of the potential for and benefits of developing demand response in the context of decarbonisation. The next section looks at the participation of demand in electricity markets, either on the load side or on the generation side. The final section describes how price-based demand response, such as dynamic pricing, could be further developed.

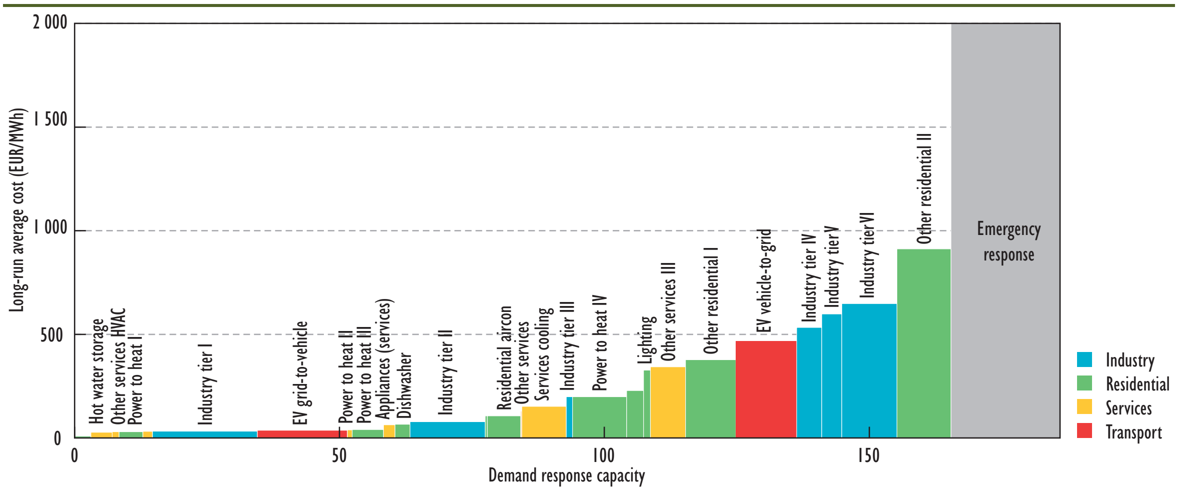
6.1. Benefits of demand response

Demand response is seen as an opportunity to adjust load according to system conditions. Loads can be reduced when there is less wind and sun, and conversely, demand can be increased when generation is abundant. This demand-side flexibility could facilitate the integration of larger shares of variable generation sources, which will be a key challenge for future decarbonised power systems.

A potential game changer

Thanks to the development of new technologies, such as smart grids, today's consumers have the possibility of changing their consumption patterns by using automatic programming and energy management systems. While demand response is not a new concept, ongoing structural changes to our societies tend to reduce the role of traditional sources of demand response, such as the reduction in the share of industry in Organisation for Economic Co-operation and Development (OECD) economies. Fortunately, changes in electricity demand paths, the deployment of new metering technologies and the development of behind-the-meter generation potentially create a new and favourable context for the large-scale deployment of demand response.

Figure 6.1 • Modelled demand response and supply curve in the European Union in 2050



Demand response potential typically amounts to around 15% of peak demand. The International Energy Agency (IEA) assessed that the potential could exceed 150 gigawatts (GW) by 2050 in the European Union (Figure 6.1), even though this capacity corresponds to different product definitions with regard to duration and frequency of response. Demand response can be deployed at four distinct levels, with an impact proportional to the scale of consumption:

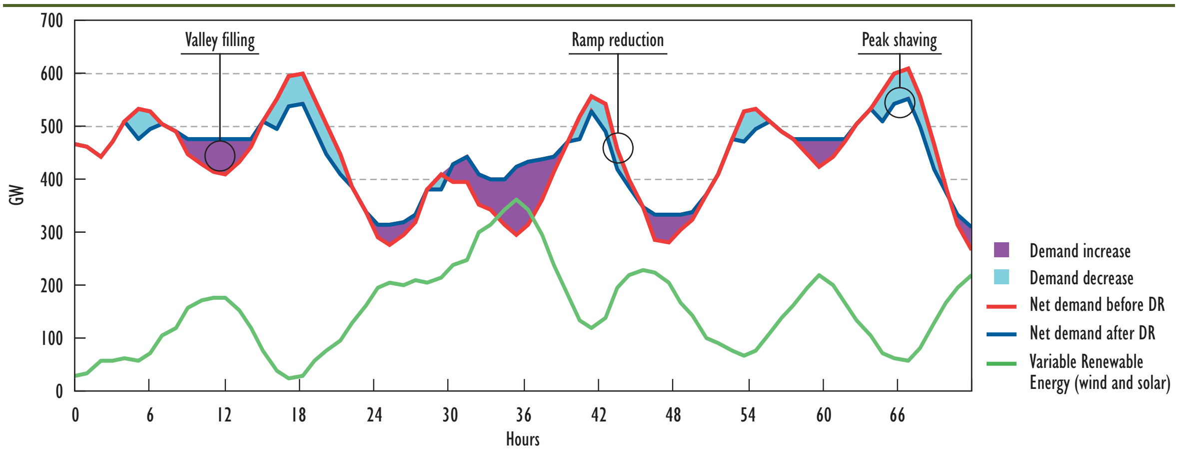
- at the industrial level, when large manufacturing plants have the flexibility to adjust production processes to electricity prices to decrease their energy costs
- at the services level, typically through automated solutions to manage air conditioning or lighting systems, also to decrease energy costs
- at the residential level, with innovative commercial services offering consumers energy savings with minimal impacts on daily life, for example via smart appliances
- at the transport level, with the deployment of electric vehicles.

The fact that electricity consumers do not typically change their consumption patterns in accordance with electricity prices has always been an issue for the design of electricity markets. Until recent years, no proper physical or market infrastructure has been in place to enable this kind of demand response. Now, however, the progressive diffusion of information and communications technologies (ICT) throughout the economy is reshaping electricity demand. Smart meters, smart appliances and the development of energy management software are driving down the transaction costs associated with the optimisation of the timing of electricity consumption.

Demand response can be used for different objectives (as illustrated in Figure 6.2) and can increase the flexibility of the load in different dimensions:

- peak shaving: reducing peak consumption during tight system conditions so as to release pressure on generation and grid capacity needs. This also reduces the need for investment in peak generation assets
- valley filling: increasing or shifting consumption to hours of ample generation of wind and solar power
- ramp reduction: reducing the steep ramping needs at peak time with the shifting of load at a time when the system is under less constraint.

Figure 6.2 • The different roles of demand response with high share of renewables (simulation)



Looking ahead, demand response technologies have the potential to become a game changer for electricity markets. Demand response can also solve the adequacy problem discussed in Chapter 4. If consumers are able to, and interested in, being responsive to prices, the electricity market should always clear at a price that reflects the value that consumers place on electricity consumption. In this case, the market could potentially always balance supply and demand. In principle, this flexibility should reduce the volatility of electricity prices, with the result that electricity would become much more like other commodities, such as gas, for instance.

Finally, demand response is also evolving because electricity generation is becoming more decentralised. As consumers invest in behind-the-meter generation, they will increasingly have access to back-up generation in case of system failure, or if the market price of electricity is too high. Energy management systems will optimise grid demand for electricity depending on wholesale prices, network charges, local storage and the fuel cost of back-up generators. This will change the merit order for the power generation mix and the value that consumers place on uninterrupted supply of electricity from the grid.

Impact on networks

From a network perspective, demand response technologies can bring more security to the system and contribute to solving stress conditions in the transmission and distribution grids, contributing to security of supply. The transmission system operator (TSO) can also use demand response for balancing purposes to ensure frequency stability and proper equilibrium between produced electricity and demand. Increasingly, the distribution system operator (DSO) will become in charge of managing the communication flow between the active consumer, the TSO and the generator.

Targeted demand response programmes could also be an alternative to investment in network capacity upgrades to address congestion. In the case of the United Kingdom, it has been estimated that the cost of network reinforcement could be around one-third less in a system with optimal demand response combined with 100% penetration of electric vehicles and heat-pump space heating (Strbac G., 2008). The benefits of demand response for network investment could be reflected in the network tariff (see Chapter 9).

Energy storage as a source of demand response

From the system perspective, energy storage works similar to load shifting. Energy storage technologies can be classified into two types, electricity storage and thermal storage (IEA, 2014), and additionally by duration of storage, either short-term, long-term or distributed battery storage.

For a controlled supply of electricity, long-term applications that enable electricity to be stored for many hours or even weeks are the most valuable. The availability of seasonal storage is extremely limited, the technology typically used for this purpose being pumped-storage hydropower (PSH), which is also today's most mature and widespread option. The IEA estimated in 2014 that 99% of electricity storage capacity is PSH, with at least 140 GW of PSH connected to the grid worldwide. Compressed air energy storage has also been successfully used in the past in the United States and Europe, but on a smaller scale.

Finally, other storage technologies, such as batteries, are beginning to offer an additional opportunity to store electricity and make it available when needed. Battery technologies have suffered from a series of challenges, ranging from energy density to safety, recyclability, battery-to-grid connection and other issues. Although battery technologies have only recently begun to be deployed on a large scale, their usage is becoming increasingly relevant with the deployment of variable renewables such as wind and solar. For example, new models of wind turbines now include a battery system that enables short-term storage. While previous systems would have relied on expensive farm-level battery storage installations, this new technology embeds the battery in the turbine system itself. This technology is associated with software applications that enable power producers and wind turbines to access real-time data and provides predictable power for the short term.

Challenges nevertheless lie ahead for the large-scale deployment of energy-storage technologies, starting with the relatively high cost of implementation, for which reliable cost recovery mechanisms would need to be put in place (Think, 2012). This has implications for market design; as energy storage can serve multiple purposes, from generation adequacy to grid stability, the regulatory framework should provide a clear classification of storage assets as generator and/or load. This clarity could ease certain technical and economic issues, such as eligibility for grid tariffs and the ownership of the storage assets in a context of unbundling (Strbac, 2008).

Energy efficiency and reduction of carbon emissions

Demand response and energy efficiency are interlinked but have different ultimate objectives. While demand response is mainly about shifting power demand by means of a price or incentive signal, energy efficiency is about maximising the output from the use of energy. The related technologies therefore serve different purposes, but appear to be compatible from a usage and regulatory point of view.

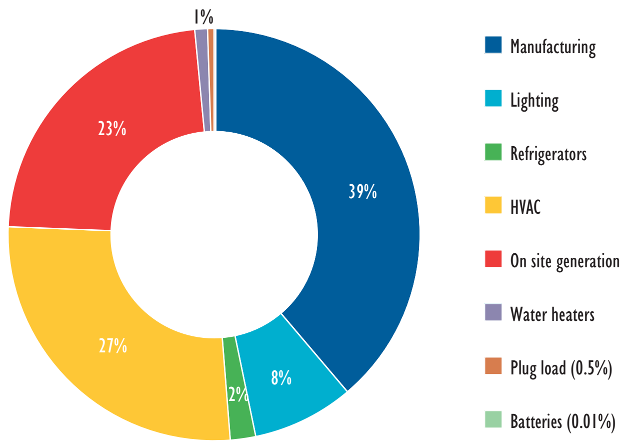
Demand response technologies can also provide environmental benefits by contributing to carbon dioxide (CO_2) emissions reduction – by switching energy demand from a time when the system relies on gas or coal, to a period when more renewables or less carbon-intensive capacity is available in the system. For instance, the displacement of 1 000 megawatts (MW) of generation from a subcritical coal power plant to a nuclear power plant could result in the avoidance of 710 00 tonnes of CO_2 emissions, representing a gain of USD 19 per megawatt hour (MWh).[24] Such gains have to be understood in the broader context of a meaningful CO_2 price.

Automation and coupling of energy sectors

Large-scale deployment of demand response will inevitably require the development of automated solutions as part of new energy management technologies. These can be linked to the price signals embedded in supply contracts. For example, automatic curtailment of consumption can be based on predefined signals sent by the TSO or the aggregator. Key technologies include building management solutions, such as on-site generation, water heating, digitally controlled thermostats, automated lighting systems, and manufacturing processes and remote-control pumps that can make subtle changes in intensity (Figure 6.3). These adjustment technologies can have a very limited impact on daily life, and while particularly relevant for commercial centres or large industrial sites, can become increasingly meaningful on an aggregated level for the residential sector.

While traditional systems have collected user data on a planned basis, for example on a fixed date each year, new smart meters allow data to be collected hourly. This amounts to a tremendous change (Cukier, 2015). The corollary of the large deployment of demand response and smart grid technologies is that this inevitably leads to an increase in the granularity of the data collected and the growing collection of data. A central challenge is to properly collect and handle these data, so as to maximise the efficiency of such technologies. "Big data" will require the development of appropriate analytical software for a better assessment of current market needs and consumer behaviour. Data management will also assist with the integration of distributed generation sources by identifying local producer behaviour and its impact on the distribution grid.

Figure 6.3 • Source of demand response in PJM (2014/15 delivery year)



Notes: HVAC = heating, ventilation and air conditioning; a medium-sized utility in the United States with average customer base of 500 000.

Source: PJM, 2015.

[24] Based on the assumption of 1 000 MW of energy from a subcritical coal power plant being substituted with a nuclear power plant operating for 800 hours, meaning almost 710 000 tonnes of CO_2 avoided, assuming a CO_2 price of USD 22 per tonne.

The management of big data requires the deployment of metering systems on the one hand, and on the other dedicated software capable of managing large quantities of data and extracting valid and useful information for the user. This kind of software has been developed by specialised companies rather than utilities, thus introducing an additional player to the management of a decarbonised power system.

Privacy and cyber security

The secure transmission and protection of consumer data are a vital prerequisite for the wide deployment of demand response, especially with smart grid systems. In order to be effective, smart metering and demand response will need to have access to a very granular level of information on final energy consumption, which can be of potential interest to commercial institutions and law enforcement bodies – and also malevolent bodies. Ensuring the maximum protection of these data must therefore be a priority for regulators. Demand response data will be handled between generators, utilities, final consumers, DSOs and possibly TSOs. To ensure the proper functioning of this relationship, end consumers will have to agree to disclose data about their electricity consumption, and likewise DSOs and TSOs will need to disclose data about network capacities and generators about their supply and related cost of production.

Two potential risks for misuse of personal data would most affect consumer trust:

- network attack, either malicious or non-malicious
- non-malicious cyber security events, such as equipment failure and user/administrator errors.

Currently the majority of cyber security events are non-malicious, which may result from natural phenomena such as hurricanes, tornadoes, floods and solar activity. Regardless of the source, the impact is often the same. It is therefore important to develop a risk assessment process.

It is also essential to establish a clear way for the consumer to consent to sharing information – either through an “opt in” solution or a default solution with detailed information provided to the consumer and the ability to opt out. Once the owner has agreed to the level of confidentiality for disclosing their data, the regulator must draw up rules for dealing with infringements and designate an appropriate regulatory body (CEER, 2012). In Europe, certain countries have opted for the creation of an independent platform in the case of smart grids, while others have delegated this competence to DSOs.

6.2. Participation of demand in electricity markets

Electricity consumers can react to variations in electricity prices according to different mechanisms (Hogan, 2010):

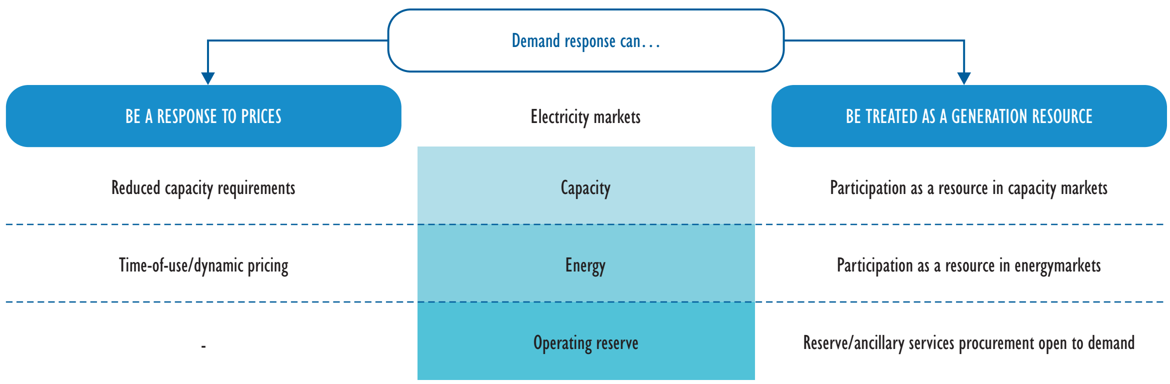
- Through dynamic pricing, where the final consumer is actively adjusting its consumption to prices.
- Through explicit contracts in which the consumer purchases a fixed quantity of electricity and re-sells its surplus (not consumed).
- Through “imputed demand response”, where an individual’s consumption is estimated on a baseline and the demand response calculated on that basis participates in markets as a source of generation.

The first two forms correspond to a response to electricity prices (Figure 6.4, left-hand column) where consumers agree to pay the marginal pricing for their consumption. While these approaches are straightforward, they have proven to be slow to develop in restructured electricity markets and have had mixed results.

This has led market designers to consider the third option: treating demand response as a generation resource (Figure 6.4, right-hand column). This might be necessary in particular where short-term market prices are not properly pricing in scarcity and real-time constraints (see Chapters 3 and 4), or where the physical and market infrastructure to implement dynamic pricing are not in place or are under development. In this case, the market design has to be updated in order to accommodate the different energy resources available, including the requirements of demand response.

This approach – treating demand response as a generation source – has already been implemented in several markets to accelerate the development of demand response. It does, however, lead to a considerable increase in the complexity of the overall market design.

Figure 6.4 • Approaches to demand response



Participation in wholesale energy markets

The straightforward solution to developing demand response is to enable customers to participate in electricity markets as loads and respond to real-time and dynamic electricity prices.

This is already the case for large industrial consumers. In practice, they buy their expected consumption in advance on forward markets and can respond to the evolution of wholesale prices by reselling on the short-term markets. If prices on the day-ahead or intraday markets are very high, consumers re-sell this electricity rather than consuming it. Consumers are billed based on the electricity actually metered.

For small consumers, direct participation in wholesale electricity markets is usually not possible. This role is devoted to electricity retailers who aggregate the load of many consumers before participating in wholesale markets. If a fraction of their consumers is price-responsive, retailers can re-sell on wholesale markets the electricity not consumed by these consumers.

Retailers can develop a price-responsive consumer base by offering different forms of contractual arrangements, such as real-time prices, CPP or peak-time rebates (see next section). Where consumers have smart meters, the retailers are responsible for ensuring that consumers actually respond. Otherwise, they will be subject to imbalances and will have to pay based on the aggregated load of their consumers.[25] Here again, consumers are billed based on the electricity actually metered by smart meters.

[25] Without smart meters, the consumption profile of consumers is calculated on average and retailers cannot know if consumers actually respond to electricity prices.

Box 6.1 • Demand response in France

France has been a front-runner in the implementation of time-of-use and dynamic electricity tariffs. In the 1960s, the national utility, EDF, was already proposing differentiated electricity tariffs (day/night and seasonal). The EJP (*effacement jour de pointe*) tariff is a form of CPP that developed the country's demand response capacity to 6 GW in 2000. Over the years, the availability of these tariffs has been reduced due to electricity market liberalisation and the capacity subsequently declined to 3 GW.

In France, demand response can participate in capacity, energy and balancing markets.

In France, the first demand response operators entered the commercial and industrial market in 2003, and the residential market in 2007. They offered consumers the ability to manage their electricity demand in exchange for financial compensation. The French TSO, RTE, opened up participation in the energy and balancing markets, but the minimum threshold for eligibility was 10 MW, automatically keeping out direct participation from residential consumers as well as small and medium-sized companies.

An outstanding question related to the remuneration of load shifting contracted by new aggregators in the system. To address this, in 2012 France introduced new rules called NEBEF (*notifications d'échange de blocs d'effacement*), making RTE the only intermediary through which bids and offers for load shifting can be made. When this system was introduced, the price set for electricity not consumed was not particularly attractive compared to other sources of power used to generate electricity.

With the introduction of a capacity remuneration mechanism in France in January 2015, demand response will become fully eligible to participate and therefore the regulated incentive mechanism will be cancelled. Aggregators will participate and be remunerated on the capacity mechanisms. In addition, a support mechanism was introduced under which aggregators benefit from a financial incentive, financed by the regulated electricity tariff.

The challenge for making demand response attractive enough for consumers to actively participate in such scheme lies in its remuneration. Wholesale market prices might not represent sufficient compensation for the consumer to reduce its load, because market prices are rarely high enough to justify buying a unit of energy not consumed. Consequently, revenues from participation in wholesale energy markets have not been sufficient to date to cover the (fixed) cost of investing in metering and other smart devices and/or self-generation units needed to develop demand response.

Participation in capacity markets

Capacity markets have been more effective at developing demand response than energy markets. As discussed in Chapter 5, capacity mechanisms have been developed to ensure sufficient resources are available to meet peak demand at least cost. In capacity markets, the only relevant role of demand response is its contribution to meeting system adequacy through load shaving.

Energy retailers can reduce their capacity requirement, for which they have to pay, by developing dynamic pricing. In addition to participating in energy markets, retailers can also reduce the peak demand of their aggregated portfolio of consumers by developing a base of consumers responsive to prices. Dynamic pricing can, in addition to reflecting the price of wholesale energy, also reflect the price that retailers have to pay for capacity. Demand response can lower the capacity that retailers or load-serving entities have the obligation to contract on capacity markets.

But the most effective solution has been the participation of demand response aggregators directly in capacity markets. This has enabled the creation of a predictable revenue stream to finance the upfront investment cost needed for demand response. With this approach, capacity remuneration mechanisms can be designed in such a way as to fully recognise the potential of demand response on a level equivalent to other sources.

A number of conditions need to be met to allow demand response to participate effectively in capacity markets.

First, it implies a clear definition of all the demand response products available, as well as full transparency with respect to the system operator and the contractual agreements between aggregators and consumers. The many interruptible contracts that were historically used to provide similar services also need to be accommodated.

Second, capacity remuneration mechanisms can successfully incentivise the development of demand response resources only if they provide a sufficiently high and predictable economic incentive to justify the initial investment from consumers, whether large-scale companies or residential consumers.

Third, while the value of demand response is not directly dependent on whether it comes from large-scale industry or residential customers, the participation of different customer classes in wholesale markets cannot be activated under the same rules. Aggregators play a key role here, acting as the operator in charge of organising demand response participants and redistributing market revenues. Market rules need to clearly define the role, responsibilities and duties of aggregators, and how they might affect demand response participation.

In the United States, the participation of demand response in capacity markets has been implemented not only by PJM since 2007, but also by the independent system operators (ISOs) of New England (ISO-NE), under its Forward Capacity Market since June 2010,[26] and New York (NYISO) under four different programmes.[27]

In the case of PJM, the introduction of a capacity mechanism has proven to be a strong enabler of demand response deployment. The capacity market, as defined under the PJM Reliability Pricing Model (RPM), was introduced by PJM in 2007 with the objective of introducing a long-term signal alongside the wholesale electricity price to provide certainty and visibility for investment in power system generation and infrastructure. The RPM is a forward market, based on capacity commitments three years ahead.

Over the past few years, demand response deployment in PJM has reached a potential to cover 6% of the system peak load, representing over 1 million end-use consumers. In PJM, participation of demand response in capacity markets can be activated in two different formats: either under an economic classification or under an emergency classification (Table 6.1). Any electricity consumer can participate in either or both depending on the circumstances.

[26] ISO-NE has fully integrated demand response in its forward capacity market since June 2010. ISO-NE differentiates between active demand response resources (real-time demand response within 30 minutes of receiving the ISO-NE dispatch) and passive demand response resources (summer peak hours or seasonal peak resources). As of June 2017, active demand response resources will also be eligible to participate in energy markets. A transitional period started in June 2012, where the participants can offer load reduction in response to day-ahead locational marginal pricing. Passive demand response is not eligible for energy markets in ISO-NE.

[27] In NYISO, demand response is eligible as an emergency resource when generation shortages put grid reliability at risk (in a programme termed ICAP Special Case Resource and Emergency DR Program). In this case, large consumers voluntarily commit to reduce their power consumption and receive compensation from NYISO (before or after the power cut, depending on the programme). Demand response is also eligible in the Day-ahead DR Program, where the load reduction is considered a "negawatt" and the remuneration is fixed by the market clearing price. Finally, demand response can also be a product on the ancillary services market, when a consumer can bid its load curtailment capacity into the real-time market and provide additional resources into operating reserves and regulation services. In this case, the scheduled offers are paid the market clearing price for this service.

Table 6.1 • Overview of demand response programmes in PJM

	Emergency Load Response Program			Economic Load Response Program
Market	**Capacity-only**	**Capacity and energy**	**Energy-only**	**Energy-only**
Capacity market	DR cleared in RPM	DR cleared in RPM	Not included in RPM	Not included in RPM
Dispatch requirement	Mandatory curtailment	Mandatory curtailment	Voluntary curtailment	Dispatched curtailment
Penalties	RPM event or test compliance penalties	RPM event or test-compliant penalties	NA	NA
Capacity payments	Capacity payments based on RPM clearing prices	Capacity payments based on RPM clearing prices	NA	NA
Energy payments	No energy payments	Energy payments based on submitted higher of "minimum dispatch price" and LMP. Energy payment during PJM declared "Emergency event mandatory curtailments"	Energy payments based on submitted higher of "minimum dispatch price" and LMP. Energy payment during PJM declared "Emergency event mandatory curtailments"	Energy payments based on full LMP. Energy payment for hours if dispatched curtailment

Notes: LMP = locational marginal price; NA = not applicable.

Source: Monitoring Analytics, 2014.

Emergency demand response, however, is by far the dominant product that clears on PJM capacity markets, alongside generation capacity, transmission upgrades and energy efficiency. This product is typically activated when the network system operator declares an emergency situation in respect of system capacity or reliability. It is commonly agreed that over a defined period (1 year), interruptions can happen a limited number of times (maximum 10 times per delivery year) and can last a maximum of 6 consecutive hours from May to September during non-holiday weekdays. For the 2012/13 delivery year, a total of 8 GW of demand response was activated from almost every segment (residential, commercial, industrial, government, education and industrial); this was roughly equivalent to 5% of the total peak demand experienced by PJM during this period.

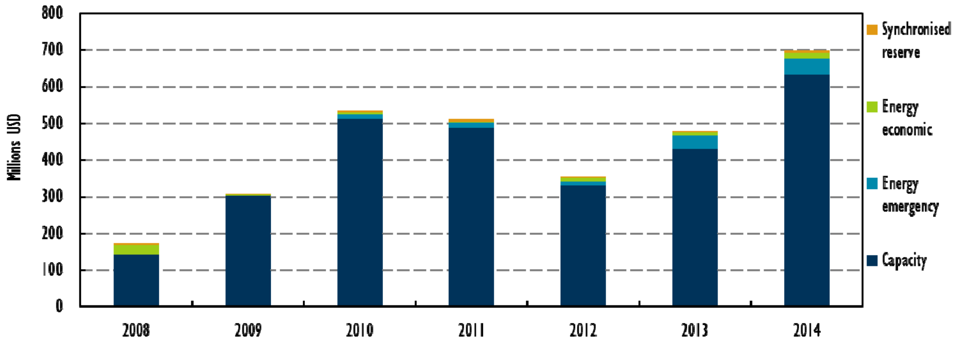
In the case of emergency demand response, the primary objective is to reduce the load on the grid system to ensure reliability of electricity supply. This scheme is considered a mandatory commitment where penalties could be applied to the customer in case of non-compliance. PJM has no direct interaction with the final electricity consumer. The interface is the Curtailment Service Provider (aggregator), who is responsible for identifying demand response opportunities for customers and installing the necessary equipment and processes.

Two other types of demand response programmes in PJM do not receive capacity payments (see Table 6.1). Emergency demand response can also be voluntary, but in this case the compensation will not be set on the capacity market basis.

The Economic Load Response Program corresponds to products activated for economic purposes on wholesale energy markets. The consumer voluntarily reduces its load during a certain time depending on wholesale prices. In this specific case, demand response is competing directly with generation resource. Under this framework, demand response can also provide ancillary services to the wholesale market under different forms: day-ahead scheduling reserve, synchronised reserve or regulation.

The introduction of the RPM has been key to the deployment of demand response in PJM. As Figure 6.5 shows, most of the revenue for demand response in PJM continues to be generated from capacity mechanisms.

Figure 6.5 • Demand response revenue by market in PJM 2008-14



Source: Monitoring Analytics, 2014.

In May 2014, the US Court of Appeals for the District of Columbia ruled on the Federal Energy Regulatory Commission (FERC) Order on Demand Response (Order 745). The court opined that FERC does not have the authority to mandate state-level payment rules for demand response providers. Though the Supreme Court has overtuned the ruling, it is an illustration of the importance of setting clear rules and responsibilities before the implementation of demand response technologies in order to avoid any form of regulatory gap.[28] The Court also criticised the FERC remuneration rule that requires PJM and other regional transmission operators to pay demand response resources market-clearing prices (full LMP), an issue to which this chapter now turns.

Box 6.2 • Level playing field for demand response illustrated in Belgium

Due to recent nuclear power plant failures and the long-term phase-out of its nuclear fleet, Belgium experienced the risk of electricity supply shortages during the winter of 2014/15. The estimated gap was around 1.2 GW of peak capacity during wintertime. Consequently, the Belgian authorities created a strategic reserve open to generation and demand response (with the requirement to contract a minimum of 50 MW from demand response). As a result of a tender, the TSO was able to contract 850 MW of capacity, of which 100 MW of demand response was retained among other sources. The majority of the 100 MW were made available by an aggregator who allied with around 50 Belgian companies, including large manufacturers. The ongoing risk of shortage is expected to persist in Belgium in future winters, and that has triggered a revision of the nuclear phase-out plan and a debate about revising the market design.

Participation in balancing and ancillary services markets

Demand response is also a solution to imbalances that electricity markets face in the very short term, and which are usually dealt with in the balancing and ancillary services markets. The balancing markets deal with short-term and temporary changes in supply/demand balance, mainly due to unforeseen changes in electricity demand or the unexpected breakdown of a generation facility or a transmission line, in particular after the markets have closed. The term

[28] PJM addressed this situation in light of the Base Residual Auction held in May 2015, by allowing load-serving entities and other Curtailment Service Providers (CSPs) to bid demand-side reductions into the auction and thereby reduce the volume of capacity resources needed by PJM in its capacity market. This so-called "stop-gap" solution proposed by PJM is based on the fact that demand response bids from load-serving entities would still be allowed in PJM capacity markets, but would stay on the demand side of the offering, and would therefore not be treated as a capacity resource. The remuneration of the capacity secured would therefore not come directly from wholesale markets. Instead the economics and incentives come directly from the avoided costs and obligations, as well as potentially from state programme and other financial incentive programmes. In other words, "if the demand response curtailment commitment is called to perform in the energy market, it may receive no additional energy market payment, but would avoid an energy payment for the demand reduced" (PJM, "The evolution of Demand Response on the PJM Wholesale market" p. 6). At the time of writing, PJM had not responded to the Supreme Court ruling.

ancillary services refers to a range of functions which TSOs contract to guarantee system security, such as black-start capabilities, fast reserve, reactive power and various other services including demand response (ENTSO-E, 2015).

Demand response relies on equipment that can react in an interval ranging from under a second to a couple of minutes. This is where the main difference lies between traditional demand response and demand response for the ancillary services market: the reduction in notification time, the speed and the accuracy of measurements (MacDonald et al., 2010).

Demand response can participate in all ancillary services markets that rely on four main types of product:

- Regulation: control of system frequency though instantaneous balance of supply and demand through automatic generation control, via a signal sent by the system operator's energy management system. For instance, industrial hydrolisers already provide such services.
- Spinning reserves: portion of remaining capacity (unloaded) from capacity units connected to the system and that can be delivered within 10 minutes.
- Non-spinning reserves: capacity that can be activated and can deliver within 10 minutes.
- Supplemental reserves: capacity that can be activated and can deliver within 30 minutes.

Nevertheless, not all forms of demand response are capable of providing reliable ancillary services. This will depend on the requirements set for the response timing, duration and whether the demand is expected to increase (RAP, 2013).

The only practical solution to providing value to demand response for its participation in the ancillary services market is to be able to dispatch and treat this resource as a generation asset (Figure 6.4). Indeed, it is difficult to imagine the possibility of having a response to real-time prices given the very short notification time (a few minutes or seconds) and the speed and accuracy of measurement needed. A series of elements need to be carefully addressed in the market design: financial compensation for demand response and market size.

Market prices have to be high enough to justify new entrants' consideration of offering demand response as an ancillary service for balancing purposes. In this regard, it has been observed that where market rules allow aggregation (which is not always the case) for participation in balancing and ancillary services markets, this can help to increase the scale of demand response such that it can compete meaningfully with other sources. On the same note, certain requirements for real-time metering (to allow the RTO to assess online demand response potential) might represent an additional and not inconsiderable expense for demand response operators, which could undermine their business case for entering ancillary services market with a large number of small customers.

Implementation challenges

The role of demand response aggregators

The large-scale development of demand response involving small customers and the participation of the residential sector call for a new form of service, which consists of adding together many individual demand response opportunities and aggregating them into products that can be offered into wholesale electricity markets. Aggregators can also act as intermediaries between the electricity supplier and the user, helping to optimise real-time electricity demand with the generation available.

