ECONtribute
MARKETS & PUBLIC POLICY

ECONtribute
Discussion Paper No. 183

Electricity Markets in Transition: A Multi-Decade Micro-Model of Entry and Exit in Advanced Wholesale Markets

Peter Cramton
David Malec
Emmanuele Bobbio
Pat Sujarittanonta

August 2022

www.econtribute.de

UNIVERSITÄT BONN
Universität zu Köln

Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany´s Excellence Strategy – EXC 2126/1-390838866 is gratefully acknowledged.

Electricity Markets in Transition

A multi-decade micro-model of entry and exit in advanced wholesale markets

Peter Cramton, Emmanuele Bobbio, David Malec, and Pacharasut Sujarittanonta[1]

September 2021 [[Latest draft]]

Abstract

Electricity markets worldwide are undergoing a many-decade transition in the way electricity is generated and consumed. The success of this transition depends critically on climate policy and market design. We model the most advanced electricity markets in the world to evaluate the impact of alternative policies on electricity market outcomes over the next 40 years, including costs, profits, social welfare, risks, and reliability. Each year, investors decide which resources enter and exit given forward-looking consistent expectations about energy profits, prices, and costs. The model is unique in modeling investment decisions at the individual unit level based on precisely calculated profits from energy, reserves, and capacity markets. These profits depend critically on the resource structure, which changes each year with investor decisions. New and essential elements of electricity markets, such as battery storage and price responsive demand are fully integrated. The model provides detailed insights into how policies such as carbon pricing impact the transition to renewable energy.

Peter Cramton
University of Cologne
Cologne, Germany 50923
pcramton@gmail.com

Emmanuele Bobbio
University of Cologne
Cologne, Germany 50923
ec.bobbio@gmail.com

David Malec
University of Maryland
College Park MD 20742
dlmalec@gmail.com

Pacharasut Sujarittanonta
Chulalongkorn University
Bangkok, Thailand 10400
pacharasut@gmail.com

[1] We are grateful to PJM Interconnection for funding and expert help. Funding also from German Research Foundation (DFG) under Germany´s Excellence Strategy–EXC 2126/1–390838866 and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program, grant 741409. The views expressed are our own and not those any organization we may be associated with.

Contents

1 Introduction

Electricity markets worldwide are undergoing a multi-decade transition in the way we consume and generate electricity. The pace and success of this transition depend critically on climate policy and market design. To evaluate the impact of policy on the energy transition, we develop a wholesale electricity market model of PJM, arguably the largest and among the most advanced electricity markets in the world. As in most markets worldwide, investors in the PJM market face a chaotic climate policy, which increases risk and uncertainty in making resource entry and exit decisions. Our model allows us to evaluate the impact of alternative policies on electricity market outcomes throughout the transition, including costs, profits, risks, reliability, and carbon emissions.

We take an investor's perspective. Each year, investors decide which resources enter and exit. We consider the market both with and without a capacity market. In the model, investors know the current state of the market and forecast the future state of the market. The market state determines profits from energy, reserves, and capacity markets. Profitable resources enter the market, and unprofitable resources exit. The equilibrium condition is that expectations are accurate. Entry and exit decisions are made three years in advance, consistent with the current capacity market and typical lead times. Investor decisions are modeled over 40 years from 2020 to 2060.

We confront two challenges. First, investments in energy resources are long-lived, so the analysis must extend for several decades. Second, the resource structure—the mix of generating resources—is undergoing substantial changes, so we cannot rely on historical profits but must explicitly model how energy, reserves, and capacity profits depend on the resource structure. Thus, we must model the energy, reserves, and capacity markets in detail, including the day-ahead market on an hourly basis and the real-time market on a five-minute basis. In this way, we can capture how profits over the year depend on the resource structure.

The electricity market design studied is the *standard market design* (FERC 2002) with several enhancements. The wholesale market consists of a day-ahead market for scheduling resources throughout the day. The day-ahead market co-optimizes energy and reserves. It is a financial market in which participants take positions for the next day consistent with their anticipated capabilities and needs. Then in real-time, every five minutes, the bid-based security-constrained economic dispatch determines real-time behavior. Both markets co-optimize energy and reserves, maximizing as-bid social welfare based on the availability of resources, the demand, and resource or system constraints.

The regulator represents consumers' preference for reliability in the real-time market with an operating reserve demand curve that reflects the value of lost load in shortage events (Hogan 2013). This administrative scarcity pricing addresses a market failure in today's market—the absence of demand response. As more and more consumers respond to the real-time price, the importance of administrative scarcity pricing will diminish. With enough price responsive demand, shortage events no longer occur because demand adjusts to balance supply. At this point, reliability is no longer an issue.

The market may have a capacity market, as PJM does today, to coordinate investments and assure adequate resources, or it may be an energy-only market, like ERCOT. The energy-only market has less assurance of sufficient resources. The regulator controls the reserve margin indirectly with the operating reserve demand curve. With a capacity market, the regulator directly determines the reserve margin through the capacity demand curve, which effectively imposes a floor on the reserve margin (Cramton et al. 2013).

Climate policy is modeled simply as a carbon price path over time. Investors form expectations about the carbon price path and assume this path when evaluating the stream of cash flows over the life of the resource. In 2017, 84 percent of utilities reported using an internal carbon price when making investments (CDP 2017). One can interpret the carbon price either as an explicit carbon price or an implicit price reflecting the financial impact of regulations, such as subsidies or emission controls. We use the carbon price as our climate policy since it is easy to interpret. It also is the least-cost instrument to achieve any desired level of emissions. The consensus among economists is well-expressed by Janet Yellen, now Treasury Secretary, "Meaningful carbon prices are the cornerstone of any effective policy packet."[2]

Our approach then identifies the least-cost transition to any climate objective. The virtue of directly pricing the social cost of carbon is well-understood. Unfortunately, the path that most easily falls out of the political process is a myriad of conflicting subsidies and command-and-control regulations. These also can be modeled but not in a coherent way.

A final motivation for using a carbon price to reflect climate policy is that it mimics what investors do in evaluating energy investments. The assumed carbon price has a first-order impact on the economics of an energy resource. Investors understand this fact.

Looking at today's planned entry and exit of resources in the United States, it may appear that the transition to renewables is happening even absent a coherent climate policy. Entry is dominated by wind and solar, and exit is dominated by coal and older gas plants. However, the existing resource structure is still dominated by coal and natural gas. Only a small fraction, currently less than four percent of capacity, changes through entry or exit in any year.

As the penetration of renewables grows, net load—the demand for energy less solar and wind production—becomes more volatile. Storage and price responsive demand will play a greater role in price formation and meeting net load. Today price formation comes primarily from coal and gas units. This role will shift to storage and price responsive demand as fossil resources exit.

Both storage and price responsive demand are explicitly modeled in the energy and reserves markets. Storage is fundamentally different than other resources (Crampes and Trochet 2019, Chen et al. 2021, Junge et al. 2021). Its behavior depends on both the current and future prices. Storage is fully integrated. It is optimally scheduled day ahead and then optimally dispatched in real-time. Storage provides both energy and reserves and participates in the capacity market like other resources. Its contribution to reliability depends on its ability to provide energy and reserves during a shortage.

Our model allows us to provide answers to many interesting questions about electricity market design and climate policy.

- How does the resource structure change over time?
- Is price formation still effective with high levels of zero-marginal cost resources?

[2] At 00:10:30 of the 8 October 2020 video announcing the G30 report on climate policy (G30 2020).

- How does the pace and cost of the energy transition depend on the carbon price path?
- What is the cost of achieving 80-percent renewable energy by 2030 and 100-percent by 2035? What is the cost of achieving other goals?
- How do an energy-only market and a capacity market compare in terms of cost and reliability? How does this depend on the operating reserve demand curve?
- What is the impact of excluding some state-subsidized resources from the capacity market?
- How does the adoption of electric vehicles impact the electricity market?
- How does the quantity and type of storage resources depend on renewable penetration?

More broadly, we develop a tool and methodology for understanding how market outcomes depend on factors that impact entry and exit decisions, such as the carbon price path, the progress of technologies (IEA 2020), and fuel costs. Our approach recognizes the essential role of the spot markets—day-ahead and real-time energy and reserves markets—in determining long-run profitability (Ela et al. 2016, Dyson et al. 2018).

Capacity market revenues are also included. A well-designed capacity market should play a secondary role. Ideally, the capacity market has full-strength performance incentives, so there is no missing money. The shortage price is the value of lost load. The capacity price then compensates the resource for the financial cost of the obligation to provide energy during a shortage. This economic cost is lost energy profits during shortages, which the capacity resource has sold forward in the capacity market. Actual capacity markets may fail to achieve this ideal (Mays 2021, Mays et al. 2019, Gramlich and Goggin 2019).

All longer-term products should be derivative of the day-ahead or real-time energy market. We see this in the ICE and CME energy futures traded in PJM and other major markets. In both cases, the futures are for one-month peak or off-peak real-time energy at a central hub or load zone. These products are traded up to five years ahead, although liquidity falls for more distant months. Forward products can improve market performance by reducing risk and market power (Jha and Wolak 2020). Cramton (2021) describes a method for efficient forward energy trading.

Many studies have modeled energy markets with high levels of renewables. This focus makes sense. The switch to renewables is an essential ingredient in any plan to reduce carbon emissions to sustainable levels. Jenkins et al. (2018) review 40 such studies. There is consensus that decarbonization must start in the electricity sector.

Further emission reductions are obtained through electrification of transportation, heating, and other carbon-intensive activities (Griffith 2020, Phadke et al. 2020, Murphy et al. 2021, ACPA 2020, Williams et al. 2021). With a few exceptions, such as Sepulveda et al. (2018), the studies model energy markets in simplified ways. Details of actual markets are ignored, such as optimal scheduling of fossil resources with three-part bids to reflect startup and minimum energy costs in addition to marginal cost (Larson et al. 2020). Other unit characteristics such as ramp rates are also typically ignored. None of the models consider equilibrium investment and model the details of the energy and reserves markets. Our model not only models these details, it includes future market enhancements, such as full integration of storage resources and price responsive demand. These are important initiatives that U.S. system operators plan to implement in the coming years.

Early long-term models, such as Hobbs et al. (2007), analyze market outcomes assuming a steady state. The goal is to identify the equilibrium resource structure when the determinants of a resource's profits do not change. Steady-state analysis is suitable in a stable world but is inappropriate for modeling a transition to a radically different world. Our research focuses on the transition process to the new world.

More recent long-term models that do allow a process of transition and explicitly model entry and exit are myopic (Richstein et al. 2014). The investments in a year depend on current year energy profits rather than forward-looking energy profits. Investors in practice are forward-looking, as in our model. Investments are made based on the projection of cash flows over the life of the resource.

Our model provides a tool for electricity market design (Wilson 2002, Cramton 2017, Wolak 2021a,b). Although we tailor the model to the PJM market, including future enhancements, we can adapt it to any electricity market by adjusting the market optimization to conform to the desired market. For example, an energy-only market is studied by turning off the capacity market. Many "what if" questions are readily answered. What happens if the operating reserve demand curve changes? What happens if the capacity demand curve changes? Can we make the operating reserve demand curve consistent with the capacity demand curve so that the capacity price reflects the financial cost of the capacity obligation?

Our model can also help inform the design of long-term products and markets. Long-term products typically are derivatives of day-ahead energy (Cramton 2021a) or real-time energy (Gimon 2020), but other constructs could be considered as well (Corneli 2020, Pierpont 2020, Tierney 2020).

One thing is certain. Decarbonizing electricity is no small task. Given the ambitious emission goals at the federal and state levels, we need powerful data-driven tools to provide evidence on difficult policy questions. Effective policies are essential. Poor policies that are expensive or ineffective are apt to derail the efforts to achieve meaningful climate goals.

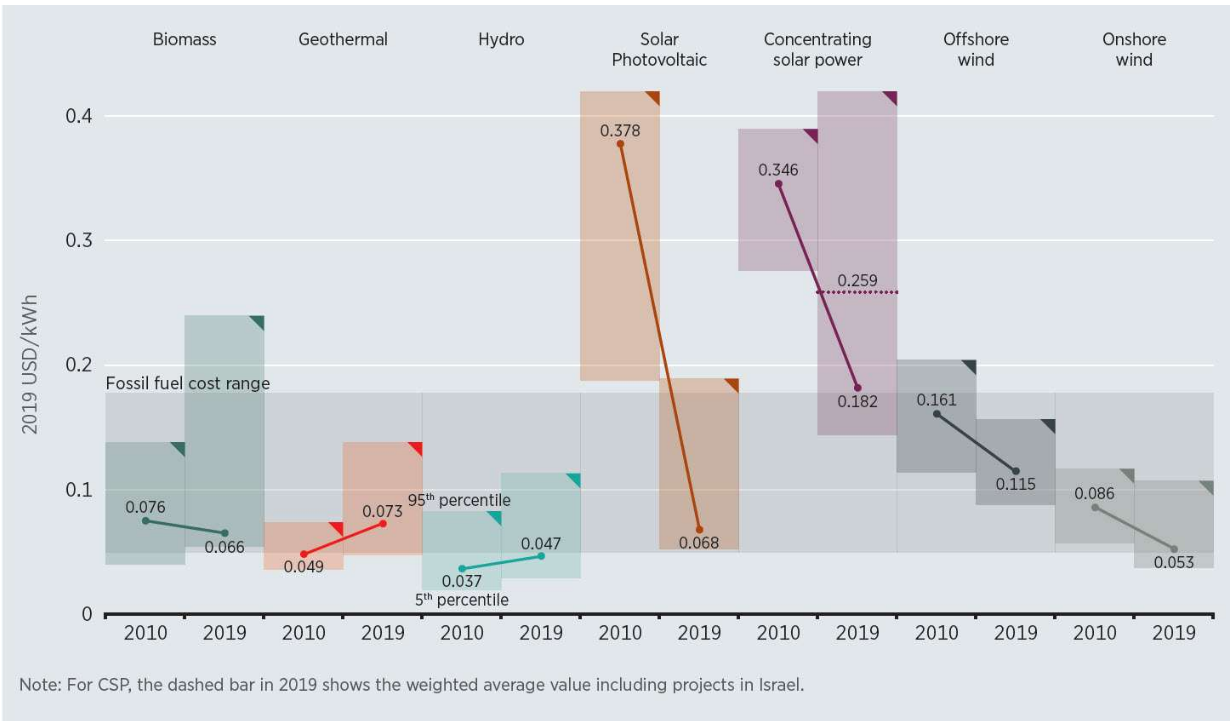
We first take a quick look at why decarbonization is a decades-long process. Section 3 describes how we forecast net load—traditional load minus the production from non-responsive resources such as solar and wind. The system operator's main task is to serve real-time net load with the supply from responsive resources. Next, we discuss the other inputs to the energy market, such as the operating reserve demand curve, forced outages, planned outages, and the bid curves for the upper segment of some gas resources. In section 5, we describe the energy market model—the core of our modeling effort. We describe the day-ahead optimization, the sequence of intraday optimizations, and finally, the real-time optimization. Section 6 presents the multi-year simulation model, where investors make annual entry and exit decisions. To make the analysis tractable, we develop an energy proxy model that econometrically estimates energy profits. The proxy model allows us to make hundreds of millions of calls to calculate energy profits from only tens of thousands of energy market runs. Each energy market run is computationally expensive. We can do tens of thousands of runs but not hundreds of millions. In section 8, we present the results.

As an illustration, the paper focuses on one policy choice that is especially relevant in PJM's market today. We compare the likely market impact of two alternative market rules to address buyer market power in the capacity market. A broad minimum offer price rule (MOPR) applies a minimum offer price to all renewable resources. A narrow MOPR only applies a minimum offer price under limited circumstances. We compare market outcomes with a broad MOPR, endorsed in the December 2019 order of the Federal Energy Regulatory Commission (FERC 2019), and a narrow MOPR (PJM 2021a,b). Cramton (2021b) discusses these two variations in detail.

2 Effective policy and good market design are needed to achieve climate goals

Someone tracking the economics of energy resources might conclude that decarbonizing electricity is easy. Solar and wind are economic in many markets and are improving at a much faster clip than fossil resources (Figure 2.1). Global weighted average levelized cost of electricity from renewables has fallen dramatically from 2010 to 2019 (IRENA 2020). Solar photovoltaic has fallen from 37.8 to 6.8 cents/kWh. Onshore wind has fallen from 8.6 to 5.3 cents/kWh. Levelized costs for fossil generation are 5 to 18 cents/kWh.

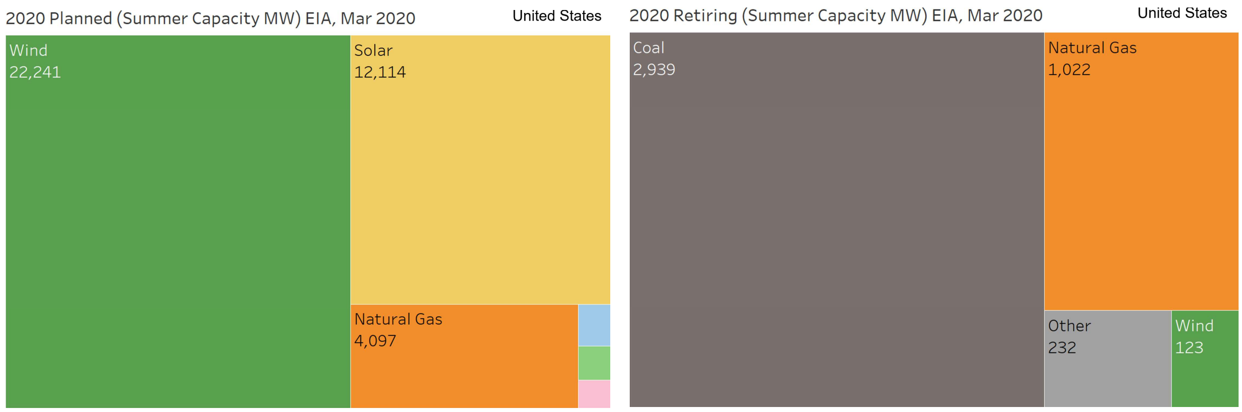
Figure 2.1: Global weighted average levelized cost of electricity from renewables 2010-19 (IRENA 2020)



Note: This data is for the year of commissioning. The thick lines are the global weighted-average LCOE value derived from the individual plants commissioned in each year. The project-level LCOE is calculated with a real weighted average cost of capital (WACC) is 7.5% for OECD countries and China and 10% for the rest of the world. The single band represents the fossil fuel-fired power generation cost range, while the bands for each technology and year represent the 5th and 95th percentile bands for renewable projects.

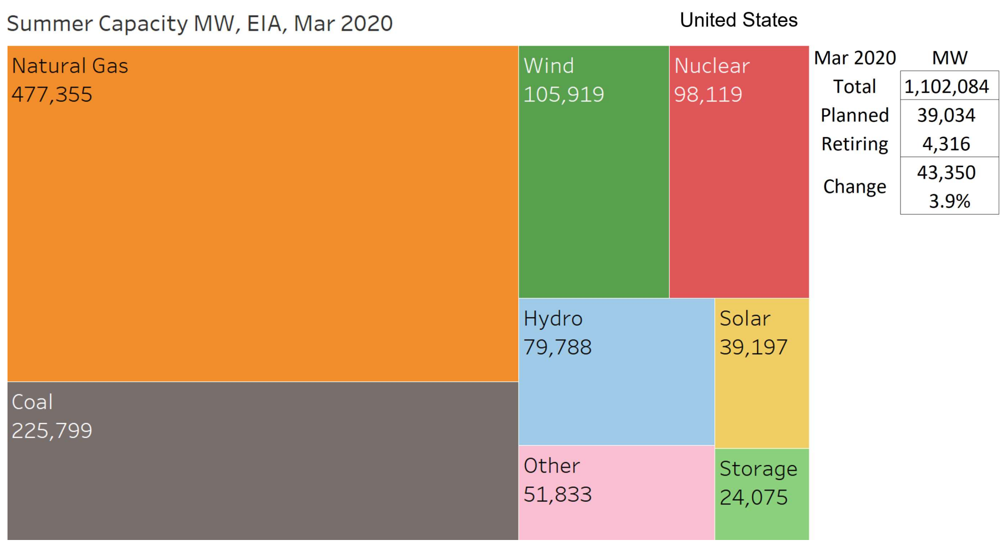
Thus, it may appear that solar and wind will enter, fossil resources will exit, and the energy transition will happen on its own without intervention. Indeed, this seems to be happening in the United States, where entry is dominated by wind and solar, and exit is dominated by coal and older gas plants (Figure 2.2).

Figure 2.2: 2020 planned entry and retirements of energy resources in the United States



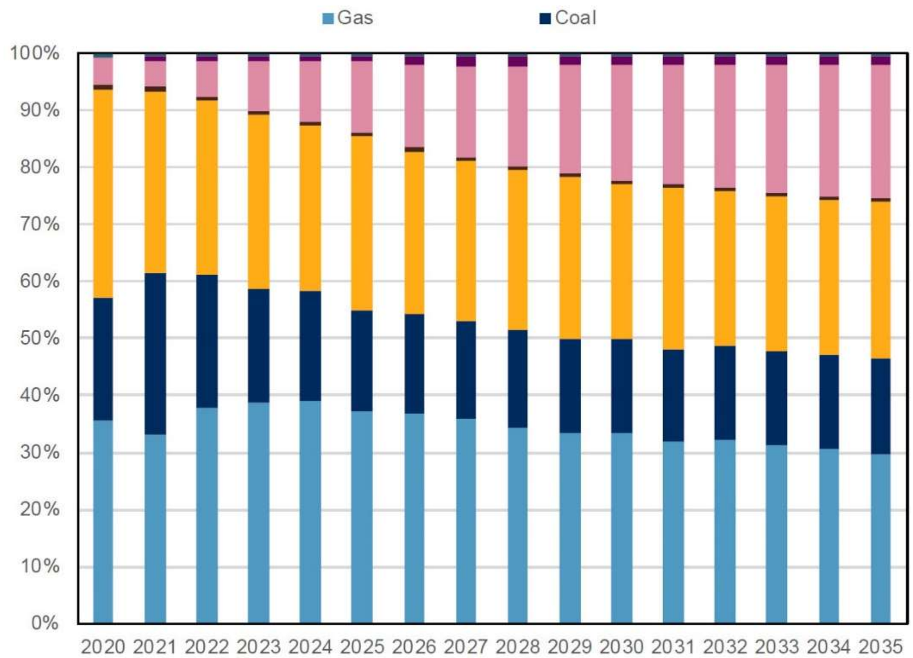
However, Figure 2.3 shows the existing resource structure is still dominated by coal and natural gas. Less than four percent of capacity changed through entry or exit in 2020. Moreover, as we add more wind and solar, the value of these resources diminishes unless complementary assets, like storage, are added.

Figure 2.3: 2020 energy resources in the United States



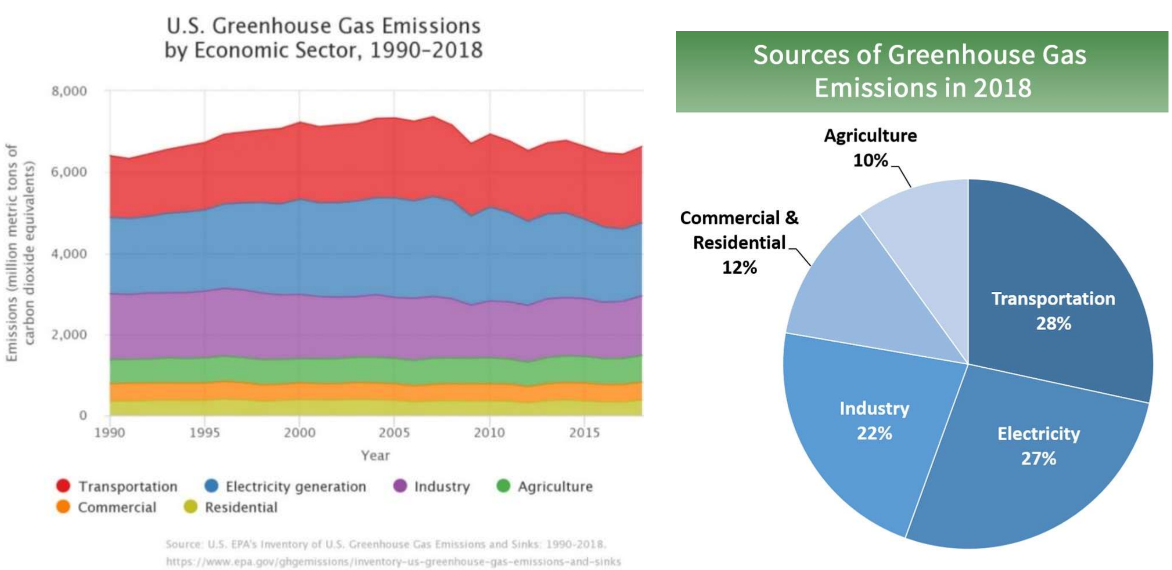
Forecasts of PJM's resource structure over the next fifteen years show a path inconsistent with state renewable portfolio standards, Biden's plan of net-zero by 2035, and the Paris goal. A recent S&P Global Market Intelligence (2020) forecast (Figure 2.4) shows coal and gas declining slowly. By 2035, they estimate coal and gas to be 46% of the market, down from 57% in 2020.