Based on contractual agreements with consumers, aggregators can therefore act to partially switch on or off certain devices without any direct action from the final user. In this way, the aggregator offers the final energy consumer the possibility of making use of demand response without the burden and complexity of following real-time pricing. Aggregators are thus key to increasing the flexibility of the power system.

Unlike generators or consumers, however, aggregators do not have a physical point of connection to the grid. Their participation in electricity markets is based on access to the network, data management and real-time metering. It requires adaptation of the market and network access rules. These adaptations are needed to lower transaction costs and are usually specific to demand response.

In addition, the aggregator will have to gain access to and manage a large flow of data, from suppliers to users. This implies on the one hand that the power system is equipped with proper data management software, and on the other that data privacy will be ensured.

Setting the baseline

When demand response is considered a generation resource, the compensation for its activation needs to be calculated based on a quantum of energy that is not consumed. This calculation requires determination of a baseline corresponding to the electricity that would have been consumed in the instance that demand response had not been activated.

The baseline is key for all markets in order to:

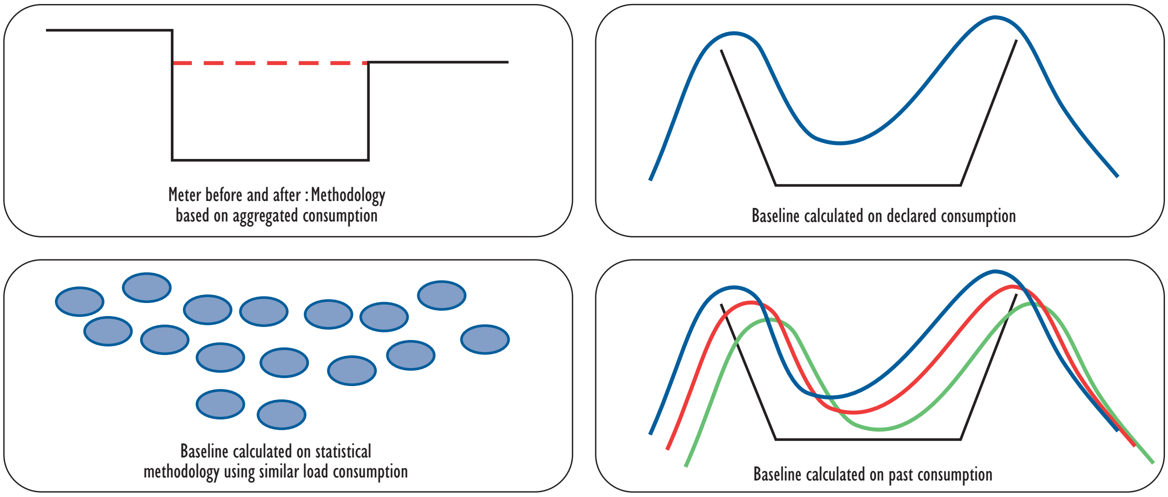
- determine the potential performance of demand responsed
- assess the actual performance of demand response in retrospect.

Several methods of calculating the baseline have been implemented or tested (Figure 6.6). While it is possible to define dozens of different methods, the most common remain simple and are based on comparisons with real metered data. The avoided energy consumed can either be estimated based on the consumption metered just before or just after. A further possibility is to use the effective consumption of a similar group of users or the same group's consumption at another time of the year under similar consumption conditions (season, temperature, weekdays vs. weekends, and so on). Another possibility might be to use declared consumption.

Some methods can lead to a relatively accurate statistical estimate of demand response, but usually require a large quantity of data and are complex to implement. Conversely, other methods are relatively simple to implement (for example, before and after or baseline calculated on past consumption). But they can lead to the over or under-estimation of actual demand response, so that consumers might be incentivised to change their consumption pattern in order to change the baseline. As quoted by Leautier (Crampes and Leautier, 2015), this was experienced with the managers of the baseball stadium in Baltimore – in order to increase their demand response volumes, and consequent remuneration, they artificially increased their baseline consumption by turning the stadium's lights on during the daytime.

In practice, the choice between these different methods remains a trade-off between the accuracy of the calculation, the availability of data and, last but not least, the simplicity of the method that needs to be explained to consumers. When choosing the methodology to calculate the baseline, it is also important to take into account the risk of gaming and wrongful behaviour. The compromise might lie in a contractual agreement between the demand response provider and the consumer to curb any incentive to overstate the baseline.

Figure 6.6 • Examples of methods used to calculate the baseline



Source: Based on RTE, 2014.

Depending on the method actually implemented, system operators also have to be aware that some demand response might not respond as expected. Aggregators usually address this issue by managing a large portfolio of demand response resources; statistically, is it possible to assess the probability of meeting the aggregated demand response activation with a high degree of confidence.

Demand response remuneration

The large-scale deployment of demand response as a generation resource in the energy markets will depend upon the level of financial compensation paid to the demand response provider.

To fully implement a level playing field for demand response, the reference point should, in principle, be wholesale market prices. However, particularly in the early stages of demand response, this price might not fully cover the investment necessary for substantial development of these technologies and the associated transaction costs. The question, therefore, for regulators and system operators is how to develop demand response remuneration without distorting the market with subsidies?

The double payment issue

A further potential distortive consequence of inadequate deployment of demand response is the issue of double payment and the question of missing revenue for the generator.

The double payment issue refers to the idea that the consumer might be over-remunerated for the electricity units that it has not consumed. If a consumer pays a fixed price per megawatt hour for electricity and decides to resell a portion of this electricity in the form of demand response at a peak electricity price, the benefit for the consumer is double: on one hand the retail bill savings made for the energy not consumed, and on the other a further benefit made from reselling this unconsumed energy on the market. So it potentially makes profit in two ways, by re-selling energy it does not pay for.

This can lead the final consumer to adopt behaviour that is not intended by demand response goals. In particular, consumers might reduce their consumption to receive a certain benefit while generation could be available at a lower cost.

A further consequence is the impact that this would have on the revenue of the supplier who is committed to delivering a predefined amount of electricity as per contractual agreement with the final consumer.

Compensation for the generator

When demand response is activated by an external operator or aggregator, the supplying company can indeed claim that a loss will be encountered. Based on a forecast of expected demand, the generator will produce a certain amount of energy to be delivered to the customer at a pre-defined price. If the aggregator activates demand response, that electricity will no longer be sold to the final consumer and, depending on the balancing mechanism in place, the generating company will be in imbalance and incur a cost on the balancing market.

In fact, consumers sell energy to aggregators, but in order to do that, they have to buy this energy from their supplier. Therefore the supplier has to be compensated in one way or another for the foregone revenues.

Clarification of the market rules is therefore necessary to ensure fair compensation for generators delivering electricity in a market with demand response aggregators. The compensation should not be equal to the marginal price, but instead a lower price to avoid over-remuneration of demand response (Chao, 2009).

6.3. Dynamic pricing

Dynamic pricing is a straightforward way of enabling demand response from small consumers. This section describes how retailers can provide a signal to electricity consumers that a change in their consumption pattern will have a bearing on their electricity bills. This signal is not provided by a flat tariff system. Flat tariff pricing will become less suitable in a decarbonised system that experiences greater volatility of marginal generation costs and electricity prices.

Dynamic pricing, or time-based pricing, refers to retail electricity prices that pass through at least part of the wholesale price volatility to final end users. One particular example is real-time pricing, where wholesale electricity prices are passed through to final consumers and bills are calculated based on hourly consumption that is metered. The concept behind dynamic pricing introduces the notion of linking prices to the variation in the marginal cost of generating electricity. This variation would, by definition, internalise a series of related costs, such as the variation in demand itself, the cost of generating and storing electricity, as well as balancing the system (Joskow and Wolfram, 2012).

Pricing has to be customer friendly

Traditional time of use (ToU) pricing poorly reflects the variations in wholesale electricity prices inherent in wind and solar power. Pre-specified ToU pricing is no longer able to reflect the dynamic nature of electricity prices in systems with high shares of variable renewables. Indeed, the marginal cost of production of electricity from variable sources can be quite low and therefore affect the wholesale price, creating hours of zero prices for several hours (see Chapter 3). Conversely, during times of low wind and sun, conventional sources need to run to meet the load, leading to a high price of electricity. To encourage the consumer to use electricity at a time when generating costs are lower, or to rely on shifting the demand, a dynamic price signal needs to be communicated.

In a ToU rate, fixed time periods are pre-set during which different electricity prices are applied. The classic example of ToU pricing is peak/off-peak electricity rates, where prices are lower during the night. This tariff system is intended to influence the consumption patterns of final consumers. ToU pricing, however, is simplistic in ways that poorly reflect the price differences across seasons. It is also not well suited to systems with a lot of variable renewable energy. Wind or solar output can radically change the traditional peak/off-peak price spread in wholesale electricity markets.

While real-time pricing of electricity is already the default option in several markets, it is still often thought to be too complex to introduce to smaller consumers. Real-time pricing exposes the consumer to actual wholesale prices, which can vary on an hourly basis. These prices more accurately reflect the cost of producing electricity at a specific time of day. In this regard, this rate is the purest way to provide a price signal to the final consumer, and therefore to incentivise consumption reduction at the most congested and expensive times. Real-time pricing is the default tariff option in a number of markets, for instance, for small consumers in Spain (see Chapter 9).

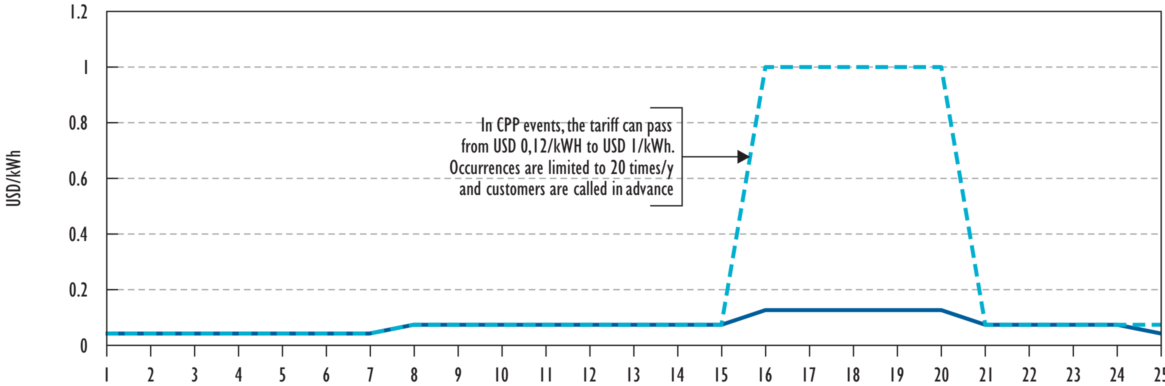
In practice, however, suppliers usually rely on simplified pricing structures, such as CPP or other forms of dynamic pricing that are easier to understand and anticipate. Consumers are not exposed to real-time prices, but are provided with simplified and easier-to-understand tariff structures. It is expected that competition between suppliers will lead to a further differentiation of their retail prices. Some suppliers might continue to offer a flat rate, while others might offer more or less sophisticated dynamic prices.

CPP is a special top-up rate at which electricity prices can substantially increase for the few days a year when wholesale prices are the highest. With this rate, utilities can incentivise the consumer to reduce energy consumption at times of peak demand. Such CPP events are usually limited in number and in duration. They are designed to respond to a risk of peak demand due to exceptional weather circumstances or to anticipate high peaks in energy (commodity) prices. During CPP events, the consumer is notified that higher peak prices will be applied during a certain period, and is invited to make the choice of reducing her consumption.

In the CPP configuration, the consumer either consents to subscribe to such a scheme or, when the tariff system is applied by default, has to pro-actively opt out of the system. In both cases, if the consumer is part of the project, he or she is notified several hours in advance that a CPP tariff will be applied. This scheme is particularly relevant at residential level with automated demand response technologies.

Figure 6.7 represents the electricity tariffs applied to the customers of DTE Energy in Michigan, the United States. Dynamic prices need to be ten times higher than normal prices to trigger a reaction. (This represents a price of USD 1 000/MWh.) Peak price rates can either be uniform or can vary from one peak to another, reflecting the fluctuation of wholesale prices and therefore acting as variable peak prices.

Figure 6.7 • Electricity tariffs during weekdays in DTE Energy

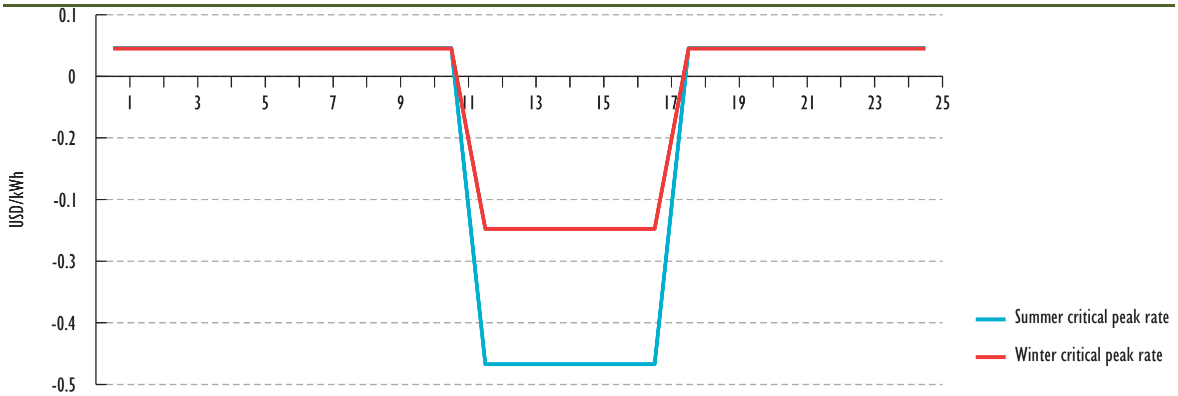


Note: kWh = kilowatt hour.

Source: IEA, DTE Energy website, accessed 2015.

The peak-time rebate rate is a variant of CPP. This retail tariff structure offers the possibility of the consumer being remunerated for load reductions during peak periods. For example, Figure 6.8 shows the tariff model for the customers of Xcel Energy in Colorado. As with CPP, such events are occasional. Calculating this rate can be subject to some uncertainties, however, considering that the financial compensation for not consuming electricity during peak periods assumes a baseline as its benchmark. A discussed previously, the complexity therefore lies in defining the right baseline to represent the "regular" electricity consumption at the same time of the day during another year.

Figure 6.8 • Peak-time rebate tariff in Colorado with Xcel Energy (simulation)



Source: IEA, Xcel Energy website, accessed 2015.

The potential for dynamic pricing remains limited, but there is no need for all consumers to respond to prices. Analysis suggests that even if fewer than 20% of customers opt for dynamic pricing, this would reap most of the benefits of demand response (Borenstein, 2011).

Challenges with dynamic pricing in restructured electricity markets

Despite their theoretical appeal, dynamic pricing offers are not well developed in practice, even in electricity markets with retail competition. Worse still, restructured electricity markets can lead to a decline in demand response brought about by historical CPP programmes (Box 6.1).

Several reasons explain why dynamic pricing fails to develop: wholesale price volatility is relatively low; there are no (or very few) extreme peak or negative prices; or the network tariff structure does not always allow dynamic prices.

Price volatility and infrequent extreme prices

Most electricity markets present a limited day/night price spread. One of the reasons for that is the development of solar PV, which reduces net demand and wholesale prices during the day, thereby reducing the day/night price differential. As a result, the possibility of shifting consumption from high-price hours to low-price hours within one day or one week offers limited gains. But this situation is expected to be transitory, and wholesale price volatility will increase with larger shares of variable renewables.

Demand response to high-peak prices is also hindered because price spikes are very infrequent and not high enough in most markets. In Europe, the last peak price recorded occurred in France for two hours in February 2012, reaching a level of EUR 2 000/MWh. In PJM, prices reached USD 3 000/MWh during the polar vortex of winter 2014.

The implementation of administrative scarcity pricing could contribute to the development of price-based demand response. As discussed in Chapter 4, better scarcity price formation will be an essential building block of well-functioning markets. If the price reached a peak of USD 10 000/MWh, a residential consumer could save USD 10 if demand were shifted or the load reduced by 1 kW for one hour, corresponding for example to the consumption of an electric heater. But as discussed in Chapter 4, these prices rarely materialise.

Similarly, negative or zero prices also rarely occur. In Germany in 2012, the day-ahead market registered negative prices for 56 hours over 15 days, and 41 hours over 10 days on the intraday market. To date, this frequency has been insufficient to justify investment, for instance, in electric boilers that could capitalise on these low price episodes.

Indeed, an important implementation issue relating to dynamic pricing is that customers need some predictability on the frequency of extreme price episodes. In order to invest in equipment that can respond to high electricity prices, consumers need to be able to assess the revenues or possible savings each year on average. If extreme prices materialise too rarely, it is likely that the demand response will be uneconomic and will not develop.

Suppliers do not usually expose consumers to real-time prices, and when they implement dynamic pricing, they rely on simplified tariff structures. To work properly, such dynamic pricing would have to be somehow predictable for consumers and occur on a regular basis. Suppliers can offer a tariff option that smooths the variability of electricity prices, and offers a predictable number of hours during which consumers will be able to actually reduce their bill, even if this does not correspond to system needs during a specific year. Acting in this role of intermediary between wholesale markets and retail markets, suppliers can, for instance, help define dynamic pricing options by adopting a statistical approach.

Unbundling and allocation of network costs

The currently limited deployment of dynamic pricing also results from increased difficulty in co-ordinating the price structure of different segments of the power sector in unbundled electricity systems. The final retail price has to be calculated as the sum of the network tariff, the wholesale price and the supplier margin/benefit. As network tariffs do not usually reflect the time variation of generation costs, the dynamic price can only come from the time variation of wholesale prices, which reduces the incentive for consumers to participate in dynamic pricing.

On average, wholesale electricity typically represents less than 50% of the final electricity bill. In order to incentivise demand response, dynamic pricing can only reflect the energy component of retail electricity tariffs. The resulting differential between expensive and inexpensive hours is not high enough to incentivise customers to react to prices.

The structure of network tariffs is a further reason why dynamic pricing is less developed in competitive electricity markets. The paradox is that vertically integrated regulated utilities have much more flexibility to allocate network costs and so can implement demand response more easily. Regulators, however, consider that such allocation of network costs can create cross-subsidies in favour of responsive consumers, and are usually reluctant to define special network tariffs for dynamic pricing and CPP that correspond to generation.

To address this issue and restore price signals favourable to demand response, it might be necessary to enable suppliers to choose the allocation of network costs across different categories of consumers. These aspects should be carefully assessed during the revision of tariff structures and analysed in relation to future investments in network capacity (see Chapter 9).

Conclusion

The path to decarbonisation will require the deployment of new technologies such as demand response. These technologies will, on the one hand, help the system to be reactive to the variability of wind and solar power, and therefore support their integration, and on the other also empower the final consumer to benefit from their flexibility.

To create a market infrastructure where demand becomes reactive to wholesale electricity prices, the final consumer should have the option to react to these prices, either directly for large customers or indirectly via the intermediation of suppliers. Therefore, the implementation of dynamic pricing should be encouraged.

In addition, the participation of demand response in capacity markets has been effective in providing revenue certainty for demand response aggregators. In this instance, demand response can be "dispatched" as if it were a generation power plant, although this has raised complex implementation issues and defining the right regulatory framework has proved to be difficult.

Lastly, consumer rights and protection should also be carefully assessed in order to build trust and understanding from the consumer, who will be the decisive actor in a successful system featuring large-scale demand response.

References

Borenstein S. (2011), "Effective and equitable adoption of opt-in residential dynamic electricity pricing", National Bureau of Economic Research, www.nber.org/papers/w18037.pdf, accessed on 13 October 2015.

CEER (Council of European Energy Regulators) (2012), *CEER Benchmarking Report on Meter Data Management Case Studies*, Council of European Energy Regulators, www.energy-regulators.eu/portal/page/portal/EER_HOME/EER_PUBLICATIONS/CEER_PAPERS/Customers/Tab3/C12-RMF-46-05_BR_MDM_07Nov2012.pdf.

Chao H. (2009), "An Economic Framework of Demand Response in Restructured Electricity Markets", Harvard Kennedy School, Harvard University, www.hks.harvard.edu/hepg/Papers/2009/Demand%20Response%20in%20Restructured%20Markets%2002-08-09.pdf, accessed on 13 October 2015.

Crampes C. and Leautier T. L. (2015), "This disappearing demand response bonus", Toulouse School of Economics Debate, http://debate.tse-fr.eu/article/disappearing-demand-response-bonus?language=en, accessed on 13 October 2015.

Cukier K. (2015), "Big data and the future of business", *MIT Technology Review*, www.technologyreview.com/view/538916/big-data-and-the-future-of-business/, accessed on 13 October 2015.

DTE Energy (2015), "Rate programs", Payment and Usage, www.newlook.dteenergy.com/wps/wcm/connect/dte-web/home/billing-and-payments/residential/rates-and-meters/rate-programs.

ENTSO-E (2015), "Balancing and ancillary services markets", www.entsoe.eu/about-entso-e/market/balancing-and-ancillary-services-markets/Pages/default.aspx.

Hogan W. (2010), "Demand response pricing in organized wholesale markets", www.hks.harvard.edu/fs/whogan/Hogan_IRC_DR_051310.pdf, accessed on 13 October 2015.

IEA (International Energy Agency) (2014), "Technology Roadmap Energy Storage", *Energy Technology Perspectives*, OECD/IEA, Paris, www.iea.org/publications/freepublications/publication/TechnologyRoadmapEnergystorage.pdf, accessed on 14 October 2015.

Joskow P. L. and Wolfram C. D. (2012), "Dynamic pricing of electricity", http://faculty.haas.berkeley.edu/wolfram/papers/AEA%20DYNAMIC%20PRICING.pdf, accessed on 13 October 2015.

MacDonald J. et al. (2010), "Demand response providing ancillary services: A comparison of opportunities and challenges in the US wholesale markets", The GridWise Architecture Council (GWAC), www.gridwiseac.org/pdfs/forum_papers12/macdonald_paper_gi12.pdf, accessed on 13 October 2015.

Monitoring Analytics (2014), *PJM Market Monitoring Report 2014*, www.monitoringanalytics.com/reports/PJM_State_of_the_Market/2014.shtml

PJM (2015), "The role of demand response in electric sector decarbonisation", presentation at EPRI-IEA Challenges in Electricity Sector Decarbonisation Expert Workshop: Integrating New Technologies While Maintaining Resource Adequacy, 28 September, Paris, www.iea.org/media/workshops/2015/esapworkshopvi/Helm.pdf.

RAP (2013), *Demand Response as a Power System Resource: Program Designs, Performance, and Lessons Learned in the United States*, Regulatory Assistance Project, Synapse Energy Economics Inc.

RTE (2014), "Demand response settlement and market design", presentation by F. Galmiche at EDF Workshop on Demand Response, Clamart.

Strbac G. (2008) "Demand side management: Benefits and challenges", *Energy Policy*, Vol. 36, Issue 12, December 2008, Pages 4419–4426.

Think (2012), "Electricity storage: How to facilitate its deployment and operation in the European Union", Final Report, June 2012, European University Institute www.eui.eu/Projects/THINK/Documents/Thinktopic/THINKTopic8online.pdf, accessed on 14 October 2015.

Xcel Energy (2015), "RPTR residential – Peak time rebate (closed), webpage: "Xcel Energy – Colorado", www.whatsmypower.com/locations/80161/rates/84143.

Chapter 7 • Interconnected transmission networks

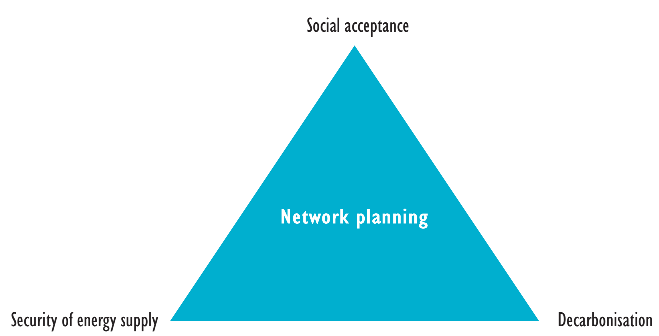
HIGHLIGHTS

- Transmission grids and new cross-regional interconnectors are often a cost-efficient alternative to investing in other resources when seeking to reduce the amount of new conventional generation needed to balance a system with high shares of wind and solar power.
- To facilitate cross-border investment, stronger co-ordination and co-operation in cross-regional resource adequacy assessments and network development plans are essential.
- Given the scale and complexity of large interconnected regions, resource adequacy assessments and network planning require processes that are both bottom-up and top-down.
- The identification and quantification of costs and benefits need to be developed jointly in the respective regions and involve all relevant market participants, including generators, demand response operators and consumers.
- The principles for allocating the cost of inter-regional network investments should be carefully crafted, as different jurisdictions will usually look at their own benefits rather than the overall welfare of a broader area.

Meeting decarbonisation targets requires a strong and reliable transmission network to connect load centres with wind and solar resources, as well as to transport electricity between countries or states. Although renewables are not a decentralised resource *per se*, in many cases good wind and solar sites are not located close to load centres, so that tapping these resources necessitates new electricity transmission lines.

Networks are also important to ensure security of supply and to generate electricity at least cost (Figure 7.1). While variable resources such as wind and solar now generate an impressive amount of electricity, they are typically not as dependable as conventional energy generation. Interconnectors remain by far the most cost-efficient solution to combining a high share of variable renewable energy (VRE) with the need to maintain the highly secure electricity supply that is enjoyed in countries of the Organisation for Economic Co-operation and Development (OECD). Transmission lines reduce the need for new generation capacity to balance the system.

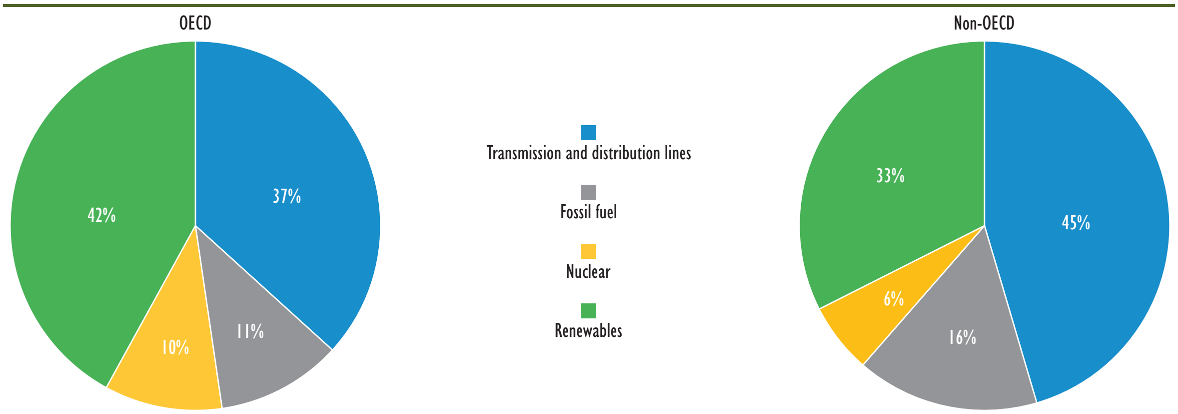
Figure 7.1 • Trilemma of transmission and distribution network planning



It is notable that networks will need to attract a scale of investment capital similar to that which is required for renewables. Under the New Policies Scenario (NPS), the IEA *World Energy*

Investment Outlook 2014 (IEA, 2014a) envisages that USD 546 billion will need to be invested in transmission networks in OECD countries by 2035. About 55% of this will be used for refurbishment, while less than 15% should be specifically related to renewables. Costs for transmission and distribution lines amount to 37% of the total investment envisaged in renewables, conventional generation and networks in OECD countries (Figure 7.2).

Figure 7.2 • Investment under the NPS 2015-40 (%)



Source: IEA, 2015.

It is often said that the electricity market should function without any constraints or congestion. In principle, in such a market, the network would enable the hourly price for electricity to be level over a large geographic area. In practice, however, reaching this goal would be extremely expensive and difficult to implement because of the large number of transmission lines needed. Costs and benefits of transmission lines need to be carefully assessed, especially in light of overall policy goals.

Several factors affect the unfettered expansion of transmission networks. While most OECD countries have been developing mechanisms to foster investment in national transmission lines, building new lines typically takes six to ten years, and sometimes longer, mainly due to permitting and licensing processes. Governments often face local opposition to building new transmission lines (by the so-called NIMBY and BANANA effects).[29]

The slow process of network reinforcement may also create increasing difficulties for meeting decarbonisation targets. In some regions, electricity generation from renewable sources already needs to be reduced at certain times (curtailment) because the electricity cannot be distributed to a wider region. Curtailment is deemed to be efficient to a certain degree. In Germany, for example, onshore wind operated at its maximum capacity for 220 hours, comprising just 10% of total annual wind generation (see Chapter 8). Nevertheless, the need for curtailment should be considered in network planning and not be an effect of slow network licensing.

Investment in transmission already poses a huge challenge within countries and states, but an even greater challenge lies in improving interconnections between them. This is the focus of this chapter. Interconnection projects involve multiple parties, each of which in many cases takes the perspective of its own region or country rather than that of overall efficiency and social welfare. While encouraging developments in fostering cross-border investment can be seen in OECD regions, there is room to improve interregional network planning and regulation to achieve decarbonisation at least cost.

[29] NIMBY is an acronym for Not In My Back Yard, and BANANA stands for Build Absolutely Nothing Anywhere Near Anything.

7.1. Networks as the backbone of the electricity market

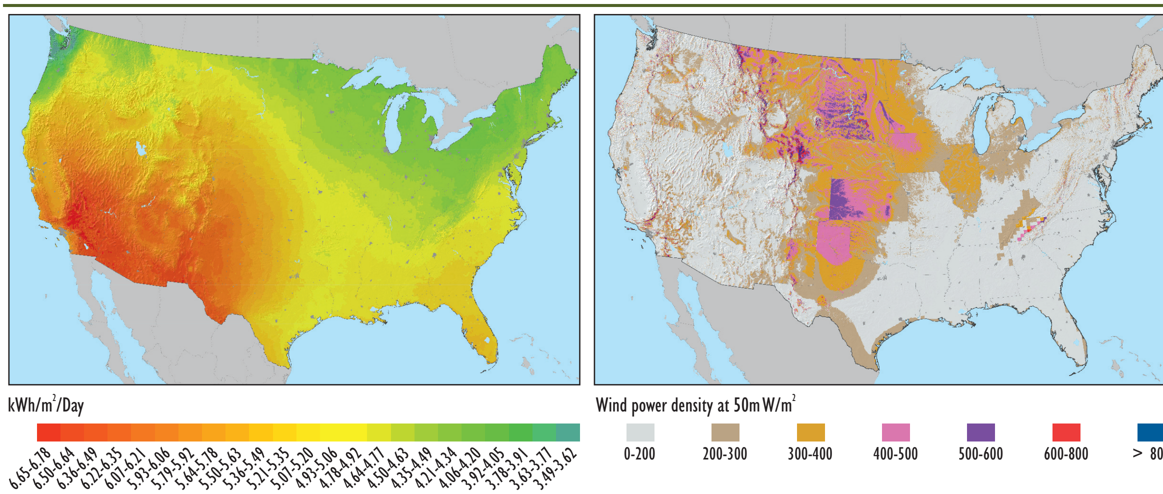
Transmission networks are fundamental to the development of electricity markets over large geographic areas and the further integration of markets across borders. The degree of market integration differs widely across OECD countries, reflecting diverse institutional and regulatory frameworks. The case for the construction of new transmission lines typically rests on reliability and efficiency considerations. Building on previous publications, *Seamless Power Markets* (IEA, 2014b) and *Electricity Networks: Infrastructure and Operations* (IEA, 2013), this section lays out the main arguments for increased interconnection in the majority of OECD countries.

Decentralised vs. centralised energy production

Contrary to common opinion, the development of low-carbon renewable generation is not decentralised *per se*, but depends on the location of natural resources such as wind and sunshine. Conditions for wind and solar power are not evenly distributed, and countries tend to deploy variable renewables more rapidly in windy and sunny locations, which are not necessarily close to load centres. In Europe, for instance, transporting the power generated along the shores of the North Sea to major load centres in the respective coastal states brings a significant need for investment in transmission. These patterns increase the benefits of market integration over larger geographic areas because cross-border trade contributes to reducing the overall cost of the electricity system by exploiting the complementarities between demand patterns and cost differences between electricity systems.

As shown in Map 7.1, conditions for photovoltaics (PV) in the United States are particularly good in the southwestern parts of the country, including parts of Texas, New Mexico, Arizona and California. For onshore wind, parts of Texas and the Midwest offer the greatest potential resources. Various studies envisage greater use of transmission lines and interconnectors to utilise these resources at least cost, for example a scenario by the US National Renewable Energy Laboratory (NREL) whereby wind energy could contribute 20% of US electricity supply (NREL, 2008).

Map 7.1 • Comparison of solar (left-hand map) and wind (right-hand map) resources in the United States



This map is without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries, and to the name of any territory, city or area.

Source: NREL, 2012.