Figure 2.4: PJM projected resource mix, 2020-35 as of 7 Dec 2020 (S&P Global Market Intelligence 2020)



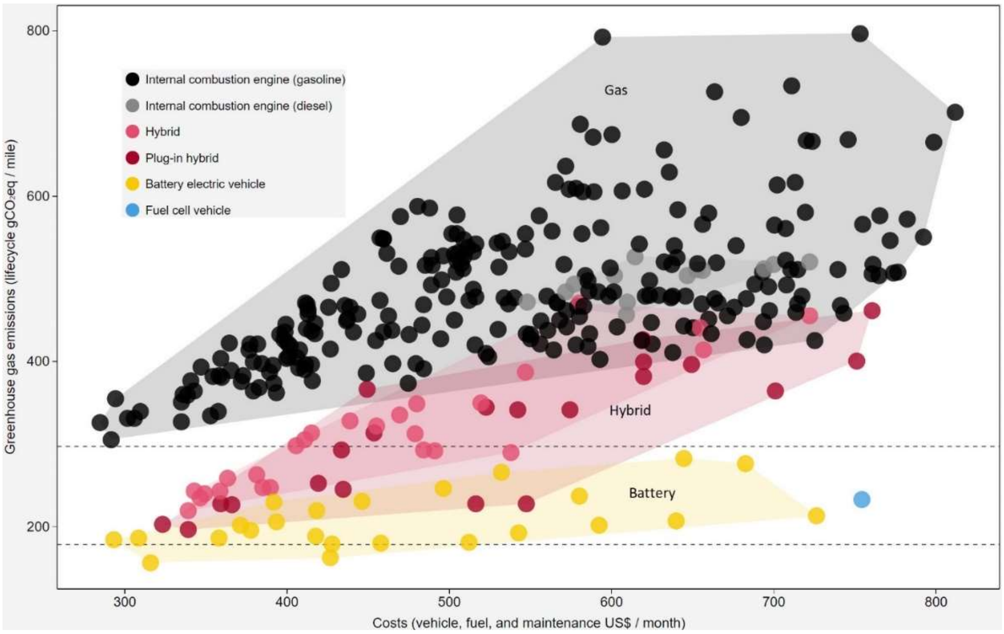
Climate policy will be needed to achieve climate goals. Effective policy and good market design will play an essential role in managing the cost of achieving these goals. Electricity must be the centerpiece of these efforts. We cannot decarbonize the transportation, heating, and industrial sectors without decarbonizing electricity. As of 2018, U.S. greenhouse gas emissions are dominated by transport, 28%, and electricity, 27% (Figure 2.5).

Figure 2.5: U.S. greenhouse gas emissions by sector (EPA 2020)



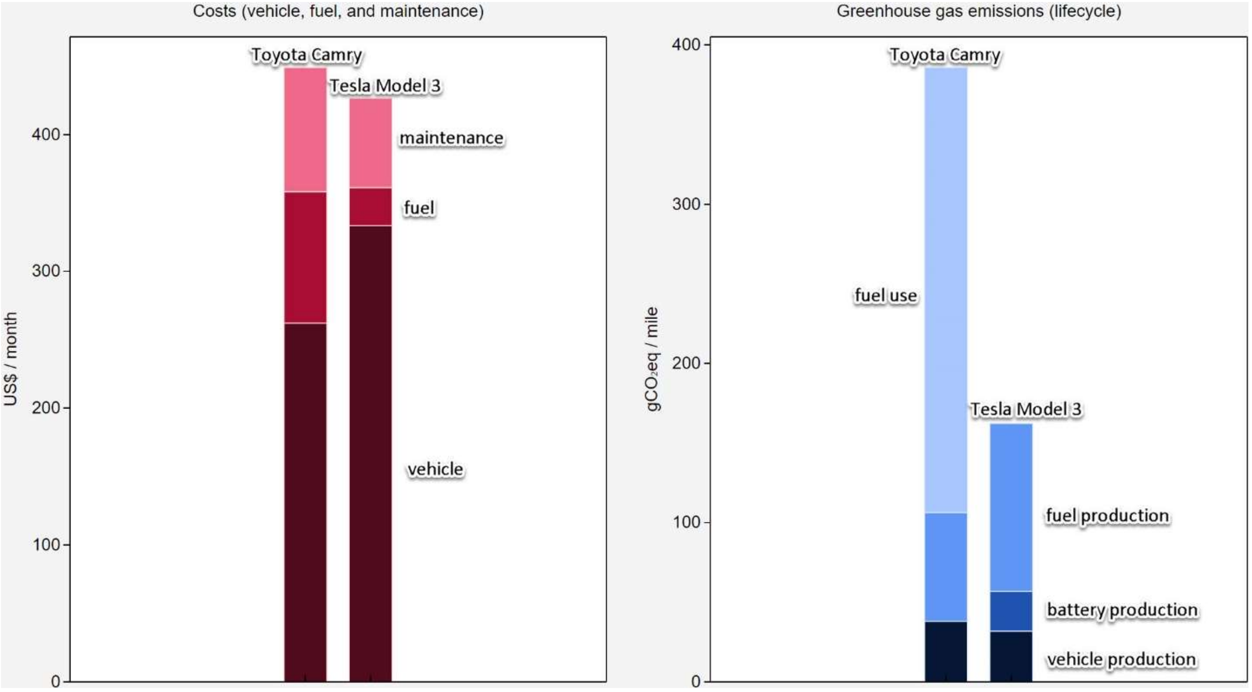
Electric vehicles are the means to decarbonize transportation, but this only happens if electricity is carbon-free. Fortunately, the progress of electric vehicles has been rapid (Tanenblatt et al. 2021). Today, electric vehicles are cost competitive with gas, diesel, and hybrid vehicles when considering the total cost of ownership. Electric vehicles have lower maintenance and fuel costs.

Figure 2.6: 2021 vehicle cost and carbon emissions (Miotti and Trancik 2021)



Two popular 2021 midsized sedans illustrate the point. A Tesla Model 3 has a monthly ownership cost of \$426 compared with \$449 for a Toyota Camry. Lifecycle emissions are 162 gCO_2eq/mile for the Tesla Model 3 and 385 for the Toyota Camry (Miotti and Trancik 2021), assuming current electricity production. If electricity were carbon-free, then the Tesla emissions would fall from 162 to less than 57.

Figure 2.6: Comparison of 2021 Toyota Camry and Tesla Model 3 (Miotti and Trancik 2021)



Seven essential questions

Our model can address many important issues of critical infrastructure in the electricity and transportation sectors.

Feasibility: Is the core wholesale market design sufficient?

The first question to answer is whether radical changes are required to the electricity market design with a high level of renewable generation. Cramton (2017) argues that incremental changes that focus on integrating storage, price responsive demand, and distributed generation within the same core structure will be sufficient. The role of price formation would shift from fossil generators to storage and price responsive demand as the share of renewables grows. Our model confirms the feasibility of incremental improvements.

Reliability: Can we avoid energy shortage?

A vital market design goal is reliable electricity service. Our model shows how reliability depends on the market design and policy. Within the framework of a capacity market, the key determinant of reliability is the capacity demand curve. In an energy-only market, reliability depends primarily on the operating reserve demand curve. The model also helps us understand when reliability events are most likely to happen and at what frequency and duration. Finally, the model can optimize the critical reliability instruments: the capacity demand curve and the operating reserves demand curve.

Price formation: Will arbitrage from storage and price responsive demand price energy?

Our model provides a detailed answer to the question of price formation. We observe the distribution of prices for energy and reserves in the day-ahead and real-time markets. This hourly price behavior day ahead and five-minute price behavior in real-time are essential outputs of the model. Understanding these prices lets us calculate the profits of different resources as the resource structure changes, and therefore is central in understanding the incentives for entry and exit.

Planning: How can transmission reduce the cost of decarbonization?

A simplifying assumption of our initial model is that transmission is built to support the entry and exit as it occurs. We, therefore, assume that transmission congestion is the result of outages and other random events. Congestion does not persist. We intend to extend the model to include persistent congestion and the potential for resolving persistent congestion with transmission investment. Transmission modeling can be done both within a market, like PJM, or across markets, such as PJM-MISO or nationwide. Co-optimizing transmission and the energy and reserves markets can identify how transmission investment fits into the transition to renewables. Historically, utilities built fossil generation near load. With renewables, this is not so easy. Often the best locations for wind and solar are far from load. Transmission investments are required (Bloom 2018, Caspary et al. 2021, Gramlich and Caspary 2021). The model helps us understand this piece of the puzzle. Our work complements state-of-the-art planning tools (Clack et al. 2020, Lombardi et al. 2020).

Policy: How does carbon pricing compare with state clean energy standards or tax credits?

Economists have long argued that carbon pricing is the best instrument for addressing climate change (MacKay et al. 2015; Cramton et al. 2017; CPLC 2017). Carbon pricing is especially effective in electricity where implementation is straightforward, and leakage is less of a concern. Our model demonstrates the power of carbon pricing in hastening the pace of the transition and measuring the cost of achieving any climate goal. The model also allows examining alternative policies such as state clean energy standards or tax credits (Pechman 2021, Spees et al. 2019). Our model can assess the cost of using these alternative instruments.

Uncertainty: What is the cost of political uncertainty?

Most everywhere in the world, climate policy has been chaotic. We have seen significant shifts in policy with each election. What are the costs of this political uncertainty? We plan to model uncertainty with a Markov-state-transition matrix. Investors recognize that policies shift over time. This uncertainty increases the volatility of payoffs, raising the cost of capital or discount rate.

Innovation: How can we harness innovation to reduce the cost of emission reduction?

Innovation has been and will continue to be essential in reducing the costs of the energy transition. Gates (2021) rightfully emphasizes innovation. Supply innovation, new technologies that can generate carbon-free energy at lower cost, and demand innovation, new technologies for lowering demand or making demand more responsive to the real-time cost of generation, are needed. Our model allows us to better understand the benefits of innovation by examining various paths of innovations for supply and demand. In this way, we can establish the relative benefits of improvements in generation cost, energy efficiency, and demand flexibility. The relative benefits and costs enable us to better allocate limited R&D funds.

3 Forecasting net load

An electricity market's primary goal is to satisfy load reliably. In a world with renewable resources, this is best thought of as serving net load—traditional load minus the production from non-responsive resources such as solar and wind. Renewable resources can be curtailed in the event of overproduction, but this is done as a last resort in response to transmission constraints. For this reason, we do not model the curtailment of renewables.

A critical input to our energy market model is the net load forecast and the net load realization. We estimate the model for gross load on historical data and use PJM's projections through 2035 to extrapolate the process to future years. For renewables, we assume stationarity and use generation data for tracking solar in PJM and the Wind Integration National Dataset Toolkit by NREL—which covers both onshore and offshore sites in PJM's footprints—to estimate the capacity factor. Using renewable penetration and efficiency, we combine these four processes and obtain the model for net-load. The model generates net load forecasts for 36.5, 35.5, …, 0.5 hours ahead, which are used to schedule resources day-ahead and during the day. The model also generates realized net load in five-minute intervals. We do this for every hour and five-minute interval for 90 years.

We now discuss some of the details of the load model.

Load model

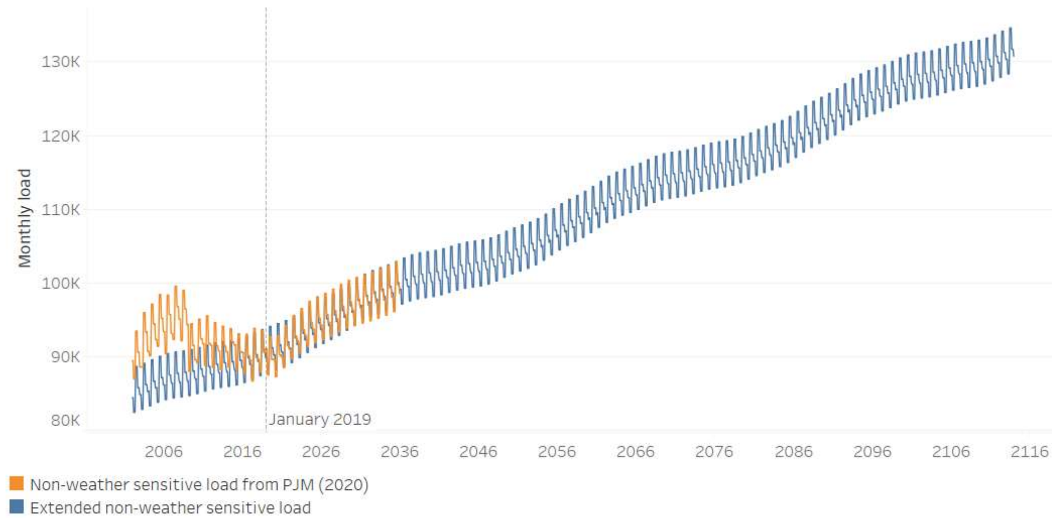
Load distribution estimation

We first need to forecast hourly load distributions for each hour in the forecast period. We use quantile regression models to estimate the distributions. The dependent variable is the hourly load, and the independent variables are the following:

- Monthly non-weather-sensitive load
- Month dummy variables
- Day-of-week dummy variables
- Hour dummy variable
- Interactions between month dummy variables and holiday dummy variables
- Interactions between day-of-week dummy variables and holiday dummy variables
- Holiday dummy variables (PJM 2020d)

The load forecast model uses the non-weather-sensitive load—the load absence cooling and heating equipment—as the main trend driver. The non-weather-sensitive load reflects what we believe about the future load. The monthly non-weather-sensitive load between 2002 and 2035 is from PJM (2020d). Then, the non-weather-sensitive load is extended to the rest of the forecast period. The independent variables are linear trend, sine trends, and month dummy variables in the baseline load profile. Figure 3.1 shows the non-weather-sensitive load from PJM (2020d) and the extended series used in the benchmark simulation.

Figure3.1: Non-weather-sensitive load



We use multiple quantile regression models to estimate first to 99th percentile with an increment of one percentile. Thus, we estimate 99 quantile regression models. Assume that the lower bound and upper bound of the distribution are equal to the first and 99th percentile, respectively.

After obtained the forecasted quantiles, we observe that the models underestimate the load distributions around the higher quantiles, especially the peak. Therefore, we scale the 51st to 100th percentiles of all forecasted distributions to match those in-sample. Mathematically, let q_{it} and $\hat{q}_{it}$ be the actual and forecasted i^{th} load quantile at time t. We aggregate the actual load quantile at the weekly level. The scaled load quantile is calculated as

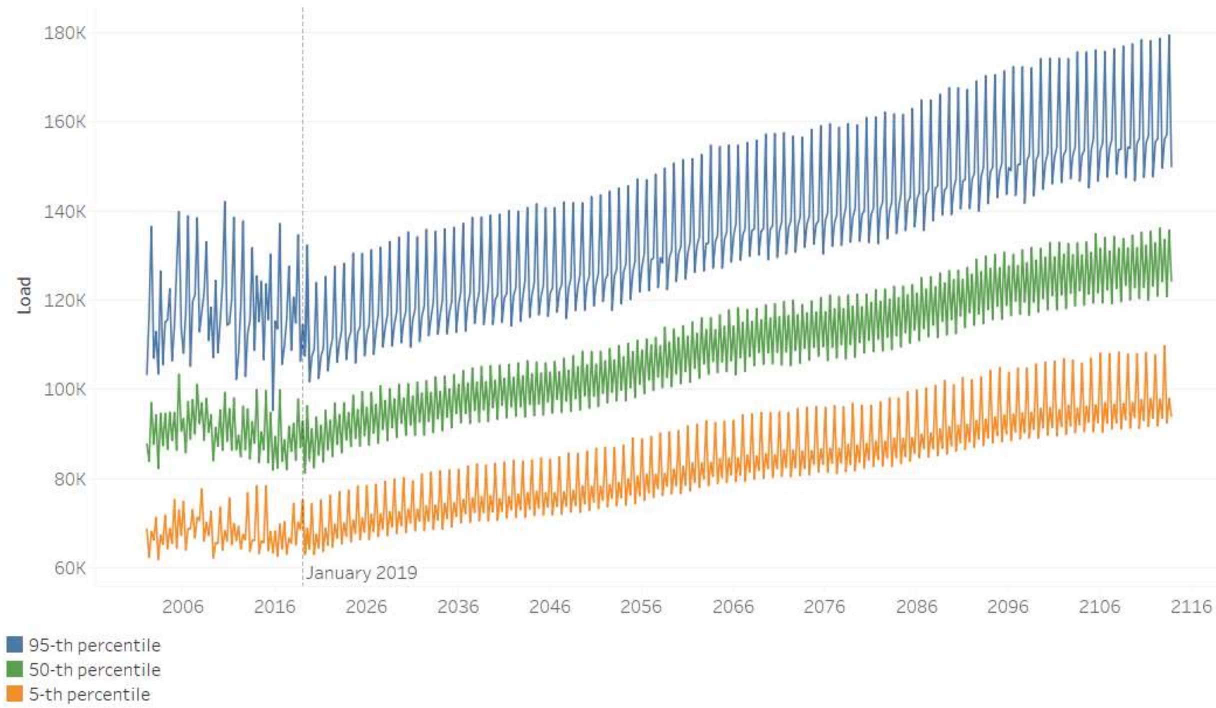
$$\tilde{q}_{it} = \begin{cases} \hat{q}_{50t} + \lambda_i(\hat{q}_{it} - \hat{q}_{50t}), i > 50 \\ \hat{q}_{50t}, i \leq 50 \end{cases}$$

The scaled factor for i^{th} quantile is defined as

$$\lambda_i = \frac{q_{100t} - q_{50t}}{q_{50t}} \cdot \frac{\hat{q}_{50t}}{\hat{q}_{100t} - \hat{q}_{50}}$$

We assume that the load is uniformly distributed between the two consecutive scaled quantiles to derive the entire load distribution. Figure 3.2 shows the forecasted load distribution from 2019 to 2113.

Figure 3.2: Load distribution



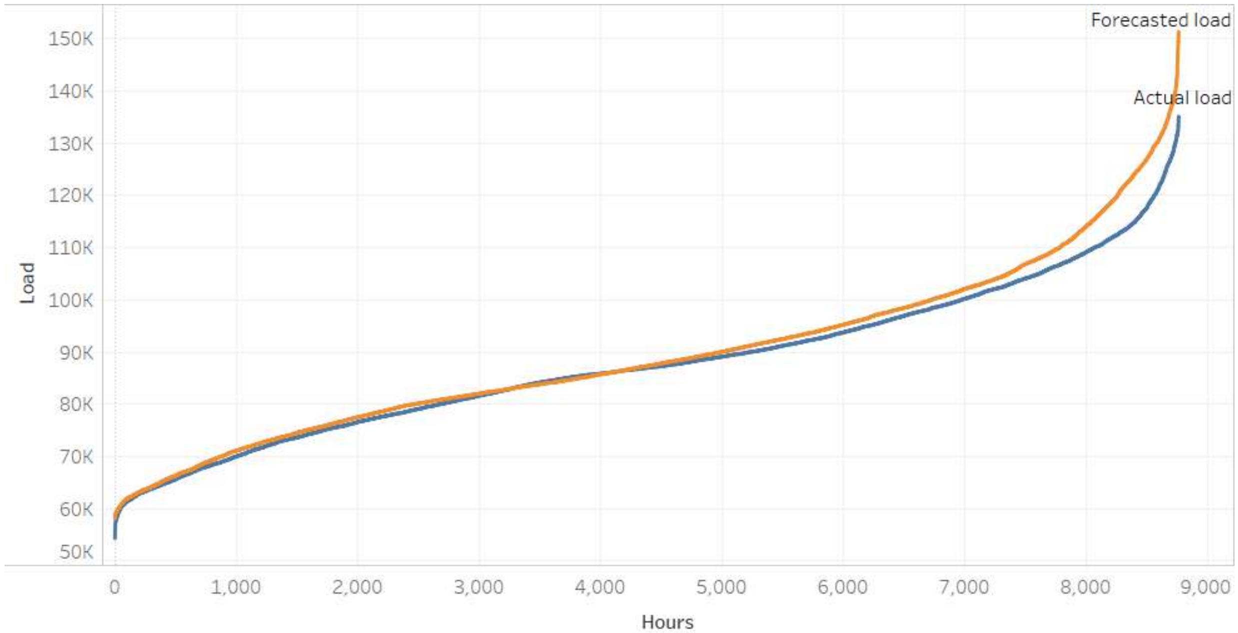
Load realization

To generate the realized hourly load, we draw the load from the scaled load distribution by the inverse transform sampling on the cumulative distribution function of load previously derived. The random process for the inverse transform sampling is defined as

$$e_t = 0.5e_{t-1} + 0.5\varepsilon_t$$

where $e_0 = 0.5$ and ε_t is normally distributed between zero and one. The realized 5-minute load is approximated using the linear interpolation between the realized hourly loads. Figure 3.3 shows the duration curve of the actual and forecasted load in 2019.

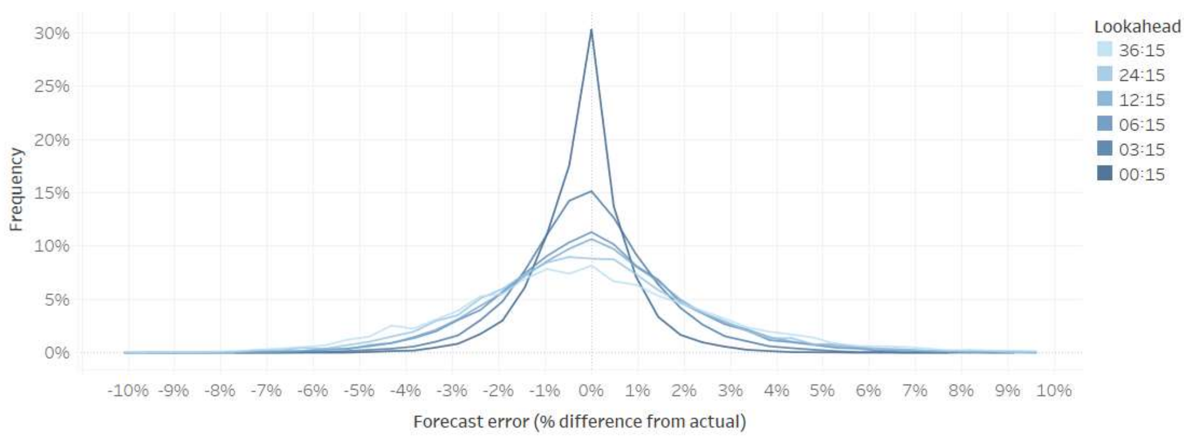
Figure 3.3: Duration curves of the actual and forecasted loads



Hourly forecasted loads

The energy market simulation requires load forecasts with 0:30, 1:30, 2:30, …, 36:30-hour lookahead. We obtain the actual load and forecast loads with 0:15, 1:15, 2:15, …, 36:15-hour lookahead between 2011 and 2019. Then, we assume that the forecast error—the percentage difference between the actual and forecast loads—is normally distributed with the mean of zero. The variance of the distribution is calculated using the actual forecast errors with the corresponding lookahead. Figure 3.4 shows the actual distribution of the forecast errors. Also, we assume that the 0:30-hour lookahead forecast is equal to the 0.15-hour look ahead, and so on. The forecasted load is the product of the forecast error and the realized load.

Figure 3.4: The actual distribution of the forecast errors



Renewable model

Solar, offshore wind, and onshore wind are modeled as the product of capacity, efficiency, and the capacity factor, reflecting weather conditions and assumed to be stationary and independent of technology. We assume the nameplate capacity and efficiency are intrinsic characteristics. The capacity factor varies at five minutes frequency based on changing weather.

The renewable models aim to explain capacity factors in the future. Thus, it is important to select renewable sites with technology relevant to the forecasting period. Then, we estimate the capacity factor distribution using a methodology similar to the load model. Finally, we apply a linear transformation to match the average capacity factor in PJM during the year and during the summer peak hours, defined as the hours from 15:00 and 18:00 (hour-ending) in June, July, and August.

Wind site selection

We select a subset of sites from NREL's Techno-Economic WIND Toolkit. We use NREL synthetic data for these sites over 2007-2013 to compute series for the aggregate capacity factor and fit a quantile regression. NREL's Techno-Economic WIND Toolkit uses detailed weather measurements at 126,691 sites in the continental United States to simulate generation at the 5-minute frequency for 2007-2013. 14,221 of these sites are offshore; the rest are onshore. We approximate PJM's footprint as follows.

We select onshore sites in P.A., NJ, MD, DE, VA, WV, OH. For IL, KY, and N.C., where PJM footprint is partial, we restrict attention to sites satisfying:

- Illinois: latitude > 40.5, , longitude > –89.5
- North Carolina: latitude > 35.3, longitude > –77.3

For offshore wind on the East Coast we consider sites satisfying:

- Latitude < 40.3
- Latitude > 35.0
- Longitude > –78.0

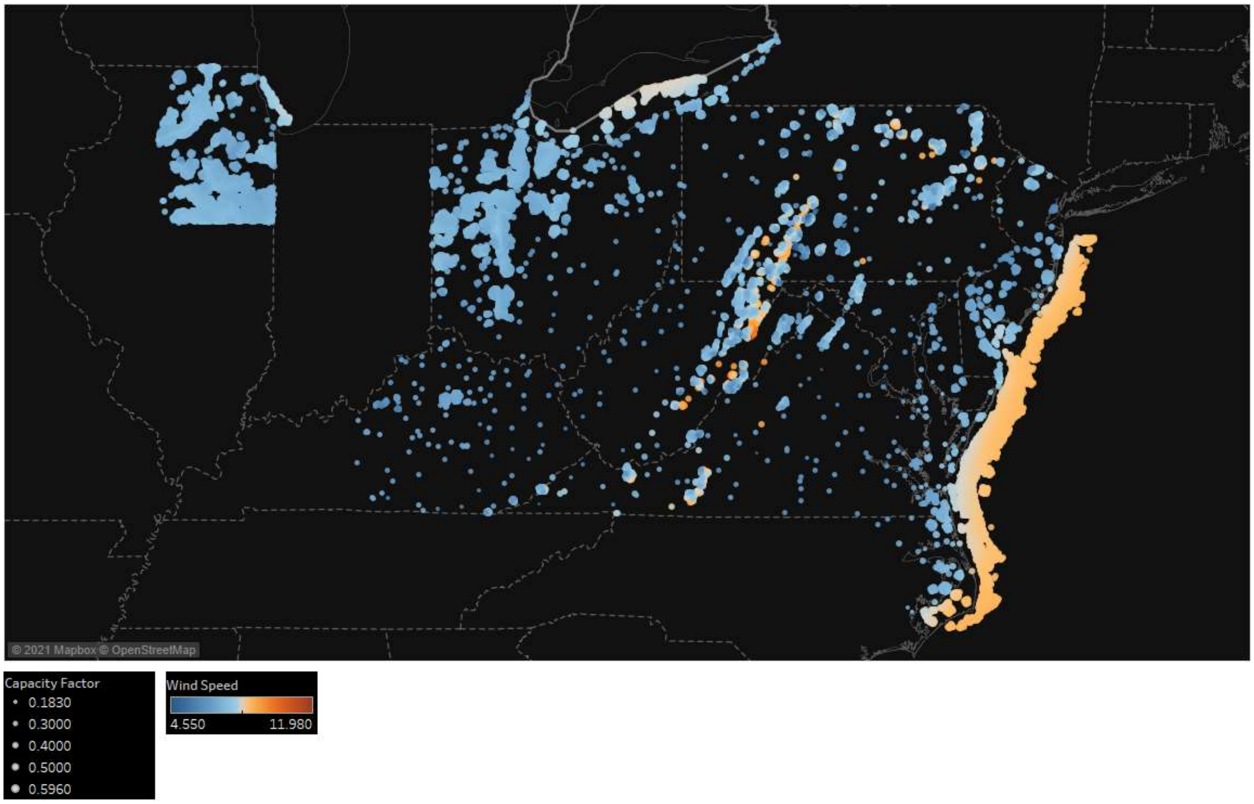
We also consider offshore sites in the Midwest by applying the conditions:

- Latitude < 42.5
- Latitude > 35.0
- Longitude < –78.0
- Longitude > –89.5

Finally, we compute the haversine distance to the closest onshore site for each offshore site (after verifying that the list of easternmost onshore sites tracks the U.S. east coast). We restrict attention to offshore sites between 10 and 65 km to approximate the region where it is possible to install fixed bottom offshore turbines.

Figure 3.5 displays the location of these sites.

Figure 3.5: Wind sites



Next, we restrict attention to sites with the 10% lowest coefficient of variation for the capacity factor. There are three motivations for applying this criterion.

First, as shown below, the capacity factor tends to be lowest during the summer peak hours, when load and energy rents are higher. Sites with more consistent wind tend to have a higher capacity factor during these hours.

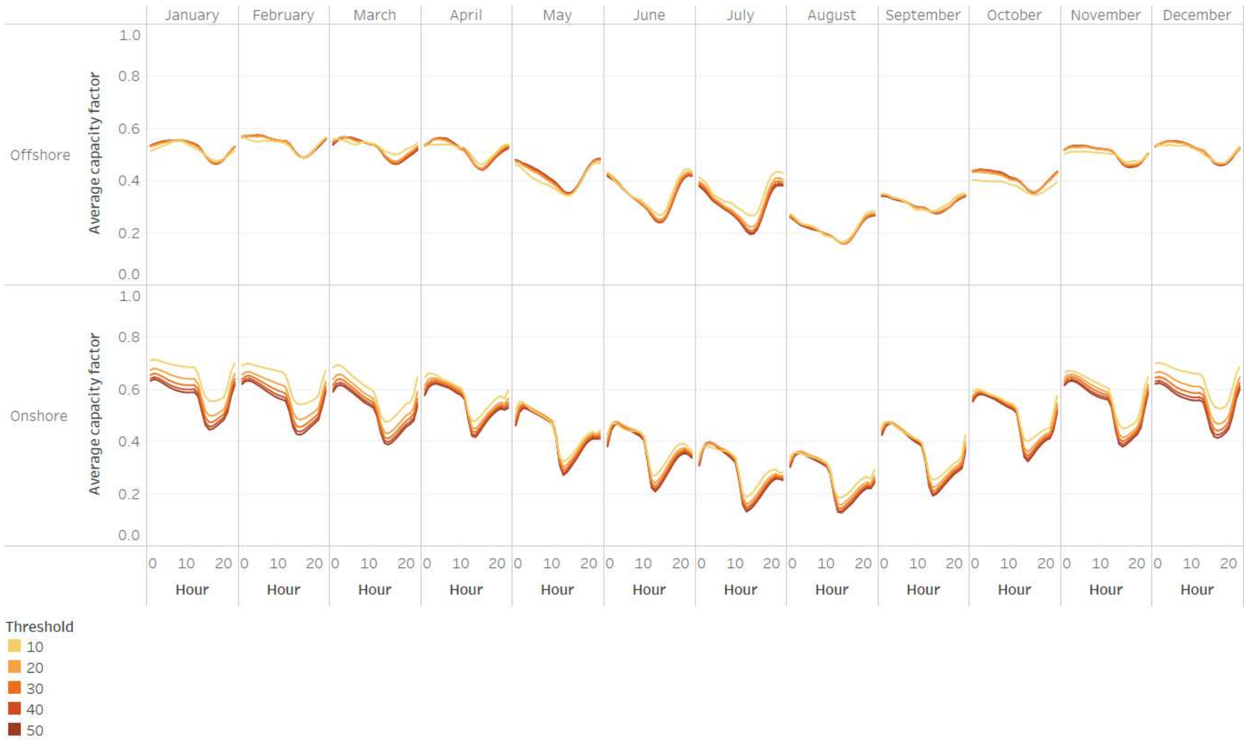
Second, investors should favor turbines with a lower specific power to perform at a higher capacity factor during these scarcity hours.

Third, technology improvements are delivering more consistent production and increasing capacity factors at low wind speed.

In sum, restricting attention to sites with more stable capacity factors should better approximate investors' choices and the development of technologies, absent a deeper wind generation model that takes weather directly as an input.

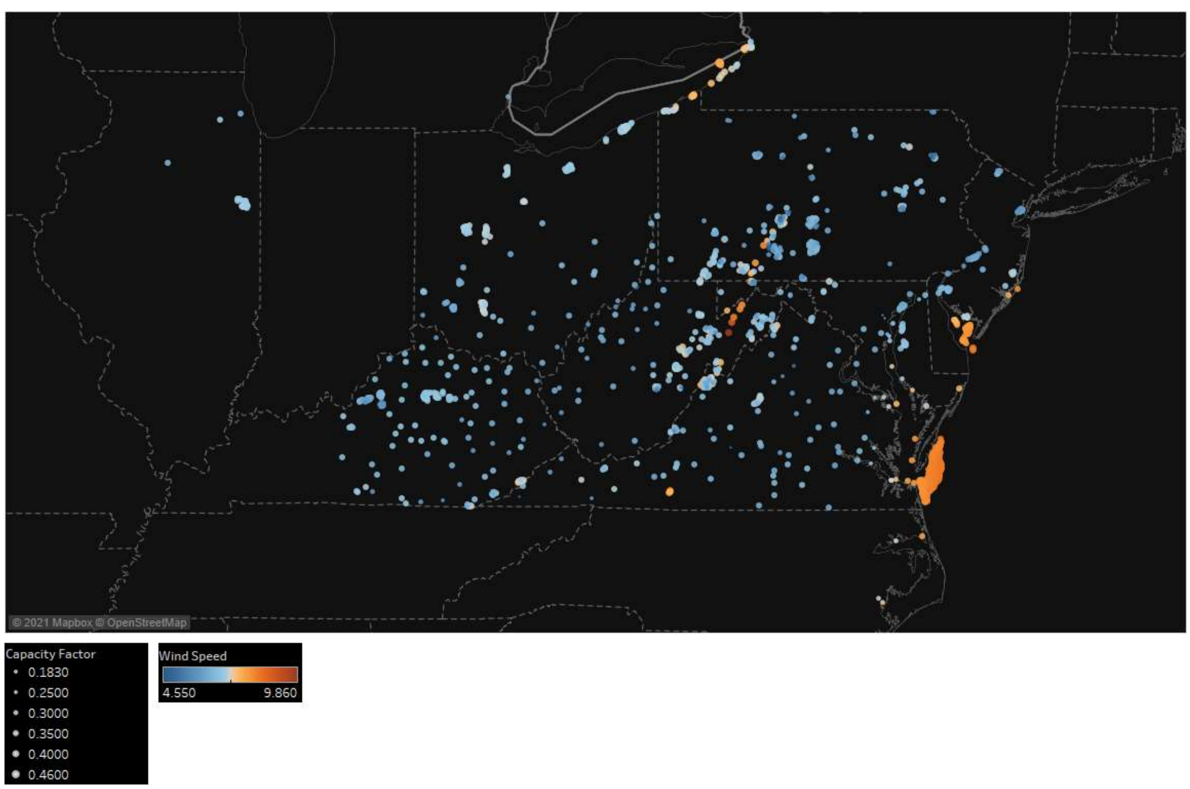
The following figure displays the aggregate capacity factor for each hour and month when selecting the 10%, 20%, 30%, 40%, and 50% of sites with the lowest coefficient of variation for the capacity factor. Figure 3.6 indicates that sites with more consistent wind are also the sites with the stronger average wind.

Figure 3.6: Aggregate wind capacity factor



Finally, Figure 3.7 displays the geographic location of the sites filtered using a 10 % threshold.

Figure 3.7: Wind sites with 10 percent threshold



Solar site selection

We select a subset of solar sites operating from 2012 to 2018. We restrict the sites to those with tracking panels that can achieve a more stable capacity factor. This solar technology can better reflect new solar investment.

Capacity factor distribution estimation, realization, and forecast error

We employ similar approaches as in the load model to estimate the distribution, generate realizations, and compute forecast errors. We use multiple quantile regression models to estimate capacity factor distribution. The dependent variable is the synthetic capacity factor, and the independent variables are the following:

- Month dummy variables
- Day-of-week dummy variables
- Hour dummy variable
- Interactions between month dummy variables and holiday dummy variables
- Interactions between day-of-week dummy variables and holiday dummy variables
- Holiday dummy variables (PJM 2020d)

We assume that the process is stationary. The capacity factor realizations are generated with the same methodology as the load model.

The forecast error is defined as the percentage difference between the actual and forecast capacity factor. The lookaheads are 0:15, 1:15, 2:15, …, 36:15-hour as in the load model. Assume that the forecast error is normally distributed with the mean of zero and the variance equal to that of the actual forecast errors. The data include PJM wind forecast errors between 3/1/2016 and 4/3/2020 and PJM solar forecast errors between 3/2/2018-4/2/2020. We draw the forecast errors of onshore and offshore wind from the same distribution.

Wind calibration

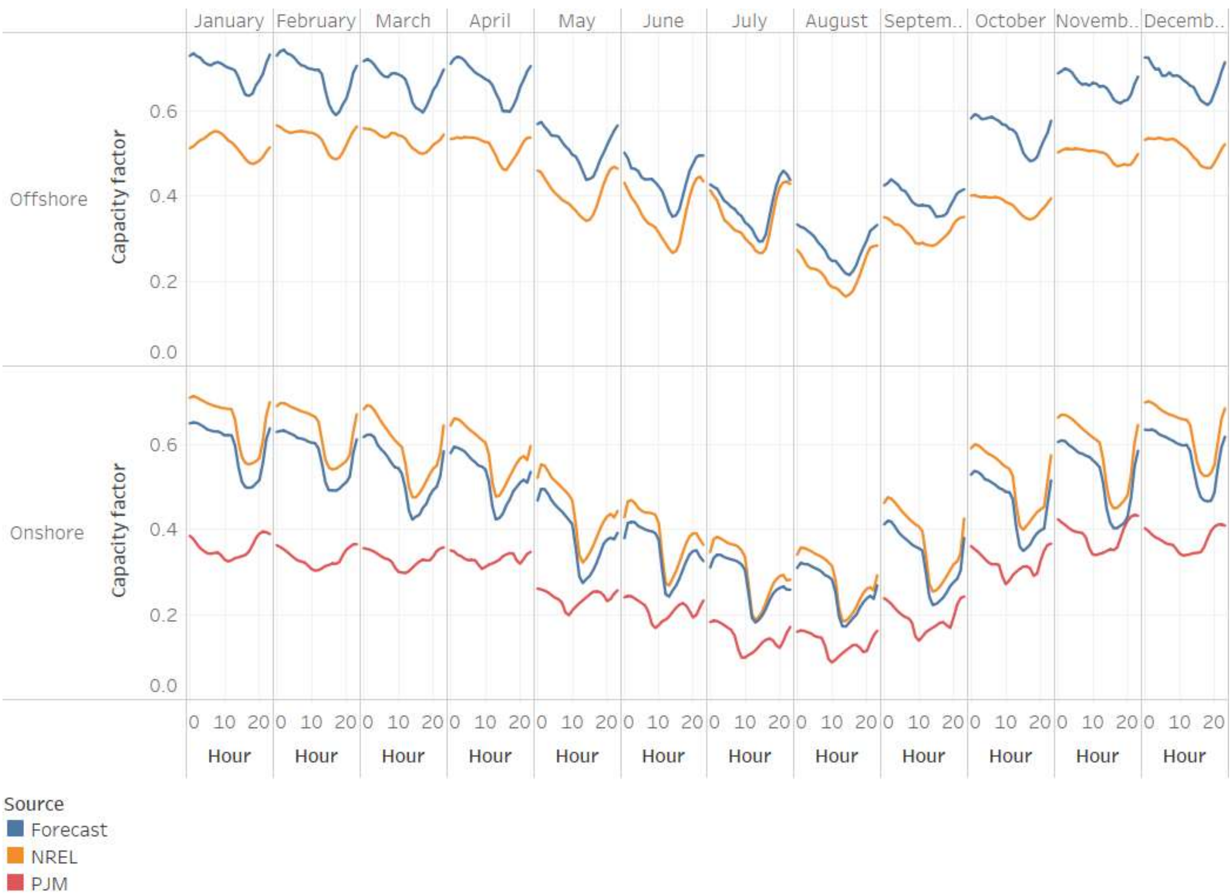
According to PJM, the capacity factor of the latest onshore turbines installed in its footprint is 45% during the year, on average, and 22% during the summer peak hours. For offshore turbines, these percentages are 50-60% and 30%, respectively. Thus, we have two targets per wind type. (We collapse the range 50-60% to a target of 55%.) We meet these targets by applying an affine transformation.[3]

For each wind type, the affine transformation is applied to all quantiles of all hours, days, and months from 2019 to 2113. The quantiles are connected by linear interpolation to form a continuous distribution. Finally, we restrict realizations to the unit interval.

[3] Specifically, for each wind type, we compute the median for each hour of each day from 2019 to 2113, compute the simple average and the overage for the summer peak hours, and use these averages together with the targets to form a system of two linear equations in two unknowns.

Figure 3.8 displays the average capacity factor by hour and month for the 10% of sites with the lowest variability ("NREL") and the rescaled process ("Forecast"). We show PJM's historical capacity factor, available for onshore wind only, for comparison.[4]

Figure 3.8: Comparison of average capacity factor



4 Other inputs to the energy market

Operating reserve demand curve (ORDC)

ORDC reflects the economic value of reserves above the minimum reserve requirement (MRR). Due to the uncertainty in forecasting net-load and generation dispatch, reserves may fall below the MRR in the real-time market. In these cases, the price is set equal to the administrative value of lost load, V. The Operating Reserve Demand Curve equals the expected value of lost load for a given level of reserves:

$$P(MRR + z) = V, \qquad if\ z < 0$$

$$P(MRR + z) = V[1 - F(z)], if\ z \geq 0$$

[4] The PJM data includes older plants, with a lower capacity factor. In the model, we impute a higher capacity factor to these plants, possibly delaying their exit and lowering energy rents for other generation units. In practice this simplification has negligible impact, because renewable resources are sponsored up to 2035 and because these older plants account for a relatively small capacity within PJM.

Here, z is the level of reserves above MRR, and F is the cumulative distribution function of the adjusted net-load forecast error. The adjusted net-load forecast error is defined as the realized net-load minus the forecasted net-load, accounting for outages of thermal units (a positive contribution) and resources available for regulation (a negative contribution). A higher level of reserves, $MRR + z$, reduces the probability that reserves fall below the MRR, $1 - F(z)$; the marginal value of reserves, P, drops as more reserves are procured ($1 - F(z)$ decreases).

The distribution of the net-load forecast error varies with the lead-time, the season, and the time of the day. Also, it depends on the resource structure. Higher penetration of intermittent resources or higher outage probability of thermal resources increase the forecast error's mean and variance.

Following PJM (2020), we initialize the value of lost load to $\$2000/MWh$, and set the minimum reserve requirement to $1450MW$, $2175MW$, and $3000MW$ for synchronized, quickstart, and supplemental reserves, respectively (PJM 2020c). The model allows for the value of lost load to increase over time. We use historical data for the years 2015-2017 to construct the distribution of the adjusted net-load forecast error by season $s \in S = \{summer, fall, winter, spring\}$, time-block $b \in B = \{23{:}00 - 2{:}00, 3{:}00 - 6{:}00, 7{:}00\ - 10{:}00, 11{:}00 - 14{:}00, 15{:}00 - 18{:}00, 19{:}00 - 22{:}00\}$ (hour-ending), and lead-time $f \in F = \{30 - minute, 60 - minute\}$, as follows.

We construct individual series at the 5-minute frequency for:

- The load forecast error, expressed in relative terms:

$$\varepsilon_t^{L,f} = \left(L_t - L_t^f - R_t\right)/L_{y(t)}^{P50}$$

where L_t is the realized load in interval t, L_t^f is the forecast at lead-time f, and $L_{y(t)}^{P50}$ is the median load in year $y(t)$; R_t is PJM's regulation requirement as reported in Table 4.1.

Table 4.1: PJM regulation requirement

Season	Dates	Non-Ramp Hours	Ramp Hours	Effective MW Requirement
Winter	Dec 1 – Feb 29	HE1 – HE4, HE10 – HE16	HE5 – HE9, HE17 – HE24	Non-Ramp = 525MW Ramp = 800MW
Spring	Mar 1 – May 31	HE1 – HE5, HE9 – HE17	HE6 – HE8, HE18 – HE24	Non-Ramp = 525MW Ramp = 800MW
Summer	Jun 1 – Aug 31	HE1 – HE5, HE15 – HE18	HE6 – HE14, HE19 – HE24	Non-Ramp = 525MW Ramp = 800MW
Fall	Sep 1 – Nov 30	HE1 – HE5, HE9 – HE17	HE6 – HE8, HE18 – HE24	Non-Ramp = 525MW Ramp = 800MW

- The intermittent generation forecast error, expressed in relative terms, for solar, offshore wind, onshore wind:

$$\varepsilon_t^{i,f} = CF_t^i - CF_t^{i,f}, for\ i \in I = \{Solar, Onshore, Offshore\}$$

- Ongoing outages started in the previous 30-minutes (or 60-minutes depending on the lead-time) as a percentage of each thermal technology installed capacity:

$$outage_t^{h,f}, for\ h \in H = \{Coal, Nuclear, CT, CC\}$$

Then, we compute the mean, $\mu^{n,\mathrm{f}}$, and standard deviation, $\sigma^{n,f}$, for each of these eight series, and their covariances, $cov^{n,m,x}$, for $n, m \in \{Load, Solar, Onshore, Offshore, Coal, Nuclear, CC, CT\}$.

We assume normality and use the expected median load and the installed capacity of intermittent and thermal resources (and, for intermittent, resources their average efficiency) to aggregate these eight series and obtain the distribution for the adjusted net-load forecast:

$$\varepsilon^{NL,f} \sim N\left(\mu^{NL,f}, \sigma^{NL\,,f}\right)$$

$$\mu^{NL,f} = \sum_{a \in L \cup I \cup H} c_y^a \times \mu^{a,f}$$

$$\sigma^{NL\,,f} = \left(\sum_{a \in L \cup I \cup H} \sum_{b \in L \cup I \cup H} c_y^a \times c_y^b \times cov^{a,b,f} \right)^{1/2}$$

with weights $c_y^i = -ICAP_y^i \times Efficiency_y^i$ for $i \in I$, $c_y^h = ICAP_y^h$ for $h \in H$, and $c_y^L = E(L_y^{P50})$.

As of 2019, PJM had no offshore wind farms in operation. We use the NREL data for the years 2007-2013. to impute the first and second-order moments as follows.

We compute the mean and standard deviation of the one hour ahead forecast error for offshore wind and impute corresponding statistics at 30-minute lead-time by assuming that the forecast error shrinks by the same factor as for onshore wind:

$$\mu^{Offshore,30} = \mu^{Offshore,60} \times \mu^{\,Onshore,30} / \mu^{Onshore,60}$$

$$\sigma^{Offshore,30} = \sigma^{Offshore,60} \times \sigma^{Onshore,30} / \sigma^{Onshore,60}$$

We use the NREL data to compute the covariance of the offshore one-hour ahead forecast error with the onshore one-hour ahead forecast error. Thus, we obtain $cov^{Offshore,Onshore,60}$.

We assume that the correlation at 30-minute lead-time is the same as at 60-minute lead-time:

$$cov^{Offshore,Onsho\quad,30} = \sigma^{Offshore,30} \times \sigma^{Onshore,30} \times corr^{Offshore,Onshore,60}$$

Finally, we assume that the correlations with the other six series (for load, the solar capacity factor, and thermal resource outages) at 30-minute and 60-minute lead times are the same as for onshore wind. Thus, we set:

$$cov^{Offshore,m,f} = \sigma^{Offshore,f} \times \sigma^{m,f} \times corr^{Onshore,m,f}$$

Tables 4.2 displays the results. The first two tables report the mean for each component of the net-load forecast error by season, time block lead-times equal to 30 and 60-minutes. The third table displays the variance-covariance matrices for the summer and winter seasons, time block 5, and 30-minute lead-time.

Table 4.2.a

Net-load forecast error components, mean, lead-time 30 minutes

Season	TBlock	Load	Solar	Onshore	Offshore	CC	CT	Nuclear	Coal
Summer	1	-7.69E-03	3.22E-03	-1.44E-02	6.39E-03	2.85E-04	5.02E-05	6.15E-05	9.81E-04
Summer	2	-7.21E-03	2.93E-03	-1.56E-02	2.02E-02	3.05E-04	5.87E-06	3.74E-06	5.84E-04
Summer	3	-9.55E-03	-3.06E-03	-1.77E-02	1.53E-02	5.34E-04	1.25E-04	7.49E-05	9.92E-04
Summer	4	-1.01E-02	-2.35E-02	-1.52E-02	3.86E-02	3.99E-04	3.88E-04	1.39E-04	1.41E-03
Summer	5	-5.69E-03	-4.13E-02	-1.83E-02	-2.43E-02	4.78E-04	4.67E-04	1.80E-04	1.36E-03
Summer	6	-8.71E-03	-6.99E-03	-2.01E-02	-5.17E-02	2.88E-04	1.46E-04	7.71E-05	1.02E-03
Fall	1	-7.71E-03	2.17E-03	-2.25E-02	-2.43E-02	3.59E-04	6.25E-05	5.93E-05	6.77E-04
Fall	2	-7.51E-03	2.13E-03	-2.38E-02	-1.42E-03	2.63E-04	6.14E-05	4.19E-05	5.56E-04
Fall	3	-7.11E-03	2.19E-03	-2.41E-02	8.39E-03	4.58E-04	2.94E-04	6.11E-05	7.76E-04
Fall	4	-6.44E-03	-2.56E-02	-1.92E-02	1.96E-02	3.69E-04	1.91E-04	5.80E-06	7.62E-04
Fall	5	-6.14E-03	-2.77E-02	-2.54E-02	1.60E-02	4.10E-04	2.63E-04	2.56E-05	7.15E-04
Fall	6	-9.37E-03	1.11E-04	-1.93E-02	-2.46E-02	2.37E-04	2.07E-04	6.73E-05	5.27E-04
Winter	1	-8.08E-03	1.45E-03	-4.15E-02	-3.23E-02	8.96E-04	7.89E-05	1.13E-04	9.42E-04
Winter	2	-8.54E-03	1.40E-03	-3.87E-02	-8.10E-03	7.07E-04	1.34E-04	1.39E-17	7.81E-04
Winter	3	-6.97E-03	6.43E-03	-3.62E-02	5.70E-03	8.18E-04	3.43E-04	9.11E-05	9.29E-04
Winter	4	-5.37E-03	1.34E-02	-3.69E-02	1.67E-02	3.99E-04	2.42E-04	6.76E-05	8.14E-04
Winter	5	-7.96E-03	3.69E-03	-4.40E-02	3.52E-02	5.28E-04	3.38E-04	7.72E-05	6.79E-04
Winter	6	-8.85E-03	1.60E-03	-3.93E-02	-2.17E-02	3.55E-04	1.86E-04	2.34E-05	7.34E-04
Spring	1	-8.58E-03	3.61E-03	-2.39E-02	-2.41E-02	3.04E-04	7.27E-05	7.08E-05	8.70E-04
Spring	2	-7.89E-03	3.61E-03	-2.71E-02	-7.06E-03	3.29E-04	1.35E-04	1.39E-17	5.89E-04
Spring	3	-6.48E-03	-2.92E-03	-2.76E-02	1.70E-02	5.67E-04	3.04E-04	1.59E-05	9.15E-04
Spring	4	-5.66E-03	-1.29E-02	-2.69E-02	4.20E-02	2.90E-04	2.55E-04	3.67E-06	8.83E-04
Spring	5	-6.59E-03	-2.69E-02	-2.77E-02	6.43E-03	3.33E-04	2.25E-04	5.35E-05	1.07E-03
Spring	6	-8.76E-03	1.45E-03	-2.94E-02	-3.10E-02	2.75E-04	2.05E-04	2.87E-05	8.31E-04

Note: series are rescaled as detailed in the text; the timeblocks 1-2-3-4-5-6 are 23:00-2:00, 3:00-6:00, 7:00-10:00, 11:00-14:00, 15:00-18:00, 19:00-22:00 (Hour Ending), respectively.