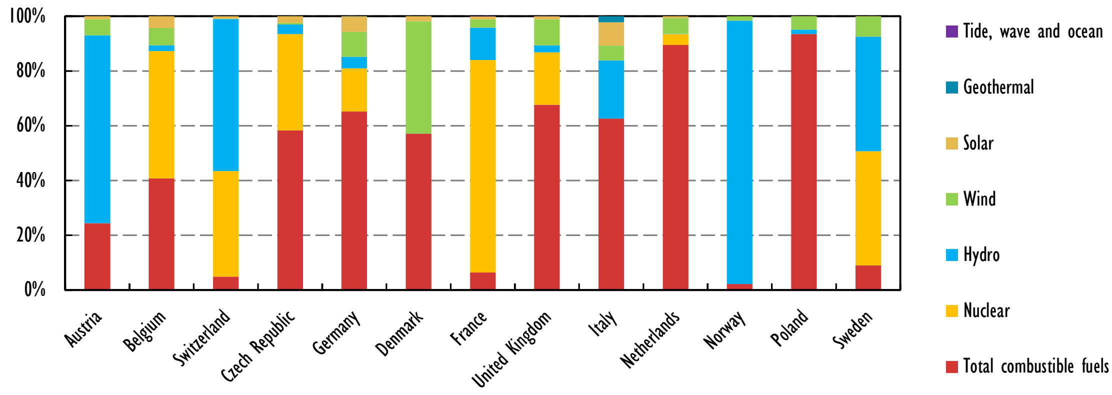
Reduction of generation costs

It is well established that market integration offers savings in overall dispatching costs. The least-cost solution to meeting demand during a certain hour is to start with the cheapest generating sources (wind, solar, run-of-river hydro and nuclear) and then call on other units in order of increasing marginal cost. The overall generation cost is therefore lower when dispatching over a broader and more diversified portfolio of plants.

Technology mixes differ substantially across Europe, a heterogeneity that creates many opportunities for trade between these countries. In Poland, the Netherlands, Italy, Great Britain and Denmark, more than 60% of energy production was generated from fossil fuel in 2013 (Figure 7.3). At the other extreme, Norway, Switzerland and Austria have considerable hydro capacity, which is used for electricity production. France, Belgium, Sweden and the Czech Republic have significant nuclear capacity. Germany and Denmark account for most wind power, and Germany and Italy for most solar power. This diverse generation profile results from differences both in national energy policies and natural endowments. And while energy policies in different countries often pursue similar objectives, the actual technology mix tends to diverge.

Transmission infrastructure can smooth out the variability of wind and solar power across large geographic areas. The aggregated load factor of renewables over large areas, in terms of percentage of peak generation, is higher than the individual load factor of one specific plant.

Figure 7.3 • Share of energy produced, December 2014



Source: IEA, 2015.

Once installed, generating capacity will last for decades, even if over time it becomes less optimal. Increasing interconnections and electricity trade to capitalise on differences in fuel costs is therefore generally beneficial, as it reduces overall costs and increases security of supply.

Looking at the diversity of power sources across North America, Canada has substantial hydro capacity, while in the United States, the Powder River Basin in the centre provides cheap coal, and the East Coast generates significant nuclear energy. The Midwest, for its part, has better wind resources, and the desert zones of Arizona and New Mexico are the best locations for solar power.

However, the development of shale gas is rapidly changing the energy arena in North America. Ubiquitous shale gas reserves, together with massive investment in the federal pipeline network, have removed bottlenecks and led to a convergence of US natural gas prices. As a result, the different electricity markets are choosing to generate electricity from gas at similarly low prices. This raises the issue of properly co-ordinating gas and electricity infrastructure since transporting gas through pipelines can be less costly than “gas by wire”.

System security

Electricity systems in all OECD regions have developed strong technical standards and norms to avoid the problem of frequency deviations, which can damage power system equipment and potentially lead to cascading blackouts. Today, as the share of wind and solar power in the energy mix increases, flows of electricity across borders tend to become more volatile and difficult to predict. This increases the complexity of managing cross-border flows and the trade in electricity.

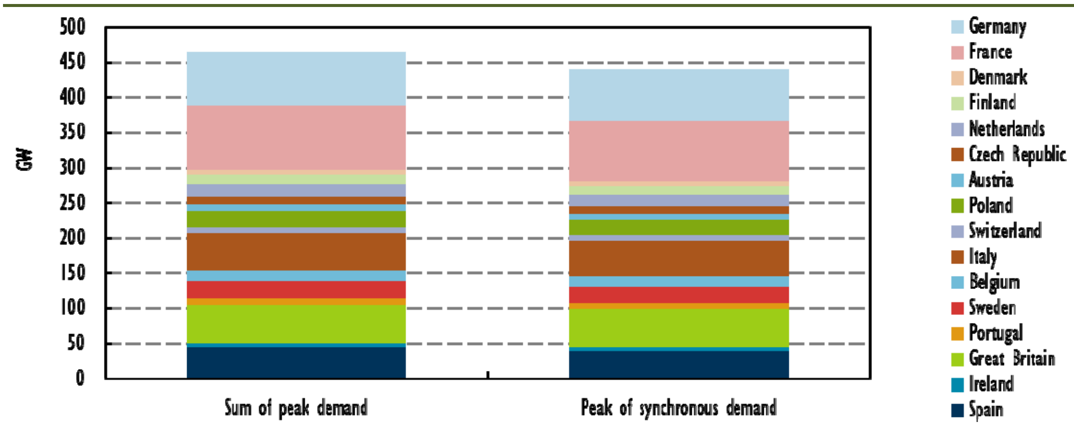
Experience shows that a lack of co-ordination among system operators is at the root of almost all major blackouts in OECD country systems. In the United States, the Great Northeast Blackout in 1965 led to the creation in 1968 of the North American Electric Reliability Corporation (NERC) to ensure the reliability of the North American bulk power system. In continental Europe, co-ordination among adjacent control areas started with the creation of the Union for the Co-ordination of Production and Transmission of Electricity (UCPTE) in 1951, which is now incorporated in the European Network of Transmission System Operators for Electricity (ENTSO-E). More recently, Coordination of Electricity System Operators (CORESO) in 2006 and the TSO Security Cooperation (TSC) in 2008 were created to support several national transmission system operators (TSOs) with wider and often closer to real-time awareness of the physical status of transmission grids across borders.

Despite some challenges associated with system security, electricity market integration offers substantial benefits from diversified supply sources and decreased costs of maintaining adequate generation capacity. Strong co-ordination among system operators is required to maintain system security over such large synchronous-frequency areas.

Regional adequacy

Interconnecting geographic areas helps to pool the expensive capacity resources required to ensure resource adequacy and maintain reserve margins. Ensuring access to a broader portfolio of power plants makes it easier to find the capacity needed to replace a power plant when it becomes unavailable due to planned maintenance, an unscheduled outage or safety concern. ENTSO-E further assumes that enhanced market integration will increase equalisation of energy prices across Europe and reduce bulk power prices by between EUR 2 and EUR 5 per megawatt hour (MWh) (ENTSO-E, 2014). This reduces the cost of maintaining adequate capacity, thereby increasing the reliability of the electricity system.

Figure 7.4 • Peak demand in ten European countries



Note: GWh = gigawatt hour.

Source: ENTSO-E, IEA.

Maximum electricity demand usually occurs at different times in neighbouring regions. Northern Europe and Canada experience peak demand in winter thanks to electric heating, whereas southern Europe and the United States experience summer peaks due to air conditioning. In central European countries, synchronous peak demand was 5% lower in 2013 than peak demand of each country taken separately (Figure 7.4), which amounts to 30 gigawatts (GW) for the region. Even though the full reduction of peak demand is not achievable, as perfect interconnection would be necessary, reinforcing interconnectors and national networks can enable the further pooling of generation.

In Europe, the Pentalateral Energy Forum[30] published an analysis of supply security for the Pentalateral region for winter 2015/16 and winter 2020/21. It clearly depicts the increasing importance of interconnection, even for an already well-interconnected area, due to the projected increase in VRE. Co-ordinating electricity resources and demand in connected areas can reduce the measured loss of load expectation (LOLE) quite significantly. In France, for example, an interconnected system reduces LOLE from 217 hours to 14 hours for 2015/16 as compared to an isolated system (Pentalateral Energy Forum, 2012).

Table 7.1 • LOLE in the Pentalateral Region in two periods

LOLE (hours) Pentalateral Region				
	2015/16		2020/21	
	Isolated	Interconnected	Isolated	Interconnected
Belgium	177	0	308	0
France	217	14	151	6
Austria	0	0	3	0
Switzerland	1 251	0	1 086	0
Germany	1	0	0	0
The Netherlands	0	0	32	0
Luxembourg	8 760	0	8 760	0

Source: Pentalateral Forum, 2015.

Regional analysis of resource adequacy assessments is also a suitable basis for network planning, to use regional effects over the long-term and ensure security of electricity supply in the connected regions. Such adequacy assessments have to properly factor in network constraints, not only on interconnectors but also within a region, so as to provide a comprehensive view of possible future requirements for grid reinforcement and generation capacity.

It is important to note that adequacy is no longer a deterministic notion in electricity systems with increasing amounts of VRE. Interregional adequacy analysis should reflect the stochastic nature of electricity demand and water, wind and solar power generation (see Chapter 4). The specific results of such probabilistic simulations depend on many assumptions about the exact shape of the probability distribution, utilisation factors and correlation of weather conditions across countries in specific hours.

[30] An intergovernmental initiative between Belgium, the Netherlands, Luxembourg, France, Germany, Austria and Switzerland.

7.2. Current state of play

Early development of interconnectors can be traced back to the 1920s. At this time it became clear that connections between utility systems could provide additional reliability with access to generation reserves in times of equipment failure, unexpected demand or routine maintenance. They also offered cost savings through reserve sharing and access to diverse and lower-cost energy resources to increase security of supply. These interconnectors were mainly built under bilateral or multilateral arrangements and long-term contracts between integrated utilities and governments. The development of these transmission lines resulted in increasingly large interconnected electricity systems with a synchronous frequency (50 hertz [Hz] or 60 Hz) (NREL, 2012).

North America

Five frequency areas currently exist in North America: Western Interconnection, Eastern Interconnection, Texas Interconnection, Alaska Interconnection and Quebec Interconnection. The different interconnections are not synchronised, precluding the use of alternative current (AC) interconnectors and limiting the level of physical interconnector capacity to direct current (DC) lines. To date, only a few DC lines with about 2 GW of interconnector capacity exist between the Western and Eastern Interconnections, and another interconnector of 2.6 GW capacity between the Eastern and Texas Interconnections.

Box 7.1 • Reinforcing the Eastern Interconnection

On behalf of the DOE, NREL assessed transmission needs across the footprint of various system operators within the Eastern Interconnection. The *Eastern Wind Integration and Transmission Study* (NREL, 2011) analyses the transmission and interconnector developments required by wind energy penetration of 20% to 30% by 2024, and the operational impact on the power system. The study was the first of its kind for the United States. As a follow-on, NREL is assessing the impact of various wind and solar deployment strategies and operational paradigms on system operation.

The *Eastern Wind Integration and Transmission Study* used three scenarios and varying wind locations and technologies to assess the interconnection requirements for integrating wind with curtailment of between 7% and 1%. Scenario 1 produced the greatest interconnector needs across multiple jurisdictions, as it aimed to capture the best onshore wind resources in the remotely located Great Plains. By contrast, Scenario 3 aimed to harness mostly offshore wind, only filling residual wind requirements with onshore wind development close to demand centres. Scenario 2 is a hybrid scenario, using a balance of wind resources located both onshore and offshore.

The study indicates the lowest power system costs under the first scenario, favouring multi-jurisdictional interconnection, transmission upgrades and increased trade flows over local scenarios and/or strong offshore wind scenarios. Calculated network economics favour an overlay grid as opposed to incremental build-out of the existing system, comprising AC circuits of up to 765 kilovolts (kV) in combination with 400 kV and 800 kV DC architecture. However, only the first two scenarios indicated positive net benefits, expressed in benefit/cost ratios of 1.22 for Scenario 1 and 1.09 for Scenario 2, whereas in Scenario 3, the production cost savings did not exceed the added transmission costs.

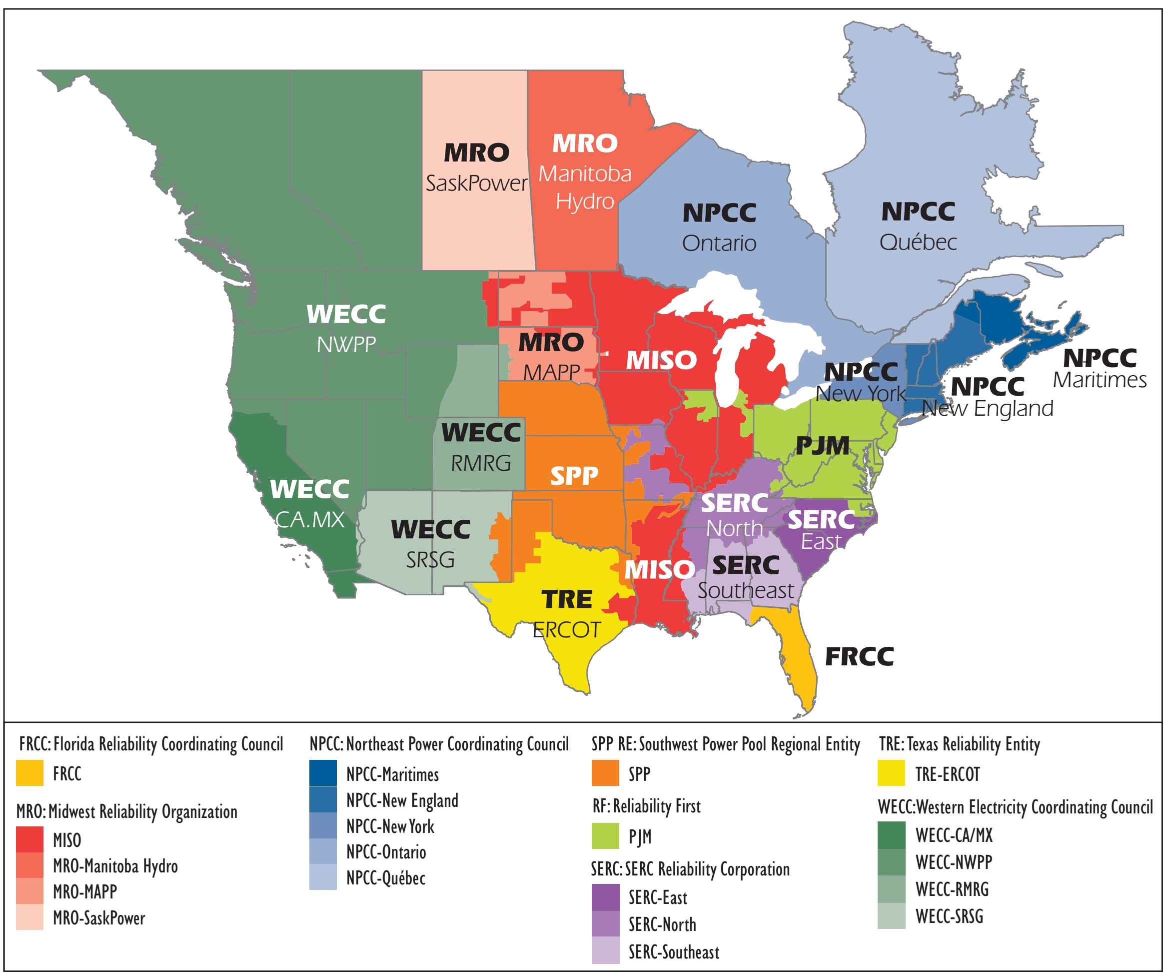
According to the US Department of Energy (DOE), investment in transmission is expected to increase in the United States. Investor-owned utilities spent a record USD 16.9 billion on transmission in 2013, up from USD 5.8 billion in 2001. The rate of investment has been low over the last two decades, at around 1 000 circuit miles per year from 1990 to 2010. But new lines account for about half of total investment, the remaining expenditure being for station equipment fixtures, towers and underground lines (DOE, 2015).

It is widely recognised that fragmented and overlapping jurisdictions threaten to impede the development of the grid. In North America, federal, regional and state institutions and regulatory structures are increasingly overlapping (Map 7.2). Compounded by the physical complexity of the grid, this creates huge institutional challenges for the development of new interconnectors.

For instance, there is no interconnector between the Texas Interconnection and the Western Interconnection. Compared with the overall installed generation capacities within each, the interconnector capacities are almost negligible.

Moreover, most interconnections in the United States and Canada are further broken down into smaller areas where a single authority is responsible for independent system operation and transmission planning. The largest areas are operated by ISOs or RTOs, often spanning multiple states and continue to develop market-based measures to supply roughly two-thirds of US and one-third of Canadian electricity demand. Transmission lines between these regions (also known as interties) have the same role and face the same issues as interconnectors, as they go beyond a single system planning jurisdiction and sometimes even span several states and power system policies.

Map 7.2 • RTOs, ISOs and NERC regional entities



This map is without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

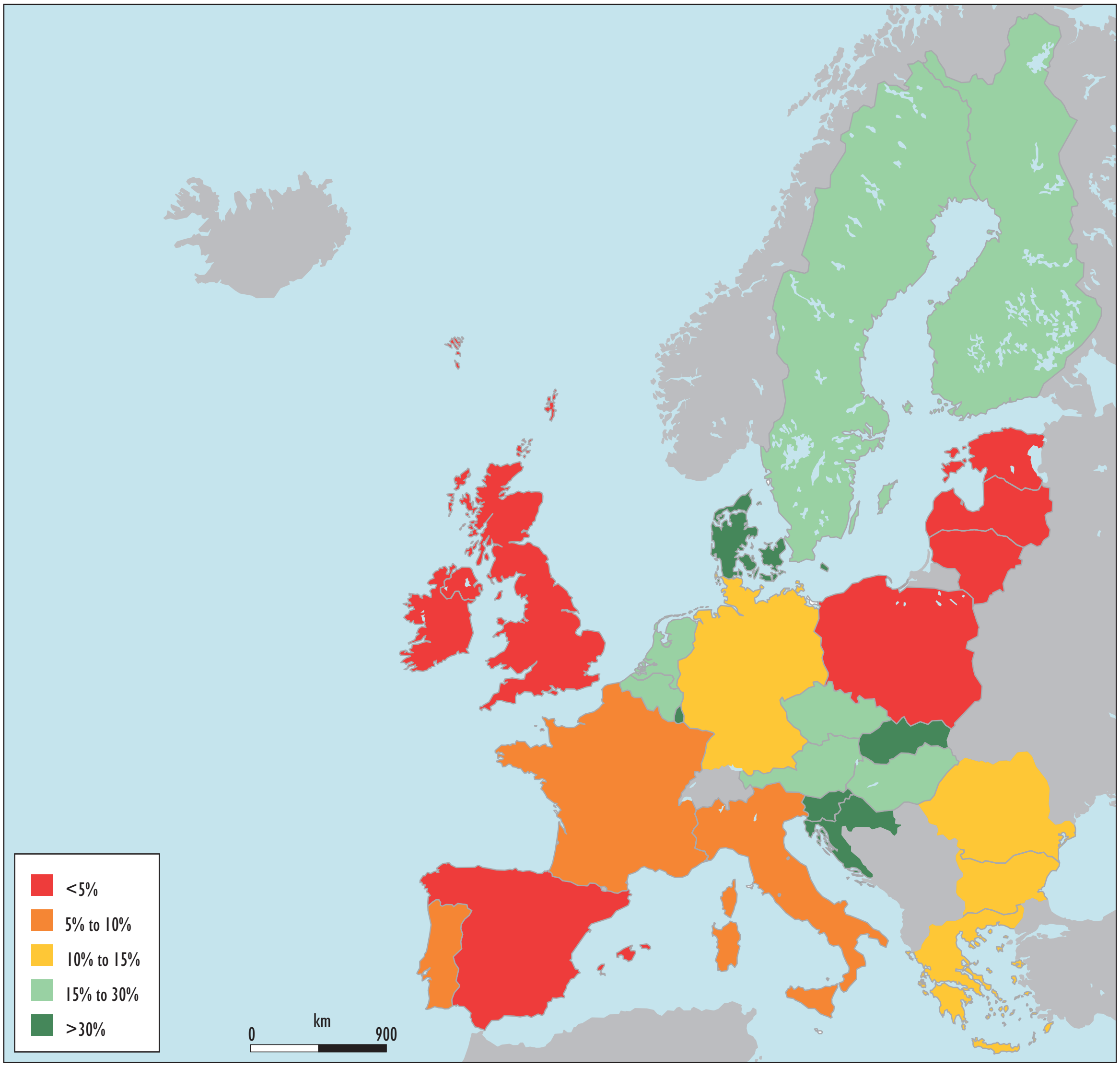
Source: IEA, 2014d.

Europe

In Europe, interconnectors between countries have created a large synchronous frequency area extending into the eastern parts of continental Europe at a frequency of 50 Hz (Map 7.3). Interconnectors amount to 11% of installed generation capacity across European countries. However, regional differences exist – in the Baltic States, for example, there is a significant need for interconnectors to increase security of supply and reduce the market power of generators. A better interconnected European energy grid would bring notable market benefits to European citizens, as consumers could save between EUR 12 billion and EUR 40 billion annually by 2030 (Booz & Co, 2013).

In 2014, the European Council discussed implementing a 15% goal for interconnection between member states in the European Union. While this goal would bring visibility to the issue, the costs and benefits of interconnectors need to be thoroughly assessed, not only from an investment perspective, but also for public acceptability and understanding, which are required for transmission lines to actually be built. Public acceptance can only be gained with thorough cost-benefit analysis to demonstrate the positives of the project. A project being built only to fulfil a percentage goal will face difficulties being accepted.

Map 7.3 • EU interconnection levels in 2020 after completion of current projects of common interest



This map is without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

Source: IEA, 2015.

Box 7.2 • Ten-year network development plan

ENTSO-E publishes a biannually updated Ten-Year Network Development Plan (TYNDP) to give an overview of the transmission expansion plans that are identified as necessary to facilitate EU energy policy goals.

The TYNDP 2014 analyses the required transmission and interconnector developments under different scenarios, termed "Visions", with renewables penetration levels of between 40% and 60% in 2030.

Vision 1 showed low interconnector and transmission needs on the basis of a low renewables penetration at 40% share. Vision 4 looked at a share of 60% renewables in the European energy system, leading to large power flows over greater distances across Europe. The vast majority of the proposed investments address renewables integration issues, either direct connection or network corridors to transport power to load centres.

The TYNDP indicates that interconnection capacity all over Europe should double by 2030 to deliver social and economic welfare, characterised by ENTSO-E as the ability to reduce congestion between two electricity markets to trade power in an economically efficient manner. Increased transmission capacity between two bidding areas reduces electricity costs for consumers in the higher-priced area and ensures that electricity is generated at cheapest cost. Of course, economic welfare depends on the scenario ENTSO-E envisages for the price for electricity.

Renewable integration is also affected by transmission grid enforcement. Depending on the scenario, between 44% and 80% of the projects in the TYNDP lead to an increase in renewables being available in the energy system, either by enabling new connections or by reducing congestion in the network.

Source: ENTSO-E, 2014.

Australia

In Australia, six jurisdictions (Queensland, New South Wales, the Australian Capital Territory, Victoria, South Australia and Tasmania) agreed to establish the competitive National Energy Market (NEM), enacted through the National Electricity Law in 1996. The NEM came into operation in 1998. Subsequently, Queensland became physically interconnected with the NEM in 2000/01, thanks to two transmission lines. Tasmania joined the NEM in 2005. In April 2006, a high-voltage DC submarine interconnector cable from Tasmania to Victoria was completed as a merchant investment project. These jurisdictions are now all physically linked by at least one interconnector. Geographically, the NEM spans 5 000 kilometres (km) and includes 40 000 km of transmission lines, making it one of the longest AC interconnections in the world.

Japan

The frequency of grid power differs between eastern and western Japan, at 50 Hz and 60 Hz respectively. This difference has an historical basis, as the Tokyo area adopted German-made generators at the beginning of the electricity business, while Osaka chose American-made ones. This frequency difference partitions Japan's national grid so that frequency converter facilities (FCFs) are necessary to connect the eastern and western power grids, known as the East-West Grid Connection. As of August 2014, three FCFs are in operation, together capable of transmitting 1.2 GW, namely Sakuma FCF and Higashi-Shimizu FCF in Shizuoka Prefecture and Shin-Shinano FCF in Nagano Prefecture.

Historically, the vertically integrated utilities were required to maintain self-sufficiency; therefore, interconnections between the ten supply areas were weak and mainly intended for operational system security purposes. This caused some regions to experience surplus capacity, while others that were directly affected by the Fukushima earthquake faced shortages. The limitations of these links have been a major problem in providing power to the areas of Japan affected by the Fukushima nuclear accident.

The capacity of East-West Grid Connection is planned for expansion to 2 100 megawatts (MW) in total by 2020. This includes an increase in the capacity of the Higashi-Shimizu FCF by up to 300 MW, completed by the Chubu Electric Power Company in February 2013.

7.3. Investments in new interconnectors

Investments in new interconnectors face the same barriers as other transmission lines, including policies and regulations, institutions, planning, utilisation rights and cost allocation (IEA, 2013).

The interregional dimension adds complexity and necessitates improved co-operation and co-ordination among all stakeholders at the planning stage, as well as at the investment and the regulatory stages. Establishing a stable regulatory framework can ensure that grid reinforcement is completed on time and in a favourable environment for parties to co-operate on future plans. A stable regulatory framework includes co-ordinated planning, a robust cost-benefit analysis methodology and a transparent and a fair approach to cost allocation.

Barriers for interconnectors

Europe and the United States have adopted different approaches to interconnection. Integrating electricity markets is very high on the political agenda of the European Commission, which has been taking steps to promote interconnector investments through the Third Energy Package policies since 2009. In contrast, the Federal Energy Regulatory Commission (FERC) in the United States relies more on bilateral discussions between balancing areas or RTOs.

While a lack of cross-border transmission lines often reflects regions' physical geography, it can also result from other barriers. Overcoming institutional differences is one of the major difficulties in integrating markets. Market integration within the same country – e.g. the United States or Australia – is already often quite challenging because of differences in state-level institutional settings and regulations. Achieving market integration across several countries – e.g. in Europe or Asia – faces even greater challenges. Governments and regulators today either have a national mandate or a mandate restricted to an individual state or province, focused on protecting the welfare of consumers and electricity security of supply in that area. Some regulators therefore propose that measures aimed at optimising social welfare at both the domestic and international level are a key factor for integrating markets.

A further consideration is that new interconnections and deeper market integration can sometimes increase prices in exporting countries. This is also a potential barrier to market integration. While increasing interconnector capacity removes congestion, it also triggers wholesale price convergence. This has benefits in terms of total welfare, but also involves local price adjustments that have important distributive impacts for consumers and producers in the different participating locations. There is strong empirical evidence that jurisdictions benefitting from cheap coal, nuclear or hydropower are reluctant to engage in electricity market integration or even liberalisation. Governments focusing on their national goals may have limited incentive to take into account their neighbouring countries.

Finally, new transmission projects often face local opposition. Populations are wary of electromagnetic fields (these fears have been proven to be unwarranted) and foresee that transmission lines might reduce the real estate value of their housing. Such resistance can require the installation of lines underground, costing five to ten times more than overhead lines. Consequently, this option weighs into the cost side of cost-benefit analysis and limits the number of projects that qualify as economic.

In these ways, interconnection and transmission capacity growth may remain constrained. Smoothing out the variability of wind and solar power across a continent would require the installation of dozens of additional gigawatts of transmission capacity. Yet it can take up to 30 years to build a single 1.2 GW line, as exemplified by the interconnector between France and Spain inaugurated in 2015. Meanwhile, as more low-carbon generation is deployed into electricity systems, the need to ensure the efficient use of existing transmission networks will grow.

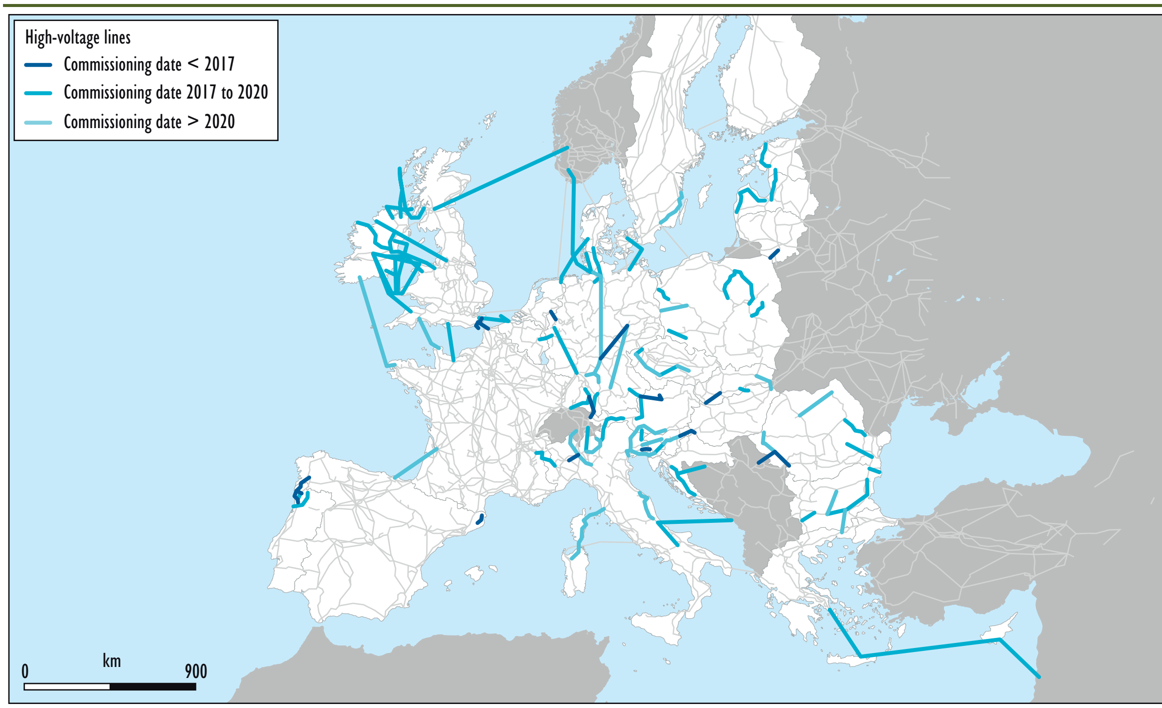
Co-ordinated planning

Co-ordinated planning of interconnectors is already underway at an inter-regional scale in Europe (with the TYNDP) and in North America (with bilateral protocols and committees). Many of these initiatives did not emerge spontaneously. Rather, they are responses to binding policy mandates from the European Commission or FERC.

The TYNDP 2014, a co-ordinated planning initiative to deliver a pan-European transmission plan within the ENTSO-E region, pinpoints about 100 spots on the European grid where bottlenecks exist or may develop if reinforcement solutions are not implemented. The interconnection capacity between the three Baltic States and their EU neighbours is predicted to need to multiply by three in all scenarios of the TYNDP. Between Ireland, Great Britain and the continent, the present capacity of 3 GW is also expected to increase, at least doubling and possibly tripling in the case of higher renewables integration (ENTSO-E, 2014).

In 2013, Europe adopted EU-wide guidelines for priority cross-border energy infrastructure projects – known as projects of common interest (PCIs) – as part of the Energy Infrastructure Package (Regulation EU 347/2013). These projects can benefit from accelerated licensing procedures, improved regulatory conditions, and access to financial support totalling EUR 5.85 billion from the Connecting Europe Facility (CEF) between 2014 and 2020. The European Union released the first list of PCIs for electricity infrastructure in October 2013. These projects of common interest are consistent with the ten-year network development plan established by ENTSO-E (Map 7.5).

Map 7.4 • Transmission lines of common interest (PCI), TYNDP 2014



This map is without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

Source: ENTSO-E, 2014.

In the United States, FERC Order 1000 (2011) set new requirements for regional planning authorities. These included exchanging data at least annually, engaging in joint efforts to harmonise model assumptions and models, and harmonising inter-regional project and cost-benefit assessments (FERC, 2011). The regional planning process must produce a "regional transmission plan" in more detail, which evaluates alternative transmission solutions that might meet the needs of the region more efficiently or more cost-effectively than solutions developed only in local transmission plans (LTPs). In addition, non-transmission alternatives must be considered on a basis comparable with transmission. In cases where alternative transmission solutions are found to be more efficient or cost-effective than options identified in LTPs, those solutions can then be selected in the regional plan for regional cost allocation.

Order 1000 has been less demanding with regard to interregional co-ordination. It has no requirement to produce interregional transmission plans or to engage in interconnection-wide planning. Instead, two neighbouring transmission planning regions must share information on their respective needs and the potential solutions to those needs, and identify and jointly evaluate interregional transmission facilities that may offer more efficient or cost-effective solutions to those regional needs.

Several US regions are planning interconnectors in a co-ordinated manner as part of – or in addition to – single operating region plans. These co-ordinated planning initiatives rest on joint agreements and planning committees, including the Northeast ISO/RTO Planning Co-ordination Protocol between ISO New England (ISO-NE), the New York ISO (NYISO) and PJM, and the Inter-Regional Planning Stakeholder Advisory Committee of PJM and Midcontinent ISO (MISO) (PJM, 2013) based on the Joint Operating Agreement (PJM and MISO, 2008).