Table 4.2.b

Net-load forecast error components, mean, lead-time 60 minutes

Season	TBlock	Load	Solar	Onshore	Offshore	CC	CT	Nuclear	Coal
Summer	1	-7.45E-03	3.22E-03	-1.19E-02	5.31E-03	5.73E-04	1.04E-04	1.20E-04	2.01E-03
Summer	2	-7.11E-03	2.93E-03	-1.44E-02	1.85E-02	5.48E-04	8.39E-06	9.95E-06	1.15E-03
Summer	3	-9.96E-03	-6.41E-03	-1.88E-02	1.62E-02	1.09E-03	2.09E-04	1.37E-04	1.95E-03
Summer	4	-1.09E-02	-2.06E-02	-1.62E-02	4.09E-02	8.07E-04	7.71E-04	2.87E-04	2.82E-03
Summer	5	-6.05E-03	-5.34E-02	-1.80E-02	-2.40E-02	9.70E-04	9.49E-04	3.61E-04	2.73E-03
Summer	6	-7.97E-03	-1.25E-02	-2.22E-02	-5.71E-02	6.00E-04	3.22E-04	1.57E-04	2.03E-03
Fall	1	-7.69E-03	2.17E-03	-1.88E-02	-2.03E-02	7.08E-04	1.25E-04	1.30E-04	1.36E-03
Fall	2	-7.60E-03	2.13E-03	-2.10E-02	-1.26E-03	5.10E-04	1.03E-04	7.62E-05	1.10E-03
Fall	3	-6.98E-03	-3.07E-03	-2.17E-02	7.56E-03	9.31E-04	5.80E-04	1.30E-04	1.57E-03
Fall	4	-6.82E-03	-2.74E-02	-1.88E-02	1.92E-02	7.39E-04	3.92E-04	1.13E-05	1.46E-03
Fall	5	-5.98E-03	-4.25E-02	-2.63E-02	1.65E-02	8.03E-04	5.26E-04	5.06E-05	1.53E-03
Fall	6	-9.03E-03	-8.26E-04	-1.77E-02	-2.25E-02	5.02E-04	4.29E-04	1.24E-04	1.00E-03
Winter	1	-8.19E-03	1.45E-03	-3.88E-02	-3.03E-02	1.73E-03	1.57E-04	2.27E-04	1.93E-03
Winter	2	-8.73E-03	1.40E-03	-3.66E-02	-7.65E-03	1.40E-03	2.25E-04	2.43E-17	1.51E-03
Winter	3	-6.35E-03	1.69E-04	-3.38E-02	5.32E-03	1.66E-03	7.06E-04	1.67E-04	1.85E-03
Winter	4	-4.55E-03	1.41E-02	-3.56E-02	1.61E-02	8.06E-04	4.90E-04	1.50E-04	1.65E-03
Winter	5	-9.00E-03	-5.23E-03	-4.47E-02	3.57E-02	1.09E-03	6.32E-04	1.54E-04	1.38E-03
Winter	6	-8.01E-03	1.60E-03	-3.55E-02	-1.96E-02	7.19E-04	4.32E-04	4.54E-05	1.44E-03
Spring	1	-9.17E-03	3.61E-03	-2.29E-02	-2.30E-02	6.02E-04	1.50E-04	1.42E-04	1.74E-03
Spring	2	-8.45E-03	3.61E-03	-2.61E-02	-6.78E-03	6.65E-04	2.32E-04	2.09E-17	1.19E-03
Spring	3	-6.09E-03	-8.83E-03	-2.76E-02	1.70E-02	1.14E-03	6.22E-04	3.13E-05	1.74E-03
Spring	4	-5.19E-03	-9.39E-03	-2.68E-02	4.18E-02	6.02E-04	5.07E-04	7.97E-06	1.84E-03
Spring	5	-6.28E-03	-3.70E-02	-2.88E-02	6.68E-03	6.66E-04	4.56E-04	1.07E-04	2.12E-03
Spring	6	-8.51E-03	-6.10E-04	-3.22E-02	-3.40E-02	5.22E-04	4.25E-04	5.73E-05	1.69E-03

Note: series are rescaled as detailed in the text; the timeblocks 1-2-3-4-5-6 are 23:00-2:00, 3:00-6:00, 7:00-10:00, 11:00-14:00, 15:00-18:00, 19:00-22:00 (Hour Ending), respectively.

Table 4.2.c

Net-load forecast error, variance-covariance matrix (timeblock 5, lead-time 30 minutes)

		Load	Solar	Onshore	Offshore	CC	CT	Nuclear	Coal
Summer	**Load**	4.73E-05	1.16E-05	2.62E-06	5.41E-06	1.91E-07	2.28E-07	8.19E-07	-7.76E-07
	Solar	1.16E-05	1.76E-03	1.47E-05	3.03E-05	1.27E-06	1.24E-06	2.41E-06	2.14E-06
	Onshore	2.62E-06	1.47E-05	1.15E-03	5.17E-04	-7.72E-06	2.12E-06	-2.11E-06	-8.10E-06
	Offshore	5.41E-06	3.03E-05	5.17E-04	2.39E-03	-1.60E-05	4.39E-06	-4.37E-06	-1.68E-05
	CC	1.91E-07	1.27E-06	-7.72E-06	-1.60E-05	5.10E-06	1.92E-09	6.85E-07	2.23E-07
	CT	2.28E-07	1.24E-06	2.12E-06	4.39E-06	1.92E-09	3.60E-06	-8.40E-08	-2.67E-08
	Nuclear	8.19E-07	2.41E-06	-2.11E-06	-4.37E-06	6.85E-07	-8.40E-08	4.62E-06	-9.90E-08
	Coal	-7.76E-07	2.14E-06	-8.10E-06	-1.68E-05	2.23E-07	-2.67E-08	-9.90E-08	9.81E-06
Winter	**Load**	4.42E-05	-1.98E-05	4.25E-06	9.69E-06	9.64E-07	6.12E-07	-2.00E-07	3.18E-07
	Solar	-1.98E-05	4.85E-03	1.88E-04	4.30E-04	-5.80E-07	1.86E-06	8.46E-07	9.46E-07
	Onshore	4.25E-06	1.88E-04	1.58E-03	5.37E-04	6.82E-06	2.84E-06	9.65E-08	5.51E-07
	Offshore	9.69E-06	4.30E-04	5.37E-04	3.61E-03	1.56E-05	6.47E-06	2.20E-07	1.26E-06
	CC	9.64E-07	-5.80E-07	6.82E-06	1.56E-05	4.82E-06	2.62E-07	-4.07E-08	2.06E-07
	CT	6.12E-07	1.86E-06	2.84E-06	6.47E-06	2.62E-07	2.82E-06	-2.11E-08	5.96E-07
	Nuclear	-2.00E-07	8.46E-07	9.65E-08	2.20E-07	-4.07E-08	-2.11E-08	1.51E-06	1.89E-07
	Coal	3.18E-07	9.46E-07	5.51E-07	1.26E-06	2.06E-07	5.96E-07	1.89E-07	5.19E-06

Note: the series are rescaled as detailed in the text; timeblocks 5 = 15:00-18:00, Hour Ending.

Finally, table 4.3 displays the mean and variance for the adjusted net-load forecast error, obtained by applying the weights displayed in tables 4.2 to the 2019 resource mix (Coal = 62097*MW*, Nuclear = 32640*MW*, CT = 28821*MW*, CC = 49957*MW*, Solar = 1331*MW*, WindOn = 7012*MW*, WindOff = 0*MW*) and to an expected median load equal to 87,000*MW*. Figure 4.1 display the corresponding ORDCs for each of the three reserve types.

Table 4.3

Net-Load forecast error, Mean and Std by season and timeblock					
		30 min. ahead		60 min. ahead	
Season	**Timeblock**	**Mean**	**Std.**	**Mean**	**Std.**
Summer	**1**	-486	691	-401	804
Summer	**2**	-463	441	-415	511
Summer	**3**	-599	575	-531	765
Summer	**4**	-603	487	-553	719
Summer	**5**	-179	690	-66	1056
Summer	**6**	-514	654	-343	984
Fall	**1**	-445	493	-405	682
Fall	**2**	-432	447	-410	532
Fall	**3**	-364	578	-279	721
Fall	**4**	-313	443	-279	647
Fall	**5**	-240	486	-122	729
Fall	**6**	-618	702	-548	1077
Winter	**1**	-297	538	-216	700
Winter	**2**	-377	532	-326	685
Winter	**3**	-243	663	-86	978
Winter	**4**	-141	539	1	747
Winter	**5**	-300	681	-291	990
Winter	**6**	-418	595	-302	887
Spring	**1**	-502	441	-486	562
Spring	**2**	-436	460	-434	543
Spring	**3**	-266	625	-135	886
Spring	**4**	-204	424	-87	604
Spring	**5**	-245	513	-107	726
Spring	**6**	-477	747	-360	1183

Note: weights (load and the energy mix) are as detailed in the text; timeblocks 1-2-3-4-5-6 are 23:00-2:00, 3:00-6:00, 7:00-10:00, 11:00-14:00, 15:00-18:00, 19:00-22:00 (Hour Ending), respectively.

Figure 4.1a: Operating reserve demand curves by time block and reserve type (summer and winter)

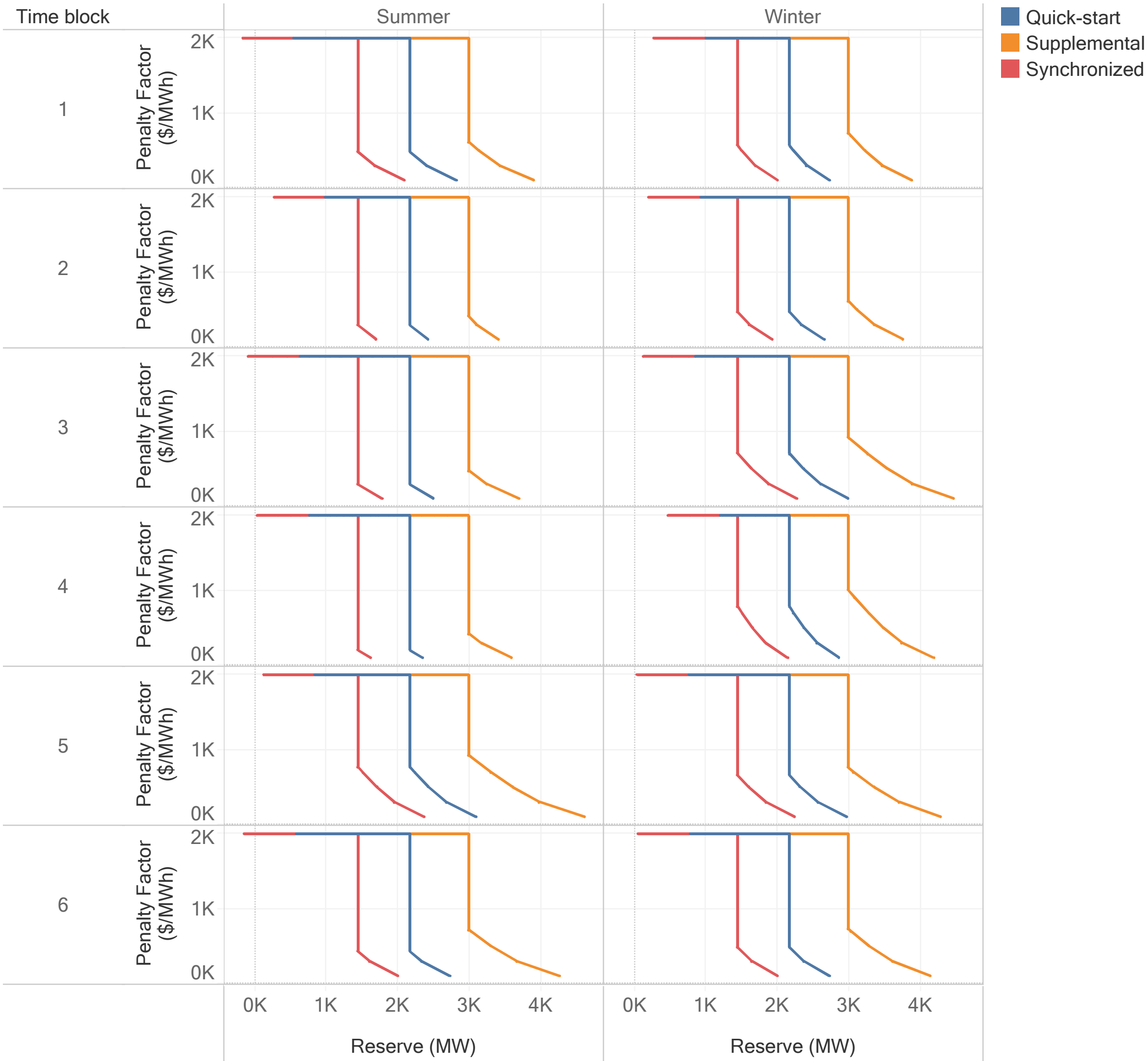
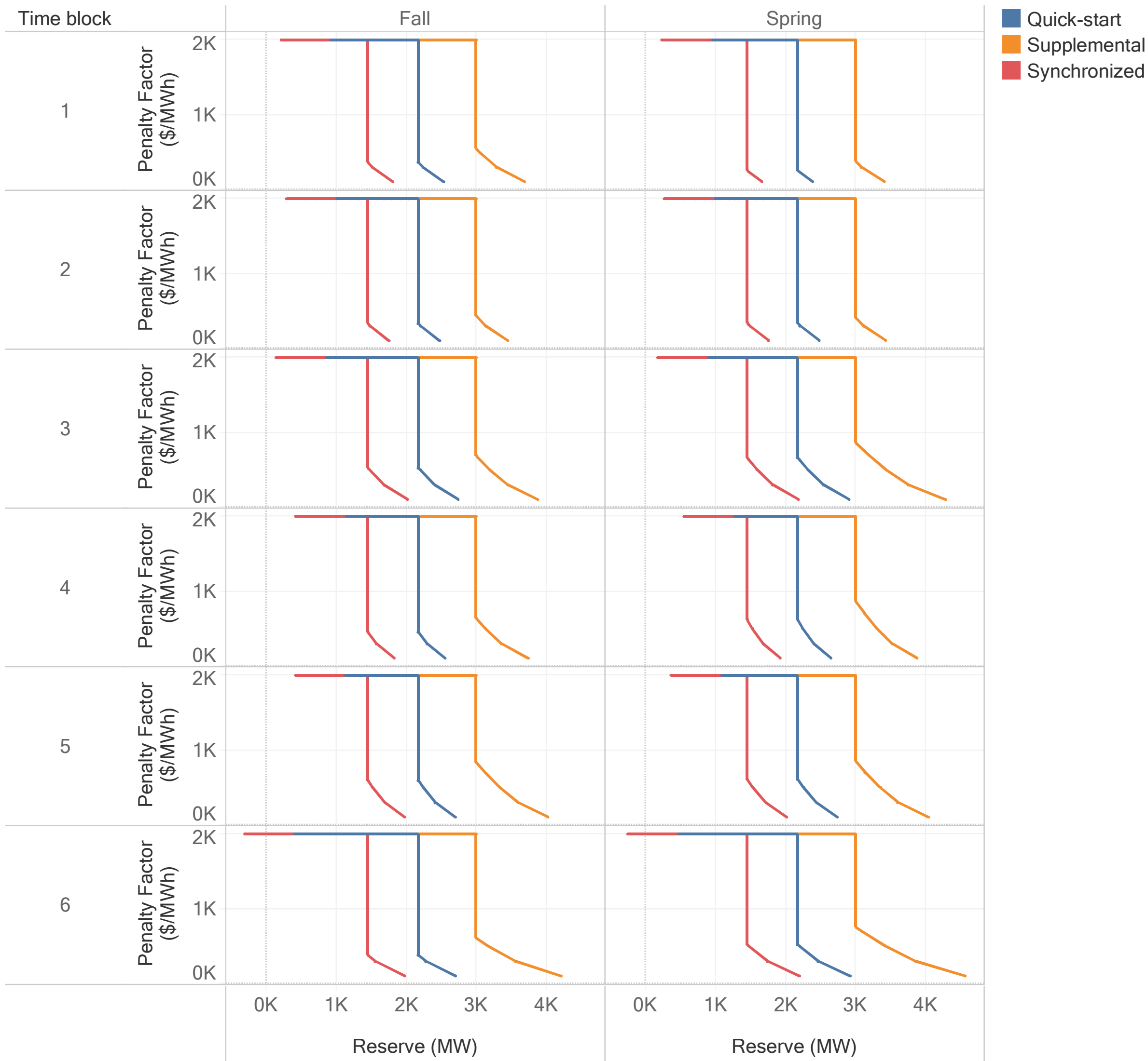


Figure 4.1b: Operating reserve demand curves by time block and reserve type (fall and spring)



Forced outages

We use data from PJM's Generator Availability Data System (GADS) for the most recent year, 2017, to estimate the outage probability for thermal units, the fraction of capacity affected by the outage, and its duration. We link GADS to the capacity market data and obtain technology class and nameplate capacity for each plant, and aggregate outages at the plant level, defined as in the Calibration subsection. The set of resources in the capacity market data extended to Fixed Resource Requirement entities defines the pool at risk.

Table 4.4 displays the 5-minute outage probability and percentiles for the size and duration of outages by technology type. Forced outages are frequent, generally affect a small fraction of nameplate capacity, except that for combustion turbines, and can last a very long time. The correlation between size and duration is modest, except possibly for combustion turbines and nuclear. This correlation needs to be precisely captured by the simulation because sizable and extended outages, though unlikely, have a significant cumulative effect on available supply.

Table 4.4: Forced outages, descriptive statistics

		CC	CT	Coal	Legacy	Nuclear
5-min. probability		3.04E-04	1.48E-04	6.79E-04	3.55E-04	1.78E-04
Fraction of percentiles	0.01	0.006	0.003	0.003	0.007	0.000
	0.05	0.021	0.022	0.007	0.015	0.001
	0.1	0.057	0.075	0.011	0.029	0.002
	0.25	0.124	0.117	0.024	0.093	0.007
	0.5	0.224	0.205	0.063	0.231	0.035
	0.75	0.390	0.397	0.198	0.488	0.222
	0.9	0.627	0.865	0.498	0.964	0.467
	0.95	0.943	0.952	0.630	1.000	0.505
	0.99	1.000	1.000	1.000	1.000	0.907
Duration, percentiles (min.)	0.01	5	4	8	11	12
	0.05	15	12	30	35	27
	0.1	24	24	50	60	55
	0.25	60	62	122	144	130
	0.5	189	209	385	494	571
	0.75	611	991	1275	1620	2282
	0.9	2484	5657	4317	3901	7200
	0.95	5828	25229	7780	7321	13240
	0.99	55975	263520	36157	34039	263520
Pearson correlation		-0.10	-0.17	0.05	-0.05	-0.09
Spearman correlation		0.01	-0.10	0.03	0.04	-0.20

We model forced outages in two stages. For each plant, we simulate the arrival of outages every five minutes using a Poisson process with a constant arrival rate, as in Table 4.4. Then, we draw the size and duration of each outage from the technology-specific empirical joint distribution.[5] We apply to combined cycle with carbon capture and storage and new nuclear the same empirical models for traditional combined cycle and nuclear.

Finally, we rescale the Poisson arrival rate by 0.7 to match the available installed capacity's aggregate time series, displayed in Figure 4.2.

Planned Outages

We link the planned outage requirement data, in weeks, to the capacity market data and assign to each unit a technology type. Then, we compute the empirical distribution of outage duration by technology type, reported in Table 4.5.

[5] Specifically, we first draw the size of the outage from the empirical marginal distribution. Then, we draw the duration of the outage from the empirical distribution conditional on outage size. We use a uniform grid with 50x50 points for combined cycle, combustion turbines, and legacy, 100x100 for coal, and 25x25 for nuclear.

Table 4.5: planned outages, empirical distribution function

Weeks	CC	CT	Coal	Legacy	Nuclear
1	0.037	0.420	0.038	0.194	0.161
2	0.100	0.606	0.170	0.323	0.484
3	0.253	0.800	0.302	0.532	0.710
4	0.526	0.905	0.443	0.581	0.903
5	0.732	0.940	0.585	0.694	0.968
6	0.868	0.963	0.670	0.774	1.000
7	0.942	0.977	0.849	0.839	1.000
8	0.974	0.991	0.915	0.855	1.000
9	1.000	0.995	0.934	0.887	1.000
10	1.000	0.998	0.972	0.903	1.000
11	1.000	1.000	0.972	0.919	1.000
12	1.000	1.000	0.981	0.919	1.000
13	1.000	1.000	0.981	0.952	1.000
14	1.000	1.000	0.991	0.952	1.000
15	1.000	1.000	0.991	0.952	1.000
16	1.000	1.000	1.000	0.952	1.000
17	1.000	1.000	1.000	0.968	1.000
18	1.000	1.000	1.000	1.000	1.000

Given a list of plants and their characteristics, we simulate planned outages as follows. For each plant, we draw the planned outage duration from the technology-specific empirical distribution. Then, we use a greedy algorithm to allocate the outage to the consecutive weeks with the lowest expected peak net-load. Specifically, we select the largest plant; take the square of the expected weekly peak net-load plus the plant's nameplate capacity; compute the rolling average, and pick the week with the lowest score as the starting week of the outage. Then, we adjust the expected weekly peak net-load series to account for the first plant's outage, consider the second-largest plant, and so on.

We initialize the expected peak net-load series using the following approximation. We take the median of the gross load distribution for each hour of the year and the median of the normalized capacity factor distribution for solar, onshore wind, and offshore wind (see section Load and Renewable Models). We combine these four series using the ICAP and efficiency of renewables in the resource pool and obtain an approximate series for the hourly expected net-load.[6] Finally, for each week, we define the expected peak net-load as the max of this series.[7]

Energy offer curves

We use offers in the day-ahead energy market for 2018 to 2020 and the capacity market data to assign the technology class to each observation. A unit generally submits different cost-based offers and possibly one price-based offer for each hour. We select price-based offers where available and one cost-based

[6] This is an approximation. The median generally differs from the mean; also, we ignore the correlations between gross load and solar and wind generation.

[7] Here we make another simplification, assuming that the max of the expectation is the expectation of the max.

offer otherwise. Among cost-based offers, we choose those with primary fuel consistent with the technology class (e.g., natural gas for combined cycle) and the higher schedule ID, if many are available.

We normalize the quantity offered at a given price by eco-max and discretize offers into buckets using a uniform grid with a step size equal to half a percentage point. If the quantity on the last segment is below eco-max, the price on this segment is extended uniformly to eco-max, consistently with PJM rules. Finally, offers can be piecewise linear or piecewise constant and are treated accordingly. For example, a piecewise constant offer of $20/MWh for 60% of capacity and $25/MWh for 20% becomes $20/MWh for the first 120 buckets and $25/MWh for the remaining 80 buckets. Thus, we obtain a balanced panel and for each month and bucket compute the simple overage across observations. The energy offer curves are stable across months, except for January, when fuel costs are higher. Accordingly, in the model, we distinguish between January and other months. Figure 4.3 displays the resulting energy offer curves. Similarly, we normalize the no-load cost by eco-max and average it. Table 4.6 reports the no-load costs.

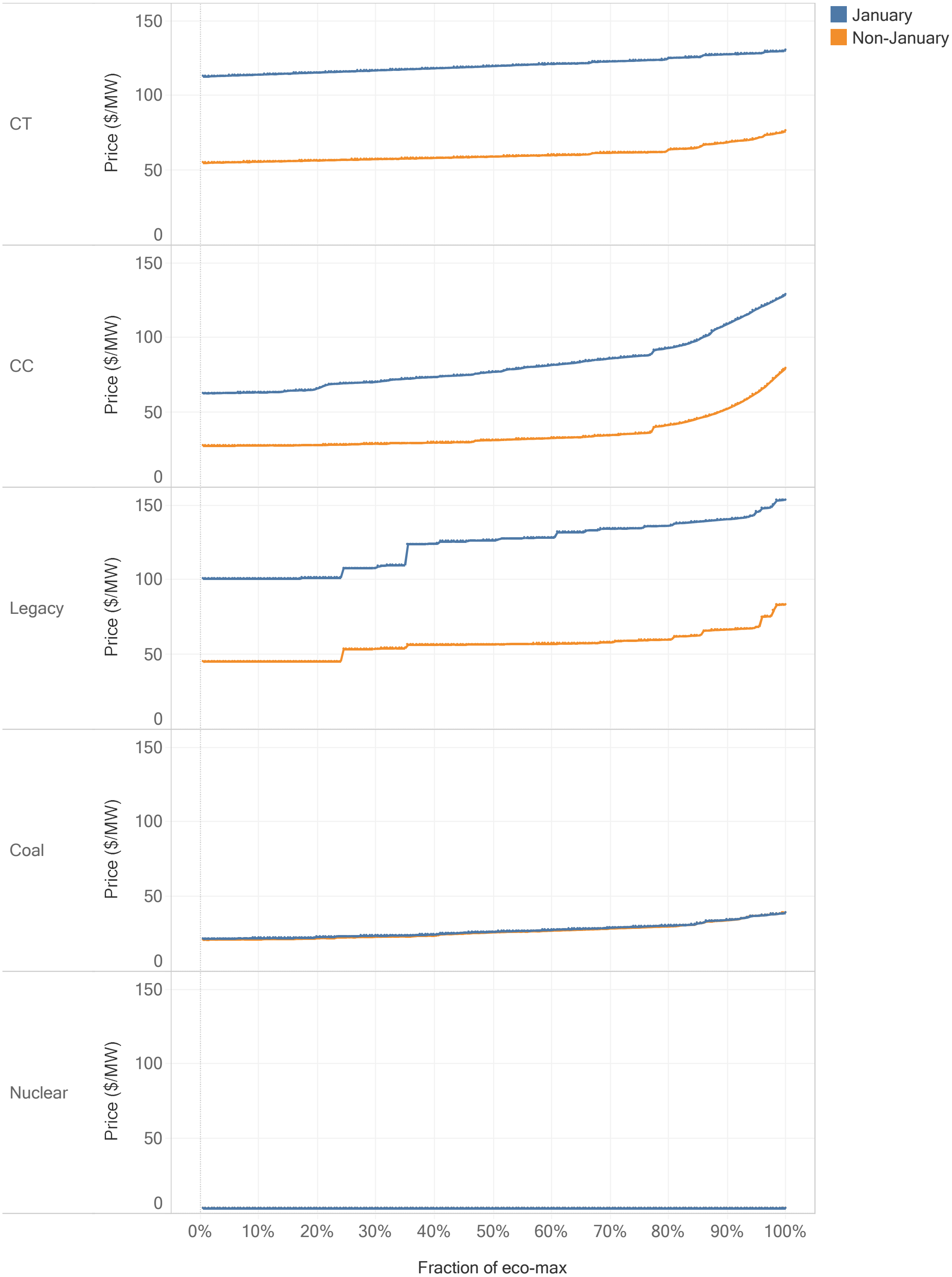
Different plants in different years have varying marginal costs depending on the technology class, vintage, fuel prices, and carbon price. We construct offer curves for the model by multiplying the marginal cost times a technology and month-specific coefficient, and the corresponding empirical offer curve re-indexed at one. These coefficients and no-load costs are reported in Table 4.6 and set to match observed total cost curves given the 2019 state of technology, fuel prices, and carbon price.[8]

Table 4.6: Marginal cost multipliers and no-load costs

	Marginal cost multiplier		No-load
	non-January	January	cost
Coal	0.86	0.89	3.41
Nuclear	0.34	0.35	0.28
NewNuclear	0.34	0.35	0.28
Combustion turbines	1.95	4.01	10.17
Combined cycle	1.33	3.03	3.94
Combine cycle-ccs	1.33	3.03	3.94
Legacy	2.17	4.85	4.71

[8] Total cost curves are equal to no-load cost plus the integral of the energy offer curve.