Co-ordinated interconnector planning is also proceeding in Australia under the Australian Energy Market Operator (AEMO) as national transmission planner (AEMO, 2014). AEMO publishes an annually updated national development plan to provide a national strategic perspective for transmission planning and co-ordination in the national energy market of Australia. AEMO also goes a step further and assesses non-network options as alternatives (AEMO, 2014).

The existence of multiple neighbouring system planners requires strong co-ordination of the planning process. The scale and complexity of the modelling of large interconnected regions requires the use of bottom-up and top-down processes simultaneously. Bottom-up planning enables integration of local or regional transmission plans that are based on a detailed knowledge of regional conditions. Top-down planning involves a central body that identifies possible intra-regional lines. Both approaches alone have shortcomings. A bottom-up approach only will be unsuitable for identifying intra-regional projects, while a top-down approach risks missing important aspects of the regional networks. Therefore, a combination of both approaches appears to be the best solution.

Modelling a transmission network for the future involves assessing modifications to complex network structures in an uncertain and changing world, choosing from many possible choices with multiple dimensions under huge uncertainty. Unbundling and ensuring the separation of transmission and generation planning increases this uncertainty. Furthermore, policies, load patterns and technologies, and with them the generation mix of the future, will certainly not maintain a steady state over the 50-year lifetime of transmission investments. To enable a robust network design, planners need to identify "least regret" investments.

One of the most important variables for network planning is the share of renewables, in particular VRE, envisaged in the system and their location. Countries with clear goals on renewables therefore ease the network planning process as they reduce uncertainty. The same is true for other network enforcement criteria, such as security of supply, flexibility, expansion of wholesale markets and mitigation of market power. The clearer the criteria and goals, the easier it will be to design a network plan that meets the future requirements of the energy system. Nonetheless, a robust and transparent methodology should be applied to determine necessary grid extensions. Robust in this context should be seen as planning under different scenarios in order to clearly identify projects that are needed in any case.

Many jurisdictions have made progress in implementing integrated planning frameworks. The following aspects need to be co-ordinated to support an efficient cross-border planning procedure:

- the use of consistent data sets
- convergence of different planning models
- harmonisation of reliability requirements.

In Europe, the regulation on access conditions (Regulation EC 714/2009) seeks to harmonise the relevant rules for developing networks and interconnectors. The objective is to ensure co-ordinated and sufficiently forward-looking planning and sound technical evolution of the transmission system (including interconnectors) in the European Union.

Co-ordinating infrastructure planning and regulatory approval in a timely and consistent manner requires extensive harmonisation of cost-benefit analyses, e.g. through the development by ENTSO-E of a single cross-border assessment methodology required by the European guidelines for the implementation of European energy infrastructure priorities (EU, 2013). In Australia, a single regulatory investment test is performed (AER, 2010). Moreover, since interstate projects need greater co-ordination and are often difficult to license, some regulators seek to incentivise them by offering a higher return on capital. In the United States for instance, FERC allows a substantially higher return on investment than the typical remuneration for regulated investments served by state public utility regulatory commissions. In the European Union, Italy provides for higher rate of return on investments (premiums) to signal investment priority for interconnectors.

Cost-benefit analysis

The need for careful assessment of the costs and benefits associated with new transmission lines calls for the application of a rigorous approach to cost-benefit analysis (CBA).

A proposal's economic benefits may include, *inter alia*:

- shared reserves
- higher reliability and supply security
- enhanced competition
- production and operational cost savings
- capacity savings due to capacity requirements
- recovery of (partly) stranded investments
- environmental impact reductions, such as lower carbon emissions
- lower congestion costs.

Its direct costs will include investment costs for the assets, with the indirect costs including the social and environmental costs of the transmission investment.

From the perspective of power market participants, network infrastructure developments often yield multiple changes. One of the most straightforward changes is the addition of a new connection between two regions with initially different electricity prices. On the generation side, the connection brings benefits to the generators in the lower-price region, as their product can be sold at a higher price to load in the wider region. At the same time, marginal generators from the higher-price region might dispatch less often and this could reduce their revenues. Provided there is sufficient network capacity, this supply change will lead to the alignment of prices between the regions. On the demand side, winners and losers can be found in the opposite direction: customers from the initial higher-price region benefit from reduced supply costs and customers from the lower-price region face higher supply costs.

Only if the benefits of power flows exceed the costs for the new transmission line can the investment be regarded as economically justifiable.

There are two especially important principles for most new network investments:

- Net benefit assessments, comprising both benefits and costs, should generally recognise full-scale market impacts of new investments.
- *Ex ante* investment cost allocation commensurate with identified beneficiaries can mitigate financing uncertainties and enhance project acceptance.

If the results from the *ex ante* identification of beneficiaries are fed into the investment cost allocation, this reduces the need for cost socialisation and is likely to enhance project acceptance.

In parallel with the assessment of investment needs as discussed above, CBAs should be developed in consultation with all market participants, requiring the availability of resources in the various stakeholder groups, as well as a determined set of rights and responsibilities.

The inclusion of CBA into the planning framework can facilitate transparency and consultation among all market players, which is likely to result in mutually acceptable assumptions on important factors that trigger future costs and benefits.

The co-ordinated development of such assumptions on future conditions is essential, as any investment planning can only be based upon expected developments. However, the assumptions should also be accompanied by risk assessments, (for example, as done by MISO as part of its annual Transmission Expansion Planning [MTEP] exercise) (MISO, 2014), as uncertainties in the assumptions can alter benefits. Risks can generally be regarded as price risks and/or quantity risks for all relevant assumptions such as demand, fuel sources or supply capacities.

The projection of benefits into the distant future increases the level of uncertainty. New transmission lines will often be capable of yielding long-run benefits over 40 to 60 years. However, applying such a long-term planning time frame will inevitably increase planning uncertainties, creating the risk of under- or overestimated benefits. The inclusion of adequate measures to assess long-term benefits and risks in economic planning principles is often underdeveloped in regulatory decision making.

The *ex ante* cost-benefit calculation should also be specific to local circumstances, because any investments will have a local influence, and loads and generators can demonstrate diverse characteristics. Depending on the market design, the inclusion of generator and load-specific conditions is also relevant as, for example, contractual arrangements for some generators or loads might exclude their benefiting from additional investments. The full inclusion of all market players is likely to exceed the information handling capabilities of any one single network planner, which demands a voluntary and not mandated participation of beneficiaries, as introduced by FERC (2010).[31]

[31] It is notable, however, that FERC has indicated that it is willing to designate a particular entity as a beneficiary of a transmission project, even if that beneficiary has not entered into a voluntary arrangement with the ISO or RTO, in order to avoid potential free-rider problems. For example, FERC has required that MISO and PJM develop a joint cost allocation method for projects in one service territory that impact one another (FERC, 2010).

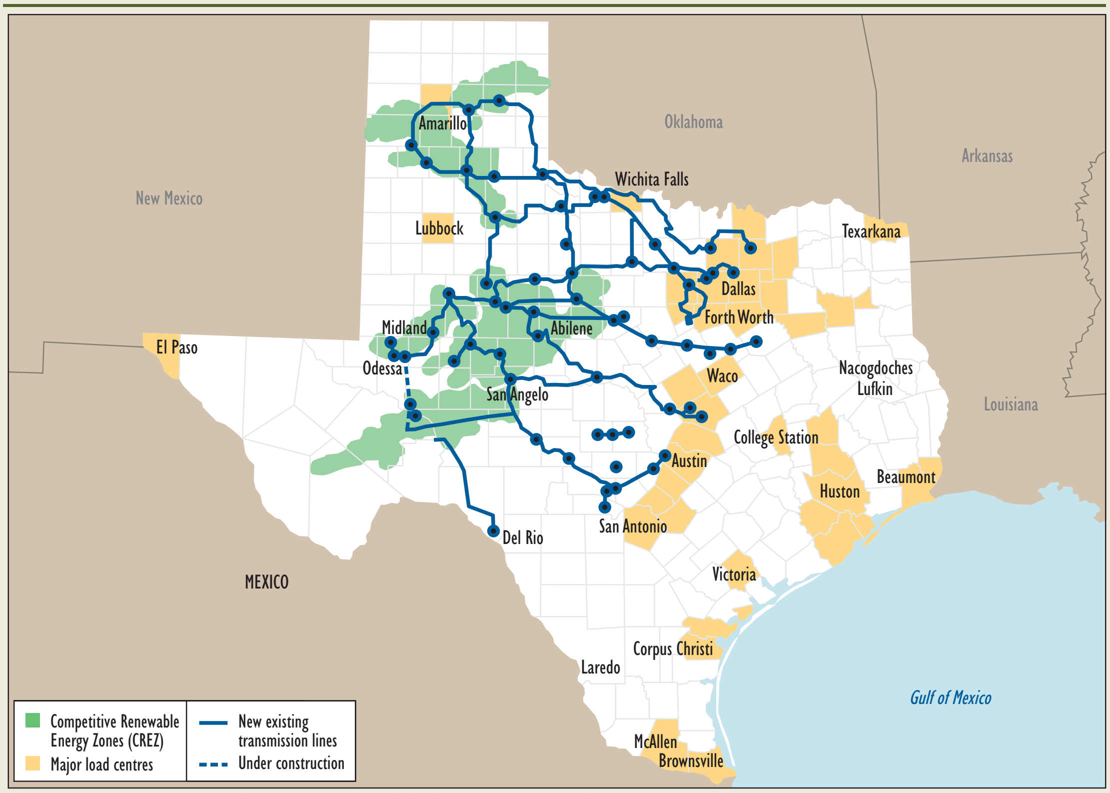
Box 7.3 • Renewable integration and grid planning in Texas

The Electric Reliability Council of Texas (ERCOT) has opted for a system that at first sight appears to socialise the costs of new transmission infrastructure, rather than following the "beneficiary pays" principle. ERCOT, due to the isolation of the state's electricity system, does not fall under FERC jurisdiction. In order to deploy the immense wind resources located in west Texas and the Panhandle region, new transmission lines have been needed to transfer electricity to the load centres, such as Houston. Competitive Renewable Energy Zones (CREZs) were created that fulfil the following criteria according to the Texas Administrative Code:

- Renewable energy resources and suitable land areas are sufficient to develop generating capacity from renewable energy technologies.
- A set level of financial commitment by generators.
- Other established criteria, for example, the estimated cost of constructing the transmission capacity necessary to deliver electricity to consumers, the envisaged energy output of renewable energy resources and the benefits of renewable energy produced in the zone.

ERCOT generated an optimisation study to determine the most cost-effective transmission investments to deliver electricity from the remote CREZs to the load centres (ERCOT, 2008). Finally, a plan was created with estimated costs of USD 4.93 billion to provide about 18 4756 MW of wind generation from CREZs. The plan involves 2 334 miles of 345 kV transmission lines in more than 100 transmission projects to provide about 64 031 gigawatt hours of wind generation annually. Transmission revenues will contribute to the investment costs and will be reimbursed by customers through socialised costs. However, the term "socialised cost" might be a little misleading in this case since one could argue that, due to the huge overall effect of the transmission investments on consumers across ERCOT, the "beneficiary pays" principle also applies to this case. ERCOT considers the benefits of the new transmission from the CREZs to extend to the entire region. It might be difficult to copy this approach on a multi-state or multi-country level, although, with some adaptations and qualifications, other regions apply rather similar approaches.

Map 7.5 • CREZs in Texas



This map is without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

Source: Public Utility Commission of Texas, 2014.

Cost allocation

Alongside the application of an accurate CBA, finding an acceptable approach to cost allocation remains one of the fundamental barriers to efficient and acceptable new investments.

The issue of cost allocation is of particular relevance to inter-regional network investment, as various parties from different states or countries are involved, which are usually more focused on regional or country issues than on the overall welfare of a larger area. In particular, lines with significance for more than one region tend to have a material effect on market conditions, and market participants might have a major incentive to free ride.

Interconnector cost allocation often causes difficulty by precluding incentives for network investors and prompting a lack of consumer or regional acceptance. FERC notes that inaccurate cost allocation represents a significant cross-border investment barrier (FERC, 2011). Consequently, clear *ex ante* rules for regulatory approval of these investments and clear rules for cost allocation between regions or countries are necessary. Transmission costs are mainly recovered from network users. A wide variety of cost allocation principles are currently used in most OECD countries. How the regions allocate the costs further down to network users is less relevant to the cross-border planning process, but is critical to gaining public acceptance of network investments.

In the United States, FERC recognised that regional differences may warrant different methods of regional and interregional cost allocation. FERC Order 1000 defines a principles-based approach which requires all transmission providers to demonstrate compliance with six cost-allocation principles – focusing mainly on regional cost allocation, but partly also applying to inter-regional cost allocation:

1. Costs are to be allocated roughly commensurate with benefits.
2. There must be no involuntary allocation of costs to those who do not benefit. A region that receives no benefit from an interregional transmission facility located in that region must be allocated no costs for that facility.
3. A benefit-cost threshold, if one is used, must not exceed 1.25. A benefit-cost threshold is not mandatory, however.
4. Costs must be allocated to those solely within a transmission planning region or regions, unless those outside that region or regions voluntarily agree to bear such costs.
5. The cost allocation methodology must be transparent with respect to determining benefits and the identification of beneficiaries.
6. Different cost allocation methodologies may be used for different types of transmission facilities (e.g. reliability, economic and public policy).

With respect to inter-regional cost allocation, FERC Order 1000 requires that interconnected, neighbouring transmission providers develop interregional co-ordination and cost allocation procedures (FERC, 2010). However, FERC has not mandated a particular methodology, but rather has required the principles-based approach described above, in recognition of the fact that different regions may require different solutions for determining the appropriate allocation of costs. Furthermore, the interregional cost allocation method may differ from the allocation methods that each transmission provider uses internally. While co-operation is mandated, transmission providers such as PJM, NYISO and ISO-NE have themselves recognised that co-ordinated planning is important if all potential benefits are to be realised (ISO-NE, NYISO and PJM, 2014).

It is also important to keep in mind that investments that do not cross administrative boundaries can still have a huge impact on cross-border capacity. For example, increasing the German network and the interconnection at the German-Austrian Border will have a huge impact on the Polish and Czech transmission networks, and the cross-border interconnection between Spain and France requires the French and Spanish networks to be able to transport the additional electricity within the countries.

Therefore, cost-allocation principles need not be limited to networks crossing regions or countries, but should be applied to network reinforcements within a region or country, thus enabling neighbouring regions or countries to be better connected. This is supported by evidence from the Australia Energy Market Commission (AEMC), which found that approximately two-thirds of all internal transmission constraints contain an inter-regional portion (AEMC, 2011). In Europe, cost allocation between countries is usually undertaken by mutual agreement between the countries affected by the transmission line, which does not necessarily need to be a line physically connecting two sides of a border, but can also be a line within one country that heavily influences the networks of neighbouring countries. With the introduction of the Third Energy Package of the European Commission and the adoption of the Network Code on Capacity Allocation and Congestion Management,[32] the European Agency for the Cooperation of Energy Regulators (ACER) gained the right to decide on cost allocation if member states of the Europe Union could not mutually agree.

In Australia, costs of new transmission projects are also partly allocated between regions according to the benefits. AEMC published its final determination and rule on inter-regional transmission charging in 2013 (AEMC, 2013). The new arrangements better reflect the benefits of transmission in supporting energy flows between regions, but do not aim to affect total revenues from each transmission business. The major expected benefits derive from enhanced incentives for businesses to pursue transmission investments, whose costs fall predominantly in their own region but whose benefits fall in neighbouring regions, since they can recover some of the investment costs from consumers in those regions. Further, the prices consumers pay for transmission services should better reflect the actual costs.

By now, there is a common understanding that costs should be allocated in proportion to the benefits. One of the advantages of this approach, together with common planning, is the prevention of overinvestment in projects that do not provide enough cumulative benefits for all jurisdictions. Another advantage of the “beneficiary pays” principle is that it ensures that each group of stakeholders is fully informed during the planning and assessment process, facilitating acceptance of new interconnector investments.

One of the difficulties of this approach is identifying all beneficiaries. Failing to do so could cause valuable projects not to be built, as insufficient benefits have been factored in to cover the cost. For this reason the inferior principle of socialisation of costs continues to be commonly used to spread costs not only within a region, but also between countries. However, socialisation reduces the locational signals for grid investment and might induce wasteful investments and reduce cost discipline (Kaplan, 2009). Conversely, with all the uncertainty related to decarbonisation, the estimation of benefits contains a high level of uncertainty and is driven by a wide range of assumptions on future technological development. Network investments made today with an average lifespan of about 40 years will experience huge changes in energy consumption and production. Even though, ideally, the “beneficiary pays” principle should be applied, in the end it has to be accepted that the result might be similar to a cost socialisation approach.

[32] EU cost allocation is also included in TYNDP, and ACER has legal competence, as set out in the Regulation (EU) No. 347/2013 of the European Parliament and of the Council of 17 April 2013 on guidelines for trans-European energy infrastructure and repealing Decision No. 1364/2006/EC and amending Regulations (EC) No. 713/2009, (EC) No. 714/2009 and (EC) No. 715/2009.

7.4. Market-based network infrastructure investments

Market parties in areas with long-term price differences between markets may be interested in investing in new interconnection capacity in order to capture the congestion rents or use the interconnection themselves. These interconnections are commonly referred to as "merchant interconnections" (De Jong and Hakvoort, 2006). In other cases, system operators organise competitive auctions for the construction of new assets.

Merchant investments

Merchant investors in transmission infrastructure rely on an established set of preconditions to govern their investment: competition, free entry and market-based pricing of transmission services. Merchant investors are allowed to collect congestion revenues in return for their investment in additional transmission capacity, equal to the difference in energy prices associated with the incremental point-to-point transmission capacity that the investments create. The congestion revenues represent the return merchant investors receive to cover the investment's capital and operating costs, and provide the financial incentive that drive "market-based" transmission investment (Joskow and Tirole, 2005).

Nowadays, the construction of new network infrastructure is mostly supported by regulation. Regulation offers a solution to the barriers that efficient, market-based network infrastructure investments face, namely uncertainty and potential revenue shortfall. In most IEA member countries, transmission investors are compensated based on revenues calculated by the regulator. Investment planning and regulatory revenue calculation and allocation seek to reduce uncertainty and potential revenue shortfall. Further barriers to merchant investments may include continuous market power-induced price spreads, rising transaction costs or incumbent rights to refuse third-party investment proposals (Joskow and Tirole, 2005; Littlechild, 2011; IEA, 2013).

However, each of these barriers appears to be country-specific – or even project-specific. The academic debate is ongoing about the experiences of real-life projects and their resulting policy implications. Observation of real-life projects has identified imperfect market information as a major hurdle to merchant transmission investments, where expected price spreads between interconnected nodes proved to be lower in practice (Littlechild, 2011). In Australia, two interconnectors, Murraylink and Directlink, which initially started as merchant projects, were transferred into the regulatory regime when the required price spreads between regions ceased to exist. Basslink remains a merchant interconnector in the Australian NEM. Despite such examples, it appears to be worthwhile to continue identifying – and possibly eliminating – barriers to merchant-based transmission investments. Any resulting increase in such investments will reduce the need for regulatory intervention to assist the process. This can mitigate some of the potential failures inherent in all regulatory processes and produce enhanced economic efficiency, innovation, technological neutrality, delivery and financial resources (Joskow, 2010).

Competition for the construction of new interconnectors

Another form of competition for transmission lines and interconnectors are auctions organised by system operators for the construction of new assets. Once a need for additional transmission lines or interconnection has been identified through planning, the ISO organises a tendering process to build and own the new asset. Unlike merchant strictly investments, the revenues are largely regulated for the specific asset.

Table 7.2 • Jurisdictions with processes for merchant investments

Jurisdictions with competitive processes	
Brazil	All transmission projects are auctioned (since 1999)
United Kingdom	Tenders for offshore grid projects
US regional planning	Competitive energy zones in Texas Various forms of competitive process in FERC jurisdiction
Ontario	One tender for transmission to date
Alberta	Competitive process developed

Competition between transmission owners is well developed in the Americas (see Table 7.2). The main advantage of such a competitive process is that alternative potential transmission owners can propose more cost-effective solutions while the system planning is undertaken by the ISO. In a European context, however, system operators are also largely transmission owners and all new transmission assets are decided and primarily built by TSOs, under the supervision of regulators. Nevertheless, Directive 2009/72/EC foresees the possibility of tendering for necessary transmission lines in cases where the transmission owner builds a line not foreseen in the TYNDP and is liable for it. Whether this option is actually feasible in Europe is questionable, as first it is difficult to prove that a delay of a line was caused by a TSO. In addition, the unity of transmission ownership and system operation creates an additional barrier to third parties. Therefore, in the European context, a competitive tender process has so far only been developed in the United Kingdom for the connection of offshore wind farms.

In the United Kingdom, since 2010, the Office of Gas and Electricity Markets (Ofgem) has provided a legal framework for tenders (UK Gov, 2010) to determine the competitive basis upon which an offshore transmission licence can be granted. The reason for using a tendering process for building the offshore network infrastructure for offshore wind farms was to ensure that the right quality lines would be built at the right time and at the right cost. Ofgem's competitive tender process (Ofgem, 2009) selects the offshore transmission owner (OFTO) after first analysing estimates for significant cost reductions to generators and consumers. The tendering rules allow for asset development either by wind farm investors or by independent network developers, which provides flexibility for generators as to who constructs the assets. If the generator chooses to build the connection, it has to transfer the assets to an OFTO upon completion of construction. The OFTO will then have upfront clarity over their revenue stream over the 20 years of depreciation, paid by National Grid Electricity Transmission (NGET), the onshore transmission system operator, and there will be no additional revenue regulation. The revenue stream includes all relevant costs for financing, designing/constructing (if applicable), operating, maintaining and decommissioning of the transmission assets. Through network charges, NGET will allocate these costs to all network users.

Results from these tenders show success in attracting investors, with new entrants and new sources of finance demonstrating interest in the sector. Funding of up to almost USD 6.4 billion was offered in relation to the USD 1.75 billion invested in 43 assets in the first tender (Ofgem, 2012). This led to estimated cost savings of about USD 300 million to USD 385 million relative to price control-based solutions (CEPA/BDO, 2014).

In Brazil, the Ministry of Mines and Energy determines transmission expansion based on studies conducted by the federal energy planning company, EPE, and the national grid operator, ONS. Subsequently, the national electricity regulator, ANEEL, implements the decisions and conducts auctions for new projects. All facilities (≥230 kV) that are required to meet system needs are auctioned to determine who builds, operates and owns them. The auction process starts with

ANEEL setting the maximum annual allowed revenue, Receita Anual Permitida (RAP), and the winner being the bidder offering the lowest RAP. The concession for each transmission line is granted for 30 years. After 15 years the RAP payment is reduced by 50%. Incentives to complete the project ahead of schedule and maintain high availability (thereby seeing an increase in revenues) are also in place (RAP, 2013).

In Brazil, to date over 50 000 km of new transmission lines (≥230 kV) have been built using auctions, with a total investment of USD 28 billion. The proposed revenue requirement would have been USD 4.45 billion per year, whereas the actual revenue requirement is USD 3.35 billion per year (The Brattle Group, 2014).

Conclusion

Networks are becoming increasingly important to guarantee security of supply and enable low-carbon electricity generation at least cost. While variable resources such as wind and solar now generate an impressive amount of electricity, they are generally not as dependable as conventional energy generation. To maintain high security of supply with a high share of VRE, interconnectors are frequently the most cost-efficient solution, depending on population density and landscape, reducing the amount of new conventional generation needed to balance the system.

Cross-border projects involve multiple parties who, in many cases, take the perspective of their region or country rather than looking at overall efficiency and social welfare. This often leads to undervaluation of cross-border transmission projects and their positive impact on regional transmission systems. To foster interconnection investments, common network planning and a mutual agreement on CBA and cost allocation are essential. The scale and complexity of the modelling of large interconnected regions requires bottom-up and top-down processes simultaneously. Effects on regional and local networks have to be taken into account and network reinforcement needs to be jointly planned.

Only where the net benefits of power flows exceed the costs of a new transmission line can the investment be regarded as economically justifiable. The exact determination of all relevant benefits, as well as their quantification methodology, should be used as a tool to inform all relevant market participants. The “beneficiary pays” principle is a good tool for these projects, with the added complexity that their benefits are widespread. Similarly, the accurate allocation of costs is of particular relevance for interregional network investment, where parties may naturally be more focused on regional or country issues than on the overall welfare of a wider area. In view of all the uncertainties related with decarbonisation, the estimation of benefits involves a high degree of uncertainty and is driven by a wide range of assumptions on the development of future technologies.

References

AEMC (Australian Energy Market Commission) (2015), “The regulatory investment framework for transmission in Australia”, AEMC, presentation to the IEA ESAP workshop on Regional Resource Adequacy, 15 January 2015, www.iea.org/media/workshops/2015/esapworkshopv/Smith2.pdf.

AEMC (2013), “Rule changes: Completed – inter-regional transmission charging”, www.aemc.gov.au/electricity/rule-changes/completed/inter-regional-transmission-charging.html.

AEMC (2011), “Transmission frameworks review – First interim report”, www.aemc.gov.au/Media/docs/First%20Interim%20Report%20Overview% 20Report-13e3ed2c-888f-4f78-b7d3-4586b0b1efb9-0.PDF.

AEMO (Australian Energy Market Operator) (2014), *National Transmission Network Development Plan*, AEMO, www.aemo.com.au/Electricity/Planning/~/media/Files/Electricity/Planning/Reports/NTNDP/2015/2015%20NTNDP_FINAL.ashx.

AER (Australian Energy Regulator) (2010), "Regulatory investment test for transmission", AER, www.aer.gov.au/system/files/Final%20RIT-T%20-%2029%20June%202010.pdf.

Booz & Company et al. (2013), "Benefits of an integrated European energy market", Booz & Company, report prepared for DG Energy, retrieved on 11 February 2014 at http://ec.europa.eu/energy/infrastructure/studies/doc/20130902_energy_integration_benefits.pdf.

The Brattle Group (2014), "Competition in transmission: Planning and development", www.brattle.com/system/publications/pdfs/000/004/977/original/Competition_in_Transmission_Planning_and_Development.pdf?1391196850.

CEPA/BDO (Center for European Policy Analsis) (2014), "Evaluation of OFTO tender round 1 benefits", Ofgem, www.ofgem.gov.uk/ofgem-publications/87717/cepabdotr1benefits assessmentfinalreport.pdf.

De Jong, H. and R. Hakvoort (2006), "Interconnection investment in Europe: Optimizing capacity from a private or a public perspective?" www.e-bridge.de/fileadmin/templates/downloads/DE/interconnector-investment-in-europe--final.pdf.

DOE (Department of Energy) (2015), "The quadrennial energy review", Department of Energy, http://energy.gov/sites/prod/files/2015/07/f24/QER%20Full%20Report_TS%26D%20April%202015_0.pdf.

ENTSO-E (2014), "Ten-year network development plan", www.entsoe.eu/major-projects/ten-year-network-development-plan/tyndp-2014/Documents/TYNDP%202014_FINAL.pdf.

ERCOT (Electricity Reliability Council of Texas) (2008), "Competitive renewable energy zones (CREZ) transmission optimization study", http://pbadupws.nrc.gov/docs/ML0914/ML091420467.pdf.

FERC (Federal Energy Regulator Commission) (2011), "Order no. 1000 - transmission planning and cost allocation", www.ferc.gov/industries/electric/indus-act/trans-plan.asp.

FERC (2010), Transmission Planning and Cost Allocation by Transmission Owning and Operating Public Utilities, www.ferc.gov/whats-new/comm-meet/2010/061710/E-9.pdf.

IEA (International Energy Agency) (2015), *Electricity information 2015*, OECD/IEA, Paris.

IEA (2014a), *World Energy Investment Outlook 2014*, IEA/OECD, Paris. www.iea.org/bookshop/477-World_Energy_Outlook_2014.

IEA (2014b), *Seamless Power Markets: Regional Integration of Electricity Markets in IEA Member Countries*, Featured Insight, IEA/OECD, Paris. www.iea.org/publications/freepublications/publication/seamless-power-markets.html.

IEA (2014c), *Energy Policies of IEA Countries: European Union*, OECD/IEA, Paris.

IEA (2014d), *Energy Policies of IEA Countries: United States*, OECD/IEA, Paris.

IEA (2013), *Electricity Networks: Infrastructure and Operations – Too Complex for a Resource?* www.iea.org/publications/insights/insightpublications/name,43756,en.html.

ISO New England, New York Iso and PJM (2014), "2013 northeastern coordinated system plan", www.nyiso.com/public/webdocs/markets_operations/services/planning/ipsac/2013_Northeastern_Coordinated_System_Plan.pdf.

Joskow, P. (2010), "Market imperfections versus regulatory imperfections", Alfred P. Sloan Foundation and Massachusetts Institute of Technology, http://economics.mit.edu/files/5619.

Joskow, P. and J. Tirole (2005), " Merchant Transmission Investment", *Journal of Industrial Economics*, Vol. 53, No. 2, pp. 233-264, http://economics.mit.edu/files/1179.

Kaplan S. (2009), "Electric power transmission: Background and policy issues", http://research.policyarchive.org/18743.pdf.

Littlechild, S. (2011), "Merchant and regulated transmission: Theory, evidence and policy", Cambridge Working Paper in Economics, www.repository.cam.ac.uk/bitstream/handle/1810/242029/cwpe1160.pdf?sequence=1&isAllowed=y.

MISO (Midcontinent Independent System Operator) (2014), "Transmission expansion plan 2014", www.misoenergy.org/Library/Repository/Study/MTEP/MTEP14/MTEP14%20Full%20Report.pdf.

NREL (National Renewable Energy Laboratory) (2012), "Renewable Energy Technology Resource Maps and Technical Potential for the United States", www.nrel.gov/gis/docs/resource_maps_201207.pptx.

NREL (2011), "Eastern wind integration and interconnection study", www.nrel.gov/electricity/transmission/eastern_renewable.html#ewits.

NREL (2008), "20% wind energy by 2030", NREL, www.nrel.gov/docs/fy08osti/41869.pdf

Ofgem (Office of Gas and Electricity Markets) (2012), "Offshore transmission coordination project – conclusions report", www.ofgem.gov.uk/ofgem-publications/51614/20120103otcp-conclusions-report.pdf.

Ofgem (2009), "Offshore electricity transmission: Final statement on the competitive tender process", www.ofgem.gov.uk/ofgem-publications/51252/offshore-electricity-transmission-final-statement-competitive-tender-process.pdf.

Pentalateral Energy Forum (2015), "Generation adequacy assessment", www.benelux.int/files/4914/2554/1545/Penta_generation_adequacy_assessment_REPORT.pdf.

PJM (2013), "Inter-regional planning stakeholder advisory committee – MISO", PJM interconnection.

PJM and MISO (2008), "Joint operating agreement between the midwest independent transmission system operator, Inc. And PJM Interconnection, L.L.C.", 11 December, www.pjm.com/media/documents/merged-tariffs/miso-joa.pdf.

Public Utility Commission of Texas (2014), "CREZ Progress Report No. 15 (April Update)", www.texascrezprojects.com/page29605717.aspx.

RAP (2013), "Regulatory framework and cost regulations for the Brazilian national grid", www.raponline.org/document/download/id/6962.

UK Gov (United Kingdom Government) (2010), "The electricity (competitive tenders for offshore transmission licences) regulations 2010", www.legislation.gov.uk/uksi/2010/1903/pdfs/uksi_20101903_en.pdf.

Chapter 8 • Regulation of distribution networks

HIGHLIGHTS

- The role of distribution grids is changing, with network operations moving from a passive role to being at the centre of the energy transition process. Grids will shape the future deployment of distributed energy resources, such as solar photovoltaics (PV), electric vehicles, micro gas-fired turbines, distributed storage and demand response.
- Active operation of distribution networks can enable reductions in or postponement of network investments. If the regulatory framework for distribution networks fails to evolve, this opportunity may be lost.
- A network regulation 2.0 is needed to establish a smart distribution network. New regulation should also enable the creation of market platforms where distributed energy resources can participate, while keeping a neutral role vis-à-vis competition.
- Decisions to refurbish the distribution grid, affecting where and when, and the degree of automation, should only be taken at the local level. Network regulation therefore needs to give balanced negative (cost efficiency) and positive (technology investment) incentives to distribution network operators and owners.

Although much of the discussion on power sector decarbonisation has focused on liberalised generation, the performance of the regulated network is also very important. Distribution typically represents 20% to 40% of electricity bills. In member countries of the Organisation for Economic Co-operation and Development (OECD), around 30% of all investment needs in the power sector to 2035 will be in distribution networks, assuming adoption of the approaches, policies and regulations signalled in the New Policies Scenario of the *World Energy Investment Outlook* (IEA, 2014a). Moreover, regulated elements of the network greatly affect the performance of competitive activities, as the first provides the infrastructure platform upon which the second relies.

At present, the management of the distribution network differs substantially from that of the transmission level. From the perspective of transmission network operators, the distribution networks have traditionally served mostly as passive load centres that channel electricity from the transmission level to end customers. Until recently, distribution-level questions were low on the agenda of regulators and policy makers.

In a changing electricity system, three challenges stand out for managing distribution networks:

- integrating new and variable renewable generation sources today (and electric vehicles)
- enhancing customers' market activity
- co-ordinating transmission and distribution networks.