Figure 4.3: Empirical energy offer curves technology and month



Notes: authors' calculations on PJM day-ahead energy market data, 2018-2020

5 Energy market model

We make three simplifying assumptions. First, transmission is assumed to be sufficiently robust so that transmission constraints never bind. Ample transmission is reasonable in a long-run model since transmission can be built to eliminate persistent constraints. Second, bids in both the capacity and energy markets are assumed to be competitive except that in the energy market, some resources bid above marginal cost near capacity, consistent with observed behavior. See Section 4 for details on the derivation of these offer curves. Our bidding assumption better replicates the "hockey stick" bidding observed in practice. Third, a carbon price path is assumed. A carbon price is a transparent way to model either an explicit carbon price or an implicit carbon price that results from climate policy, whether a carbon tax, cap-and-trade, or other regulation.

Energy profits—the profits earned in the energy and reserves markets over the year—are determined by explicitly modeling the day-ahead and real-time energy and reserves markets. The model of the wholesale markets is as accurate as possible. We model every unit at the five-minute level. Units, including battery storage, are optimally scheduled based on expectations and dispatched in real-time. Both the day-ahead and real-time markets are modeled. Shortage and near-shortage conditions typically are observed only in real-time. The day-ahead market is a financial market that optimizes unit commitments for each of the 24 hours in the next day based upon the day-ahead forecast. During the day, the forecast is revised, resulting in changes in unit commitments. In real-time, the optimal dispatch is determined based on the available resources and demand realizations. Both day-ahead and real-time markets co-optimize energy and reserves to maximize as-bid social welfare.

The greatest computational challenge of our approach is computing the energy profits over the year for a fixed set of resources (the resource structure) and the associated parameters of supply and demand. The day-ahead market is a large mixed-integer program (MIP) that determines a tentative schedule for resources and prices for the next day. The problem is hard because of non-convexities coming primarily from startup costs and times and physical limits in output and ramping. During the day, there are shocks to supply and demand from outages or weather that require adjustments throughout the day. These shocks require running the MIP every hour to adjust schedules in response to new information. Real-time dispatch is a linear program (LP), which dispatches online resources.

Determining the market outcome of a single day requires solving many hard problems. To compute the energy profits over the year, we must do this for many days as electricity demand and supply vary considerably throughout the year. We do this detailed calculation for ten days of each month from which we extract one representative week per month. We use the first two days of the ten to prime the market and the last to avoid end-game effects. This yields a direct calculation of a representative week in each month. From this, we calculate energy profits and other market outcomes for the year, given a fixed set of resources and other inputs.

Using 22 high-end computational servers and a state-of-the-art solver (Gurobi 9.0), we can compute about 1000 annual energy market runs in 24 hours. Unfortunately, to determine optimal entry and exit decisions over many decades based on consistent expectations, we need about 80 million calls to the energy market model. Direct computation of energy profits would take 219 years to compute a single multi-decade scenario with our 22 dedicated servers. An alternative approach is needed.

Further details of the energy market model

Our energy market simulation has three core components:

- A mixed-integer program (MIP) schedules each unit throughout the day.
- A linear program (LP) dispatches energy and reserves of online units ten minutes ahead.
- A market state tracks the status of every unit in real-time.

During a single market simulation, we use each of these pieces repeatedly in a cyclic fashion. An overview of the core loop is shown in Figure 5.1, where box (A) corresponds to the market state, box (B) corresponds to the LP, and box (C) corresponds to the MIP.

Figure 5.1: Core energy market simulation loop

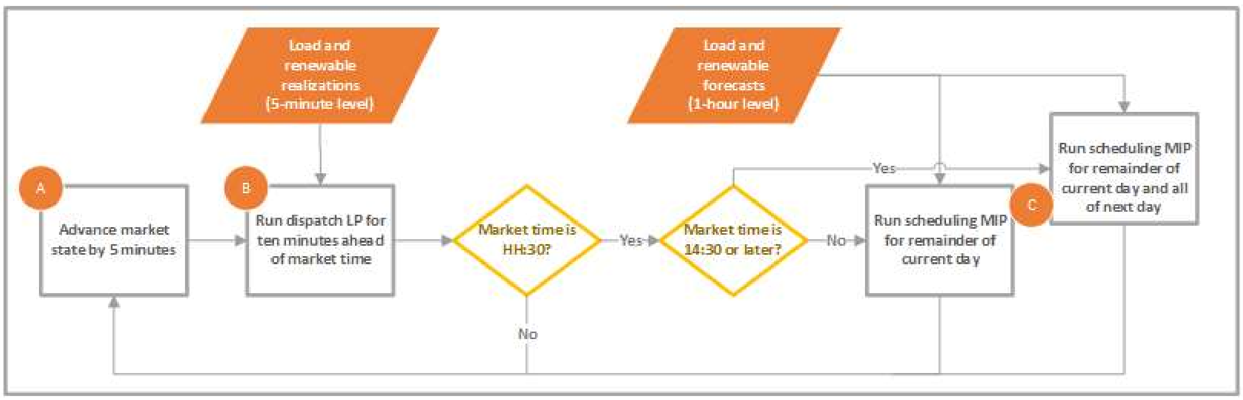


Figure 5.1 captures the overall cadence of our simulation process. The dispatch LP is run every five minutes, while the scheduling MIP is run once an hour on the half-hour. The scheduling MIP includes all remaining hours in the current day; starting at 14:30, the MIP is expanded to include all hours of the following day. Once all necessary LP and MIP runs have been performed for a particular five-minute timepoint, the market state is advanced by five minutes. Updating the market state includes several types of tasks: processing events such as outages, processing instructions to start up and shut down, and determining how units react to dispatch instructions. The following subsection discusses each of the three main components in more detail.

All three components make critical contributions to determining market outcomes. The scheduling MIP first includes a given day at 14:30 the prior day. This first inclusion plays a special role: it determines the day-ahead prices and unit commitments for the energy market. The scheduling MIP will be re-run once an hour thereafter and will include hours from this day until 22:30 on the day itself. These subsequent runs have no price impact and do not create financial commitments. Their role is to adjust unit schedules to best serve the market's needs under current conditions and information.

On the other hand, the dispatch LP determines how the currently online units can be best used to meet realized demand with a ten-minute lead time. For example, when dispatch is run at 10:25, it optimizes unit output in the window 10:35-10:45. Each run of the dispatch LP determines the real-time prices for the first five minutes of the period it optimizes. Finally, the market state determines how individual units behave, notably how they react to market instructions. While all units make a best effort to follow these instructions, events like outages may interfere. Thus, it is the market state which determines real-time quantities.

Market outcomes are the result of all three components. The revenue of units in the market is determined via a two-stage settlement process. Units are first compensated for their day-ahead commitments at day-ahead prices. Units are then paid or charged for their real-time deviation from these day-ahead commitments at real-time prices.

Figure 5.2: Simplified unit-commitment MIP

Maximize:

$$\sum_t \mathrm{Benefit}_t(PriceResp_t) + \sum_t \mathrm{ORDC}_t(Reserve_t) - \sum_t \sum_g \left(\mathrm{Cost}_g(gen_{g,t}) + \mathrm{StartCost}_g \cdot start_g\right)$$

Subject to:

$$NetLoad_t + PriceResp_t = \sum_g generate_{g,t} + \sum_s (discharge_{s,t} - charge_{s,t}) \quad \forall t \quad (1)$$

$$generate_{g,t} \geq \mathrm{MinGen}_g \cdot on_{g,t} \quad \forall t, \forall g \quad (2)$$

$$generate_{g,t} \leq \mathrm{MaxGen}_g \cdot on_{g,t} \quad \forall t, \forall g \quad (3)$$

$$generate_{g,t} \geq generate_{g,t-1} - 60 \cdot \mathrm{Ramp}_g \quad \forall t, \forall g \quad (4)$$

$$generate_{g,t} \leq generate_{g,t-1} + 60 \cdot \mathrm{Ramp}_g \quad \forall t, \forall g \quad (5)$$

$$start_{g,t} \geq on_{g,t} - on_{g,t-1} \quad \forall t, \forall g \quad (6)$$

$$on_{g,t} \geq \sum_{\Delta=0}^{\mathrm{MinOn}_g} start_{g,t-\Delta} \quad \forall t, \forall g \quad (7)$$

$$on_{g,t} \leq \sum_{\Delta=0}^{\mathrm{MinOff}_g} \left(1 - start_{g,t+\Delta}\right) \quad \forall t, \forall g \quad (8)$$

$$charge_{s,t} \leq \mathrm{MaxCharge}_s \quad \forall t, \forall s \quad (9)$$

$$discharge_{s,t} \leq \mathrm{MaxDischarge}_s \quad \forall t, \forall s \quad (10)$$

$$store_{s,t} \leq \mathrm{MaxStore}_s \quad \forall t, \forall s \quad (11)$$

$$store_{s,t} - store_{s,t-1} = \mathrm{Eff}_s \cdot charge_{s,t} - discharge_{s,t} \quad \forall t, \forall s \quad (12)$$

$$Reserve_t = \sum_g reserve_{g,t} + \sum_s reserve_{s,t} \quad \forall t \quad (13)$$

$$reserve_{g,t} \leq 10 \cdot \mathrm{Ramp}_g \cdot on_{g,t} \quad \forall t, \forall g \quad (14)$$

$$reserve_{g,t} + generate_{g,t} \leq \mathrm{MaxGen}_g \quad \forall t, \forall g \quad (15)$$

$$reserve_{s,t} + discharge_{s,t} \leq \mathrm{MaxDischarge}_s + charge_{s,t} \quad \forall t, \forall s \quad (16)$$

$$reserve_{s,t} \leq store_{s,t} \quad \forall t, \forall s \quad (17)$$

$$generation, charge, discharge, store, reserve \quad \geq \ 0 \quad (18)$$

$$on, start \quad \in \{0,1\} \quad (19)$$

Unit commitment MIP

The MIP used to schedule units is based on a standard unit-commitment MIP framework. In our market, the MIP plays two roles: in addition to performing day-ahead unit commitment, it also provides intra-day updates to unit schedules to address changing circumstances and information. Note that we only ever run a single MIP at a given time, and thus the MIP will serve both purposes simultaneously. At 14:30, a single MIP is run to perform intra-day scheduling for the remainder of the current day and determine prices and commitments for the following day. Figure 5.2 gives a simplified formulation of the scheduling MIP.

The MIP maximizes as-bid social welfare, the benefit from satisfying price-responsive demand and the operating reserve demand curve (ORDC) less the cost incurred to meet load. Constraint (1) captures our most important requirement: clearing the market. Demand is the realized load less any renewable production and demand reduction from price; supply is generation output plus storage discharge less storage charging. Supply and demand balance for every time interval.

Most of the remaining constraints are focused on ensuring that units' operational parameters are respected. Constraints (2)-(3) ensure that the output of generation units lies within their operating range and is zero when they are offline. Constraints (4)-(5) ensure that units' ramp rates are respected. Constraints (7)-(8) enforce any requirements units have for the minimum time between consecutive online or offline periods. Constraints (9)-(10) ensure that storage units' charge and discharge rates remain within their operating range. Constraints (11)-(12) ensure that storage units' stored energy levels are kept consistent and within their operating range from one time to the next.

Finally, several constraints control how we procure reserves. Constraint 13 aggregates the reserves procured from both generation and storage units. Constraints (14)-(15) ensure that generation units can provide reserves if called upon without exceeding their operating parameters. Constraints (16)-(17) ensure that storage units can provide reserves when called. In constraint (17), a storage unit can provide reserves is by reducing a charge operation. For brevity, our presentation here shows a single class of reserves. Our simulations include three tiers of reserves consistent with the PJM market.

Dispatch LP

The LP used to dispatch units is a simplification of the MIP presented in Figure 5.2. The derivation involves two primary changes: (1) the program is simplified to include only a single ten-minute period, and (2)units' operating states are taken to match the current market state, fixing the corresponding binary decision variables. Note that change (2) is precisely why our problem is an LP rather than a MIP.

While the dispatch LP is a simplified version of the scheduling MIP, it does introduce a new complication in how storage units are handled. Recall that, since the MIP spans multiple periods, it naturally captures the ability of storage units to time-shift demand. Since the LP covers only a single period, we must adjust how storage units are handled. We address this issue by having storage units bid in costs to the LP; consistency is achieved by basing these costs on the shadow price for stored energy in the scheduling MIP. More details are provided in a later subsection.

Market state

The overall energy market state must be carefully maintained and updated as the simulation runs. When we advance the time in the market from one dispatch run to the next, we must carry out multiple processes correctly and harmoniously. Specifically, we must:

- process all startup and shutdown instructions produced when scheduling generation units,
- ramp all generation units according to current state and best-effort attempts to follow dispatch,
- apply any forced outages that occur for the appropriate duration, and
- update the stored energy of every storage unit.

Although performing the above requires care, it does not require computational sophistication on the same level as the other components; it amounts to bookkeeping.

Payments

As previously mentioned, payments are calculated using a two-stage settlement. Day-ahead prices and quantities are derived from the scheduling MIP. The first time the MIP includes a given day determines the unit commitment, fixing the day-ahead prices and quantities for settlement. Real-time prices and quantities are determined by the dispatch LP and units' best efforts to follow dispatch instructions, respectively. Finally, uplift is introduced in the form of make-whole payments. Any unit which fails to honor its day-ahead commitment due to a forced outage is disqualified from receiving an uplift payment. We give additional details on each of these points below.

The derivation of prices from the scheduling MIP follows standard practice. Once we run our optimization, we fix all binary decision variables at their optimal values. This reduces our MIP to an LP, allowing us to compute shadow prices for each constraint. The market-clearing constraint (Constraint 1) provides the energy prices, while the reserve aggregation constraint (Constraint 13) provides the prices for reserves. The derivation of real-time prices is analogous, except that dispatch already simplifies the optimization from a MIP to an LP to compute shadow prices directly.

While day-ahead quantities come directly from the scheduling MIP's variables, real-time quantities require more effort. The dispatch LP contains the relevant variables, but there are two complicating issues: first, the dispatch LP lacks any variables for offline reserves; second, units may not always perfectly follow the instructions from the dispatch LP for power output. Reserves are dealt with as follows: online reserve quantities come from the dispatch LP, while offline reserve quantities come from the most recent run of the scheduling MIP. Double charging is explicitly prevented. Reserve quantities are capped at maximum output less any realized power output.

Real-time quantities for power are more complicated. While units make a best effort to follow dispatch instructions, it is not always possible to follow them exactly. This is not a flaw of the model but rather a fundamental aspect of the market. Specifically, dispatch chooses a target output for each unit ten minutes into the future. Thus, every *five* minutes, dispatch effectively tells each unit how to ramp for the next *ten* minutes. Every five-minute period is subject to two different ramping instructions. These instructions may be incompatible—for example, a sudden, dramatic change in the load forecast could result in a unit being asked to ramp both up and down. In our model, units handle conflicting instructions by focusing on the output targets. At any time, each unit has a current output and two targets: five minutes out and ten minutes out. The unit's next five minutes of ramping will be chosen to minimize its combined deviation from these two output targets, with ties broken in favor of minimizing the maximum deviation from either target. Payments are based on these realized outputs rather than the (potentially infeasible) dispatched outputs.

Outages and shortages

Our market model incorporates generation unit outages—both planned and forced. We also introduce robustness to outages into the model by allowing for emergency-start of combustion turbine (CT) units in response to prices. This last decision allows the market to be more robust to unexpected demand spikes, whether caused by outages or load forecasting errors.

Planned outages

Planned outages are incorporated into our model as follows. Each unit has an associated annual planned outage requirement scheduled to occur sometime outside the peak months. Planned outages have durations that are a multiple of one week in our model. Since each run of the energy market is split into single-week blocks, we check for each block which units had a planned outage during that week. These units are then kept offline for the entirety of the simulation for the relevant block.

Forced outages

Forced outages are incorporated into our model as follows. Each forced outage has a randomly drawn start time, duration, and output reduction factor. Until a forced outage begins, all components of our energy market simulation remain oblivious to the outage. In particular, the scheduling MIP is not provided information about it. The outage affects the relevant unit's operation in two ways: first, the unit's maximum and minimum output are both scaled down by the output reduction factor; second, if the unit is offline and has a reduced *maximum* output below its *minimum* output, it cannot be started until the outage completes. When the outage concludes, all operational parameters are returned to normal.

Forced outages impact both dispatch and scheduling procedures. Outages can result in a unit having output that is no longer in its feasible range. Whenever this happens, the unit is forced to ignore any dispatch instructions and ramp as quickly as possible to return its output to the correct range. Once this has been done, the unit returns to standard dispatch control. Both the scheduling MIP and dispatch LP are adjusted to be consistent with the unit's reduced output range for the outage duration. More precisely, the adjustments to the MIP and LP are made for the duration of the outage, but in the case of the MIP, the adjustment assumes the outage will continue indefinitely.

Emergency-start combustion turbine units

One challenge in our model is that the market can be slow to respond when additional units are needed to meet load. Specifically, the only method for starting units in the core model is via the scheduling MIP. The MIP, however, is run only once per hour, with a half-hour lead time. Thus, on average, we can expect it to take almost an hour for additional units to be started in response to new information. This delay often leaves us unable to handle sudden, unexpected spikes in demand adequately.

To address this, we introduce an emergency-start capability into the model for combustion turbine units. Whenever the real-time price exceeds a specified threshold, we consider all combustion turbine units and select the one that would require the fewest minutes to become profitable at the current price. We limit the number of emergency starts to one per dispatch period, one every five minutes. Note that while the emergency-start process is manual, we leave stopping the units up to future runs of the scheduling MIP.

Modeling storage

Storage technologies differ from generation technologies at a fundamental level. They face completely different costs when providing power: generation units face reasonably well-defined costs as they convert fuel to power, whereas storage units face a stochastic opportunity cost. This cost is determined in the scheduling MIP: while the objective function includes multiple terms to capture the costs of generation units, there are no related terms for storage units. Opportunity costs (and benefits) for storage units are instead implicit in the objective function. Maximizing social welfare naturally means storage units should time-shift supply to meet demand; they should charge when demand is low and discharge when demand is high.

In the scheduling MIP, Constraints (11)-(12) couple a storage unit's decisions to charge and discharge across periods and ensure that prior charges can account for any discharges. The scheduling MIP captures tradeoffs storage units face across a full day, as welfare is maximized when storage units charge when energy costs (and hence prices) are low and discharge when energy costs (and hence prices) are high. While this addresses how storage units behave from one hour to another during a single day, it does not address how they should be dispatched, nor how much stored energy should carry from one day to the next. Both issues must be dealt with for storage to behave reasonably within our energy market simulation. Assigning an explicit price to stored energy proves to solve both issues, as we will describe.

Storage behavior in real-time

In real-time, we capture storage unit tradeoffs for charging or discharging by bidding a cost function into the dispatch LP. This approach is quite different from the scheduling MIP, where such cost functions are only present for generation units. We derive cost functions for storage units from the shadow price corresponding to constraint (12) in the MIP for the current period. Since this constraint is responsible for maintaining the unit's current energy level from one time to the next, the shadow price naturally corresponds to the unit's marginal value on the stored energy. While this provides a starting point for our cost functions, there are some subtle but essential issues using this hourly marginal value in a real-time context.

Since the scheduling MIP deals with full hours, the shadow price corresponds to an *average hourly* price. Thus, if we apply it as a sharp threshold to the real-time price, we very well may buy or sell much less than expected. For example, suppose the average price in an hour is just slightly over the shadow price. In that case, minor variations in the real-time price from one dispatch period to another make it probable the unit transacts only half as often as the MIP would suggest. We alleviate this issue by smoothing out the threshold in our cost function.

Storage behavior across multiple days

Another challenge for modeling storage is ensuring good end-of-day behavior. As stated, the scheduling MIP places no inherent value on a storage unit carrying energy over from one day to the next. If left unaddressed, this consistently leads to a price crash in the final period of every MIP as storage units liquidate "worthless" stored energy. A specific price crash self-corrects as soon as the MIP is extended to include more periods, but the overall problem is merely deferred as the new final period suffers the same problem. To avoid such inconsistent and irrational behavior, we must ensure the MIP explicitly considers the value storage units place on having stored energy as time continues past what is included in the MIP.

Our approach centers around finding, for each storage unit, a continuation value to capture what each stored MWh of energy will be worth at the start of the next day. Summarizing the continuation value with a single number in this way is a significant simplification. As we store more and more energy, the marginal benefit from each MWh should decrease, not remain constant. The issue runs even deeper as other units' energy stored should also affect this marginal benefit. This interdependence is quite complex, as it will be informed by unit-specific characteristics such as energy storage limits. We do not address this complexity. Instead, we settle for ensuring that the numbers we choose for each unit are consistent and representative of a plausible market state.

Our process for calculating continuation values is straightforward. The first time we need to optimize a given day, we first run an optimization for the following day. We assume that the set of online generation units will remain the same into this following day but remove any restriction on their initial power output levels. For storage units, we replace our usual start-of-day constraints on energy levels with a new constraint. Each storage unit must end the day with at least as much stored energy as it started. Note that this constraint will always be tight since storage units realize no benefit from retaining energy at the end of the day. After optimizing the MIP, we compute the associated pricing LP and consider the shadow price associated with our new constraint. These shadow prices capture the marginal value for more energy *at an optimal stored energy level*, and so they should provide a reasonable approximation of optimal behavior. Furthermore, since they are derived from a specific MIP solution, they will be consistent with each other.

Once calculated, we introduce the continuation values into the MIP as follows. We add a new term to the objective, which sums each storage unit's stored energy in the final period weighted by that unit's continuation value. Once the optimization has been run, we make two changes to the computation of the associated pricing LP: first, we drop this new objective term; and second, we fix the end-of-day energy level for each storage unit, just as we fix binary decision variables. The pricing LP is then optimized as usual.

We make two additional observations. First, it would be natural to ask why, once we have computed optimal end-of-day energy levels for each unit, we bother with shadow prices rather than simply enforcing these end-of-day energy levels as a hard constraint. We do so to avoid a range of feasibility issues. Since real-time dispatch outcomes can vary from the behavior suggested by the scheduling MIP, such an end-of-day constraint can go from feasible to infeasible between one run of the MIP and the next, especially as the end of the day grows closer. Second, this deviation between real-time and scheduled outcomes raises an issue for continuation values: how are they incorporated into real-time behavior? To this point, we note that the end-of-day continuation values will directly influence the shadow prices used for real-time costs, especially near the end of the day. As such, we can be assured that real-time bidding will also reasonably value end-of-day stored energy.

6 Multi-year simulation

Environment

Time is discrete, and the frequency is annual. Technology, demand, input prices, the carbon price, the fraction of price responsive load, and the ORDC vary exogenously over time, making some plants unprofitable and opening opportunities for new plants to enter the electricity market. Investors make optimal entry and exit decisions based on expected profitability in the energy, ancillary service, and capacity markets. These decisions drive the evolution of resource structure—the set of existing and planned resources.

Every year the grid operator runs an auction to procure capacity *h* years forward, with a commitment period of one year. The administrative demand curve depends on the forecasted load and the annualized cost of entry for the marginal technology, net of energy offsets, and adjusted for the capacity value. Energy offsets are a function of past profits in the energy and ancillary service markets, henceforth "energy profits" and "energy market," respectively. The capacity value is a function of performance, defined as the ratio of generation and reserves to nameplate capacity during near-scarcity events. In the model, we use double exponential smoothing.

Entry and exit decisions are announced *h* years in advance in the capacity auction and are irreversible.[9] All existing and planned resources offer their capacity competitively in the auction. A plant's capacity in the auction is equal to nameplate times the capacity value. The equilibrium capacity price clears supply and demand. Thus, the capacity price, entry decisions, and exit decisions are codetermined in the capacity market. We ignore resales of capacity obligations.

A new plant embodies the latest technology vintage. Technical characteristics remain constant during the life of the plant. Investors finance the construction of new plants with equity and borrowed money. They pay the annualized cost of entry every year, starting with construction until the end of the repayment period or decommissioning, whichever occurs first. The investor also pays fixed operation and maintenance costs which depend on the age of the plant.

The model accounts for 15 technology types: coal, nuclear, combustion turbines (CT), combined cycle with and without carbon capture and sequestration (CC and CC-ccs respectively), gas-legacy (henceforth legacy), onshore wind, offshore wind, solar, pump-storage (also referred to as hydro), and battery storage with durations of 1, 2, 4, and 8 hours.

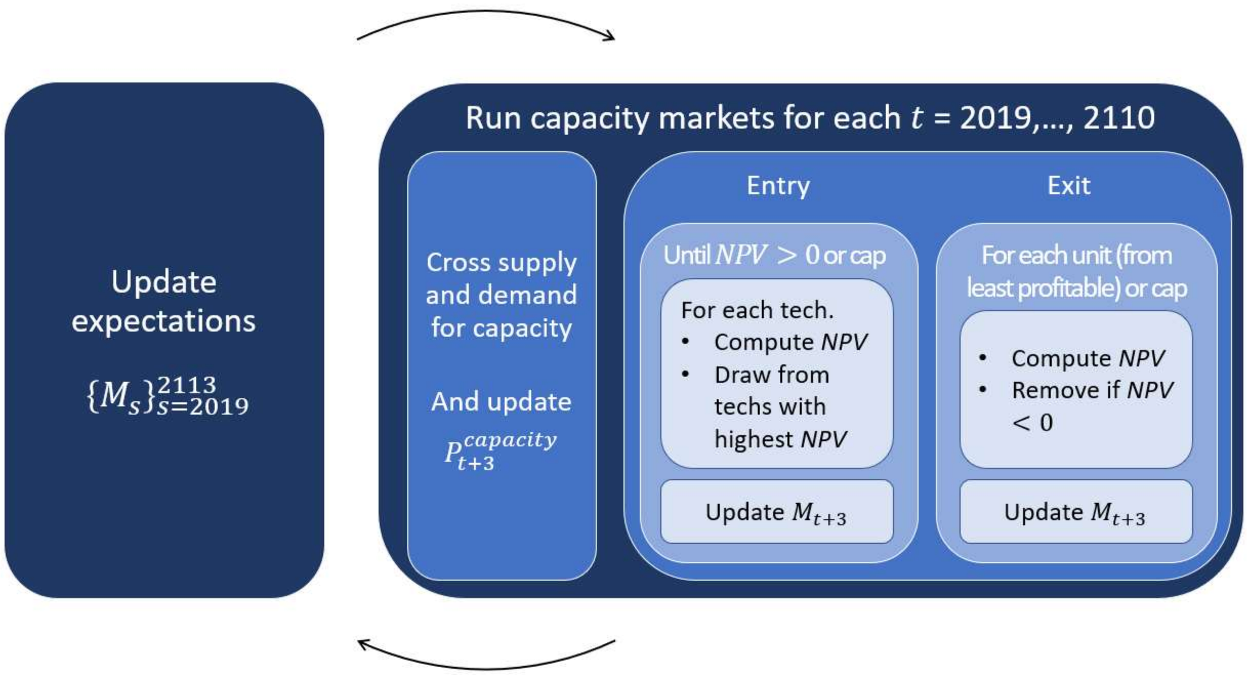
The model is sketched in Figure 6.1 and consists of three nested fixed-point problems. *i*) Starting with the innermost fixed-point, investors make entry and exit decisions consistent with the capacity price in the current auction and expectations on the evolution of the market state. *ii*) The equilibrium capacity price balances the supply resulting from the entry and exit decisions with demand. *iii*) Expectations are consistent with the evolution of market structure following the sequence of capacity auctions and entry and exit decisions every year.

Finally, we use the multi-year simulation results to optimize the double exponential parameters for updating energy offsets and capacity values by minimizing the mean squared forecast error four years

[9] In the empirical implementation we set *h* equal to three years, the planning horizon in PJM's capacity auction. We disregard the time for development, permitting and engineering when calibrating the model. These activities have modest financial costs and do not bear significantly on the irreversibility of entry decisions. The plant construction time is three years or shorter for all representative technologies except that for traditional nuclear which has a construction time of four years (EIA 2020b). We assume that investors postpone construction so that the plant goes online exactly after three years if the plant construction time is shorter. Setting the construction time for traditional nuclear to three years biases results in favor of this technology. In practice, we do not observe entry of traditional nuclear in the model. We also consider next generation nuclear which has an anticipated construction time lower than three years due to simpler design and modularity.

ahead.[10] We run the multi-year simulation repeatedly until the optimized weights are sufficiently close to those used to run the latest simulation.

Figure 6.1: Multi-year simulation



Notes: M_t is the market state in year t.

In the remainder of this section, we introduce the notation, characterize expectations for the evolution of the electricity market, and how they affect the net present value of existing and planned resources. Then, we provide details on the mechanics of the capacity market, where exit and entry decisions are made based on net present values, and where the capacity price is determined. Next, we discuss how to accommodate states' sponsored resources and minimum offer price floors in the capacity market. Finally, we outline the model calibration strategy and describe the side scenarios. The formal description of the model and details on data sources and methods are in the Mathematical Annex and the Data Annex, respectively.

Notation

The market state M_t in year *t* consists of the resource structure S_t, the frontier technologies available to new entrants J_{t+1} the following year, the vector of fuel prices P_t^F, the carbon price P_t^O, the ORDC penalty parameter $\overline{P}_t$, load D_t, the fraction of price responsive load PRD_t, capacity prices $\left\{P_{t+k}^C\right\}_{k=0}^h$, and the capacity market parameters B_{t+1}. Each existing plant and new project *i* in S_t has a capacity market state $G_{i,t}$. It consists of the plant's capacity obligations for years *t* to *t+h* and a dummy. This dummy equals one if the plant has cleared at some previous date and zero otherwise. The set of capacity market

[10] The capacity market planning horizon is three years, but PJM sets energy offsets and capacity values before the capacity auction when profits and performance are know up to the previous year.

parameters B_t includes expected peak load, capacity values, energy offsets, minimum offer price floors for each technology, and reference technology. We use $Z_t = \{J_t, P_t^F, P_t^C, P_t^O, \overline{P}_t, D_t, PRD_t, \{P_{t+k}^C\}_{k=0}^h, B_{t+1}\}$ to denote the set of aggregate quantities other than the list of resources S_t and write $M_t = \{S_t, Z_t\}$.

Expectations and net present value

Exit and entry decisions basis on net present value calculations (NPV). The net present value depends on the stream of future energy profits and capacity payments. Capacity payments are proportional to size. We assume that costs and energy revenues are also proportional to size so that power plants have constant returns to scale. The assumption that energy revenues are linear in size is reasonable because PJM's electricity market is large relative to the minimum efficient scale of any generation and storage technology.

With these assumptions, entry and exit decisions depend on the asset's net present value per MW of nameplate capacity, henceforth NPV. We use $\pi^{j,v}(M_t, g_t)$ to denote the profits per MW earned by a plant operating vintage *v* of technology *j* with capacity obligations g_t in market state M_t:

$$\pi^{j,v}\left(M_t, g_t\right) = EP^{j,v}\left(M_t\right) + 365 \times g_t \times cv_t^j \times P_t^C$$

EP denotes energy profits per MW and *cv* the capacity value. Capacity is a daily product, hence the multiplication by 365. The capacity value is set at the technology level consistent with PJM's practice, but the model can easily accommodate capacity values set at the plant level.

We make three simplifying assumptions making the model computationally tractable.

1. Investors form approximate expectations on the future set of resources. We refer to this simplified representation as to the "resource mix." The resource mix is a relatively small set of statistics useful for forecasting energy profits. It contains, for example, the total installed capacity and the mean vintage for each technology class. We use $\tilde{S}$ and $\tilde{M} = \left\{\tilde{S}, Z\right\}$ to denote the resource mix and the "approximate market state."
2. Investors use a simple statistical model mapping the approximate market state into expected energy profits and performance. We refer to this model as to the "energy proxy model"—discussed below in detail.
3. Investors ignore the approximate nature of the energy proxy model and perfectly anticipate the path of exogenous quantities.

These assumptions make the model computationally tractable. Also, assumptions 1 and 2 arguably reflect heuristics used by investors to forecast profitability in practice. As discussed below, the model can accommodate fundamental uncertainty, but we ignore it in this work and focus on perfect foresight equilibria.

Using these assumptions, we construct investors' expectations as follows. The capacity market is run at the beginning of the year when the state of the market is M_{t-1}.

- Investors know the exogenous statistical process governing technology, fuel prices, the carbon price, the ORDC penalty parameter, load, and price responsive demand and make exact predictions for these quantities.
- They use retirement and entry decisions announced up to *t-1* to form expectations for the resource mix in years *t* through *t-1+h*. The capacity price is also known up to year *t-1+h*. As discussed below, entry and exit decisions and the capacity price for year *t+h* are codetermined. Thus, when firms make exit decisions and entry decisions in year *t*, they know the resource mix and capacity price for year *t+h*. For the years after that, expectations are updated by simple exponential smoothing using the resource mix and capacity price from the latest run of the multi-year simulation. We run the multi-year simulation repeatedly until these expectations are reasonably accurate.
- Investors use expectations for the resource mix and exogenous quantities and the energy proxy model to predict energy profits and performance in future years and update expectations for the capacity auction parameters *B* dynamically.[11]

This completes the characterization of investors' expectations for all the quantities in the reduced market state $\tilde{M}$.

- Finally, the investor can use expectations for the reduced market state to predict the plant's state in future capacity auctions *G* (consisting of the capacity obligations sold in the auction and the dummy indicating if the plant has cleared in the capacity market at some previous date).

The net present value per MW of nameplate capacity for plant *i* going into the capacity market is then readily computed by backward induction using the pair of functional equations:

$$
\begin{aligned}
X^{j,v,a,l}\left(\tilde{M}_{t-1}, G_{i,t-1}\right) &= \sum_{k=0}^{l-1} \delta^k \left[\tilde{\pi}^{j,v}(\tilde{M}_{t+k}, g_{i,t+k}) - C^{j,v,a+k}\right] + \delta^l \phi^{j,v,a+l}, \text{ for } l \le h \\
V^{j,v,a,l}\left(\tilde{M}_{t-1}, G_{i,t-1}\right) &= \max\left\langle X^{j,v,a,h}\left(\tilde{M}_{t-1}, G_{i,t-1}\right),\right. \\
&\qquad \left. \tilde{\pi}^{j,v}(\tilde{M}_t, g_{i,t}) - C^{j,v,a} + \delta V^{j,v,a+1,l-1}\left(\tilde{M}_t, G_{i,t}\right)\right\rangle, \text{ for } l > h
\end{aligned}
$$

a is the age of the plant, *l* the residual life, δ the discount factor, $\tilde{\pi}$ the energy proxy model for energy profits, *C* is the annualized cost of entry plus fixed operation and maintenance costs, ϕ is the scrap value. If the residual life of the plant is less than or equal to the planning horizon *h*, the plant operates for another *l* years and then exits. If it is greater than *h*, the investor can retire the plant at *t+h* or run it for another year and revisit the "exit-or-stay" decision in *t+1*. If the investor announces the plant's retirement, the residual life is shortened to *h*. This decision is irreversible.

[11] Suppose we are in year *t* and need to formulate expectations on energy offsets, capacity values, the net-cost of entry and the net going forward cost for year *t+k* capacity auction (with delivery year *t+k+h*). We use expectations for market conditions in *t+k-1* to run the energy proxy model and obtain energy profits, and performance for that year; then, we update energy offsets and capacity values for year *t+k* capacity auction. Finally, using these quantities and expectations for technology we can estimate the net cost of entry and the net going forward cost. We start with *k=0* and build expectations iterating forward.

Capacity Market and Exit and entry Decisions

Investors enter the capacity market with expectations for the evolution of the market state and behave competitively.

A capacity price is announced, starting at some value $\ddot{P}^{C}_{t+h}$—the "dots" indicate that the price is provisional. We select an existing plant at random and compute the net present value by value function iteration. If the exit branch has a higher value than the continuation branch, the investor announces the plant's retirement in year *t+h*, and we update the expectation for the reduced market state $\tilde{M}_{t+h}$. We repeat these steps until all existing plants are selected or the number of retiring MW is above a threshold—a model parameter.

Then, we consider entry decisions. We draw the project's nameplate capacity from the corresponding distribution for each technology, update expectations for the resource mix *h* years ahead, and compute the project's net present value. We draw the new plant from the set of *n* projects with the highest positive NPV, using the square of NPVs as weights, and update the reduced market state $\tilde{M}_{t+h}$. We repeat these steps until all technologies have negative NPV or the planned MW are above a threshold, or the ratio of capacity to peak load exceeds a threshold. These thresholds and *n* are also parameters of the model.

We continue alternating between existing and prospective units until no plant wants to revise its decision. When this happens, we check the capacity offered in the market against the capacity demand curve. If the two curves cross at a price within a critical range of the announced price, $\ddot{P}^{C}_{t+h}$, the algorithm stops. Otherwise, we restore the list of energy resources to their initial configuration and update the provisional price using bisection.

Minimum Offer Price Rules

With a minimum offer price rule (MOPR), resources may not clear in the capacity market and still operate in the energy market. For new plants and resources that have never cleared in the capacity market, the price floor is equal to Net-CONE. For other resources, it is equal to Net-ACR. Net-CONE is the annualized entry cost, plus fixed operation and maintenance cost, minus energy offsets divided by the capacity value. Net-ACR is the fixed operation and maintenance cost, minus energy offsets divided by the capacity value.

The expectation routine remains fundamentally unchanged with MOPR. Investors anticipate changes in technology, fuel prices, carbon price, ORDC penalty parameter, load, and price responsive demand and have expectations for the resource mix. They use the energy proxy model to estimate future energy profits and performance and formulate predictions for energy offsets and capacity values. Then, they update expectations for offer floors. Given expectations for capacity prices and MOPR, investors can assess when their assets will clear in the capacity market and the capacity payments they will receive.

Next, consider the capacity market. A price is announced. An existing unit is selected at random. The investor checks if the price is above Net-ACR or Net-CONE if the resource had not cleared in previous auctions. If the price is above MOPR, the unit receives capacity payments in *t+h*, and its status changes to "cleared" if it had not cleared in previous auctions. Then, the investor updates the NPV and decides whether to retire or keep the plant in the market. The procedure for potential entrants is similar except that the status is "not cleared," and the MOPR is equal to Net-CONE.

With MOPR, the system operator ignores the capacity of units not clearing in the auction, even if they plan on operating in the energy market at *t+h*. Thus, the system operator over-procures capacity.

Modeling fundamental uncertainty

The model can accommodate randomness in technology, prices, and demand, with no changes, if the error term is uncorrelated over time. The model equilibrium remains unaffected because investors are forward-looking and risk-neutral.

Fundamental uncertainty—for example, on the implementation of climate change policies—is of great importance for the energy transition because assets have a long capital recovery period. The model can accommodate this type of uncertainty by limiting uncertainty to a narrow set of alternative scenarios. For example, we could assume that a federal carbon tax will be implemented in 2024, 2028, or 2032 and take a low, medium, or high trajectory. Then, the investor would face nine scenarios (three starting dates times three paths). These nine scenarios would correspond to nine equilibrium trajectories for the reduced market state $\tilde{M}$. Investors would weigh these nine scenarios when forming expectations for NPV. The market equilibrium would consist of a fixed-point problem for these nine sequences for the expected market state.

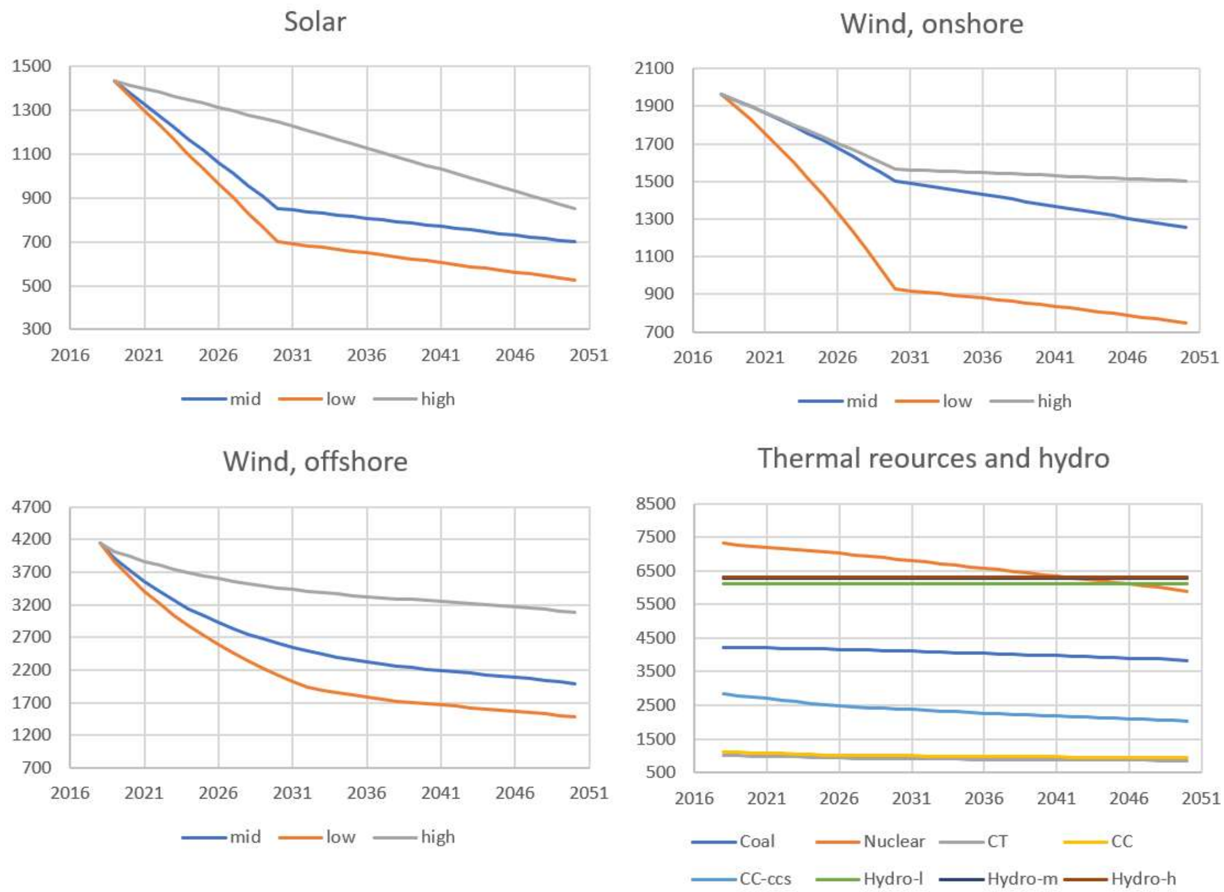
Model Calibration

We simulate the model from 2019 through 2040. We extend the simulation horizon to 2094 to mitigate end-of-game effects.

The representative technologies for thermal resources are described in EIA (2020b) and for renewables and storage in NREL (2020b). For entry costs (CAPEX), fixed operation and maintenance costs (FOMC), variable operation and maintenance costs (VOMC), and heat rates, we use projections up to 2050 from NREL (2020b). These series correspond to those underlying the 2020 Annual Energy Outlook (EIA 2020a) for thermal technologies. For legacy, we use the same values as for combined cycle. We inflate prices from 2018 to 2019 dollars using the consumer price index for the United States. Startup costs are significant for combustion turbines. We use EIA (2020b) cost breakdown to exclude this cost component from VOMC for combustion turbines. NREL (2020b) also reports capacity factor series for renewables. We use these series to construct technical efficiency indexes for renewable resources of different vintages. Then, we set a renewable plant generation equal to the baseline capacity factor—defined as in section 3—times the plant's technical efficiency and nameplate capacity.

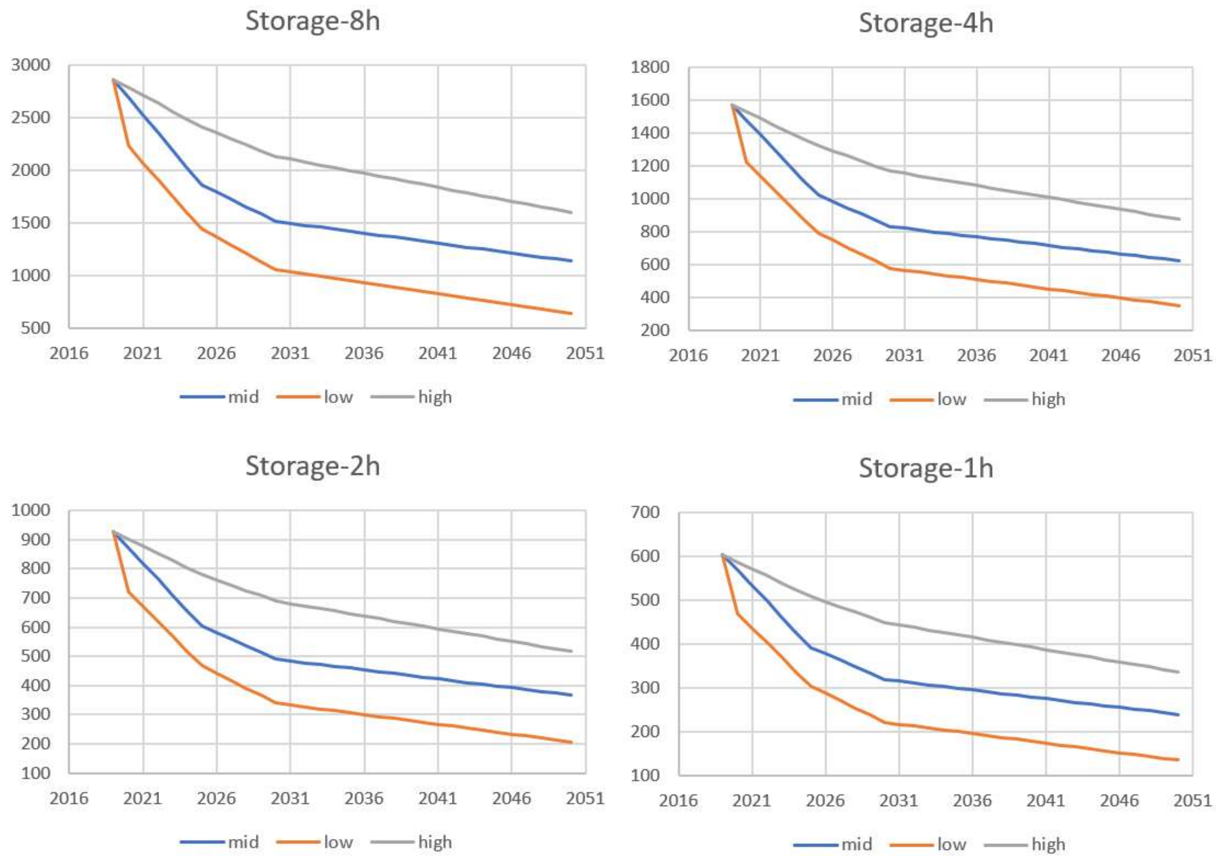
Figure 6.2-3-4 display the series for CAPEX, FOMC, and efficiency used in the model. Following NREL (2020b), the fixed operation and maintenance cost for battery storage is set equal to 2.5% CAPEX—not reported. Due to technical progress, these series move significantly over time, especially for renewable technologies. Table 6.1 reports values for VOMC and heat rates. Heat rates change modestly, while VOMCs are constant.

Figure 6.2.a: Assumptions for CAPEX



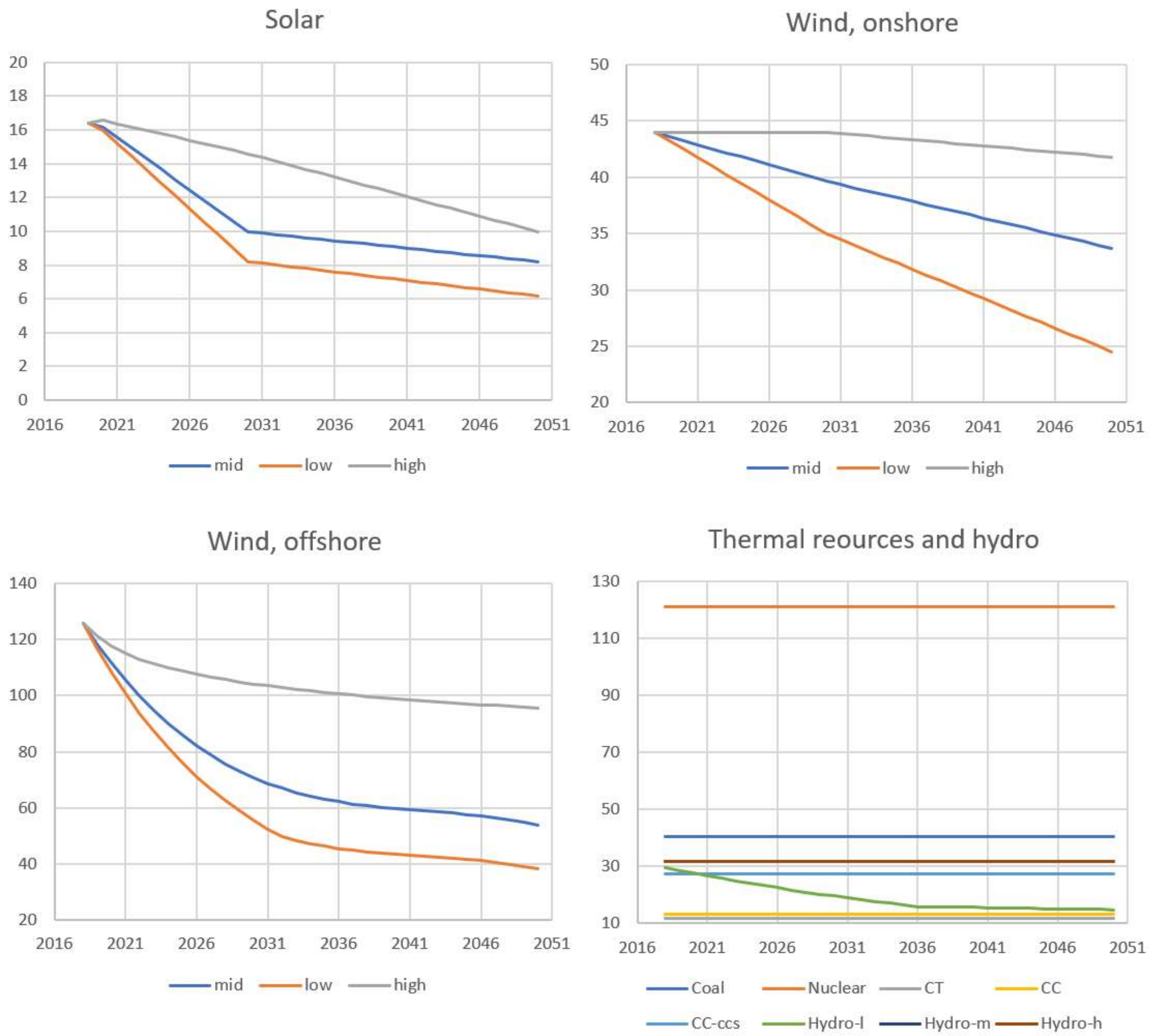
Notes: CAPEX is inflated to 2019 $/kW-year using the consumer price index for the United States. Source, NREL (2020b).