Massive investments are needed in distribution networks (Table 8.1), but they often lack detailed regulatory oversight, perhaps because tracking multiple individual network developers' investment plans would pose a time-consuming challenge for regulators. Regulatory frameworks will have to be designed to handle as efficiently as possible a large number of heterogeneous distribution networks, each presenting specific challenges (IEA, 2013).

After reviewing the new challenges for distribution system operators, this chapter presents the essential elements of a network regulation 2.0.

Table 8.1 • Anticipated distribution network investment costs

Investment in distribution in the New Policies Scenario 2014-35 (2012 USD billion)				
		Additions		
	Total	New demand	Renewables	Refurbishment
OECD	1 635	521	53	1 062
Americas	696	245	18	433
United States	564	183	16	365
Europe	590	157	23	409
European Union	516	105	23	388
Asia Oceania	350	119	11	219
Japan	199	44	10	146

Source: IEA, 2014a.

8.1. Distributed resources call for a rethinking of regulation

The increasing penetration of distributed energy resources (DER) – such as distributed demand response, generation (Table 8.2) and storage, smarter technologies and active management techniques for the distribution grid – creates new challenges for the management and operation of distribution networks, as well as for their regulation.

DER in the range of 3 kilowatts (kW) to 50 megawatts (MW) are already connected to the distribution network in locations close to consumers, or even connected to electrical installations behind the meter. These consist of a range of smaller-scale and modular devices designed to provide electricity (and sometimes also thermal energy), and include fossil and renewable energy technologies, energy storage devices (e.g. batteries and flywheels) and co-generation systems.[1] Distributed generation offers solutions to many of the challenges associated with today's electrical power system, including blackouts and brownouts, energy security concerns, power quality issues, tighter emissions standards, grid bottlenecks and greater control over energy costs.

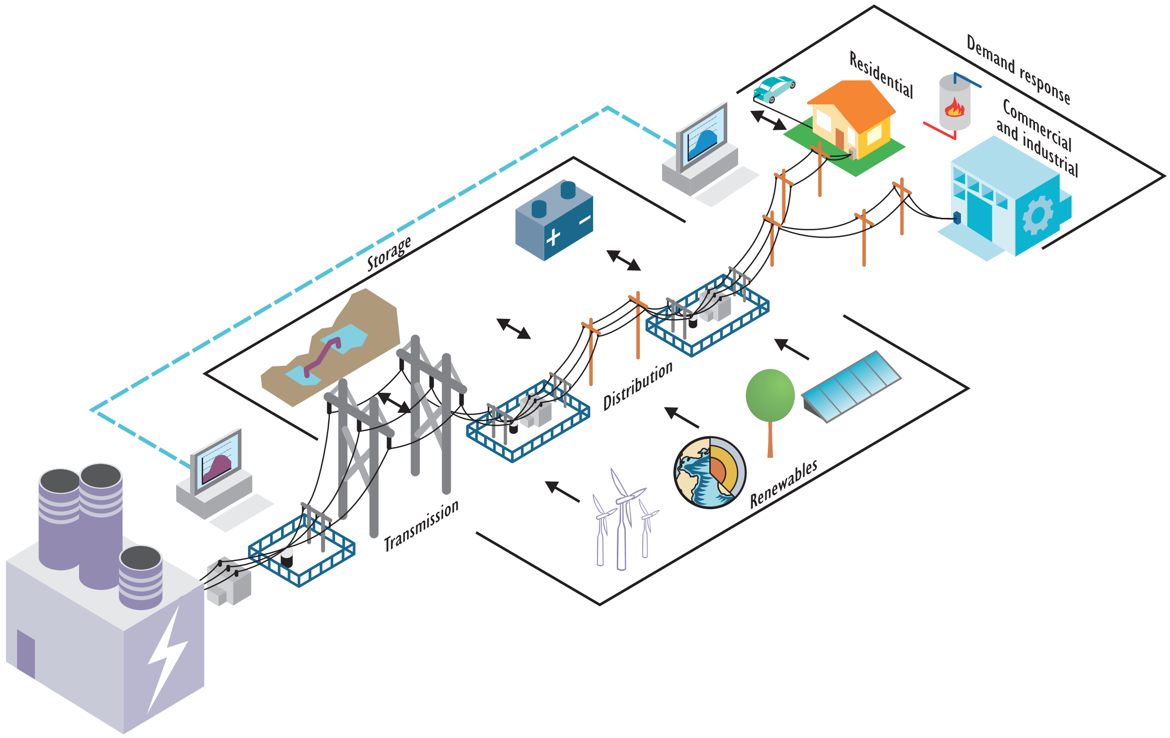
Table 8.2 • Examples of distributed generation technologies

Renewable	Solar PV Onshore wind Small hydroelectric Wood Municipal solid waste (renewable component)
Non-renewable	Natural gas-fired fuel cells Small reciprocating engines Natural gas-fired small and microturbines

In the future, the further deployment of these technologies, demand response and electric vehicles will all play an important role. Their expansion and location define the requirements for future distribution networks, and therefore influence future decisions on the deployment and remuneration of investment. Consequently, regulatory frameworks for distribution network investment and management are not only a question of network regulation, but also influence the future of the entire energy system (Figure 8.1).

[1] Co-generation refers to the combined production of heat and power.

Figure 8.1 • Network and power flow structure of the future



Dispatchable distributed generation increases reliability

Electricity consumers can use back-up diesel or natural gas generators to increase their security of electricity supply, for instance in hospitals, data centres or large buildings. Standby power is used where there is a high incidence of supply interruptions, or the provider is slow to restore supply after an interruption. With increasing incentives for co-generation and renewables, and increasing volatility of electricity prices, further objectives are cost reductions and lower emissions. Many combinations of technologies and fuel options are available (Box 8.1).

Box 8.1 • Distributed conventional energy generation

- Diesel engine gensets* are a cost-effective, reliable and widely used technology manufactured in a wide range of sizes, from about 1 kW up to about 10 MW. They can either be cycled frequently to operate as peak-load power plants or used as load-following plants; they can also be run in baseload mode in off-grid systems. Their major drawbacks include very high levels of emissions and the need to muffle loud engine noise.
- Dual-fuel engine gensets consist of a diesel-cycle engine being able to use a mixture of natural gas and diesel fuel. The small amount of diesel fuel allows the use of compression ignition, and the high percentage of natural gas in the mix results in much lower emissions compared to a diesel engine. In most other ways, dual-fuel engines are comparable to diesels. Natural gas engine gensets comprise a reciprocating natural gas-fuelled engine. In most other respects, natural gas engines perform similarly to diesels and dual-fuel engines, but have the potential for the lowest emissions. They are available in sizes from a few kilowatts to about 5 MW.
- Combustion turbines burn gas or liquid fuel. They usually take a few more minutes to get up to speed in comparison to reciprocating engines. Gas turbines are usually used for peaking and load-following applications and for baseload operation in larger sizes. Installed costs are higher than those of reciprocating engines, and maintenance costs lower. Turbines are efficient and relatively clean. They are available in sizes ranging from about 300 kW to several hundred megawatts.

Box 8.1 • Distributed conventional energy generation (continued)

- Microturbines are smaller, less efficient versions of combustion turbines, in the range of about 30-250 kW. Microturbines are targeted at the small industrial and commercial market and designed to be compact, affordable, reliable, modular and simple to install.
- Fuel cells produce direct current electricity by a thermochemical process. The power is inverted to alternating current for grid operation. By-products are heat, water and carbon dioxide, making fuel cells one of the cleanest sources of power at the place of generation. Unless it is transported to the site, the hydrogen comes from reforming a fuel such as natural gas or propane, a process that may produce environmental emissions. Fuel cells are efficient, quiet and modular. They are available in sizes ranging from a few watts to 200 kW.

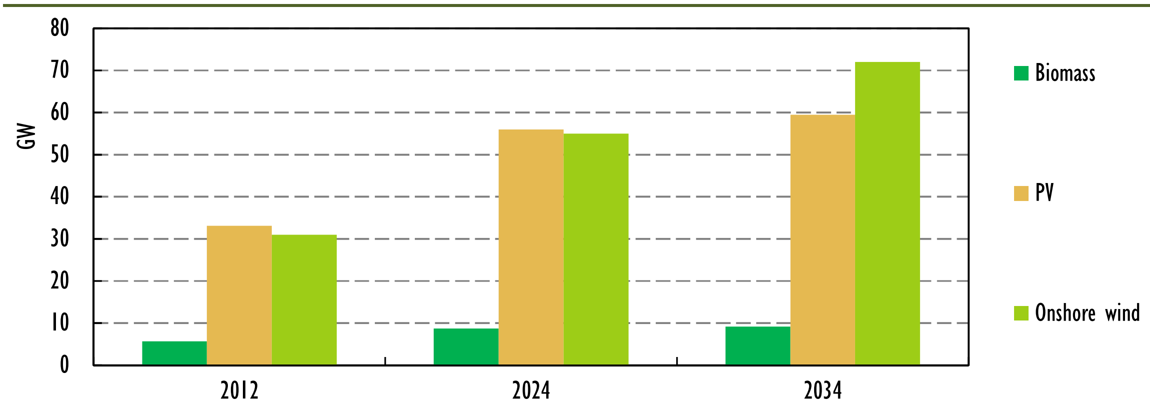
* A genset is an engine-powered machine used to generate electricity.

Source: DOE, 2002.

Variable renewable distributed generation and hot spots

Integrating large shares of variable renewable energy (VRE) puts new demands on network services. In the 450 Scenario of the *World Energy Outlook 2014* (IEA, 2014b), in which policies to achieve the internationally agreed goal of limiting the average global temperature increase to 2°C are fully implemented, the share of renewables in the overall energy mix increases to 30% by 2040. In electricity sectors, this share could rise to 51% in 2040. Country- or region-specific targets often exceed these levels.

Figure 8.2 • Projected generation of typically decentralised renewable energy in Germany



Note: Corresponds to Scenario B of the Scenario Framework for Network Development Planning in Germany; this is considered to be the leading scenario, as it is described as being close to the government's renewables targets.

Source: BNetzA, 2013.

Several countries have already brought significant amounts of renewable generation into their distribution networks. In Germany, for example, distribution operators have connected about 61 gigawatts (GW), which amounts to about 98% of the overall renewables connected to the German network. The record of successful deployment of VRE in Germany reflects both a legal obligation to connect any new generator to the network, and efficient monitoring and dispute resolution management. Additionally, this uptake has mostly taken place within distribution companies that do not own or operate their own conventional generation assets, thus avoiding conflicts of interest.

The installed capacity of renewables can vary greatly between regions or countries. According to forecasts, installed renewables generation is set to more than double by 2032 to meet policy goals in Germany, and might increase threefold if each federal state within Germany were to meet its own targets (Figure 8.2)(BMWi, 2014).

Legislation also influences the deployment of renewables, as municipalities and local authorities frequently need to identify specific areas for wind generation, which leads to concentrations of VRE in "hot spots". For example, a small number of network operators in Germany currently have to deal with an installed capacity of renewables greater than peak demand at that point of connection.

Distributed energy storage can reduce local peaks

Utilities typically use batteries to provide an uninterruptible supply of electricity to substation switchgear and to start backup power systems. Storage can also enable load levelling and peak shaving over a period of hours, and increase power quality and reliability for residential, commercial and industrial customers by providing backup during power outages.

Domestic hot water tanks constitute one of the most common thermal energy storage technologies installed today. Other technologies, such as ice and chilled water storage, play an important role in several countries, including Australia, the United States, China and Japan, as utilities seek to reduce peak loads and consumers seek to lower their electricity bills. Underground thermal energy storage (UTES) systems are frequently found in Canada, Germany, and many other European countries (IEA, 2011).

Electric vehicles could increase peak demand

Electric vehicles (EVs) and plug-in hybrid vehicles (PHEVs) will take some time to significantly penetrate markets, but they are emerging as new sources of demand. According to government policy targets, roughly 20 million electric cars will be on the road by 2020 across major economies (ICCT, 2013).

This fleet will soon start to affect the existing distribution network infrastructure, especially in states and regions with high penetration levels, such as California with 1.5 million zero-emission vehicles expected by 2025 (California Energy Commission [CEC], 2013). EV charging is currently being discussed in at least three modes, each differing in the required charging time. Slow charging at level 1 will cause up to 2.4 kW of capacity demand, level 2 already demands 19.2 kW, and a third, yet to be fully defined level, is likely create demand between 20 kW and 250 kW (NREL, 2010). For comparison purposes, regular peak capacity demand of the average household is in a range between 3.5 kW and 5 kW.

The impact of EV charging is largely determined by the time of peak use. Household electricity demand normally peaks in the hours after work, with people coming home and switching on electrical appliances. It is generally expected that, if uncoordinated and without fast-charging devices, EVs will contribute to that peak electricity demand since most drivers will return home and plug in their EV for recharging over several hours. If, in 2020, all EVs in California start charging at the same time, this could cause a capacity demand surge of between 3.6 GW (level 1) and 30 GW (level 3, lower value) in a 52 GW peak demand system. Assuming that this capacity demand is aligned with predominant peak demand in distribution networks, it would add to the system's peak demand and require significant distribution and transmission network (and generation) infrastructure upgrades.

Real-time pricing for EV charging can significantly improve this situation. Research (CEC, 2009) has calculated an incremental peak demand of up to 200 MW for California's distribution networks by 2020. Rolling out the charging infrastructure, particularly at the household level, can be designed to reduce peak impacts but leave customers with the final choice.

Another benefit of real-time price signals is prevention of massive peak impacts on distribution and transmission systems, although this would require technical equipment to be installed on the connection side of the EV. Depending on the local penetration rate, price signals can ensure freedom of choice and customer flexibility at the same time. Real-time prices do not rely on centrally controlled charging to avoid peak impacts. Under a market framework, network operators would be able to shed loads from EVs at peak times to avoid congestion and reliability problems. This measure, however, is likely to increase customers' uncertainty about potential charging patterns, and might affect the attractiveness of EVs in general (IEA, 2013).

Demand response has to address distribution system needs

Demand response services provide benefits both to the electricity market and to distribution networks. For example, if excess renewable generation and the resulting low market prices trigger additional demand, this added demand could exceed today's network capabilities and thus trigger a need for additional investment in distribution networks (DENA, 2012). This example illustrates the potential need for co-ordination between the market-based and network-based benefits of demand response.

A similar impact can be expected from the introduction and use of various storage technologies. Co-ordination of demand response and storage-providing services to the electricity market and to the network are achievable by introducing two-tier, real-time price formation, with different prices for the network and the electricity component.

8.2. Design of market platforms for distributed resources

Unlocking participation of distributed resources in the market

Microgrids and potential of new technologies

The possibility of expanding the market model to local electricity systems is receiving increasing attention. In contrast to decades of prioritising economies of scale and centralisation, today more generation plants are connected at medium and low voltages, while demand response and storage are becoming available. Distribution network operators and regulators are seeking new ways to operate their network.

Free-standing microgrids are not a new concept, but remain a niche market. They are intended to replace a grid connection for improved reliability and have been developed mainly for the military, academic campuses or industrial uses. Retrofitting existing equipment for these off-grid solutions is, however, relatively costly. Costs can be lower where completely new districts or industrial cities are being built, but their numbers remains low in slow-growth OECD economies. In addition, microgrids are mainly technical solutions to allow central control of electrical equipment at a local level, and are not intended to be market platforms.

Meanwhile, regulators and policymakers are also actively promoting the idea of energy independence at the margins of their jurisdiction. This can be done at the level of a municipality, a district, a department or a state, depending on the extent of the local authority. In these cases, electricity systems are much larger than microgrids. The rapid development of solar PV is presented as an opportunity to exchange excess electricity with neighbours during sunny hours and store power by using batteries.

The best-known initiative has been promoted by the Public Service Commission (PSC) of New York under the name of Reforming the Energy Vision (REV). The REV aims to speed up the transition to energy efficiency and renewables, both by overhauling the regulations that govern utility companies and by designing new energy markets at the distribution level. As part of this plan, the utility acting as a distributed system platform provider (DSPP) will actively co-ordinate customer activities. The goal is to give consumers more control over their energy use and engage them as energy producers. The function of the DSPP will be complemented by competitive energy service providers; both generators and retailers of electricity are encouraged to expand their business models to participate in DER markets, co-ordinated by the DSPP. One driver of this REV initiative is to minimise costs by delaying costly network investments through the more active participation of demand and local resources, such as batteries, solar PV and back-up generators.

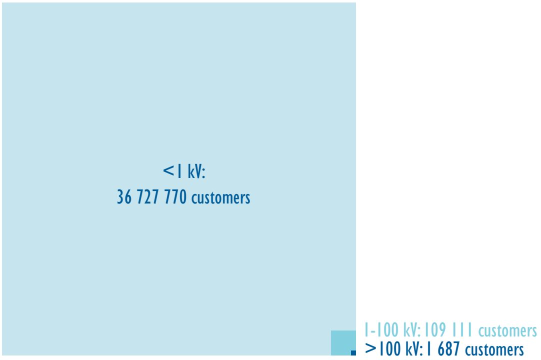
Similar visions have been developed in Europe by the Council of European Energy Regulators (CEER), which has put forward the notion of distribution system operators (DSOs) as neutral market facilitators to enable the development of new market-based services and to ensure secure system operation (CEER, 2015). In order to do this, DSOs will need to manage their networks actively by employing smart grid solutions and creating innovative investments. Network co-ordination matters will be co-ordinated between DSOs and transmission system operators (TSOs). Additionally, DSOs will need to adapt their networks to meet new demands, such as EV recharging stations and compressed natural gas filling stations. Cyber security, data management and data privacy have been identified as critical issues that should not be overlooked.

The Internet of Things plays a key role in these visions. There is little doubt that the internet, and more generally information and telecommunication technologies, are spreading into all sectors of the economy. In this respect, the electricity sector, while a promising candidate, is also a latecomer. Inexpensive two-way communication systems are now available for almost all electric appliances, from home appliances to heaters or even lighting. New business models are also based on micro-transactions of a few cents on mobile phones for so called free-to-start games. The electricity sector can expect to see massive change in the coming decades.

Simplifying market participation for small consumers and producers

Mass business models capable of handling these technologies are yet to be found. Indeed, the solutions developed to co-ordinate, for instance, 300 generators and 1 000 large consumers cannot be transposed to the mass market of 30 million customers connected to the low-voltage network with a dozen controllable electrical appliances, batteries or solar PV systems each (Figure 8.3).

Figure 8.3 • Number of consumers connected at different voltages in Italy



Note: kV = kilovolt.
Source: Massimiano, M. (2015)

Transaction costs remain a key barrier to the participation of distributed resources in existing market arrangements. Wholesale markets usually limit participation to bids above a couple of megawatts, to keep the computational system tractable and reduce transaction costs. It is clear that a small rooftop solar PV generator will never submit five-minute bids on the day-ahead market. Electricity markets are far too complex, even for full-time market participants. Simple solutions are needed.

Furthermore, the gains associated with market participation are usually relatively small. For example, active demand response participation can save perhaps a few percent on bills, say USD 50 maximum, but most consumers will not expend much time or energy to achieve this saving. Simple and scalable solutions have to be found so people can limit the time they spend on this to a couple of hours a year.

More importantly, consumers or small generators will not participate if the gains associated with participation are not easy to understand and the returns sufficiently certain to justify investment. Understanding complex and uncertain gains with little time is simply not feasible. Here again, direct participation in complex hourly market platforms is not possible and simple solutions have to be implemented.

Against this background, retailers, distribution companies and aggregators have a key role to play in providing a connection between complex electricity markets and consumers or small producers. Such solutions are likely to involve long-term contracts to justify investment, with simplified propositions such as interruptible contracts, contracts with terms similar to critical peak pricing with red hours and green hours, or automated solutions. A more sophisticated approach could be to translate these into energy services, for instance higher or lower comfort levels. Lower comfort temperatures would be less expensive but the temperature of the apartment would be lower on winter days. A more detailed discussion is provided in Chapter 6 on demand response.

Co-ordination of distributed market platforms with short-term wholesale markets

Added complexity: Unbundling between the distribution network and retailers

The development of a market platform for distributed resources can be facilitated by the deployment of advanced metering infrastructure (AMI), a meter data management system (MDMS), customer web portals, and an outage management system (OMS). These systems build on the functionality of existing smart grids to provide customers with previously unavailable options. For example, AMI supports potential implementation of time-based rate programmes that can help customers reduce peak loads and lower their monthly bills. The integration of AMI with the new OMS provides improved outage management and restoration services. This combination of advanced technologies and new data analysis capabilities enables more efficient design and operation of the electricity distribution system, resulting in differentiated quality of service.

In jurisdictions with no retail competition, the distribution company is at the same time the retailer of energy. It is not surprising that most experiments in the United States have taken place in states with no retail competition; vertical integration makes the co-ordination of scattered small gains easier. Initial roll-outs covered, for example, several hundred thousand customers in Colorado, South Dakota, Maine and Wyoming. Funding for many of these smart grid projects was provided by the American Recovery and Reinvestment Act of 2009.

By contrast, the unbundling of distribution companies and retailers increases complexity. Distributed resources can provide gains for distribution companies (for example, by delaying investment) and retailers (by lowering energy bills). It is therefore necessary *ex ante* to quantify the share of benefits that will accrue to distribution companies (in terms of delayed investment) and the share of benefits for retailers (in terms of lower electricity costs). In practice, this raises issues for the design of network access tariffs and the definition of commercial offers by retailers – they need to be simple while at the same time capturing as far as possible the expected benefits from a wholesale market perspective and from a distribution network perspective.

Once agreed and deployed, the distributed resources have to be visible and should be controllable by system operators in case of emergency situations. This increases the cost of information technology (IT) systems and can create conflicts of use. For example, the definition of hours when distributed resources should be activated can differ from a network and a retailer perspective. This is a source of potential conflict or inefficiency.

Adding other market participants, such as demand response aggregators, further complicates the picture. Even if contractual arrangements can, in principle, solve these issues, regulators are likely to have to define the rules regarding the distribution grid and participation of aggregators.

Coupling the distributed market platform and wholesale markets

A distributed market platform would, in effect, create a local market price with a geographic resolution even higher than a nodal pricing system (see Chapter 3). At this local level, it appears that the traditional locational marginal pricing model with central dispatching of resources would be difficult to realise. Despite increasing computational power, central computers are limited in what they can control in real time.

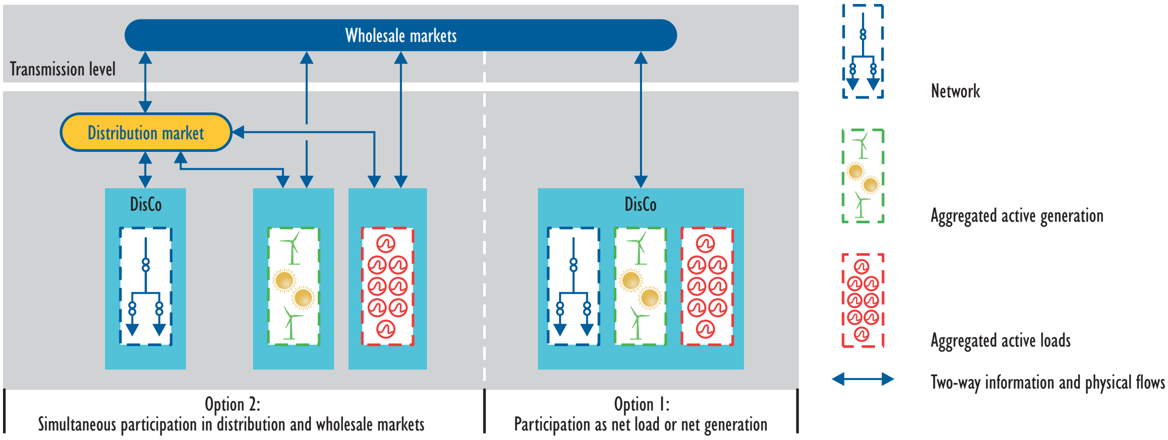
Several possible models could circumvent this.

One option is for each distributed market platform to aggregate local generation and load, and only to participate in the wholesale market according to the net load or net generation at its points of connection. A distribution company would be seen to be a significant resource from a market perspective, and could submit generation and demand bids on wholesale electricity markets for its load at each point of connection. This model, however, does not seem compatible with retail competition.

Another possibility is to enable simultaneous participation on a wholesale market platform and a local distributed services platform. This model is more complex but is compatible with retail competition, and has been analysed in a European context by CEER. A distribution market platform ensures that capabilities developed by retailers and aggregators can be activated to meet the needs of distribution networks, in addition to their participation in wholesale electricity markets (Figure 8.4). In practice, this means that bids have to be submitted on electricity markets with indications of location. Aggregation can take place either at a market-wide scale or at a local scale in order to allow activation for distribution network purposes. The bids are not plant specific, but the portfolio of resources has to be location specific, if necessary down to the level of electric feeder.

Co-ordination between distribution and wholesale levels is becoming more complicated. In general, distribution is passive and load forecast depends on temperature and weather conditions. Once distribution systems become active, this can change load forecasts at the wholesale level. When a distribution company reduces load on its service area, this also reduces load for the wholesale market. This creates the need for continuous information updates. Conversely, distributed resources can contribute to addressing wholesale system needs in the intraday and day-ahead timeframes. The response of retailers or aggregators to wholesale prices can create local congestion problems on the distribution network.

Figure 8.4 • Distribution market-wholesale market interface



Note: DisCo = distribution company.

Transparent short-term markets are useful for the secure co-ordination of distributed resources, distribution networks and the wholesale system. Retailers and distribution companies have to be able to react to the evolution of intra-day prices by activating their portfolio of resources, and such activations should be translated into updated schedules on the short-term markets. A necessary condition for these decentralised reactions is for intraday prices to be transparent and to reflect the marginal cost of resources with the highest resolution possible, in particular geographically, so as to fit with the local nature of distributed resources.

The co-ordination of distributed and wholesale short-term markets involves striking a balance between centralised dispatch and decentralised updates of generation and demand schedules. Transparent location-specific intraday prices would be important for ensuring price co-ordination of centralised resources, distributed resources, aggregators, distribution companies and retailers.

8.3. Towards DSO regulation 2.0

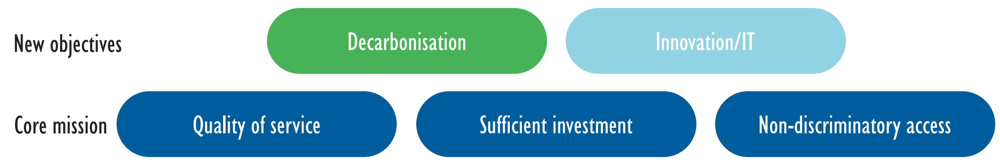
Regulation is not only about cost minimisation – providing an incentive for investment is also becoming increasingly important and is an additional goal. Quality of service, the introduction of smart technologies and innovation are also gaining importance.

Regulatory incentive mechanisms or price caps were introduced in the 1990s. The aim of such measures is to substitute market incentives by making use of energy firm's information advantage and profit motive to bring improvements in efficiency. In this way, the regulator exerts less control over behaviour but instead rewards outcomes (Vogelsang, 2002). Although it concerns a regulated segment, incentive regulation has proved to be an important benefit of the electricity sector restructuring that took place in the 1990s (Joskow, 2008).

It was anticipated that regulatory incentive mechanisms would increase efforts by regulated firms to reduce costs, improve service quality in a cost-effective way, stimulate the introduction of new products and services, and stimulate efficient investment in, and pricing of access to, regulated network infrastructure services. It has generally proved successful, especially in relation to efficiency gains and cost reductions in transmission and distribution network operations.

Numerous new challenges face distribution networks and therefore their regulation (Figure 8.5). In many OECD countries, distribution networks are ageing and technological change brings huge uncertainties for the future of the networks and their regulation. At this stage, the future energy technology mix is quite difficult to foresee and will depend to a significant extent on policy barriers or incentives and regulation. The actual amount of DER to be integrated in the network might vary and could change quickly. Additionally, network tariffs are set to increase in the future. In Germany, for example, the anticipated overall cost increase in distributed systems will amount to at least 10% by 2032. This may lead to significant increases in costs in individual dustribution networks. Crucially, discrete investment decisions are unlikely to bring a steadily rolled-out increase in network capacity that precisely fits the overall needs of the network. Lumpiness of investment often leads to temporary asymmetry.

Figure 8.5 • Core mission and new objectives of distribution network regulation



Network regulation can either facilitate or hinder the construction and use of decentralised energy resources. Regulation needs to shift away from exclusively focusing on incentives for reducing operating expenses without paying attention to investment. Distribution investment is becoming an increasingly important part of regulation. An efficient regulatory framework has to aim for an optimum balance between operating and maintenance costs and investment, and to take into account the fact that networks enable the functioning of the electricity market. Consequently, a system-wide perspective is needed when making framework changes.

Network regulation encompasses several activities that can facilitate the decarbonisation of the electricity market, including:

- defining the (new) activities that network operators are allowed to carry out, and how they should co-ordinate with transmission system operators
- monitoring network planning and approving investments (and connections)
- setting regulated revenues and the corresponding levels of network tariffs.

Establishing a sound system of regulation can provide suitable incentives for network owners and operators to foster the integration of distributed energy resources and new technologies.

Do market platforms represent a new role for DSOs?

Decentralised resources are prompting the evolution of the scope of activities of regulated distribution companies, as solar PV systems, batteries and demand response capabilities are connected to the distribution grid. Regulators now have to redefine the frontier between regulated activities, subject to price regulation, and market activities so as to spur the deployment of new technologies while also preventing cross-subsidy from regulated activities to market-based activities.

The role that DSOs should play as "neutral market facilitators" has been widely discussed, both in the United States and in Europe (CEER, 2015). A wide range of participants, including retailers, information and communication technology (ICT) companies and metering companies, are now conducting business in the balancing, ancillary services and power markets. With the increase in distributed energy resources, the number of new market participants will similarly increase, and new business models will emerge with market participants offering different kinds of energy services and aggregating from various sources. Consequently, the complexity of the system will increase and the role of the DSO should be clearly defined to facilitate entry by new market players, by creating a level playing field.

Without doubt the DSO's core tasks remain investing, maintaining and operating the distribution network, and these activities clearly need to be regulated due to their monopolistic nature. For the other activities, such as acting as a market platform or managing data, however, several forms of organisation are emerging.

Several of these new tasks could be fulfilled by market participants or the DSO, or by third parties, regulated or non-regulated. Such tasks include ownership and management of meters, data handling and the charging structure for EVs. Here, decisions on IT become an important issue as they can enable market participants to actively take part in local energy markets close to real time. Defining the responsibilities of different participants is important to facilitate the wide deployment of several of the envisaged technologies.

The distributed system platform provider (DSPP) proposal currently being discussed in New York is one of the first practical steps to regulating the active participation of generators and consumers at the distribution level. The DSPP is envisaged to be an intelligent market platform that provides safe, reliable and efficient electricity services by integrating diverse resources to meet customers' and society's evolving needs. The DSPP is intended to foster a wide range of market activity that monetises system and social values, by enabling active customer and third-party engagement while being aligned with the wholesale market and bulk power system.

Box 8.2 • The future role of DSOs: a European perspective

The European Union continues to discuss the future role of the DSO, with common agreement that increasing volumes of DER offer new ways for DSOs to fulfil their role. DER not only offers solutions for grid operation, but can also have positive effects on longer-term planning for grid investment, being used to manage congestion in the short run and avoid or postpone future grid investments.

Overall, they provide DSOs with a wide range of additional instruments. But DSOs need to have an active role in order to efficiently use DER. This can be made possible through procurement processes similar to those facilitated by the TSOs for balancing or ancillary services. DER technologies can also provide services to other market agents, such as retailers or aggregators, who could either use them or resell them in the ancillary service market of the TSO. The DSO is therefore competing directly with other energy market participants for DER services, which presents the risk that DSOs might abuse their role as market facilitator. Therefore, clear regulation is needed to encourage the DSO to start a new market platform to procure clearly defined system services at the distribution level in a transparent way.

Two-way co-operation between DSOs and TSOs

At a high level of DER penetration, a structured exchange of information between TSOs and DSOs will become necessary. In countries where renewables (especially variable wind and solar) are concentrated in certain areas, the feed-in of electricity from the distribution to the transmission grid increases burdens on the transmission network. Enhanced co-ordination and communication between DSO and TSO systems is therefore needed to enable TSOs to take into account DSOs' planning and management.

Centralised generators send schedules to the TSO in order to balance the system, but at the distribution level, DSOs do not yet have systems to acquire equivalent data from distributed generation. As the DSOs gain information about the forecasts and schedules from distributed generation, they will become able to manage local network constraints. The action then taken by the DSO on the basis of this information might either help transmission network management, or could in some cases increase the operational difficulties at the TSO level.

Similarly, flexible resources participating in the balancing market of the bulk power system can cause unexpected congestion at the DSO level. Hence, strong co-operation is needed to manage short-term markets, so as to avoid action by one party creating new obstacles for the other.

The two-way interface between transmission and distribution necessitates an increased degree of co-operation both for system planning and for short-run operations.

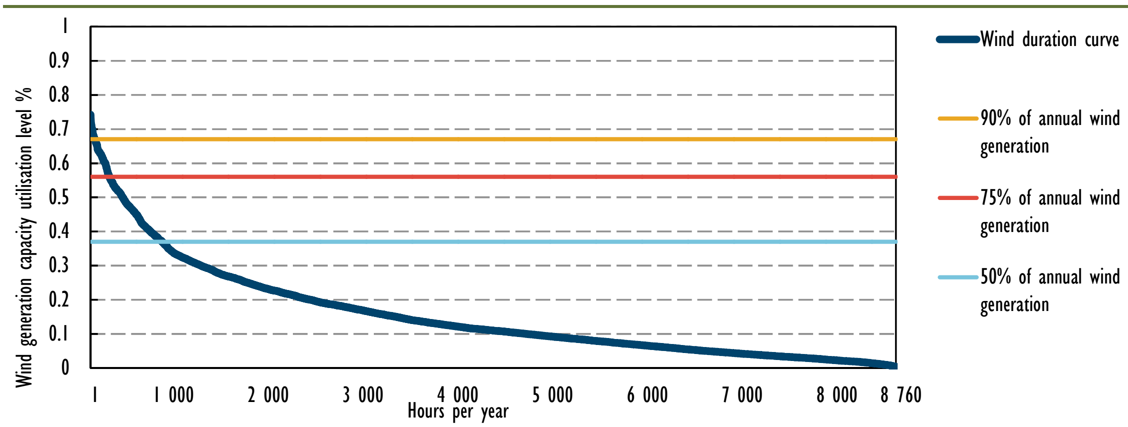
Planning network investment with renewables

One important dimension of network planning is the ability of renewables generators to decrease their production or to curtail. For a distribution company, having the ability to control distributed resources and curtail, to a limited extent, can considerably reduce infrastructure reinforcement costs, and therefore accelerate the integration of wind and solar power.

Some regulatory frameworks oblige network developers to design renewable generator connections and to mitigate system-wide network congestion in order to fully integrate renewable electricity into the network, as well as to compensate renewable generators for any foregone revenues. Whilst giving additional investment certainty to renewable generators, this approach is likely to lead to further uneconomic network investment decisions and result in excessive costs to the electricity system and its users.

At the European level, network infrastructure investment usually results from a requirement to enable 100% feed-in of wind and solar generation. In Germany, the annual onshore wind generation pattern amounts to roughly 25% of average generation capacity utilisation (hours of full load). The annual generation capacity utilisation curve (wind duration curve) of the total installed wind generation capacity (Figure 8.6) shows a steep decline, which implies that capacity utilisation shares above 50% happen in fewer than 500 hours per year. Nevertheless, this steep ramp down also implies that the 220 hours with the highest capacity utilisation share contributed only 10% of the overall annual wind generation. In view of this, the question of efficient network infrastructure investment, as determined by cost-benefit assessments, is clearly important.

Figure 8.6 • Annual generation wind duration curve in Germany



Source: EEX transparency.

Management of generation of renewables such as wind and solar should already be considered in network planning to avoid network reinforcement for the last kilowatt hour. Extending the network infrastructure to meet production in these few hours leads to investments whose benefits are inadequate in relation to costs.

In network planning, a reduction in feed-in requirements for wind and solar generation should be considered. In the management of energy networks, operators of the distributed system should also be able to further decrease variable renewable generation in order to give the network operator a certain degree of flexibility in its daily management. To ensure that such curtailments are efficient, and to avoid disproportionately frequent impacts on individual generators, the chronology of shut-downs should follow technical as well as economic criteria so as to deliver security of supply at low cost.

The regulatory regime also needs to take into account differences between network operators. The network investments spurred by the expansion of wind and solar are not equally distributed throughout the system; therefore the impact on network expansion necessarily varies. Decisions on the configuration of the network and its management through the use of smart technologies should remain at a local level with the distribution operator.

Achieving a suitable combination of network planning, management of generation and new installations of variable ratio transformers can reduce the need for network reinforcement. In Germany, for example, studies suggest reductions on the scale of 20% for costs and 60% for network expansion (DENA, 2012). In view of overall cost efficiency, regulation should therefore foster both integrated planning and smart technologies. By using smarter planning concepts, overall costs decrease while operating costs increase. Consequently, any regulatory regime incentivising network operators to focus on short-term gains through reduction of operating costs will not be helpful.

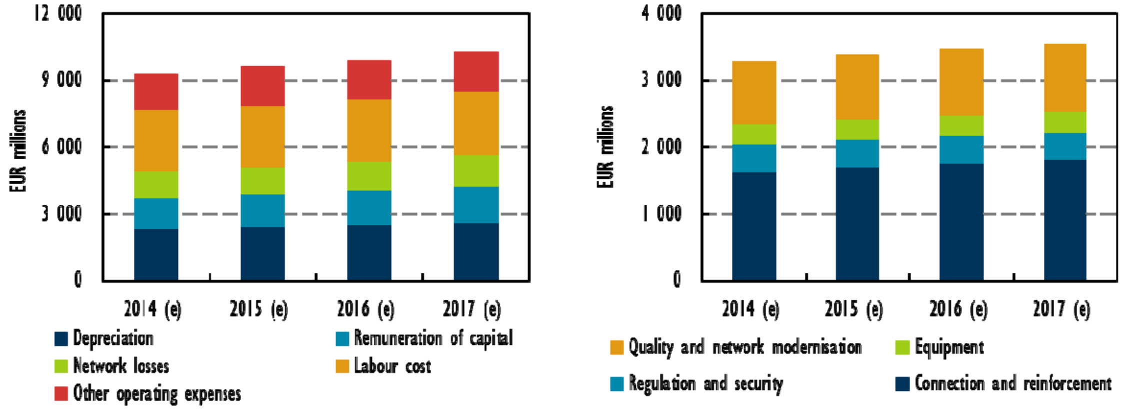
Regulating investments, however, remains a difficult task for regulators. Distribution network investments are expensive, lumpy and small in scale, causing information asymmetry on investment costs between the regulator and the regulated. A further complication comes from the fact that investments have to be made in anticipation of a prospective future deployment of distributed resources.

Regulation of allowed revenues

After the definition of regulated activities and the approval of investment plans, the key building block of economic regulation is the determination of allowed revenues. Regulated revenues are used to calculate network prices, either by the regulator (France, Spain) or by regulated companies (Germany, the United Kingdom). As illustrated in Figure 8.7 below, revenues are related to network investment costs, in terms of depreciation of assets and the remuneration of capital. This should not be surprising given that managing networks is a very capital-intensive activity. In practice, regulators approve investments under a position of information asymmetry *vis-à-vis* the regulated company, which is a difficult task and can lead to overinvestment that will increase the allowed revenues and the corresponding network tariff for decades to come.

Current network tariffs typically represent 20% to 40% of electricity bills in most countries, depending on the scope of network activities, geography (per unit cost is lower in more densely populated countries) and age of the network. Existing distribution networks in OECD countries were mainly built in the 1960s and 1970s meaning that the assets are largely depreciated, which keeps tariffs relatively low.

Figure 8.7 • Evolution of the allowed revenues (left) and projected investment (right) of the Electricité Réseau de Distribution France (ERDF) for the regulatory period 2014-17

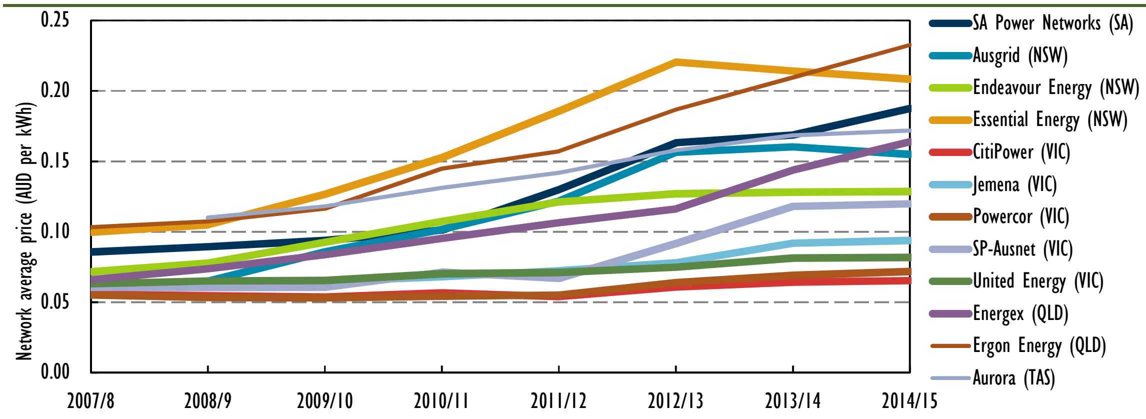


Note: (e) = estimate.

Source: based on CRE, 2013.

Looking ahead, massive investment in networks will significantly increase electricity bills. In Australia, for example, network charges for consumers in South Australia, Queensland and New South Wales more than doubled between 2008/09 and 2014/15 (Figure 8.8). This price increase has taken place at a time when electricity consumption from the grid is mostly stable but often declining, due to the combined effects of lower economic growth, improvements in energy efficiency and the development of behind-the-meter generation that further reduces the billing base of distribution companies (see Chapter 9). At some point, the regulator's determination of allowed revenues might need to take into account the feasibility of raising these revenues from consumers, to avoid a non-negligible risk of seeing stranded network investments.

Figure 8.8 • Network tariff development in selected states of Australia



Notes: NSW = New South Wales; QLD = Queensland; SA = South Australia; TAS = Tasmania; VIC = Victoria.

Source: CME, 2015.

The increasing role of DER, especially from variable wind and solar energy, also affects DSOs' total costs. Connecting DER requires network investment to cope with new flows and with the volatility in flows and demand fluctuation, although ICT structures and DER together promise a new set of instruments to allow better operation of the network to minimise local congestion.

In this context, DSOs' balance between the operating and capital expenditure (OPEX and CAPEX) is also set to change. The services provided by DER are likely to enable a decrease in unit OPEX costs, as they can replace more costly internal operations. With regard to CAPEX, on the one hand the use of DER might limit the level of grid investment needed, by reducing length of lines. On the other hand, smarter distribution grids also require an additional layer of investment in smart devices.

In traditional incentive-based regulation, DSO remuneration is usually regulated for a certain period for a specific range of services. The corresponding CAPEX and OPEX can be treated either individually or together as total expenditure (TOTEX). One practical issue faced by regulation in the light of future investment in grid technology is that investments are being undertaken in one regulatory period, while their benefits can only be fully realised during the following regulatory period. This accounting problem might potentially hinder innovative investment. This calls for an adaptation of regulation, including changing the length of the regulatory period, adapting the benchmarking process and increasing the role of output-based regulation.

Innovative approaches to assessing investments in distribution grids have been introduced in Spain, using reference network models to estimate efficient distribution costs and to plan distribution networks. This approach can be an alternative to adapting the traditional benchmarking of OPEX and CAPEX. Alternatively, in the United Kingdom output-oriented regulation and an increase in the regulatory period have been introduced to incentive DSOs to invest in smart infrastructure.

Benchmarking

Benchmarking DSO costs remains an important part of designing regulation to provide the right incentives – known as “incentive regulation”. Gathering information for benchmarking also provides an opportunity for regulators to gain deeper insight into the cost structure of DSOs.

As DSOs face contrasting challenges from renewables and distributed generation, the choice of indicators and metrics for benchmarking deserves attention. In particular, investment in smarter distribution network equipment needs to be considered to allow benchmarking to present a realistic picture of the overall efficiency of a DSO. A simple increase in the length of a network does not provide an adequate picture, as high-technology alternatives might be less costly, reduce network enforcement and induce the efficient use of the network.

Regulatory frameworks for gas and electricity distribution networks today remain simple and rather similar, being focused on evaluating efficiency and how this is converted into revenue. The sound implementation of incentive regulation mechanisms mainly depends on information gathering, auditing, and accounting – benchmarking of costs is therefore fundamental. Incentive regulation mechanisms have this in common with traditional cost-of-service or rate-of-return regulation (Joskow, 2008).

With benchmarking, a DSO’s performance is compared to the performance of other comparable DSOs, and penalties or rewards are assessed based on the relative performance. For instance, the regulator might identify a number of comparable operators, compare their cost efficiency and regulate them as follows:

- The most efficient operators set the benchmark and are allowed to recoup 100% of these identified costs, and at the same time can reap profit increases from cost reductions.
- Less efficient operators are allowed to recover a decreasing share of their initial costs during the following regulatory period.

One of the main challenges for benchmarking network operators is therefore the different environments they face; it is important to make sure that the operators' situations are similar and to use statistical techniques to adjust for any quantifiable differences over which the operators have no control.

For small DSOs, certain network regulations in OECD countries allow for exemption from and simplification of the benchmarking process, in order to reduce regulatory costs. Nevertheless, it is important to ensure fair treatment; the scope of exemptions and simplifications should be limited.

Increasing the duration of the regulatory period

The length of the regulatory period is an important factor in incentive regulation. On the one hand, the regulatory period must be long enough for network owners to implement initiatives to reduce cost and enjoy the resulting profits over a reasonable period of time. On the other, the longer the regulatory period, the longer customers wait to share in the benefits of over-performance. Therefore, regulatory periods in OECD countries tend to be in the range of four to five years.

In countries with a longer history of network regulation, the main ways to achieve efficiency gains have already been undertaken. In such countries, extending the regulatory period could provide several benefits. A slightly longer timeframe better aligns the incentives of distribution companies, as they can benefit from longer-term investments and management decisions, in the form of higher profits. Upping regulatory certainty can also improve network planning, and so could help to foster innovative investment.

Input- vs. outcome-based regulation

Incentive regulation has delivered efficiency gains in the form of cost reductions after periods of inefficient management. Most existing regulatory regimes today focus on comparing DSOs with one another and with their own past performance. The focus of network regulation needs to evolve beyond this. In the future, DSOs will be expected to carry out new functions and cope with changes in distributed generation. DSOs need to invest to maintain the efficiency and security of the system at least cost. The majority of current regulatory frameworks, however, are not well equipped to incentivise DSOs to exploit all technical and managerial options.

To date, success has rarely been measured as optimal performance. In some regulatory regimes, a quality element has been added to an incentive regulation to ensure high network quality is maintained, but this element usually plays a minor rule, as poor quality was not a problem in most OECD countries.

Regulation should provide a greater incentive for DSOs to increase their focus on output. An output-based model would shift the focus of DSOs to striving for long-term results, such as customer value or quality of service. The regulatory regime has to mirror the attributes demanded by society. With this perspective, the outputs against which the performance of DSOs is measured should include a broad base of quantifiable indicators that are transparent and easy to assess and measure. Table 8.3 provides an indicative list of such output indicators.

Table 8.3 • Indicative list of outputs to measure the performance of Distribution Network Operators (DNOs)

Environment	Help to minimise the environmental impact of DNO operations and of users by: ensuring low-carbon technologies can be connected at a reasonable price and in a timely manner; managing losses; minimising business carbon footprint; minimising material non-carbon emissions; using a stakeholder-based approach to visual amenity in areas of outstanding natural beauty or national parks; potentially encouraging any role the DNOs might play in local authorities' integrated energy schemes.
Reliability	Maintain operational performance for existing and future customers by improving existing reliability indices and expanding these output measures to include network risk and criticality; incentivise reliability using the interruptions incentive scheme and guaranteed standards, and have regard to the worst-served customers.
Connections	Connect users and suppliers of energy in a timely and cost-effective manner and provide high quality information, in a transparent way; incorporate the societal costs of delays to connection in the output baseline and/or incentive mechanism.
Customer service	Maintain levels of customer satisfaction through the broad measure of customer satisfaction.
Safety	Maintain compliance with Health and Safety Executive requirements.
Social obligations	Could include initiatives to target the fuel poor.

Source: Ofgem, 2012.

This outline recognises that developing output parameters is a difficult task in the absence of a track record for these indicators. Additionally, a wide variety of parameters needs to be created in order to avoid gaps in incentives. Otherwise, DSOs might focus on certain specific areas and metrics and neglect other important areas where no incentive has been incorporated.

It also has to be recognised that a purely output-based approach would place too much risk on the side of the DSO. Given the important function of the DSO and the sector's huge investment needs, inputs also need to continue being regulated. It is important to keep a balance between input parameters and output parameters.

Managing the transition from incentive-based regulation towards a more output-based regulation creates further uncertainty for any regulated company. To smooth the transition, the introduction of output performance could be progressively included as an element in the regulation of network service quality.

Symmetry of incentives: A crucial factor

By and large, incentives in incentive-based regulation have been set on the assumption that a DSO has an obligation to manage and invest in the distribution grid, and is therefore given the opportunity to recover its costs and earn a sufficient return on its investment. The incentives are unidirectional and negative, decreasing revenue in case certain requirements are not met. Under this approach, regulators perceive positive incentives as creating unnecessary profits, as the DSO is already able to gain a sufficient return on equity. The outcome of this approach is that the DSO cuts its spending as much as possible to fulfil its basic obligation and any implemented quality regulation.

In the past, this approach has led to good results by increasing efficiency in transmission and distribution networks. However, it provides no incentive to provide a superior service and meet overall economic goals with an increase in the "smartness" of the network. With a changing scope of distribution networks, especially the necessity to invest in new technologies, it will be necessary to consider a more symmetrical approach to incentives. DSOs should be rewarded if they achieve better results in areas of innovation and customer services. Regulation should optimise the level of inputs to achieve policy outcomes overall and not only secure near-term reductions in expenses.

Capital costs vs. operating costs, regulation for innovation

At present, regulation remunerates capital investment with a reasonable rate of return. With this approach, DSOs have no incentive to invest in sustainable and smarter distribution networks, but instead continue to invest in traditional network assets to increase the capacity of lines even if cheaper solutions could be found. This runs counter to the efficient utilisation of distributed energy resources in the future. Therefore, regulation should consider approaches that encourage the DSO to use an efficient allocation of capital and operating expenses to bring forward regulatory objectives.

DSOs will generally invest in innovative smart technologies only where the risk of failure or a decrease in efficiency is sufficiently low and has no influence on the benchmarking used in the regulatory process. This presents the problem of DSOs taking investment decisions only where a sufficient return can be made during the regulatory period. In order to foster further investment in smarter network technologies, additional incentives or support might be necessary.

Box 8.3 • Output regulation in Great Britain

British distribution network operators (DNOs) distribute energy to about 28 million customers across 242 000 square kilometres. The distribution network of up to 132 kV with a length of about 767 000 kilometres. In Britain, 14 DNOs manage the grid, which are owned and operated by 6 private companies. The size of DNOs varies between 600 000 clients and 3.3 million clients.

Britain has extensive experience with incentive regulation, with the so-called RPI-X regulation, which it implemented in the early 1990s, improving monitoring and benchmarking of it during the following decade. Initial incentives for output and add-ons for new investments were implemented to encourage innovation and social and environmental responsibility.

In 2010, the RIIO (which stands for Revenue = Incentives + Innovations + Outputs) process was started to increase focus on innovation and investment, in addition to efficient network operations.

The extended regulatory period of eight years aims to incentivise the DNOs towards long-term planning. In conjunction with a greater reliance on output indicators, it should increase the signal to DNOs to strive for the articulated policy objectives. A safety net has also been implemented in the form of the possibility of annual reopening of the revenue cap, and pass-through costs for uncontrollable costs, uncertainty and investment shortfalls. In addition to output factors, the process includes incentives that can generate positive income for DNOs. This increases the focus on these outputs. The total expenditure (TOTEX) approach aims to create an equal incentive to reduce overall costs. In addition, Ofgem can grant an increase in WACC if higher investment risk due to innovation can be proven (Ofgem, 2010).

The most crucial change is to the output-oriented incentive scheme. If the DNO fulfils the required targets in these output categories, it can gain both monetary and non-monetary rewards. The following output parameters are taken into account.

A non-monetary reward would be, for example, publication of information that might cause an increase or decrease in reputation. Monetary incentives derive from an increase or decrease in the revenue cap or an increase or decrease in the allowed rate of return. Reliability of supply, for example, is calculated by the system average interruption frequency index (SAIFI) and system average interruption duration index (SAIDI). The possible increase or decrease of return is between +/- 1.5% and 2% of the WACC.

Because the first regulatory period in the new regime started only in 2013 for electricity and gas TSOs, and in 2015 for electricity DNOs, costs and results cannot yet be evaluated. Due to the relatively small number of network operators, Ofgem will later be able to carefully assess each network operator.

Incentives to innovate can take the form of tenders for financial support for innovative projects and field studies. If public money is spent, the results of these studies could also be disseminated to other DSOs. In the United Kingdom, innovation is being supported with two mechanisms. First, network operators can increase their return on equity by between 0.5% and 1% if they can prove an excellent innovation strategy with the national regulator, Ofgem (Office of Gas and Electricity Markets). Second, network operators can apply for tendered innovation subsidies through Network Innovation Competitions (for DSOs and TSOs). For the first two years of the current regulatory period, GBP 90 million per year is available. Network operators can submit up to three projects and gain up to GBP 10 million per project.

Even some of the more mature technologies might still be more risky for DSOs than investment in business-as-usual technology, and here also implementation and deployment may require additional incentives. In 2011, Italy introduced additional remuneration above the weighted average cost of capital (WACC) for modernising distribution networks, to incentivise the deployment of control, regulation and management of load and generating units and EV charging systems.

Conclusion

The role of distribution networks is changing. Historically, DSOs have mostly provided one-way channels for electricity to flow from the transmission level down to end customers. In the future, distribution networks face three main challenges: integrating new and variable renewable sources today and EVs in the future; enhancing customers' market participation; and interfacing with bulk transmission networks and the wholesale market.

To address all of these challenges, regulatory frameworks will have to be designed to handle efficiently a large number of heterogeneous distribution networks. The introduction of ICT is crucial, and enables several of the new tasks associated with managing distributed resources to be fulfilled either by market participants, the DSO or even third parties. The possibility of controlling and managing the output of variable renewable generation, such as wind and solar, should be considered in network planning to avoid expensive network reinforcements.

However, the investment decisions of distribution companies will continue to be shaped by the incentives provided by regulation. Assessing the suitability of any local investment is a very difficult task for a regulator, and therefore the decision should remain at the DSO level. With that perspective, introducing additional output parameters in network regulation could foster the necessary changes to the network infrastructure and provide the incentive for further deployment of smart technologies.

With the growth of DER, reforming the design and structure of network tariffs is becoming a pressing challenge, an issue that is further discussed in Chapter 9.

References

BMWi (Federal Ministry of Economy and Energy) (2014), "Moderne Verteilernetze für Deutschland" [Modern distribution networks for Germany], study commissioned by the Federal Ministry of Economy and Energy (BMWi], www.bmwi.de/DE/Mediathek/publikationen,did=654018.html.

BnetzA (2013), Genehmiung des Szenarioahmens fuer die Netzentwicklungsplanung und Offshore- Netzentwicklungsplanung, www.netzausbau.de/SharedDocs/Downloads/DE/Charlie/Szenariorahmen/Szenariorahmen2024_Genehmigung.pdf?__blob=publicationFile.

CEC (California Energy Commission) (2013), *Integrated Energy Policy Report 2013*, www.energy.ca.gov/2013publications/CEC-100-2013-001/CEC-100-2013-001-CMF.pdf.

CEC (2009), "Transportation energy forecasts and analyses for the 2009 integrated energy policy report", www.energy.ca.gov/2010publications/CEC-600-2010-002/CEC-600-2010-002-SF.PDF.

CEER (Council of European Energy Regulators)(2015), *The Future Role of DSOs: A CEER Conclusions Paper*, www.ceer.eu/portal/page/portal/EER_HOME/EER_PUBLICATIONS/CEER_PAPERS/Cross-Sectoral/Tab1/C15-DSO-16-03_DSO%20Conclusions_13%20July%202015.pdf.

CME (2015), Network tariffs applicable to households in Australia: Empirical evidence, http://cmeaustralia.com.au/wp-content/uploads/2013/09/FINAL-PAPER-1-EMPIRICS-.pdf.

CRE (2013), Projet de décision relative aux tarifs d'utilisation d'un réseau public d'électricité dans le domaine de tension HTA ou BT, Délibération de la Commission de Régulation de l'Energie du 13 Novembre 2013 [Draft decision on the tariffs for the use of a public electricity network in the area of voltage HTA or BT, Deliberation of the Committee of Regulation of energy from 13 November 2013], www.cre.fr/documents/deliberations/decision/turpe-4-hta-bt/consulter-la-deliberation.

DENA (German Energy Agency) (2012), "Integration der erneuerbaren Energien in den deutschen/europäischen Strommarkt" [Integration of renewable energies in the German/European electricity market], Deutsche Energie-Agentur GmbH, www.dena.de/fileadmin/user_upload/Presse/Meldungen/2012/Endbericht_Integration_EE.pdf.

Eurelectric (2013), Power Distribution in Europe Facts & Figures, Union of the Electricity Industry www.eurelectric.org/media/113155/dso_report-web_final-2013-030-0764-01-e.pdf

ICCT (International Council on Clean Transportation) (2013), "Electric vehicle grid integration in the U.S., Europe, and China: Challenges and choices for electricity and transportation policy", www.theicct.org/sites/default/files/publications/EVpolicies_final_July11.pdf.

IEA (2014a), *World Energy Investment Outlook*, OECD/IEA, Paris.

IEA (2014b), *World Energy Outlook 2014*, OECD/IEA, Paris.

IEA (2013), "Electricity networks: Infrastructure and operations – too complex for a resource?" OECD/IEA, Paris, www.iea.org/publications/insights/insightpublications/name,43756,en.html.

IEA (2011), *Technology Roadmap: Energy Efficient Buildings: Heating and Cooling Equipment,* OECD/ IEA, Paris, France.

Joskow, P (2008), Incentive Regulation and Its Application to Electricity Networks, http://economics.mit.edu/files/3623.

NREL (2010), "Plug-in electric vehicle infrastructure: A foundation for electrified transportation", www.nrel.gov/docs/fy10osti/47951.pdf.

Massimiano, M. (2015), Autorità per l'Energia Elettrica, il Gas e il Sistema Idrico, personal communication.

Ofgem (Office of Gas and Electricity Markets) (2012), Open letter consultation on the way forward for the next electricity distribution price control review – RIIO-ED1 www.ofgem.gov.uk/ofgem-publications/47182/riioed1launchopenletter.pdf

Ofgem (2010), "RIIO: A new way to regulate networks", Factsheet, www.ofgem.gov.uk/ofgem-publications/64031/re-wiringbritainfs.pdf.

Vogelsang (2002), "A 20-year perspective on incentive regulation for public utilities", http://link.springer.com/article/10.1023%2FA%3A1019992018453#page-1.

Chapter 9 • Retail pricing

HIGHLIGHTS

- Consumers make consumption and investment decisions based on price signals they receive through retail electricity tariffs, and not wholesale electricity prices. In many countries, low-carbon policies are financed by levies, thus increasing the final price of electricity. The energy price component represents only a proportion of total electricity costs – for example, 43% in Europe.
- While competition between retailers has been introduced in many markets, barriers to retail competition still have to be minimised to encourage innovation in service and pricing.
- Meanwhile, distributed resources such as rooftop solar photovoltaics (PV) and storage can be installed behind the meter at a rapid pace and be financed by savings on electricity bills. These self-consumers are still connected to the grid, but may contribute less to network costs. This creates distributive effects and can affect the sustainability of the power system.
- Consequently, retail price reform is urgently needed. Retail prices should give the right incentives to both network users and distributed energy resources, in a timely and location-specific manner.
- In particular, network tariffs need to cover the costs of infrastructure, should send a signal for efficient use of the network, should allocate costs to all users and should be calculated using a simple and transparent methodology.
- Development of real-time pricing reflecting local power production should be encouraged to give the right signals to invest and operate distributed resources.

Electricity retail prices take into account the aggregate of a range of variables: the market prices for producing and delivering energy, the margin for retailers, network tariffs, metering and billing costs and energy taxes. In the context of decarbonisation, they carry the relevant cost signals to the final consumer. Retail prices should ideally guide consumption and investment decisions, thereby contributing to lower expenditure and more efficient investment.

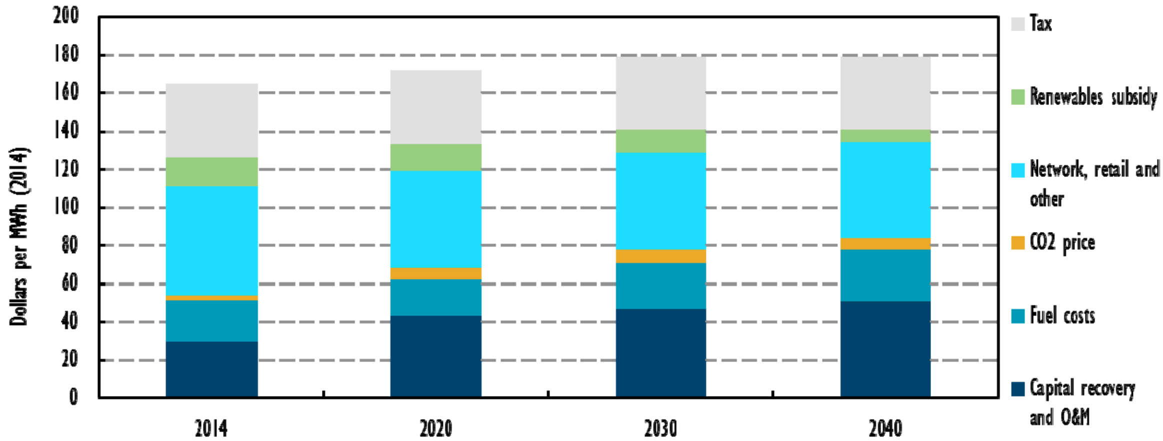
With the declining cost of behind-the-meter distributed generation and demand reponse, the level and structure of retail pricing have to evolve to reflect increasing consumer choice and responsiveness. It is thus important to move away from a paradigm where consumers are treated as inelastic bill payers. Instead, retail pricing should reflect the fact that the value of electricity fluctuates to send efficient signals to generation and storage decisions behind the meter. This chapter starts with a description of the state of play of retail competition and the evolution of retail prices. The second section calls for a reform of retail pricing, including the introduction of dynamic pricing and the reform of energy taxation and network tariffs.

9.1. Retail prices, competition and behind-the-meter generation

Evolution of retail prices

Electricity prices are expected to increase with decarbonisation, both for industrial consumers (Figure 9.1) and households. Even if decarbonisation tends to reduce wholesale prices (Chapter 2), in Europe the electricity taxes and levies that finance low-carbon policies will increase (Figure 9.1). In the United States, however, where there is currently no price on carbon dioxide (CO_2), where the price of natural gas is low and where renewables benefit primarily from tax credits, electricity prices are expected to remain lower than in most European countries.

Figure 9.1 • Average electricity prices in the industry sector by cost component in the European Union in the New Policies Scenario (NPS)



Notes: MWh = megawatt hour.

Wholesale electricity prices in the short run may not fully reflect the underlying costs, which can result in insufficient levels of capital recovery, as is the case in the European Union today.

Source: IEA, 2015a.

There are major disparities in end-user electricity prices. In general, prices can be broken down into four main components: energy, distribution, energy taxes and value added tax (VAT).[1] The relative size of each component varies between countries. In 2012, for example, the energy price component including retail margins represented a mere 43% of the average total bill in Europe, with distribution at 30%, energy taxes at 13% and VAT at 14% (Vaasa ETT, 2013). Under such circumstances, any variation in the wholesale energy price or suppliers' costs has a relatively small influence on the price paid by the customer.

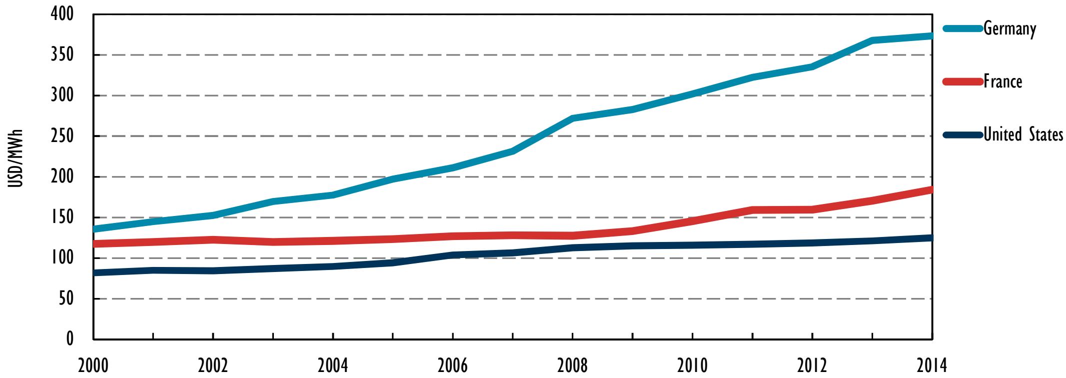
Consequently, retail prices may be significantly affected by the tax system, which varies from country to country as well as by consumption sector. For instance, as shown in Figure 9.2, Germany, with substantial taxes and surcharges, has much higher prices than the United States, where electricity consumption is not subject to tax. Historical trends also show that the incremental rate is higher in the German system.

In the United States, depending on the state, certain renewable portfolio standard (RPS) costs and other renewable energy investments are passed through to consumers in the form of higher electricity rates, although the impact on bills has generally been limited due to federal subsidies and RPS cost-containment measures. Many federal policies and state-based incentives lower the cost of renewable energy without directly passing costs on to bill payers.