Figure 6.2.b: Assumptions for CAPEX



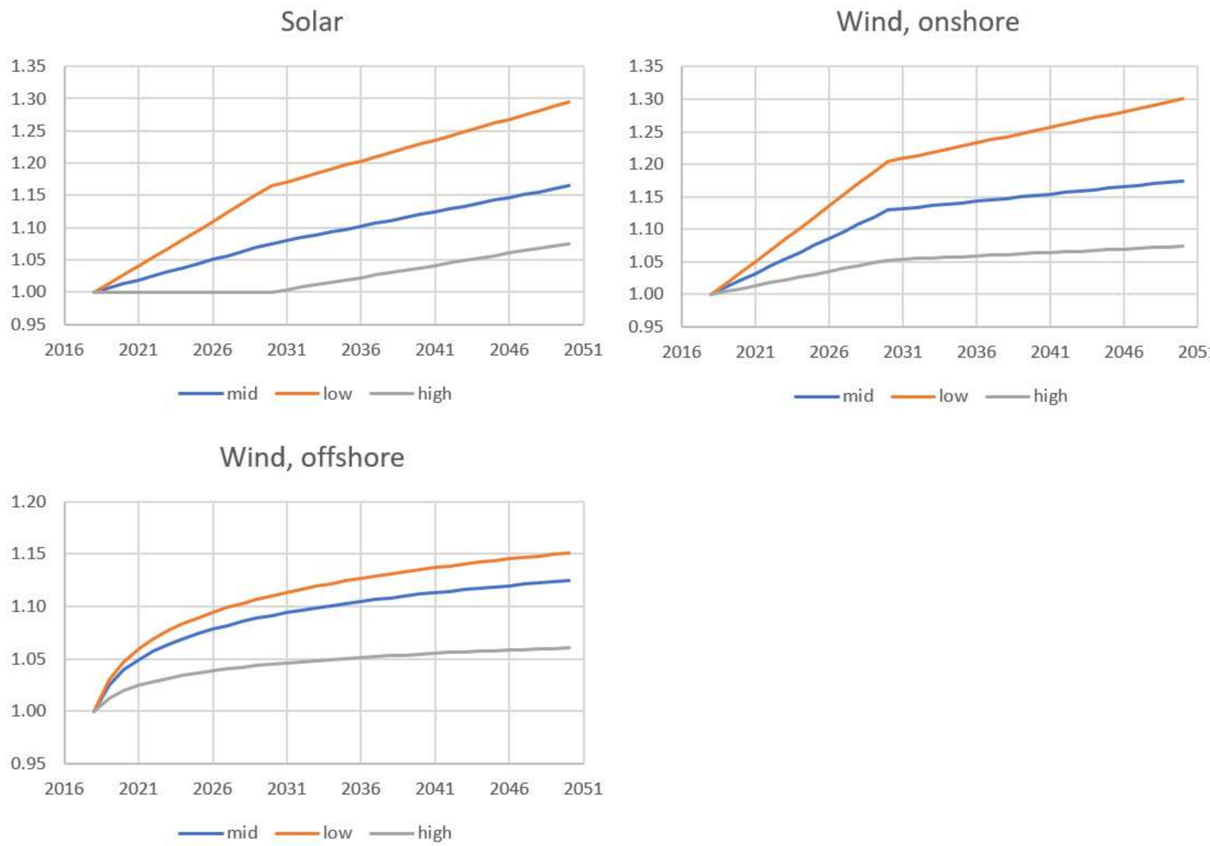
Notes: CAPEX is inflated to 2019 $/kW-year using the consumer price index for the United States. Source, NREL (2020b).

Figure 6.3: Assumptions for fixed operation and maintenance costs (FOMC)



Notes: FOMC is inflated to 2019 $/kW-year using the consumer price index for the United States. Source, NREL (2020b).

Figure 6.4: Technical efficiency for solar, onshore wind, and offshore wind



Notes: capacity factor series normalized to be 1 in 2019. Source, NREL (2020b).

Table 6.1: Variable Operation and Maintenance Cost (VOMC) and Heat Rate

	VOMC		Heat Rate	
	2019	2050	2019	2050
Coal	4.48	4.48	8.64	8.64
Nuclear	2.36	2.36	10.46	10.46
Combustion Turbine	0.60	0.60	9.51	9.51
Combined Cycle	2.20	2.20	6.40	6.40
CC with carbon capture and storage	5.82	5.82	7.53	7.53

Notes: VOMC is inflated to 2019 $/MWh using the consumer price index for the United States; Heat rates are in Btu/MWh. Source, NREL (2020b).

Dispatch characteristics (min and max-generation, ramp rates, startup times and costs, cooldown times, minimum run and downtimes) for thermal technologies basis on PJM experts' advice and the day-ahead energy offer data from July 17, 2019, hour ending at 18:00. The peak load hour for that year was July 19, hour ending at 18:00. We link this data to the capacity market data from the Base Residual Auction for

2021/2022. The capacity market data contains information on the technology class of each unit. We classify steam units using coal as "coal" units and other steam units as legacy gas since natural gas accounts for 90% of these observations' fuel. Next, we aggregate units into plants based on the resource's name in the capacity market data. For example, two units named "plant-name 1 Nuclear" and "plant-name 2 Nuclear" become a single plant called "plant-name." We follow PJM's recommendations and aggregate units at a more granular level for combined cycle by grouping together the turbine, boiler, and steam generator. Then, we compute medians by technology class for each dispatch characteristic; we use the mean for nuclear due to the limited sample size. For combustion turbines and combined cycle, we also rely on PJM's expert advice.

Baseline dispatch characteristics for thermal technologies are in Table 6.2. For combined cycle with carbon capture and sequestration, we use the same baseline characteristics for combined cycle. Dispatch characteristics are assumed to remain constant across vintages, except for combustion turbine and combined cycle technologies. We suppose that combustion turbines become more flexible over time, and combined cycle becomes more efficient. Specifically, combustion turbines' no-load cost, startup times and costs, ramp rates, and minimum run and downtimes improve by 2% per year; combined cycle's VOMC and heat rate decrease by 2% per year. We set a plant's max-generation equal to nameplate capacity and rescale min-generation, startup costs, the no-load cost, and ramp-rate by the ratio of nameplate capacity to max-generation for the technology class.

Table 6.2: Dispatch Parameters

Dispatch characteristics by technology class

	Generation		Startup cost ($)			Startup time (H)		
	Min	Max	Cold	Warm	Hot	Cold	Warm	Hot
Combined cycle	0.52	1.0	2.1	1.3	0.7	3.0	2.0	2.0
Combined cycle-ccs	0.52	1.0	2.1	1.3	0.7	3.0	2.0	2.0
Combustion turbines	0.69	1.0	71.4	71.4	71.4	0.17	0.17	0.17
Coal	0.50	1.0	95.1	57.5	41.1	16	11	8
Legacy	0.64	1.0	0.0	0.0	0.0	0.10	0.10	0.10
Nuclear	0.46	1.0	35.0	23.3	13.8	11	5	4
New nuclear	0.46	1.0	35.0	23.3	13.8	11	5	4

	Cooldown time (H)		Min. time (H)		Ramp-rate
	To cold	To warm	Down	Run	(%/H)
Combined cycle	56	28	4.0	4	120
Combined cycle-ccs	56	28	4.0	4	120
Combustion turbines	2	1	1.0	2	1029
Coal	48	12	17.5	15	39.7
Legacy	47	12	7.0	8	88.0
Nuclear	7	7	48.0	24	13.4
New nuclear	7	7	48.0	24	13.4

Notes: PJM day-ahead energy offers data from July 17, 2019, hour ending at 18:00 and PJM experts' advice. Startup, cooldown, minimum run and downtimes, and ramp rates are in hours; minimum and maximum generation are in MW.

We set the discharge efficiency to 100% for all storage technologies, the discharge rate equal to nameplate capacity, and the depth of discharge to 0%. Based on PJM data, we calibrate the charge rate of pump storage to 50% of nameplate capacity and the charge efficiency to 75% for units that entered before 1975, and 65% and 85% respectively for units built after 1975. The duration is calibrated to 9 hours for all pump-storage units, regardless of their entry date. For battery storage, we set the charge rate equal to nameplate capacity and initialize the charging efficiency to 89%, based on Tesla (2020); we assume that the charging efficiency improves by 0.5 percentage points per year up to a maximum of 95%.

The size distributions for coal, nuclear, legacy, and pump storage are estimated using the capacity market data's nameplate capacity. We calculate technology-specific deciles and approximate the cumulative distribution function by linear interpolation. Table 6.3 reports the deciles. The other technologies are assumed to be modular, with module size equal to 1000-MW for combined cycle, 350-MW for combustion turbines, 200-MW for renewables and battery storage.[12]

Table 6.3: Size distribution of plants (MW of unforced capacity)

	P10	P20	P30	P40	P50	P60	P70	P80	P90
Coal	80	111	240	510	627	1014	1273	1490	1700
Nuclear	1213	1689	1779	1813	1892	2207	2268	2297	2400
Legacy	24	34	45	50	56	78	140	404	767
Hydropower	445	471	504	544	585	779	973	1459	2237

Sources: PJM, Base Residual Auction delivery year 2021/2022, cleared MW of nameplate capacity.

Plants' construction times and technical lives are from EIA (2020b), with some exceptions. For storage, we assume a technical life of 15 years, consistent with PJM (2020b); for coal, nuclear, and pump-storage, we suppose longer life spans because plants belonging to these technologies in PJM are generally older than the technical lives in EIA (2020b). On the other hand, we assume that coal's and nuclear's FOMCs increase by 2% per year to account for technical obsolescence. Table 6.4 displays these parameter values.

[12] The minimum efficient scale for intermittent resources and battery storage is smaller. Setting it to a larger number reduces computation time, because there are fewer units in the market and fewer exiting and entering units. This choice does not affect results significantly since the size of the market is large, approximately 180,000-MW in 2019, net of Fixed Resource Requirement (FRR ≅ 13,200-MW) and Energy Efficiency resources (EE ≅ 3,900-MW).

Table 6.4: Other technology characteristics

	Operating life (years)	Construction time (years)	Repayment period (years)	Initial resources (MW)	Initial capacity values	Min. capacity value	Initial energy offsets ($/MW)	2019 Cost of entry ($/MW-day)
Coal	100	3	20	51226	93.3%	93.3%	15,695	1,306
Nuclear	150	3	20	32640	98.5%	98.5%	188,705	2,393
NewNuclear	50	3	20	0	98.5%	98.5%	188,705	1,562
Combustion turbines	40	2	20	28793	91.6%	91.6%	17,520	314
Combined cycle	40	2	20	50666	97.3%	97.3%	61,320	346
Combine cycle-ccs	40	2	20	0	97.3%	97.3%	61,320	865
Legacy	40	2	20	10870	93.8%	93.8%	61,320	346
Hydropower	200	3	20	5574	97.7%	20.0%	95,265	1,862
Solar	30	1	20	3127	60.0%	10.0%	67,525	450
Wind, Onshore	25	1	20	7834	17.6%	10.0%	87,600	666
Wind, Offshore	25	2	20	0	26.0%	10.0%	123,005	1,431
Battery-8h	15	1	10	0	100.0%	40.0%	84,680	1,375
Battery-4h	15	1	10	0	90.0%	40.0%	42,340	755
Battery-2h	15	1	10	0	85.0%	40.0%	21,170	445
Battery-1h	15	1	10	0	75.0%	40.0%	10,585	290

Notes: The operating life and construction time are from EIA (2020b), except for coal, nuclear, and hydropower; energy offsets are from PJM (2020a); energy offsets for hydropower, 8, 2, and 1-hour storage are obtained by rescaling energy offsets for 4-hour storage, based on relative duration. The entry cost is annualized CAPEX plus FOMC divided by 365. For new nuclear, the number is for 2027, the first year the technology becomes available.

Based on historical data, we set the maximum amount of MW that can exit or enter the market each year to 7.5 and 9-GW. Also, we set the ceiling for the ratio of capacity to peak load equal to 1.18. This number is high since capacity is nameplate times the capacity value. The cap does not bind in equilibrium and only limits the fluctuations of the model as it converges.

The MW amount of state-sponsored resources entering each year is based on PJM estimates and reported in Table 6.5. State-sponsored resources, including nuclear, are assumed to stay in the market regardless of their economics.

Table 6.5: State-sponsored resources (MW)

	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036...
BTM Solar	1057	1241	1280	1365	1039	782	740	702	722	766	849	906	935	996	1020	1166	0
Solar	429	920	1058	1137	815	1296	1305	1273	1699	1739	1768	1557	1600	1629	1636	1666	0
Offshore Wind	12	18	700	400	1280	2080	828	400	1300	400	800	1400	500	1000	1000	1800	0
Onshore Wind	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Notes: Authors' calculations from PJM estimates.

Energy resources are initialized based on those that participated in the Base Residual Auction or Fixed Resource Requirement alternative for 2021/2022. We ignore demand response and energy efficiency resources. Table 6.4 reports total nameplate capacity by technology class for 2019, along with initial values for energy offsets and capacity values. PJM had no offshore wind farms and battery facilities in 2019. Energy offsets are from PJM (2020b), except for pump storage and 1, 2, and 8-hour batteries. For these technologies, we set energy offsets proportional to those for 4-hour batteries, based on duration.

Capacity values of thermal technologies and pump storage are initialized at one minus the EFORd computed from the capacity market data. We use the values from PJM (2020b) for renewables, and for batteries, we rely on simulations of the energy market model. Finally, for legacy and CC-ccs, we use energy offsets and capacity values as for combined cycle. We initialize the double exponential smoothing parameters to {0.4, 0.4} for both energy offsets and capacity values. For capacity values, we impose the floors listed in Table 6.4. For profits, we set a ceiling equal to six times the annualized cost of entry in 2019.

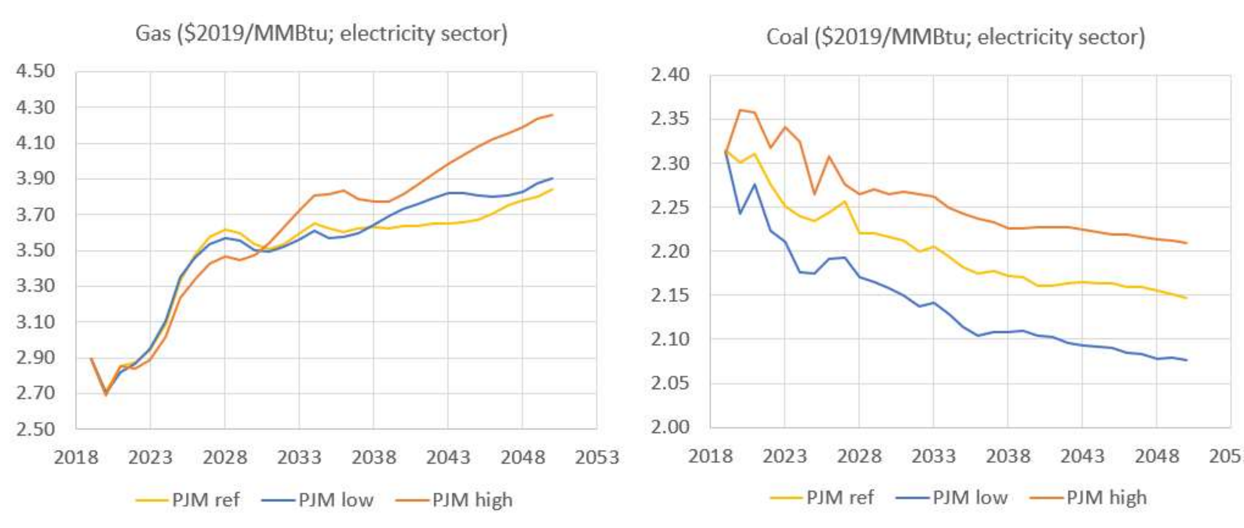
The discount factor is set equal to 8.2%, the after-tax weighted average cost of capital from PJM (2020b).

We set the planning horizon for the capacity auction to three years, as in PJM, and use the 2022-2023 Base Residual Auction demand curve.[13] We shift the demand curve by a constant to match PJM's 2021 capacity price of $140/MW-day.

We assume that price responsive demand is isoelastic with an elasticity parameter of *–0.1,* as in Alcott (2012). The fraction of price responsive load is set to 0% in 2019 and kept constant in the baseline scenario.

Fuel prices are from EIA (2020a) and expressed in 2019 USD. We select series for the power sector and aggregate prices for the East North Central, East South Central, South Atlantic, and Mid-Atlantic regions using the distribution of plants across states in PJM footprint. Figure 6.4 displays our series for coal and gas. The uranium price is national and forecasted to increase by approximately 0.2% per year in all scenarios.

Figure 6.4: Gas and coal price scenarios



Notes: Authors' calculations from EIA (2020a) and the distribution of nameplate capacity across PJM states.

[13] The ceiling is *1.5 x Net-CONE*, the knots are *(–1.2% RR, 1.5 Net-CONE)*, *(1.9%RR, 0.75 Net-CONE)*, and *(7.8%RR, 0)*, where RR is the Reliability Requirement adjusted for the fixed resource requirement alternative (*FRR*) and equal to *RR = E(D) x (1+IRM) x (1-EFORd) – FRR*, *IRM* is the Installed Reserve Margin, and *E(D)* is the forecasted peak load. As in the 2022/2023 Base Residual Auction we set *IRM = 15.7%*, and *EFORd = 5.9%*. We set FRR to zero because the model's set of resources encompasses fixed resource requirement entities. These parameters are kept fixed throughout the simulation, and let *Net-CONE* and *E(D)* vary over time.

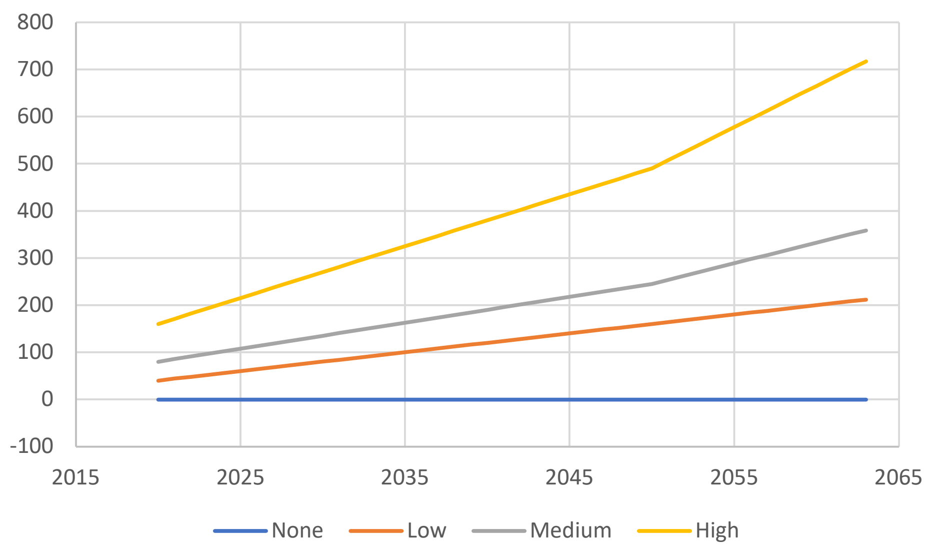
Series for CAPEX, FOMC, heat rates, fuel prices, and load are extended from 2050 to 2113 by assuming a constant growth rate, using the average growth rate for 2041-2050. Carbon price series are piecewise linear and extended from 2070 to 2113 using linear interpolation.

Scenarios

We consider three scenarios for renewable technologies' costs (mid, low, and high cost) and three for fuel prices (mid, low, and high oil price). These are official scenarios from EIA (2020a) and NREL (2020a, 2020b). Figures 6.2, 6.3, and 6.4 display these series.

For the carbon price, we consider four scenarios. A no-price scenario, a low-price scenario (low end to achieve Paris agreement, CPLC, 2017), a medium price scenario (low end to achieve a 1.5-degree Celsius rise in temperatures, IPCC, 2018), and high price scenario (twice the medium price scenario). We use 100-year Global Warming Potentials from IPCC (2014) and emission rates from EPA (2014) to translate heat rates into emissions and apply the carbon price. Figure 6.5 and Table 6.6 display carbon price series and emission factors.

Figure 6.5: Carbon Price Scenarios (2019$/tCO2e)



Notes: "Low" = low end to achieve Paris agreement, CPLC, 2017; "Medium" = low end to limit the rise in temperatures to 1.5-degree Celsius, IPCC, 2018; "High" = twice the "Medium".

Table 6.6: Emission Factors

	kg CO2/MMBtu	g CH4/MMBtu	g N2O/MMBtu	kg CO2e/MMBtu
Coal (1)	95.5	11	1.6	96.3
Natural Gas (1	53.1	1	0.1	53.1
GWP (2)	1	28	265	

Sources: EPA (2014) and IPCC (2014).

The baseline scenario has no price-responsive demand and a constant ORDC. We also consider two side cases (mid and high) where the fraction of load that responds to prices grows by 1 and 2 percentage points per year, and another two side cases (mid and high) where the ORDC increases by $50/MW and $100/MW per year.

We model a side scenario where next-generation nuclear becomes commercially available in 2030. Next-generation nuclear is assumed to have a smaller efficient scale and lower construction and operation and maintenance costs than standard nuclear. Based on NuScale (2020), we set the minimum efficient scale to 300-MW, CAPEX to $4386/kW (as opposed to $7278/kW for nuclear), and FOMC and VOMC 20% below values for nuclear. Dispatch constraints are set equal to those for traditional nuclear. We also consider a low-cost next-generation nuclear scenario, where FOMC and VOMC are 50% below values for nuclear.

Finally, we consider two market designs. The reference scenario has the narrow MOPR proposed by PJM (2021a,b). The narrow MOPR applies case-by-case if the system operators or the independent market monitor detect instances of buyer market power or if states provide resources with conditional support based on a resource's capacity market outcomes. In the model, we assume that the narrow MOPR effectively deters buyer market power and that resources do not receive conditional state support. Thus, the model is as if there were no MOPR. Alternatively, we consider the broad MOPR established by FERC (2019). In this scenario, the minimum offer price applies to all state-sponsored resources starting operation after 2021.

7 Energy proxy model

Investors' expectations of future profits drive the multi-year simulation. In principle, we could use the energy market model to simulate annual profits. In practice, this is infeasible. Even restricting the energy market simulation to one week per month, solving the model for one year takes approximately 30 minutes on a single core. One multi-year simulation makes 80 million calls to the energy market model, on average, which would require 4500 years of computation time.

Our solution strategy is to use the electricity market model to populate a database of known instances and estimate an econometric model on this database to predict energy profits and performance. This approach is consistent with the behavioral assumptions made in the multi-year simulation and arguably reflects heuristics investors use to assess future profitability. Investors entertain some reasonable yet approximate expectations on the future direction of the electricity market, as described by the sequence of approximate market states $\tilde{M} = \{\tilde{S}, Z\}$ in the multi-year simulation. Also, they broadly understand the complementarities and substitution possibilities between different technologies, and the effect of market-wide variables, such as the carbon price, on energy profits and performance. They use these expectations and the approximate model to forecast profits and inform their investment and retirement decisions.

The training and testing database

We initialize the database of known instances by restricting attention to a relatively small portion of the market state. We start with a time horizon of 40 years, assuming no price responsive demand, and keeping the ORDC penalty parameter constant at $2000/MWh. We restrict attention to 2-hour battery storage, ignoring batteries with a different duration. Also, we suppose that all units have maximum charging efficiency, 95%. This assumption biases profits of early vintages upward but is essentially harmless because batteries do not enter in the first few years of the simulation in equilibrium. Similarly, we assume that older pump-storage units have the same characteristics as newer units. This assumption biases profits of older units upward, but we suppose that state-sponsored resources and pump-storage do not retire.

We use our prior for the complementarities and substitution possibilities between different technologies and guess several possible market paths. For example, high renewable penetration requires a large amount of storage or combustion turbines; nuclear is inflexible and complementary to storage but somewhat substitutable with combined cycle. Also, we postulate no entry of coal, traditional nuclear, legacy, and pump-storage units.

Each path is a sequence of instances. We perturb these instances to grow the dataset further and limit our priors' effect on the results. We expand the energy mix to a corresponding set of plants and generate inputs for forced outages, planned outages, and ORDC curves, as detailed in section 4. Then, we run the energy market model for each instance and obtain an initial set of observations to estimate the energy proxy model.

We gradually expand our knowledge of the model space by allowing the fraction of price responsive load and the ORDC penalty parameter to vary. We keep growing the database with instances queried by the multi-year simulation and random instances. Including random instances limits the collinearity between observations generated by the simulation and balances the exploration-exploitation tradeoff. As the database becomes more extensive, we can increase the flexibility of the energy proxy model, increasing the accuracy of the multi-year simulation. This improvement is accomplished algorithmically with the machine learning techniques described below.

The baseline model

The following observations inform the specification of the baseline econometric model.