In Europe, in line with European Directive 2009/28/EC, many countries have added a surcharge or an energy tax on to retail prices to finance feed-in tariffs and other support schemes, such as renewable obligations.

[1] The United States has no VAT component.

Figure 9.2 • Average household retail tariffs in United States, France and Germany

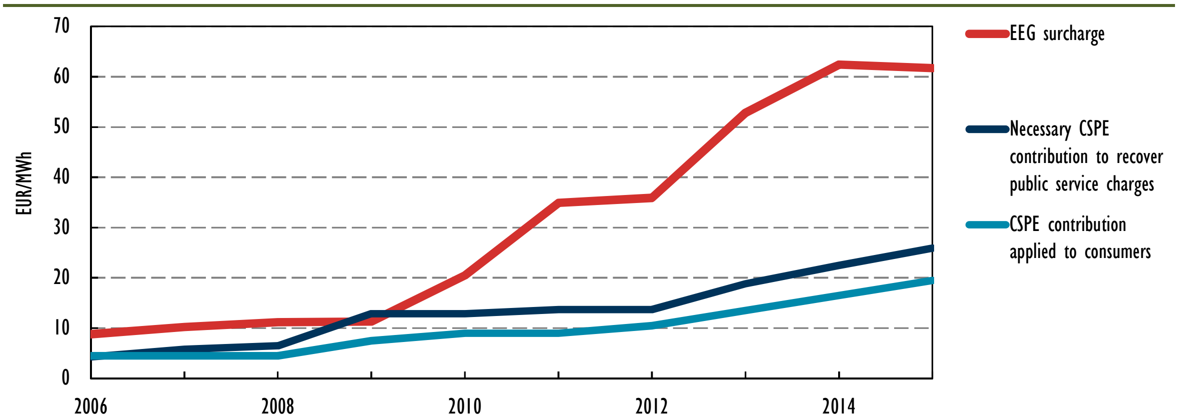


Source: IEA, 2015b.

For example, according to the German Renewable Energy Act (*Erneuerbare-Energien-Gesetz*, EEG) the so-called EEG surcharge recovers the costs of feeding renewable energy into the grid. This act stipulates that German power supply companies have to purchase energy generated by EEG installations. Figure 9.3 shows that the EEG surcharge has been on the rise; however, the large amount of renewable energy flowing into the German market has actually caused electricity costs on the wholesale market to sink. The rise in retail prices could indicate that household consumers are not benefiting from these declining prices.

In France, in order to enable electricity companies to recover costs incurred in performing their mandatory public service duties, a Public Electricity Service Contribution (*Contribution au Service Public de l'Électricité,* CSPE) is imposed on customers' bills. Of the total CSPE, 71% contributes to the financing of renewable energy production and co-generation,[2] 5.5% for social electricity tariffs, while 23.5% goes to the nationwide equalisation of tariffs (U. Lomas, 2013). As can be observed from Figure 9.3, the CSPE is currently limited to EUR 19.5/MWh, which is still less than the total cost of public service obligations of EUR 25.9/MWh.

Figure 9.3 • Evolution of the unitary surcharges in France (CSPE) and Germany (EEG)



Notes: CSPE = *Contribution au Service Public de l'Électricité* (Public Electricity Service Contribution); EEG = *Erneuerbare-Energien-Gesetz* (German Renewable Energy Act).

Sources: CRE, 2015; Bundesnetzagentur, 2010, 2011, 2012, 2013 and 2014.

Given the level of these elements, competition is only relevant for less than 50% of household electricity bills in the European Union. Moreover, the regulated components are expected to increase

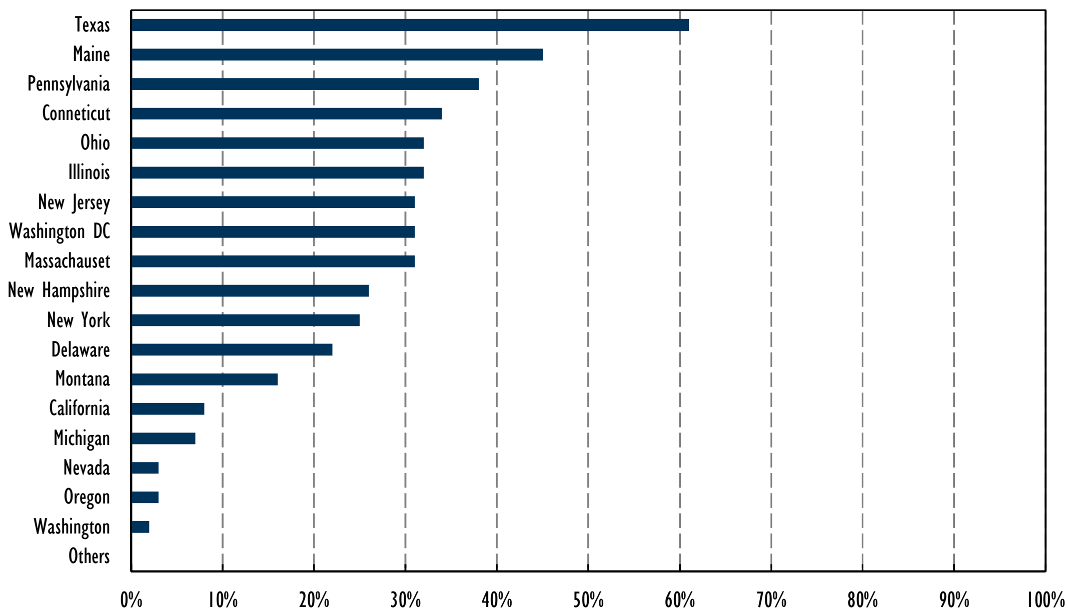
[2] Co-generation refers to the combined production of heat and power.

due to the costs of energy policy and taxes. This has implications both for competition between retailers and for an emerging form of competition – that from behind-the-meter generation.

Retail competition

The extent to which retail competition has been introduced across electricity markets varies greatly. In the United States, for example, no retail competition exists in 29 out of the 50 states (Borenstein and Bushnell, 2015). In the 21 states with retail choice programmes, most activity is clustered in the Northeast, which also has a high proportion of consumption. One notable leader is Texas, which has been one of the markets most consistently ranked for competition in the world, with over 60% of sales from retail power marketers. Figure 9.4 illustrates the fraction of total sales in each state from entities with an ownership classification of retail power marketer.[3]

Figure 9.4 • Share of retail sales from retail power marketers, United States



In Europe, retail competition is mandatory in all member states, according to EU Directive 2003/54/EC. However, the intensity of retail competition varies greatly from one country to another. Despite the efforts of the European Commission and regulators, the market share of new entrants remains relatively limited, reflecting the fact that the electricity product is narrow and homogeneous, with limited possibility for differentiation. Only 43% of the average bill represents energy prices (and retailers' commercial costs) and is therefore open to competition, so from the perspective of the consumer the gains from switching are typically modest.

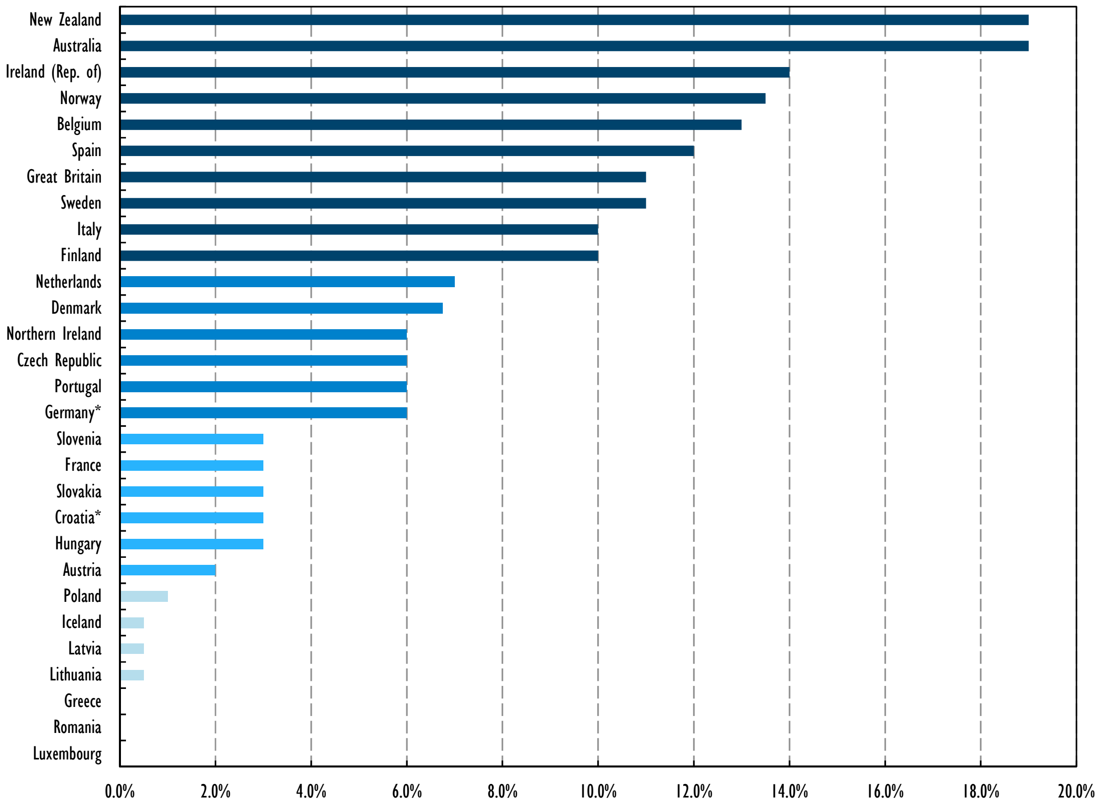
An indicator of the degree of market competition may include the degree of customer switching activity over a given period of time, because a switch normally occurs only when pricing and/or services are attractive enough for consumers to make the change. However, it should be noted that for any market, the ease of switching highly influences the switching rate. Moderate or low switching rates in markets can also reflect successful efforts by retailers in retaining their existing customers. Price-matching strategies, competitive rewards programmes and the flexibility for

[3] A retail power marketer is defined as a business entity engaged in buying, selling and marketing electricity. Power marketers do not usually own generating or transmission facilities. These entities file with the Federal Energy Regulatory Commission (FERC) for status as a power marketer (source: The Retail Energy Supply Association).

customers to negotiate contracts, for instance, are attractive because people generally prefer to avoid the perceived hassle of the switching process.

An example of switching rate tracking is presented in Figure 9.5 (VaasaETT, 2015).[4] With approximately 19% switching rate in 2014, New Zealand has been among the top ranks for many years, while Belgium takes a dominant lead in Europe. In both countries, switching activities have been bolstered by committed and supportive regulators, public awareness campaigns and active marketing activities (Metering International, 2013).

Figure 9.5 • Switching rates in selected countries, 2014



* Countries with 2013 data.
Source: VaasaETT, 2015.

Factors deterring consumer switching may include a lack of awareness of potential savings. This is especially true if the savings associated with switching remain limited to a small portion of the electricity bill. According to research (VaasaETT, 2013), households in the EU-15 countries[5] could have saved 14% on their electricity bills if they had left their standard contract and switched to the cheapest available option in 2012. Meanwhile, the perceived complexity of the switching process may also discourage consumers, even when they are aware of the potential savings that can be made.

[4] A switch is only recorded if a customer switches to a supplier other than the incumbent supplier. A switch additionally includes a re-switch, i.e. when a customer switches for the second or subsequent time, even within the same measured period of time, and a switch-back, i.e. when a customer switches back to his or her former or previous supplier.

[5] Austria, Belgium Denmark, Finland, France, Germany, Great Britain, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Spain and Sweden.

Another striking factor is consumer loyalty to the incumbent, which is especially true when municipal suppliers are among the competing retailers. Consumers tend to stick with what they know or trust, as they may have the impression that other retailers, especially new entrants, may not be able provide adequate technical assistance and service in the case of a disruption.

Phasing out regulated prices

Regulated tariffs exist alongside competitive retail prices in many liberalised power markets. For example, household end-user regulated price existed in 15 out of 28 EU countries as of December 2013 (ACER/CEER, 2013). In some electricity markets, retail competition is only available for high-voltage consumers above a certain capacity, and therefore residential customers are still regulated.

In Japan, for example, extra-high-voltage customers (above 2 megawatts) became eligible to choose their electricity supplier in the year 2000, high-voltage customers above 500 kilowatts became eligible in April 2004, and high-voltage customers above 50 kW became eligible in April 2005. Full liberalisation, to be implemented in 2016 as the second stage of ongoing electricity system reform, will abolish this market entry regulation. Additionally, retail tariff regulation of existing power companies is due to be abolished in stages around 2018-20.

The presence of regulated tariffs distorts retail competition. Very often regulated tariffs are below the cost that historical companies need to bear to offer their services, resulting in a so-called tariff deficit. Such a tariff deficit may originate from a political decision. In the case of Spain, a desire to control inflation has strongly determined the process of fixing regulated tariffs.

Competition can be promoted by removing regulated tariffs or ensuring the tariffs are on a par with those offered by retail companies. In New South Wales, Australia, for example, the government removed retail price regulation on 1 July 2014. On that date, any customers who had not switched over to a market contract were automatically transferred to a transitional tariff. Even prior to this removal, almost 2 million customers (constituting approximately 60% of households and small businesses) had already switched from a regulated electricity contract to a market contract (NSW Gov, 2014).

Retail competition and innovative retail prices

Broadly speaking, the aim of retail competition is to reduce electricity bills, encourage innovative retail pricing and provide choice to consumers through innovative services.

In some countries, real-time pricing based on wholesale market prices is the rule for all consumers. In Spain the government decided that the default tariff should reflect market prices (Box 9.1). This ensures that wholesale market prices are properly passed through to final consumers, either directly or indirectly via retailers' commercial offers.

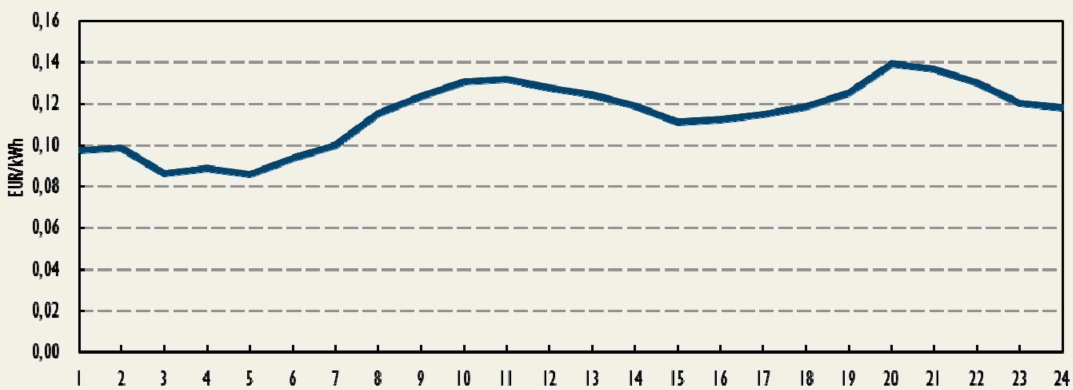
Under real-time variable pricing, the rate changes with market conditions; consequently, consumers who are willing to adjust their use of electricity accordingly may see some savings. However, most consumers are poorly informed as to when electricity prices increase, potentially leading to significant increases in their bills during certain months. For instance, in January 2014 the northeastern United States was significantly affected by the so-called polar vortex, a few weeks of extreme cold. Due to a combination of increased usage and a hike in the generation portion of their tariff, the monthly bills of some residential consumers in Pennsylvania more than tripled relative to their usual amount (McCloskey, 2014).

Consequently, most retailers offer fixed tariffs with rates that are locked in for a certain period (for example, 12 or 24 months) thereby reducing price volatility for consumers. Although fixed prices provide stability, this certainty may come at the expense of higher costs for the customer as compared to paying under variable pricing.

Box 9.1 • The Spanish Electricity Tariff Reform

In April 2014, Spain modified its electricity price formation system for retail customers. The previous system consisted of auctions organised by the Spanish Ministry for Industry. The auction results set the base price, covering most of the day, and a peak price. This price was applied to the majority of retail customers in Spain. The final price often significantly diverged from actual wholesale market prices, with variations ranging from 6% to 20%. The new regime that entered into force in 2014 is a first of its kind in Europe, exposing around 15.7 million customers directly to the electricity wholesale market price and its hourly fluctuations by making this the default pricing option (Figure 9.6).

Figure 9.6 • Voluntary prices for small consumers in Spain as of 5 February 2015



Note: kWh = kilowatt hour; "small consumers" means those with a power supply of less than 10 kW.

Source: Red Electrica de Espana, 2015.

Any consumer can nevertheless opt out of the system and subscribe to any supplier or contract structure.

This new tariff system implies that the consumer is exposed to price volatility during the course of the day, and therefore that he or she has to adapt their consumption pattern accordingly. Smart metering and automated solutions should help to facilitate such behaviour. Spain is currently progressing well in deploying its smart grids and metering devices. As of 1 October 2015, customers with smart meters are billed according to real hourly consumption, while customers without smart meters are billed according to standard consumption profiles defined by the transmission system operator, Red Eléctrica de España (REE). The full deployment of smart meters is expected to be completed by the end of 2018.

Spot price offers are also available in some markets, for instance in the Nordic countries. In these countries, a customer with a spot-price-tied contract will typically pay the average monthly spot price on the NordPool power exchange plus a mark-up (VaasaETT, 2013). For customers with this kind of contract, the mark-up difference between suppliers is the potential gain from switching.

Retailers may also offer green products to consumers willing to pay a premium. Implementation varies from one country to another.[6] The ability to guarantee green products is highly dependent on the presence of a transparent and credible system.

[6] For instance, the green products offered to German consumers are mainly Norwegian hydropower, and to a smaller extent, Austrian and Swiss hydropower (Hast et al., 2014). The reason for this is that the German support scheme for renewable energy, the EEG, forbids electricity remunerated under the EEG to be marketed as green power.

In the past, retail competition has rarely led to innovative tariff structures, a disappointment to those who have advocated retail pricing on those grounds. One possible reason lies in the perceived complexity of retail tariffs, in particular when it comes to dynamic or time-of-use pricing. In practice, the historical rate structure and time differentiation continue to shape how competitors price electricity. For instance, in France or the United Kingdom, most suppliers continue to propose a peak/off-peak tariff or the Economy 7 tariff, with rebates either on the fixed fee or the variable fee. To date, differentiation of retail prices by means of sophisticated dynamic pricing structures has failed to develop.

The development of innovative pricing offers should be encouraged. In particular, dynamic pricing could better reflect the progress of scarcity pricing rules (see Chapter 3), the increased volatility of wholesale prices and the periods of low electricity prices during hours of over-generation. Such changes are needed to send efficient signals to consumers, who are increasingly able to react to prices in real time, as well as invest in rooftop solar PV and storage devices.

A new form of competition: Behind-the-meter generation

Investment signals for behind-the-meter generation largely derive from retail electricity tariffs, both in terms of level and structure.

For household self-consumers, the break-even point for solar PV has typically been considered to be when the cost to the consumers reached socket parity – that is, the point at which the levelised cost of electricity (LCOE) falls to, or below, the per kilowatt hour cost of electricity obtained from the grid, i.e. the variable part of a consumer's electricity bill (IEA, 2014). More and more households may decide to invest in rooftop solar PV on the basis of the savings on their electricity bills, particularly in markets with purely volumetric retail tariffs and high taxes, and where net-metering is allowed.

However, while this approach reflects how consumers assess the value in investing in rooftop solar PV, it has shortcomings and does not actually reflect the competitiveness of this technology. From a system perspective, households with a PV system reduce their contribution to the system cost, shifting the burden to households without PV systems (IEA, 2013). Despite the falling cost to the individual consumer, solar PV actually reduces total system costs only in a few locations.

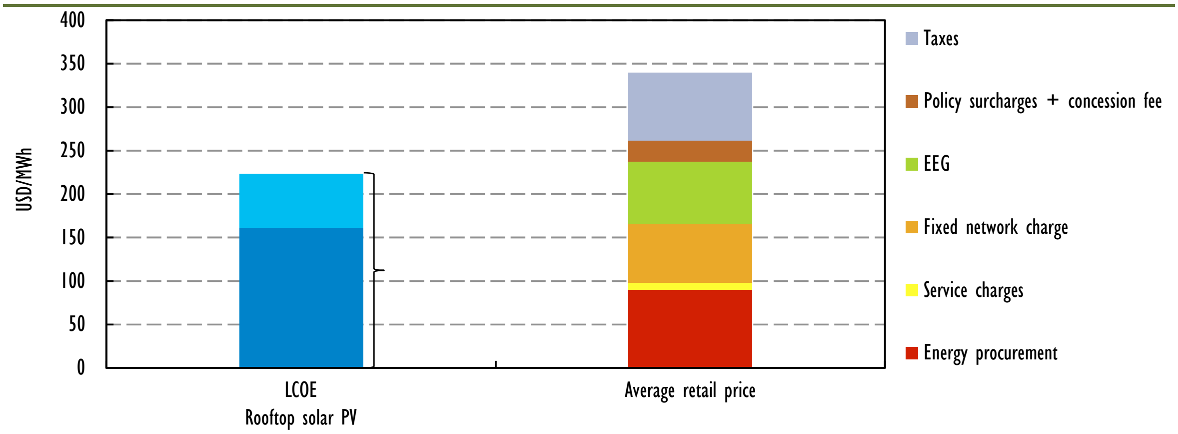
Socket parity can only be considered an accurate indication of the competitiveness of distributed technologies if retail electricity tariffs accurately reflect the cost of electricity over time and by location (IEA, 2014). However, in practice, per-kWh electricity tariffs do not reflect costs with sufficient temporal and spatial resolution, or the high share of fixed costs.

Historically, electricity rates have been designed on the basis of several considerations, including the notion that electricity is a public service and that access to it is universally needed. This has led to uniform prices across geography in some countries (for example in France, there is a *péréquation tarifaire* or tariff equalisation) and consumer categories that poorly reflect the costs of serving each user. In addition, electricity consumption has historically been inelastic to prices. With the rapid deployment of distributed resources, this is becoming less and less the case.

More importantly, since retail tariffs usually include numerous cost components, any investment signals from the wholesale market may be highly diluted and therefore provide inaccurate incentives for investment in distributed resources. As such, reaching socket parity is a poor indication of solar PV costs falling below their value to the system. In many cases the avoided costs for the system are much lower than the savings consumers make on their bill.

The discussion about retail pricing is further complicated by the fact that some regulations, such as net metering, constitute implicit support for consumers installing distributed resources. In certain jurisdictions, consumers can reduce their billed electricity consumption even when they are not consuming electricity but their rooftop solar PV system is generating (in practice, some meters turn in the other direction when they inject into the grid and this reduces their bill). Net metering is an implicit feed-in tariff set at the level of the variable part of the retail price, which is not related to the cost of solar PV (Figure 9.7).

Figure 9.7 • Comparison of LCOE for PV and the average retail price in Germany



Sources: LCOE: IEA/NEA, 2015; Average retail price: Monitoring report, 2014; Bundesnetzagentur für Elektrizität, Gas, Telekommunikation, Post und Eisenbahnen.

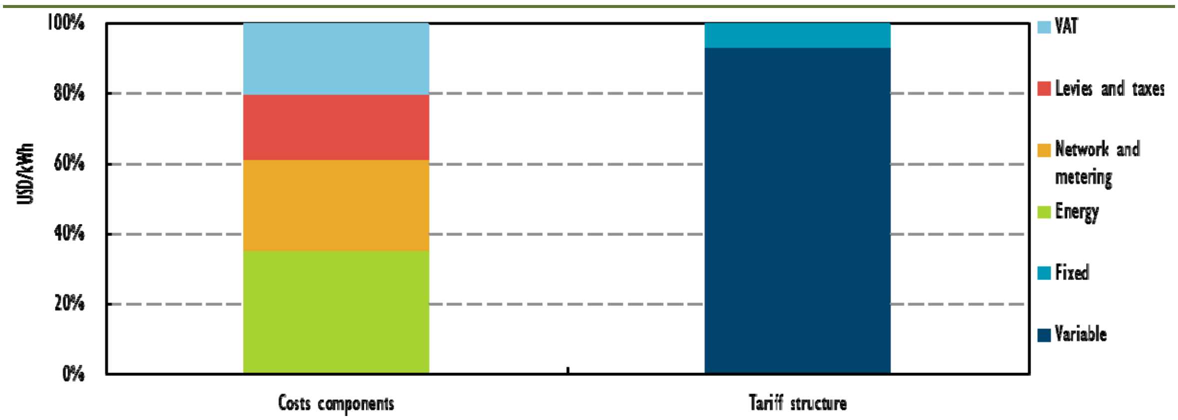
Indeed, despite a rise in the installation of distributed generation, the grid is still needed. Consumers benefit from the reliability provided by the grid at times when solar PV is not generating. Very few consumers are really ready to go entirely off-grid.

Reform of retail pricing structures is therefore necessary, including the way that the energy component is priced and a redesign of network tariffs to provide proper cost recovery and correct signals to consumers.

9.2. Retail price reform

In theory, retail pricing systems have to recover the sum of network costs and energy costs, including the supplier margin and energy taxes. At its broadest level, the design of retail prices consists of allocating these costs to different tariff components (Figure 9.8).

Figure 9.8 • Cost components and tariff structure of selected retail electricity prices (average for Paris, Berlin and Amsterdam)



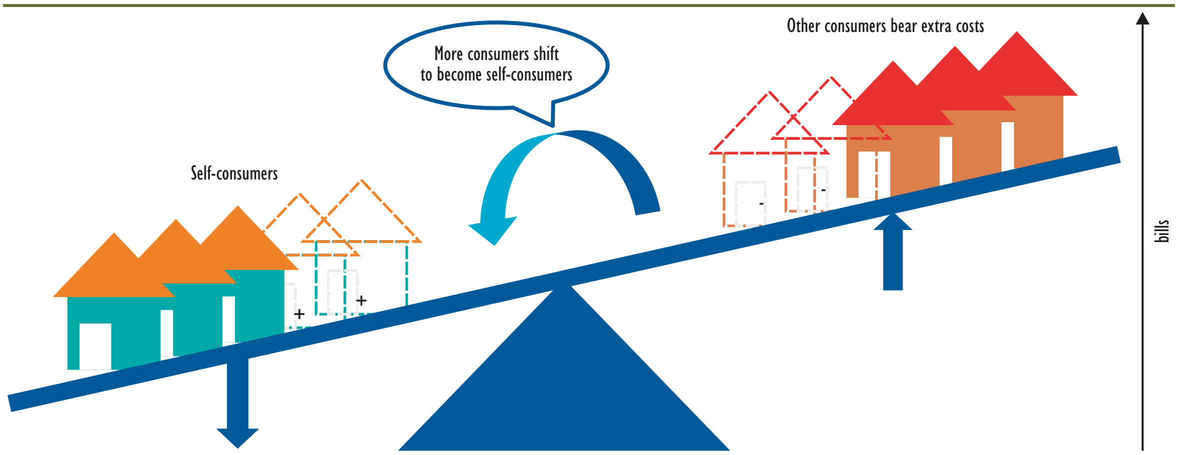
With vertically integrated regulated monopolies, retail prices tended to be relatively simple, with some markets implementing entirely volumetric charges. Other regulated markets, however, have had sophisticated retail pricing systems. In France for instance, EdF (Électricité de France) tariffs comprise various options that reflect the load factor, time of use and critical peak pricing. Integrated utilities could allocate costs to different consumer groups with a high degree of flexibility, often leading to cross-subsidies.

Electricity restructuring has introduced new constraints on the design of retail electricity pricing. Tariffs must be calculated as the stack of network prices plus energy prices, and are based on wholesale hourly electricity prices plus energy taxes, which are usually taxes in EUR/MWh falling automatically on the energy component of the retail price. Where retail competition exists, the structure of retail prices is largely the result of the underlying components.

Network tariff structure taking distributed generation into consideration

Network tariff redesign becomes essential in the context of the declining cost of distributed resources. Indeed, even though distributed generation reduces the energy withdrawn from the network, the consumer still needs the network most of the time. Unless consumers are prepared to go off-grid, distribution lines and transformers must be maintained. While some of these costs are driven by the amount of energy consumed, a large proportion is either fixed or dependent on the coincident peak demand of consumers.

Figure 9.9 • A slippery slope towards self-consumption



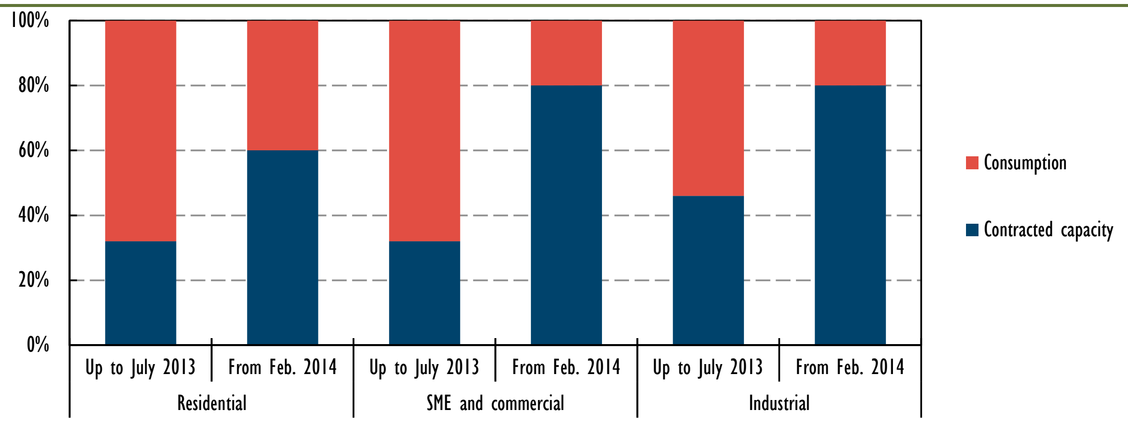
Existing methods of cost allocation tend to lead to a bias in favour of consumers that are equipped with distributed energy resources, such as solar PV or batteries. Consequently, the poor design of retail prices creates a problem of sustainability for the power industry as a whole. As consumers are incentivised to reduce their consumption from the grid and replace it with behind-the-meter generation, they avoid paying electricity taxes and contribute less towards the fixed and common costs. Confronted with a shrinking billing base, utilities have to increase their rates, which further increases the cross-over point where self-consumption is cheaper than paying the utility for electricity. This further motivates consumers to install behind-the-meter generation. This is illustrated in Figure 9.8. A slippery slope is created for customers to move to distributed energy, to avoid increasing network costs, and fewer and fewer consumers pay towards the transmission and distribution networks.

Furthermore, by concentrating fixed cost recovery on fewer households who cannot afford the installation of solar PV systems or who do not own their homes, this also creates distributive effects and a redistribution of rents between consumers (Borenstein and Bushnell, 2015). This situation is not sustainable.

Consequently, network tariffs should be rebalanced away from a volumetric charging basis towards a capacity basis in order to better reflect the cost structure of the network infrastructure. Several utilities and regulators, including in California in the United States and Spain in Europe, have already taken steps to increase the fixed component of retail electricity tariffs. Spain has already implemented this change, where the share of network costs recouped from the capacity component increased on average from 34% in 2011 to 68% in 2014 (Figure 9.10). While regulators are usually slow to implement such changes, Spain's example proves that the tariff structure can be changed rapidly.

Not surprisingly, such a move has been met with mixed reactions, as some stakeholders view it as unduly hindering the uptake of distributed energy resources. A capacity charge reduces the volumetric component of the retail tariff and therefore the savings that can be made with behind-the-meter generation. Of course, all consumers will be affected by such charges. They have sometimes been perceived as retroactive changes that affect the return on investment of small rooftop solar PV systems. In Spain and parts of the United States, certain solar PV associations and interest groups have described proposed measures as "sun taxes".

Figure 9.10 • Evolution of the Spanish network access tariff structure



Note: SME = small and medium-sized enterprise.

Source: Iberdrola, 2015.

Rebalancing network tariffs towards the fixed component is not intended to, and does not necessarily, hinder the development of behind-the-meter resources. Regulators and policy makers should seek to facilitate the development of behind-the-meter generation when it is efficient to do so. From that perspective, cost-benefit analysis (CBA) might be useful to assess conditions under which behind-the-meter generation is desirable. The answer to this question is likely to involve setting retail tariffs that reflect the long-run marginal costs of networks and generation.

In undertaking such CBA and defining the rules for the development of distributed generation, regulators should be aware that distribution companies can use their market power and information to constrain the deployment of distributed generation that challenges their business model, rather than seeking to identify an efficient framework that accommodates distributed resources.

The right regulatory framework for the efficient development of behind-the-meter generation remains largely an open question. If solar PV is indeed less expensive than the long-term cost of replacing or developing the centralised system, then the most efficient option may indeed be to deploy more solar PV. As consumers make their decisions based on retail electricity prices using billing information at the meter, electricity pricing is the key to ensuring efficient decentralised customer decisions.

Towards a network tariff structure 2.0

To ensure a sustainable regulatory framework and to give the right incentives to both network users and distributed energy resources, network tariffs should be designed according to the following principles, which supplement the usual non-discriminatory paradigm:

1) Tariffs need to cover the total costs of necessary infrastructure.

2) Tariffs should send a signal for efficient use of the network.

3) Tariffs should be cost reflective and allocate fixed and common costs to all consumers, including self-generators, instead of increasing the bills of only those consumers not able to reduce their grid consumption.

4) Tariffs should be calculated using a simple, transparent methodology.

A network tariff fulfilling the first three principles above is likely to entail the use of smart metering in all households, and should be based on a detailed calculation of operational and future investment needs using sophisticated network modelling and network management. Meeting the fourth principle (using a simple and transparent methodology) implies that tariff structures should be simplified. Some examples of possible and existing tariffs are depicted in Table 9.2.

Table 9.2• Overview of network tariffs structure

Types of network tariff		
Volumetric tariffs	Fixed tariff	Price per unit of energy (kWh)
	Time-of-use tariff	Price dependent on the time of consumption or feed-in
Capacity tariffs	Tariff varies depending on level of capacity	Definition of different capacity levels: one price per quantity of capacity
	Time-of-use tariff	Price of kWh depends on time of consumption
Multi-part tariffs	Combination of fixed, volumetric and capacity tariffs	
Other	Interruptible tariff	Reduction in network tariffs for the permission to control a certain amount of the load

As discussed above, purely volumetric tariffs combined with simple net-metering tends to lead to a bias in favour of distributed resources. Under such a tariff structure, the cost savings for customers using distributed resources may exceed the savings for the system.

In contrast, pure capacity tariffs (where the consumer pays a fixed charge for a set amount of usage) are a simple instrument widely adopted in other areas, for example telecommunications. Consumers easily understand capacity prices, but these prices do not reflect the variable components of the cost and therefore do not promote consumer engagement and can lead to higher consumption. In addition, pure capacity tariffs could excessively reward consumers who install batteries in order to reduce their subscribed capacity.

Two-part tariffs that consist of a fixed charge plus volumetric charges offer a better solution. For most distribution systems, the objective should be to define the two-part tariffs or multipart tariffs.

Network tariffs could also possibly be differentiated according to time and locational dimensions in order to reflect the actual costs of the network. Indeed, more sophisticated time-sensitive rates could even be envisaged with investment in advanced metering. In theory, algorithms could calculate the cost of reliability and overall system efficiency for all the products being transacted and for each consumer.

There are limits, however, to the degree of complexity that network tariffs can reasonably reach. Increased complexity leads to higher transaction costs, as tariffs become harder to understand, potentially raising acceptance issues for access to a public service infrastructure. It is clear that the design of a network tariff fulfilling the principles of efficient use of the network, full cost coverage, cost allocation and transparency, will require further assessment.

First connection to the grid

Connection-charging methodologies differ widely between member countries of the Organisation for Economic Co-operation and Development. In some markets, so-called “shallow charges” only cover the direct costs of connection to the nearest point of the distribution grid. In that case new customers are implicitly subsidised if the overall cost to the whole system of the new connection exceeds the connection charge. Such costs, however, are difficult to allocate.

In other markets, connection charges include the deep costs corresponding to the implications for the upstream grid infrastructure and its possible reinforcement to support the new grid user. Such deep connection charges provide a locational signal to new network users, but may cause issues with grid users entering the system at a later stage, possibly free-riding on the infrastructure paid by previous entrants. Some countries, including the Netherlands, use a mix of both charges. Units above a certain threshold pay deep charges, whereas smaller units only pay connection charges.

Dynamic/real-time pricing

The tariff system for electricity can have a direct impact on consumer consumption patterns, as explained in Chapter 6. In fact, real-time pricing is perhaps even more relevant for those who can install solar PV or storage behind the meter (prosumers). Real-time pricing can tell such consumers the market value of the electricity they do not consume from the grid or that they export to it (see chapter 2).

Traditional time-of-use (ToU) rates have historically comprised separate day and night tariffs, with a fixed number of hours each day throughout the year. Such peak and off-peak prices do not reflect variations in wind and solar generation.

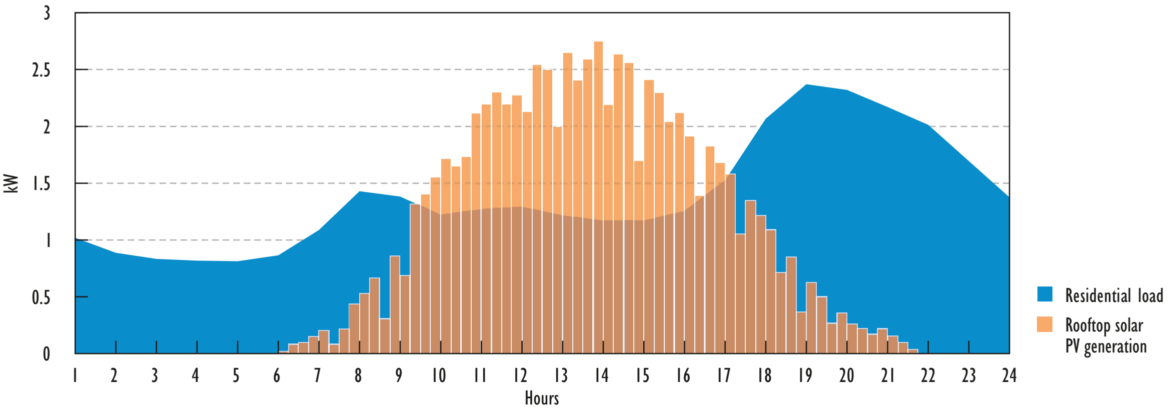
Real-time pricing represents a useful tool to factor in the value of renewable energy sources. It can accurately reflect the real-time variations in the cost of electricity generation, providing the consumer with the incentive to install storage and generation that can react to and arbitrage such prices. For example, if electricity prices are very low during the summer daytime because, for example, large amounts of solar PV energy are entering the system, real-time retail prices would also have to be low during these hours. A consumer would not be able to reduce his or her bill by adding an additional solar PV panel, but might instead – assuming a meaningful reduction from current cost levels – install a battery.

In the absence of such real-time price information, generation investment decisions are distorted because they are based on average energy market prices and average network costs and taxes, and are likely to lead to over- or in some cases under-investment.

As already discussed previously, the example of the tariff reform introduced in Spain is particularly relevant (see Box9.1). The government decided that real-time pricing will be the default option to bill consumers (and retailers). Consumers have the possibility of opting out of the real-time tariff by choosing a competitive offer that is simplified.

Smart meters are currently being rolled out in many countries and can accelerate the deployment of real-time pricing of electricity. Hourly profiles of household load and solar PV generation can be used to efficiently price behind-the-meter generation from rooftop solar PV (Figure 9.11). Real-time pricing also helps to increase levels of awareness of consumption patterns among consumers.

Figure 9.11 • An example of hourly profile of household load and solar PV generation (illustrative)



Notes: this example depicts a typical household with a 5 kWp solar PV installation; kWp = kilowatt peak.

However, directly exposing consumers to the variations of wholesale prices with real-time pricing, even if they are prosumers with solar PV and storage, significantly increases complexity. Consumers are not typically used to this price information, and usually do not have time to monitor their meter and the hourly price of electricity. In practice, even small generators up to a couple of megawatts prefer having a fixed price arrangement, in order to reduce the complexity of financial calculations and be able to quantify the profitability of their investments.

Therefore, the power supplier should be encouraged to develop "dynamic time-of-use" tariff plans, with different tariff options reflecting the production path of the local power generation mix. Intelligent devices, such as smart grids and automation systems, would also come into play in facilitating the engagement of the customer and increasing this kind of simplified exposure to pricing signals.

Future deployment of behind-the-meter generation could be subject to the following options:

- Locational wholesale price pass-through. Exposure to wholesale prices (for instance via real-time pricing) could provide incentives for customers to install behind-the-meter generation when it is efficient to do so.
- Dynamic time-of-use pricing. Suppliers offer simplified rate structures, such as dynamic pricing, which reflect the cost structure and price variations. Consumers can decide whether to invest in behind-the-meter generation based on these simplified tariff structures.
- Service-based options. Energy service companies can also play a key role in the connection between complex wholesale electricity markets and consumers. Suppliers and energy service companies can offer products and services that meet consumers' need for simplicity.

For distributed generation and storage, the menu of electricity prices offered at the meter will therefore remain a central piece of the future regulatory framework.

Electricity taxation

The level of retail electricity prices is to a great extent dependent on the tax system. The share of taxes and levies (including, where relevant, VAT) within the total price of electricity varies greatly from country to country, even within the European Union. For example, for a standard medium-sized household with annual electricity consumption between 2 500 and 5 000 kWh, VAT, taxes and levies represent 56.8 % of the final price in Denmark, 51.6% in Germany, and 41.7% in Portugal (Eurostat, 2015). Only two EU member states – the United Kingdom and Malta – have single digit tax rates in the final electricity price.

Discussions continue on the long-term direction of electricity taxation and the method to determine the appropriate taxation policy. In Europe, the choice has been made to expose consumers to the cost of electricity, including renewable deployment policies. In Germany, the allocation basis for the Renewable Energy Levy has already been extended to cover the consumption of self-generated electricity.[7] Other taxation policies being discussed include whether charges for environmental purposes (climate change mitigation, research and development, and deployment support) should be financed either by levies or from general taxation (Newbery, 2015).

More generally, high levels of taxes and levies tend to make electricity more expensive and reduce electricity consumption, which is beneficial from the perspective of energy efficiency. However, high taxes also discourage the use of electricity for transport and the electrification of heating, both key components of energy-sector decarbonisation.

Conclusion

The need for retail price reform is urgent, especially in countries where behind-the-meter generation is developing quickly, in response to increasing retail prices and the declining cost of rooftop solar PV, and to a lesser extent storage. Indeed, prices to final consumers are driving consumption and investment decisions, and the aim of reforms is to mitigate potential free-riding on the part of prosumers, who under inefficient tariff structures pay less towards the network fixed costs and reduce their contribution to renewable policy costs at the expense of other consumers.

Tariff structure, taxation and the lack of time-varying options are the main inefficiencies in existing retail price options. Network tariffs have to be rebalanced from energy charges towards fixed and capacity components. Retail prices have to reflect as far as possible this cost structure and wholesale prices, so as to incentivise consumers to participate more actively in markets.

From that perspective, service innovation and pricing options should be encouraged by phasing out regulated prices in stages once sufficient competition is reached. Retailers should be encouraged to reflect both real-time market pricing and local power production conditions in their tariffs, to convey the actual value of power production and system conditions to consumers.

References

ACER/CEER (Agency for Cooperation of Energy Regulators/Council of European Energy Regulators) (2013), “Annual Report on the Results of Monitoring the Internal Electricity and Natural Gas Markets in 2013”, ACER/CEER.

BnetzA (2013), Genehmiung des Szenarioahmens fuer die Netzentwicklungsplanung und Offshore-Netzentwicklungsplanung [Approval of the scenario framework for network development planning and offshore grid development planning], Study on behalf of the Federal Ministry of Economy and Energy (BMWi), www.netzausbau.de/SharedDocs/Downloads/DE/Charlie/Szenariorahmen/Szenariorahmen2024_Genehmigung.pdf?__blob=publicationFile.

Borenstein, S. and B. Bushnell (2015), *The U.S. Electricity Industry after 20 Years of Restructuring*, Energy Institute at HaasWorking Paper 252R, https://ei.haas.berkeley.edu/research/papers/WP252.pdf.

[7] Applicable to certain PV systems installed after 31 September 2014, whereby systems that have a capacity below 10 kW and have self-consumption of less than 10 MWh/year are excluded. Islanding systems and systems that are 100% self-sufficient and do not ask for financial compensation for electricity fed into the grid are also excluded. Furthermore, the levy is never applied to the full extent on self-consumption. Currently it is at 30%, in 2016 it will be at 35% and from 2017 onwards 40% of the Renewable Energy Levy (Bundesministerium der Justiz und für Verbraucherschutz, 2015).

Bundesministerium der Justiz und für Verbraucherschutz (2015), *Gesetz für den Ausbau erneuerbarer Energien* (Erneuerbare-Energien-Gesetz - EEG 2014) § 61 EEG-Umlage für Letztverbraucher und Eigenversorger, www.gesetze-im-internet.de/eeg_2014/__61.html.

Bundesnetzagentur (2010, 2011, 2012, 2013 and 2014), *Monitoring report (2010, 2011, 2012, 2013 and 2014)*, Bundesnetzagentur für Elektrizität, Gas, Telekommunikation, Post und Eisenbahnen.

CRE (2015), « Commission de régulation de l'énergie » [Energy Regulatory Commission], website, www.cre.fr/operateurs/service-public-de-l-electricite-cspe/montant

Eurostat (2015), Energy price statistics, http://ec.europa.eu/eurostat/statistics-explained/index.php/Energy_price_statistics.

Frontier Economics, Hast, A. et al. (2015), "Review of green electricity products in the United Kingdom, Germany and Finland", *Renewable and Sustainable Energy Reviews*, Vol. 42, February, pp. 1370–1384.

Iberdrola (2015), "Electricity tariff structure – Spanish experience, network investment and regulation", 14 January 2015, presentation by Francisco Laverón, Direction of Energy Policies, Iberdrola, IEA Electricity Security Advisory Panel (ESAP) Expert Workshop IV, Network Investment and Regulation, Wednesday, 14 January 2015.

IEA (International Energy Agency) (2015a), *World Energy Outlook 2015*, OECD/IEA, Paris.

IEA (2015b), *Energy Prices and Taxes 2015,* [3rd Quarter 2015], www.iea.org/statistics/.

IEA (2014), *World Energy Outlook 2014*, OECD/IEA, Paris.

IEA (2013), *Electricity Networks: Infrastructure and Operations – Too Complex for a Resource?* www.iea.org/publications/insights/insightpublications/name,43756,en.html.

IEA/NEA (2015) *Projected Costs of Generating Electricity* 2015, OECD Publishing, Paris.

McCloskey, T.J. (2014), "Retail choice in Pennsylvania and the impacts of the polar vortex", presentation for NARUC Subcommittee on Accounting and Finance, 18 September 2014.

Metering International (2013), "New Zealand now the most active energy market globally", www.metering.com/new-zealand-now-the-most-active-energy-market-globally/.

Newbery D, (2015), Reforming UK energy policy to live within its means, EPRG working paper 1516, http://www.eprg.group.cam.ac.uk/wp-content/uploads/2015/09/1516-PDF.pdf.

New South Wales Government (2014), "Removal of electricity price regulation (deregulation): More competition for NSW electricity consumers", Department of Industry Resources and Energy, NSW Government, www.resourcesandenergy.nsw.gov.au/energy-consumers/energy-sources/electricity/removal-of-electricity-price-regulation.

Red Electrica de Espana (2015), e.sios website, www.esios.ree.es, accessed 16 November 2015.

VaasaETT (2015), *World Energy Retail Market Rankings 2014*, VaasaETT Utility Customer Switching Research Project, VaasaETT Oy Ab Ltd, www.utilitycustomerswitching.eu/.

VaasaETT (2013), European Residential Energy Price Report, VaasaETT Oy Ab Ltd. www.vaasaett.com/wp-content/uploads/2013/05/European-Residential-Energy-Price-Report-2013_Final.pdf.

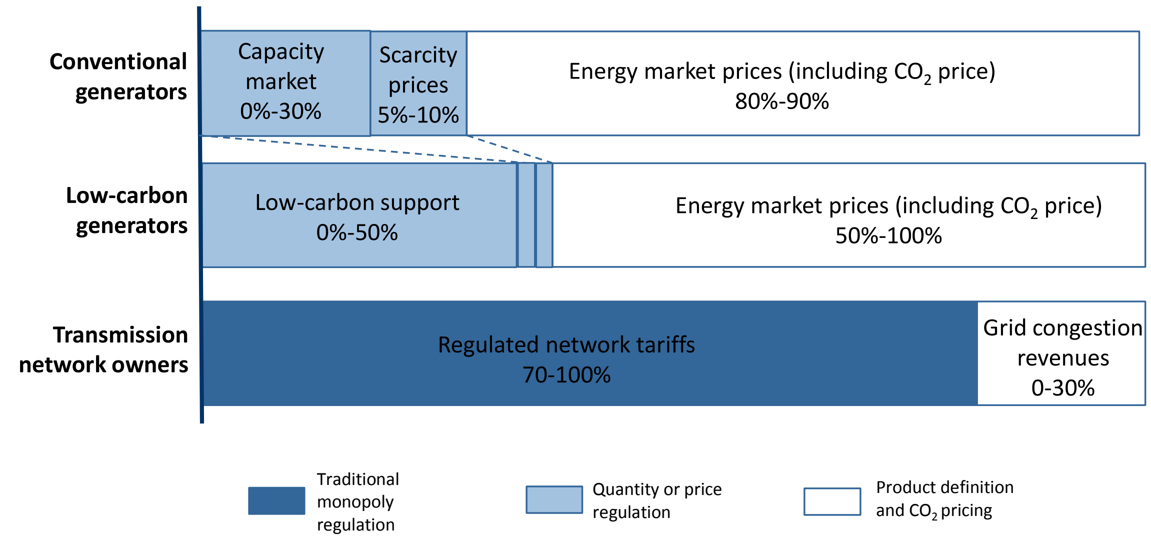
Chapter 10 • General conclusion and key recommendations

Re-powering markets implies a shift in the way most governments and regulators comprehend electricity markets. The traditional binary distinction between competitive generation and natural monopoly networks is inadequate for describing market frameworks. The transition to low-carbon power systems requires the mainstreaming and integration of carbon policies and renewable support policies into a consistent electricity market framework.

Competitive electricity market arrangements should be further developed wherever possible across all market segments. Wholesale energy market prices can be the principal source of revenues and information for many power sector decisions (Figure 10.1), providing incentives to generate electricity or be available for reliability, and revealing the value of the different resources participating in electricity markets – power plants, storage and demand response. Even the value of transmission can, to a certain extent, be revealed by the level of congestion revenues derived from market prices.

Yet, electricity markets have to be supplemented by regulatory interventions in order to ensure an effective transition at least cost. To differing extents, all segments of the electricity system combine elements of regulation and market arrangements. For example, regulators still need to define market rules and mitigate market power. Depending on how reliability is regulated, capacity markets might need to be implemented. Low-carbon support is needed to meet renewables and CO_2 emission targets, even after fully integrating them into markets (Figure 10.1), so as to mitigate carbon price risk for investor and more general market price risk.

Figure 10.1 • Potential for market revenues for conventional, low-carbon plants and transmission infrastructure, indicative values suggested by different sources (%)



Taking this into consideration, the transition to low-carbon power can be carried out through upgrades to existing market arrangements and regulatory instruments. The necessary upgrades can be identified in the best practices of existing electricity markets in Europe, in the Australian National Electricity Market, and in North America.

In the longer run, the design of markets will be shaped by technologies such as storage, demand response and consumers installing distributed resources. But this is not yet the case and, for the time being, market design requires no shift in paradigm. Keeping this in mind, the transition phase is likely to be an evolutionary process based on the interactions between technologies and market rules.

Towards a comprehensive overview of the market framework

"Re-powering" means retaining the overall market architecture, modernising the market rules, and mainstreaming low-carbon and renewable generation into the market. During the low-carbon transition, electricity markets require a set of policies that define the relationship between regulation and competitive markets (Table 10.1).

Table 10.1 • Overview of market framework for the decarbonisation

Objective	*Policy*	Type of regulation	Competitive markets
Low-carbon investments	*Carbon pricing*	○ Carbon regulation	○ Carbon price *(trading scheme)* ○ Long-term contracts
	Additional policy: Support schemes	○ Low-C long-term support	○ Auctions set support level ○ *Integration in markets*
Operational efficiency / Reliability and adequacy	*Short-term energy markets*	○ Market rules ○ Scarcity pricing ○ Reliability standards	○ Energy prices with a high *geographical resolution* ○ Energy prices with a high *temporal resolution* ○ *Dynamic pricing offers*
	Additional policy: Capacity markets	○ Capacity requirements ○ Demand response product definition	○ Capacity prices ○ Demand response participation
Network efficiency	*Regulation*	○ Regional planning ○ Network cost allocation	○ Congestion revenues ○ Transmission auctions
Consumption	*Retail pricing*	○ Network tariff structure ○ Taxation and levies	○ Retail competitive prices ○ Distributed resources

Where implemented, the market framework summarised in Table 10.1 would work as described in the following sections.:

Low-carbon investment (Chapter 2)

Carbon pricing remains a primary instrument for increasing the competitiveness of low-carbon investment. Government regulation consists of introducing a carbon tax or creating a market for CO_2 emissions. Competitive carbon trading markets set the price of CO_2 emissions. Once a carbon price is credible and efficient, market participants can enter into long-term contracts, for example private power purchasing agreements for new low-carbon investments (such as wind, solar power and nuclear). Such low-carbon investment is, however, hindered, by the long-term uncertainty associated with electricity prices, in particular due to the accelerated

Consequently, additional policies are introduced to support renewables, such as tax credits or support schemes. These support schemes provide predictability in the long term and mitigate market risks and CO_2 price risk. Auctions can be organised for larger projects to create competition for the market and set the support level needed.

Long-term support schemes are integrated into markets and supplement market revenues, rather than replace them. Low-carbon generators earn a significant proportion of their revenues from markets. These revenues provide a highly important market feedback loop on the relative value of different low-carbon technologies, in particular variable wind and solar. The level of support is modulated according to the evolution of CO_2 prices and, possibly, market prices.

Operational efficiency, reliability and adequacy (Chapters 3 to 6)

As most new low-carbon investments are in the form of wind and solar power, their integration into the grid has to be efficient while maintaining reliability and adequacy.

Short-term energy markets introduced in most countries in the 1990s continue to represent the foundation of electricity markets. These markets do not design themselves: regulators set market rules by defining the products exchanged. In order to efficiently integrate high shares of wind and solar power, the products exchanged on short-term markets are defined with a high geographical and temporal resolution. Competition between resources in energy markets leads to energy prices that reflect the marginal costs at each location with short scheduling intervals.

Maintaining reliability is another important role for short-term energy markets. Prices during capacity shortage events reflect this scarcity and are regulated according to different dimensions. Price caps are set using a value of loss load of around USD 10 000 per megawatt hour or more, and control of market power is defined *ex ante*. Regulators also define short-term prices when operating reserves are depleted (for example, the operating reserve price curve applied in the Electric Reliability Council of Texas area) in order to make sure that scarcity is priced into short-term markets.

In addition to prices, well-functioning short-term markets require an explicit regulation of reliability standards, preferably using probabilistic criteria such as the loss of load expectation (LOLE) of expected unserved energy (EUE). These standards define the probability of having to curtail load when the market does not clear, and is the basis for setting the regulation of scarcity prices.

The reflection of scarcity prices in dynamic pricing offers on the final market (retail market) is to be promoted. Retailers develop real-time pricing options that ensure the participation of demand in wholesale electricity markets, contributing to increased operation efficiency and better-developed demand response.

Nonetheless, capacity mechanisms are generally introduced, even with efficient scarcity pricing, in order to meet reliability standard at all times, in particular during an investment phase. From a regulatory perspective, capacity markets are a regulation of the quantity of capacity required (unlike regulation of scarcity pricing, which is a form of price regulation). Competitive capacity markets determine capacity prices and the level of participation of different resources.

Demand response participates in capacity markets. Product definition by regulators (i.e. setting the baseline used to quantify the demand response and remuneration rules) influences the level of demand response that clears on capacity markets.

Networks efficiency (Chapters 7 and 8)

In addition to short-term markets, the transition to low-carbon power with high shares of variable renewable energy requires efficient network development, both at the transmission and distribution levels.

Transmission expansion is planned in a way that is integrated with renewable deployment, in order to minimise the overall cost. Regulators develop regional planning across borders. Cost-benefit analysis (CBA) helps regulators allocate the cost of new transmission among different stakeholders. Efficient transmission investment results in optimal levels of congestion. Merchant transmission investments financed by revenues stemming from wholesale electricity price differences at two locations in the electricity system are possible, but rare.

Where a CBA concludes that new transmission investment is warranted, it is usually possible to rely on competitive procedures, such as transmission auctions, to determine who will build, own and manage the new asset.

The regulation of the distribution networks takes advantage of the development of distributed generation – mainly solar photovoltaics, as well as small-scale storage and demand response. Distribution networks are a new market platform. As distributed resources can delay or substitute network investment, the economic regulation of the network is output-based so as to incentivise efficient investments.

Consumption (Chapter 9)

On the consumer side, policies encourage competitive retail pricing. Regulators rebalance the network tariff structure to reflect costs better, towards a fixed component and capacity and less based on energy consumed. Retailers are competing to offer innovative services and exploit the gains from consumers' participation in electricity markets.

Consumers are more elastic to retail prices because they have the possibility of reducing their consumption with behind-the-meter generation and storage. Retail prices, including taxes, seek to incentivise the deployment of distributed resources while preventing free riding on other electricity consumers, to ensure sustainability. A new form of competition takes place, not between retailers, but among distributed resources.

A market design for the transition

The market framework described here is primarily a list of best practices that can already been found in existing markets. It integrates carbon policies and support policies in the overall electricity market framework, retaining the existing market architecture while modernising the market design.

Such a framework seems fit for purpose during the next phase of the transition to low-carbon power in most scenarios. This market framework is complex, which is unavoidable for the electricity sector. But it can evolve, if necessary, as technology progresses.

Key recommendations: A new market framework for decarbonisation

Policy recommendations on re-powering electricity policy must take the following into account.

First, the path to the successful decarbonisation of power needs to maintain security of supply and keep electricity prices affordable. In a context of increasing retail prices, continuous improvement in energy efficiency will keep bills affordable. Decarbonisation is only one dimension of the energy trilemma, along with security and affordability, and to date this has required trade-offs.

Second, electricity markets and their regulatory frameworks remain under the shared responsibility of different jurisdictions at local, state and continental levels. While the decarbonisation goal is global and electricity systems cross borders, governments and states remain accountable for electricity security and, consequently, national, institutional and regulatory frameworks matter for market design. There is “no one size fits all” solution.

Third, complexity is unavoidable when it comes to decarbonising electricity, at once the most promising and also the most complex part of the energy sector. Fortunately, great effort has already been made over the past 20 years and valuable lessons have been learned from across the OECD regions.

Market design for the low-carbon transition will be evolutionary, and involves learning by doing. Following initial changes to market design, many additional changes have been implemented in power markets in the United Kingdom, California, Brazil and France and further changes are likely in the future.

Based on the analysis developed in this report, the following recommendations should be considered to create a market design and regulatory framework fit for purpose for the low-carbon transformation of the electricity system.

1) Supplement the market revenues of new low-carbon investment with long term risk-sharing tools. While it is a standard recommendation that carbon pricing must be introduced or strengthened to encourage deployment of low-carbon generation and electricity savings, it has also to be recognised that establishing or reinforcing the credibility of carbon pricing is likely to take time and may even increase perceived market price risks. Meeting renewable policy targets will require accelerated deployment and may depress electricity prices. Therefore, low-carbon investments have to be supported during the transition to low-carbon power. Support could take some form of long-term arrangements, including provisions to modulate support when carbon pricing is strengthened. This would facilitate the integration of low-carbon investments into the market while mitigating market price risk.

2) Increase the transparency and geographical resolution of prices during the adjustment period before operations. Locational marginal prices should be transparent during the adjustment period - that is, the last few hours before operations. Intra-day and balancing/real-time markets should be better integrated into a single market platform, and intra-day prices should be transparent to signal to market participants how to adjust their schedules as forecast errors are reduced. Scarcity pricing rules have to be defined ex ante, both in price terms and to address market power issues. Pricing of over-generation should reflect the actual marginal costs.

3) Regulate reliability by setting standards, defining scarcity pricing rules, and consider capacity mechanisms to create a safety net. Reliability standards should be set by regulators, while at the same time scarcity pricing should be developed where it does not already exist. Capacity mechanisms, if well designed, can provide additional safety nets to address the uncertainties of decarbonisation and make sure that reliability standards are always met.

4) Promote efficient demand participation. For small consumers, retailers have a key role to play by offering dynamic, time-varying energy prices. “Dispatching” demand response (in a similar way to generation resources) in wholesale markets should be limited to kick-starting nascent demand response markets, because of its complexity and risk of undue subsidisation.

5) Foster regional co-ordination for interconnections. Location planning for low-carbon power generation and network developments should be done across jurisdictional borders with all resources in an integrated fashion. When regional integration and new interconnectors raise differentiated impacts, a supranational entity may be necessary to evaluate the costs of new interconnection and allocate them to all relevant parties according to their share of the regional benefits.

6) Modernise the regulation of distribution networks ("Regulation 2.0"). The regulatory system should implement output-based regulation over longer periods to enable the efficient trade-offs between operating expenses (OPEX) and capital expenditures (CAPEX) and fully tap the potential of new, distributed technologies.

7) Urgently reform retail pricing to become more cost reflective. Retail rates, including electricity taxes and surcharges, should better reflect underlying costs, including time-varying electricity prices and the fixed nature of network costs, in order to enable the efficient deployment of distributed resources.

Abbreviations and acronyms

2DS	2 degree scenario
AEMC	Australian Energy Market Commission
ANEEL	national electricity regulator (Brazil)
CAPEX	capital expenditure
CBA	cost benefit assessment
CCGT	combined cycle gas turbine
CCS	carbon capture and sequestration
CEC	California Energy Commission
CEER	Council of European Energy Regulators
CEF	connecting Europe facility
CfD	contracts for difference
CHP	combined heat and power
CM	capacity market
CO_2	carbon dioxide
CONE	cost of new entry
CORESO	Coordination of electricity system operators
CPS	current policies scenario
CREZ	competitive renewable energy zone
CSPE	public electricity service contribution
CT	combustion turbine
DC	direct current
DER	distributed energy resources
DMNC	dependable maximum net capability
DOE	Department of Energy (United States)
DSPP	distributed system platform provider
EEG	renewable energy act (*Erneuerbare-Energien-Gesetz*)
ELCC	effective load carrying capability
ENTSO-E	European Network of Transmission System Operators
EPE	*Empresa de Pesquisa Energética* (EPE) (Brazil)
EPEX	European power exchange
ESAP	Electricity Security Advisory Panel
EU ETS	European Union emissions trading system
EUE	expected unserved energy
EV	electric vehicle
FCF	frequency converter facility
FCO_2	emissions of carbon dioxide
FERC	Federal Energy Regulation Commission
FIT	Feed-in tariff
FTR	financial transmission rights
GIVAR	grid integration of variable renewables
GWh	gigawatt
ICAP	installed capacity
ICT	information and communications technologies
IRR	internal rate of return
ISO	independent system operator
ITC	investment tax credit
KWh	kilowatt hour
LCOE	levelised cost of electricity

LCSCD	least cost security constrained dispatch
LDA	locational deliverability areas
LMP	locational marginal price
LOLE	loss of load expectation
LOLH	loss of load hours
LOLP	loss of load probability
LSE	load serving entity
LTP	local transmission plans
MBTU	million British thermal units
MMS	market management system
MOPR	minimum offer price rule
MTEP	transmission expansion planning
MWh	megawatt hour
NEM	National Energy Ministry (Australia)
NERC	North American Electric Reliability Corporation
NEW	Northwestern Europe
NGET	national grid electricity transmission
NPS	new policy scenario
NREL	National Renewable Energy Lab
OCGT	open cycle gas turbine
OFGEM	Office of Gas and Electricity Markets (UK)
OFTO	offshore transmission owner
ONS	system operator (Brazil)
OPEX	Operating expenses
OTC	over the counter
P+P	point to point
PCI	projects of common interest
PHEV	plug-in hybrid vehicle
PPA	power purchase agreement
Ppm	parts per million
PTC	production tax credit
RAP	reference annual revenue
REV	reforming the energy vision
ROI	return on investment
RPM	reliability pricing model
RPS	renewable portfolio standards
RTO	regional transmission operator
SC	steam cycle
SCED	security constrained economic dispatch
SME	small and medium enterprise
SOS	security of supply
SWE	Southwestern Europe
TGC	tradable green certificates
TOTEX	Total expenses
TPS	three pivotal supplier
TWh	terawatt hour
TYNDP	Ten-Year Network Development Plan
UCPTE	Union for coordination of production and transmission of electricity
VAT	value added tax
VOLL	value of lost load
VRE	variable renewable energy
WEO	*World Energy Outlook*

SECOND EDITION, March 2016

IEA Publications
9, rue de la Fédération, 75739 Paris cedex 15
Layout in France by Jouve.

Cover design: IEA. Photo credits: Nasa

Re-powering Markets

Market design and regulation during the transition to low-carbon power systems

"Re-powering" refers to the process of replacing older power stations with ones that are more efficient and more powerful, but the term also lends itself to market design. To facilitate the transition to a low-carbon economy, electricity markets will need to be "re-powered": older market frameworks must be replaced with ones suitable for decarbonisation while ensuring a secure electricity supply. Market rules need to be modernised and better matched with low-carbon policies while keeping the same overall market architecture.

Re-powering electricity markets can be done in several ways, depending on the existing market design or regulatory framework. Changes can be as limited as increasing the temporal or geographical resolution of existing markets or putting a price on scarcity, or as extensive as creating short-term markets and incorporating policies to increase renewables and reduce carbon emissions as part of a consistent market framework.

Re-Powering Markets brings together today's best practices in new electricity market design and details the most effective and efficient ways for re-powering electricity markets to address the 21st century challenges of transitioning to low-carbon electricity.

Electricity Market Series

2016