1. The PJM market is sizable, especially relative to the minimum scale of renewables, battery storage, next-generation nuclear and combustion turbines. Thus, we can ignore indivisibilities and guess that market outcomes are scale-invariant—doubling supply and demand leaves energy profits and performances unchanged.
2. Only the total generation of renewables is relevant for market outcomes, not the individual units' efficiency and nameplate. All that matters is the aggregate production of solar and wind in each 5-minute interval. A unit's profits and performance are then proportional to its efficiency relative to the technology class'.
3. The effect of fuel prices depends on the nameplate capacity of resources in the market using that fuel—similarly for the carbon price and the impact of vintages.

Consistent with these observations, we define the approximate market state $\tilde{M}$ as comprising the total nameplate capacity of each technology, the mean vintage of combustion turbines and combined cycle plants, the mean efficiency of renewables, the vector of fuel prices, the carbon price, the ORDC penalty parameter, the median load, and the fraction of price responsive demand:

$$\tilde{M} = \left\{ \{MW_j\}_{j \in J}, \{\overline{v}_j\}_{j \in \text{vint}}, \{\overline{e}_j\}_{j \in \text{ren}}, P^F, P^O, \overline{P}, D, PRD \right\}$$

MW is nameplate capacity, $\bar{v}$ is the average vintage of combined cycle and combustion turbine plants, with *vint = {CT, CC, CC-ccs}*, and $\bar{e}$ is the average efficiency of renewables, with *ren = {solar, onshore wind, offshore wind*}.

Using observations 1 and 2 above, we define the normalized, effective installed capacity for technology j as:

$$mw_j = \begin{cases} \frac{MW_j \bar{e}_j}{D} & \text{if } j \in \text{ren} \\ \frac{MW_j}{D} & \text{otherwise} \end{cases}$$

Also, consistent with observation 3, we allow for interactions between mw_j and:

- The corresponding fuel price, if technology j uses natural gas, coal, or uranium,
- The carbon price, if technology j uses gas or coal,
- The average vintage $\bar{v}$ for combustion turbine and combined cycle technologies.

Guided by observations 1-3, we adopt the following parsimonious specification to construct expectations for average energy profits by technology class j :

$$\begin{aligned} ep_j \quad = \quad & \sum_{k \in J} \beta_{j,k}^{mw} mw_k \\ & + \sum_{k \in \text{vint}} \beta_{j,k}^{a} \bar{a}_k + \sum_{k \in \text{vint}} \beta_{j,k}^{a,mw} mw_k \bar{a}_k \\ & + \sum_{k \in \text{fuels}} \beta_{j,k}^{PF} p_k^F + \sum_{k \in \text{thermal}} \beta_{j,k}^{PF,mw} mw_k p_{f(k)}^F \\ & + \beta_j^{PO} p^O + \sum_{k \in \text{emissions}} \beta_{j,k}^{PO,mw} mw_k p^O \\ & + \beta_j^{ORDC} ordc + \beta_j^{PRD} PRD + \epsilon_j \end{aligned}$$

ep = T(EP) is some transform (for example, the natural logarithm) of energy profits per unit of effective nameplate capacity; *fuels = {coal, natural gas, uranium}* is the set of fuel types and *fossil = {coal, legacy, CT, CC, CC-ccs}* is the set of technologies producing greenhouse gas emissions; p^o, p^f, and *ordc* are some transform (for example, the natural logarithm) of the carbon price, fuel prices, and ORDC penalty factor; *f(·)* maps technologies into fuel types.

An advantage of this parsimonious specification is that it remains somewhat interpretable. For example, the coefficients $\beta_{j,j}^{mw}$ should be negative, whereas $\beta_{j,k \neq j}^{mw}$ reflects the complementarities and substitution possibilities across different technologies.

Next, we model plant-level energy profits. Coal, nuclear, legacy, pump-storage, and next-generation nuclear technologies remain the same over time. We set investors' expected energy profits equal to the technology class average for plants operating these technologies. Combustion turbine, combined cycle, renewable, and battery storage technologies improve over time. As argued above, we ignore technical progress for battery storage and set expected profits equal to the technology class average. For renewables, we note that generation is proportional to efficiency. Thus, energy profits per unit of nameplate capacity for resource i are equal to:

$$\widehat{EP}_i = e_i \times T^{-1}\left(\widehat{ep}_{j(i)}\right)$$

Where *j(i)* is the technology operated by unit *i.* Finally, for combustion turbines and combined cycle technologies, we estimate the model:

$$EP_i - EP_{j(i)} = \eta_j\left(\nu_i - \overline{\nu}_{j(i)}\right) + \xi_i$$

This specification ensures that the expected energy profits predicted by the proxy model at the plant level, EP_i, are consistent with the expected energy profits predicted by the energy proxy model for the technology class, $EP_{j(i)}$.

In the multi-year simulation, we use performance to update capacity values which we set at the technology class level consistent with PJM's practice. We estimate technologies' performance using the same specification for energy profits:

$$\begin{aligned} T\left(perf_j\right) \quad = \quad & \sum_{k\in J}\gamma_{j,k}^{mw} mw_k \\ & + \sum_{k\in\text{vint}} \gamma_{j,k}^{a}\overline{a}_k + \sum_{k\in\text{vint}} \gamma_{j,k}^{a,mw} mw_k \overline{a}_k \\ & + \sum_{k\in\text{fuels}} \gamma_{j,k}^{PF} p_k^F + \sum_{k\in\text{thermal}} \gamma_{j,k}^{PF,mw} mw_k p_{f(k)}^F \\ & + \gamma_j^{PO} p^O + \sum_{k\in\text{emissions}} \gamma_{j,k}^{PO,mw} mw_k p^O \\ & + \gamma_j^{ORDC} ordc + \gamma_j^{PRD} PRD + \epsilon_j \end{aligned}$$

Here *T* is some transform of performance, for example, logit. As for energy profits, we divide the left-hand-side variable, *perf*, by the average efficiency of renewables. For renewables, we transform the model prediction using the inverse of *T* multiplied by the average efficiency of the technology class.

A richer model, ensembles

The baseline econometric model is parsimonious and interpretable. Also, the predictor is a continuous function, which is necessary for market equilibrium in the multi-year simulation. The net present value must change gradually as investors make entry and exit decisions. However, the baseline econometric model is too simple to approximate energy profits and performance over the entire model space. First, the resource mix can be very unbalanced. For example, generation could be insufficient to serve load, resulting in frequent shortages, high electricity prices, and extremely high energy profits. Or combined cycle, combustion turbines, and storage may be inadequate to smooth net-load oscillations at high renewable penetration. These combinations of resources are not observed in equilibrium but may be queried by the multi-year simulation as it converges. Second, and most importantly, nonlinearities are an intrinsic feature of electricity markets. For example, when coal is the marginal technology, the carbon price can increase profits for gas units, whereas the relationship turns negative as coal units exit the electricity market.

To address these problems, we start by restricting the sample. We compute the ratio of energy profits to the 2019 entry cost for each technology and consider instances such that the L^p-norm is less than $n^{1/p}k$, with n the number of technologies. We set $p=2$ and $k=6$, restricting the sample to cases with energy profits approximately smaller than six times the entry cost in 2019. When running the multi-year simulation, we enforce bounds preventing the model from reaching excessively unbalanced states, as discussed in section 6. Next, we add a second layer to the model that can account for strong nonlinearities. First, we use regression trees to partition the model space into relatively homogenous regions. Then, we estimate the baseline model within each sub-region using LASSO. Thus, we allow the model to capture strong nonlinearities while retaining local interpretability and continuity. We tune the hyperparameters of the tree and LASSO estimators and select other features of the model using standard machine learning techniques. Specifically, we select among models varying along several dimensions as follows:

- We estimate energy profits directly or split energy profits into base and peak energy profits and estimate separate models for these two components. Peak energy profits are revenues earned during 5-minute intervals when the price is above $200/MWh.
- We consider specifications in logs or without logs for energy profits and probit and logit for performance.
- We allow for fuel prices, the carbon price, and the ORDC penalty factor to be in logs or not.
- We consider trees with different minimum size leaves or no tree.
- We consider a LASSO estimator or simple regression.

We split the data using 1/3 of observations for testing and use cross-validation to select the penalty parameter of the LASSO estimator in each leaf. Then, we choose the minimum leaf size and the transformation for the dependent and independent variables by computing the mean-squared error on the testing data. We consider four different minimum leaf sizes and two possible transformations for the dependent and independent variables. Thus, we use the test data to select among 16 models.

8 Preliminary results

We generate an initial dataset of 20,060 energy market instances as described in section 7. We set $p = 2$ and $k = 6$ leaving us with 18,477 observations. The test data has 6148 observations to select among 16 competing specifications.

We estimate the energy proxy model and run the multi-year simulation for the reference case defined in section 6. The carbon price increases by $3/tCO2e, the ORDC penalty factor is constant at $2000/MWh, and there is no price responsive load. The cost and efficiency of renewables are from NREL's reference scenario; combined cycle plants become 1% more efficient every year and combustion turbines 1% more flexible. Traditional nuclear, solar, onshore wind, and offshore wind are state-sponsored. Pump-storage and state-sponsored resources are assumed to remain in the market regardless of their economics. We fix the capacity values for thermal resources to one minus the forced outage rate. This simple approach is apt to be more consistent with PJM effective load-carrying compatibility method for determining capacity values. For renewables and storage, we use the energy proxy model to estimate performance. We cap energy profits to six times the 2019 entry cost to limit oscillations as the model converges to the equilibrium. The capacity market has a narrow minimum offer price rule (MOPR). We also consider the case with the broad MOPR. The broad MOPR applies to all state-sponsored resources beginning operations after 2021. Comparing the narrow and broad MOPR scenarios shows the model's usefulness in evaluating alternative market designs and addressing policy questions in general. The double exponential smoothing weights are optimized on the reference case and are equal to 0.93 and 0.0 for energy offsets and 0.95 and 0.0 for capacity values for the intercept and trend coefficients, respectively.

For each point along the equilibrium path of the simulations, we run the energy market model allowing us to evaluate the energy proxy model's fit and compute granular statistics useful to assess costs, profitability, and reliability.

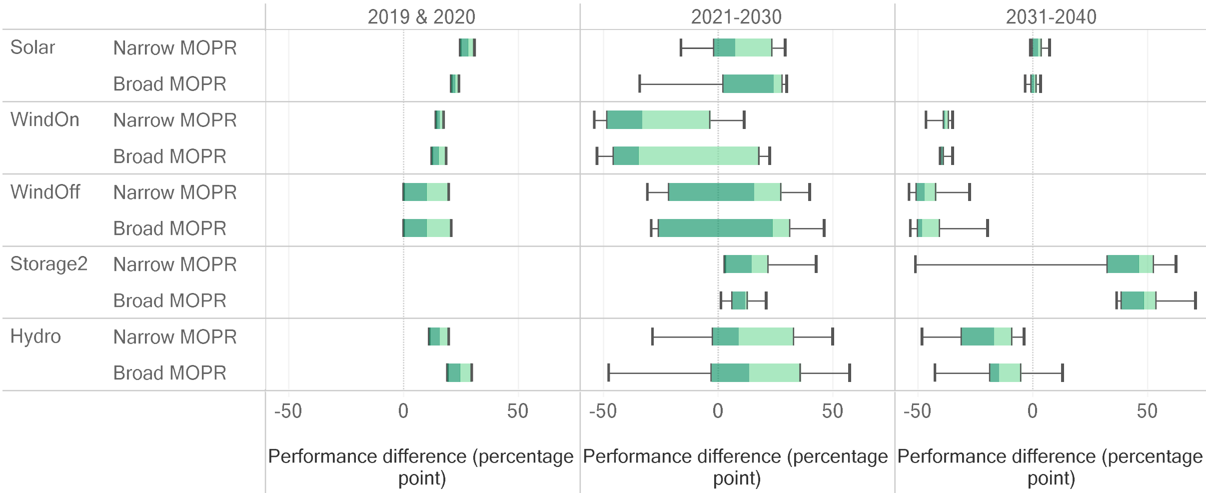
Energy proxy model fit

Figures 8.1-8.2 display the forecast error of the energy proxy model by technology for 2019-2020 and the next two decades. The forecast error for energy profits is normalized using the 2019 entry cost for each technology from Table 6.4. The energy proxy model for profits fits the data well. The forecast error slightly increases over time as the equilibrium path gets further from the known instances used to train the proxy model. Performance differences are more prominent, especially for wind technologies.

Figure 8.1: Energy market proxy fit along the equilibrium path, energy profits

Notes: the forecast error is annual energy profits predicted by the energy proxy model minus the energy profits simulated in the energy market model divide by the 2019 cost of entry along the equilibrium path of the multi-year simulation. The box plot displays the interquartile range and minimum and maximum forecast error.

Figure 8.2: Energy market proxy fit along the equilibrium path, performance



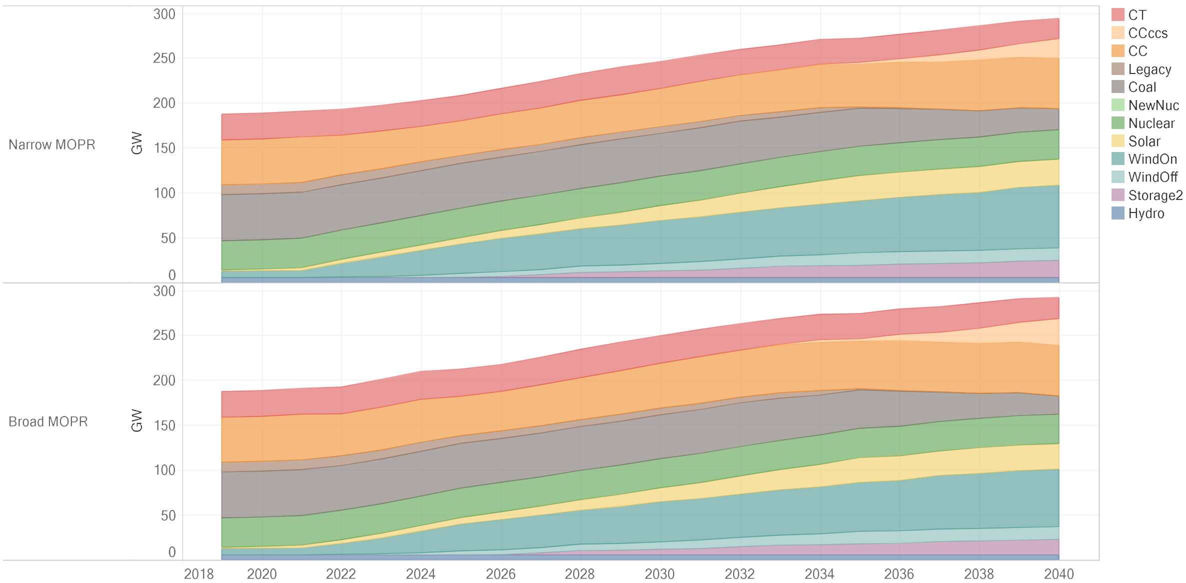
Notes: the forecast error is annual performance predicted by the energy proxy model minus performance simulated in the energy market model along the equilibrium path of the multi-year simulation. The box plot displays the interquartile range and minimum and maximum forecast error.

The deterioration of the fit over time is expected. The procedure described in section 7 limits the amount of overfitting on the data. However, this data reflects our prior and is not representative of the actual data generating process. This problem should self-correct as we iterate between the multi-year simulation and the energy market model and include in the dataset instances on and near the candidate equilibrium path.

Resource mix and entry and exit

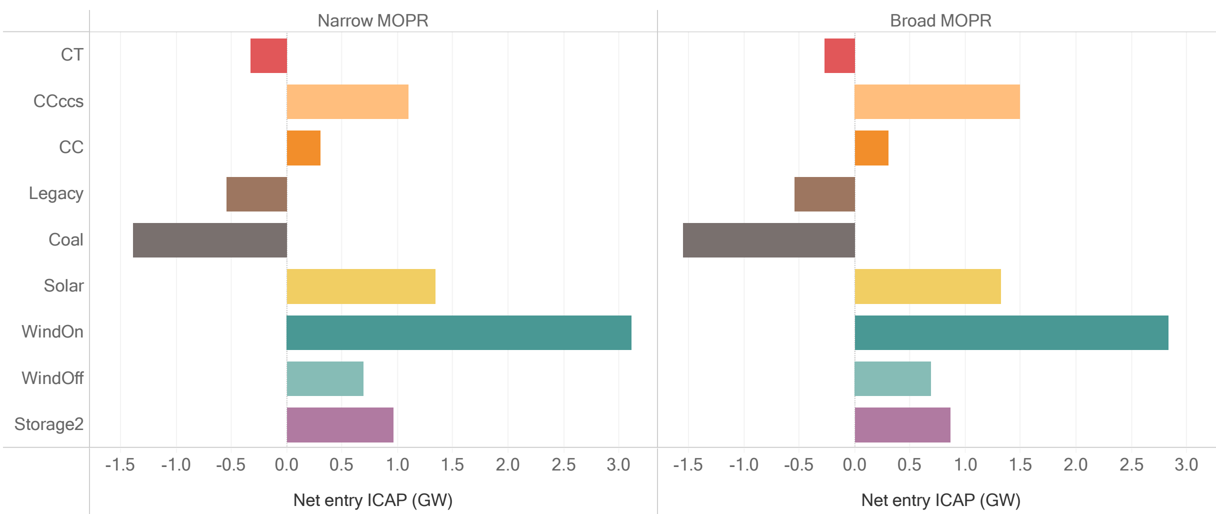
Figure 8.3 shows the evolution of the resource mix between 2019 and 2040. Changes in the resource mix result from investors' entry and exit decisions. Existing resources retire because they reach the end of their technical life or earlier if profits turn negative going forward, net of fixed costs. New resources enter based on forward-looking profit expectations or irrespective of profitability based on states' targets.

Figure 8.3: Energy mix by year, 2019-2040



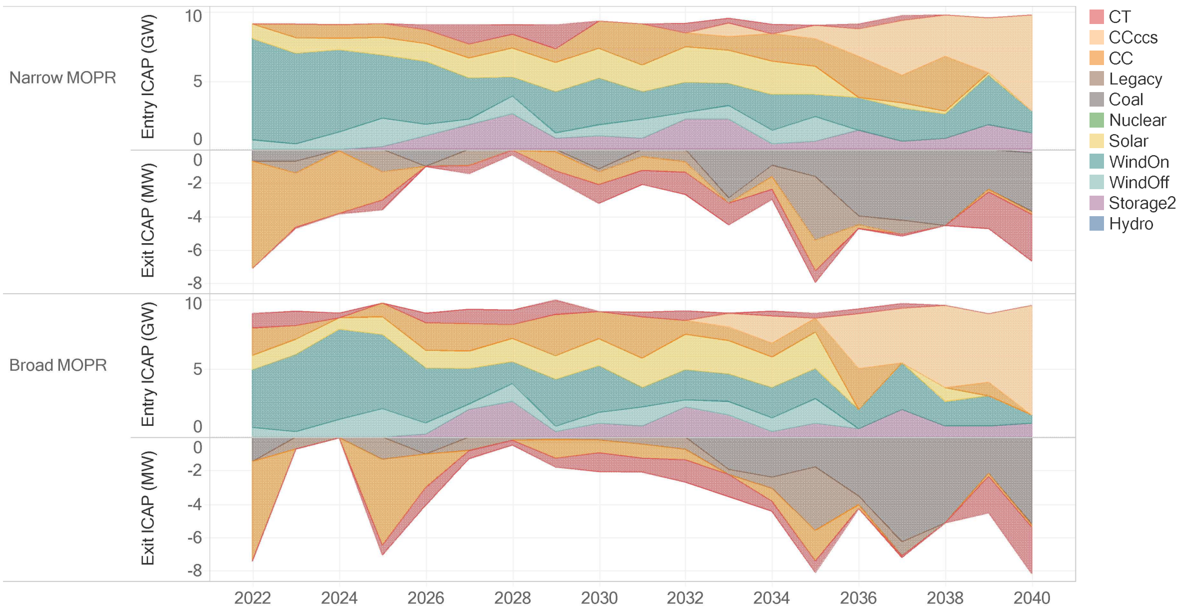
Figures 8.4 and 8.5 display net entries (entries minus exits) in GWs averaged over 2021-2040 and the detail of entry and exit by year, respectively. Figures 8.6 and 8.7 display the net present value for incumbents and new entrants underlying those decisions.

Figure 8.4: Average annual net entries, 2019-2040



Given the modest carbon price, thermal resources continue to play an important role. Coal gradually exits, as do the legacy gas resources. These resources are replaced with storage, wind, and solar. Combined cycle units remain a critical component in the mix and enter the market, offsetting the reduction in combustion turbines. In 2032 the carbon price reaches a sufficiently high level that combined cycle plants with carbon capture start replacing traditional combined cycle plants. As the carbon price continues rising and renewable resources saturate the market, combined cycle projects with carbon capture dominate the entry queue.

Figure 8.5: Entry and exit by year, 2019-2040



The vast majority of exits are economically motivated. The net present value for incumbents is positive and significant for three reasons. First, plants staying in the market have better technical characteristics. Surviving combined cycle and combustion turbine plants operate more recent vintages. The remaining coal plants are less old and have lower fixed operation and maintenance costs. Second, some of the remaining plants have finished paying the entry cost installments. Third, the exit of some plants lifts the profitability of the plants staying in the market. Fourth, the investor retires the plant if the value of announcing the retirement and continuing operations for three years exceeds the value of postponing the announcement by at least one year. Investors are forward-looking and anticipate future negative profits. Thus, negative profits rarely occur in practice. A retiring plant will generally earn three years of positive profits before terminating operations. Negative profits can become more frequent if the market hits the annual retirement cap repeatedly. However, this happens only in a handful of years along the equilibrium path. The constraint captures the real-world case in which too many plants wish to retire simultaneously, hindering reliability and triggering the system operator response to keep some plants in the market with make-whole payments.

Initially, the only new technology with a positive net present value is onshore wind. The net present value turns positive in 2021 for batteries, 2024 for solar, and 2029 for combined cycle with carbon capture and sequestration (entry follows after three years consistent with the planning horizon). Offshore wind never becomes profitable but enters the market because of states' support.

Figure 8.6: Net present value of incumbent resources by year, 2019-2040

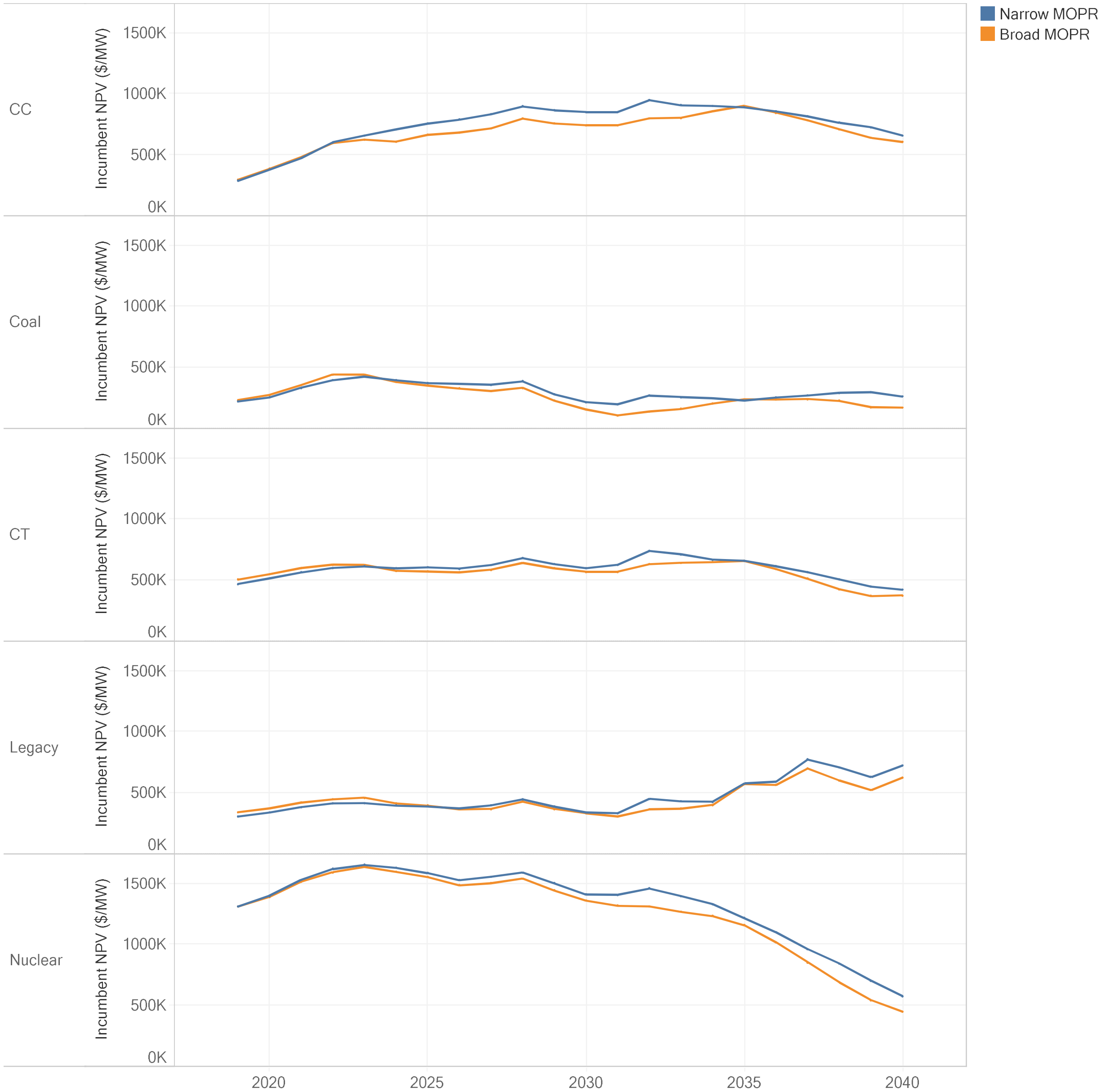
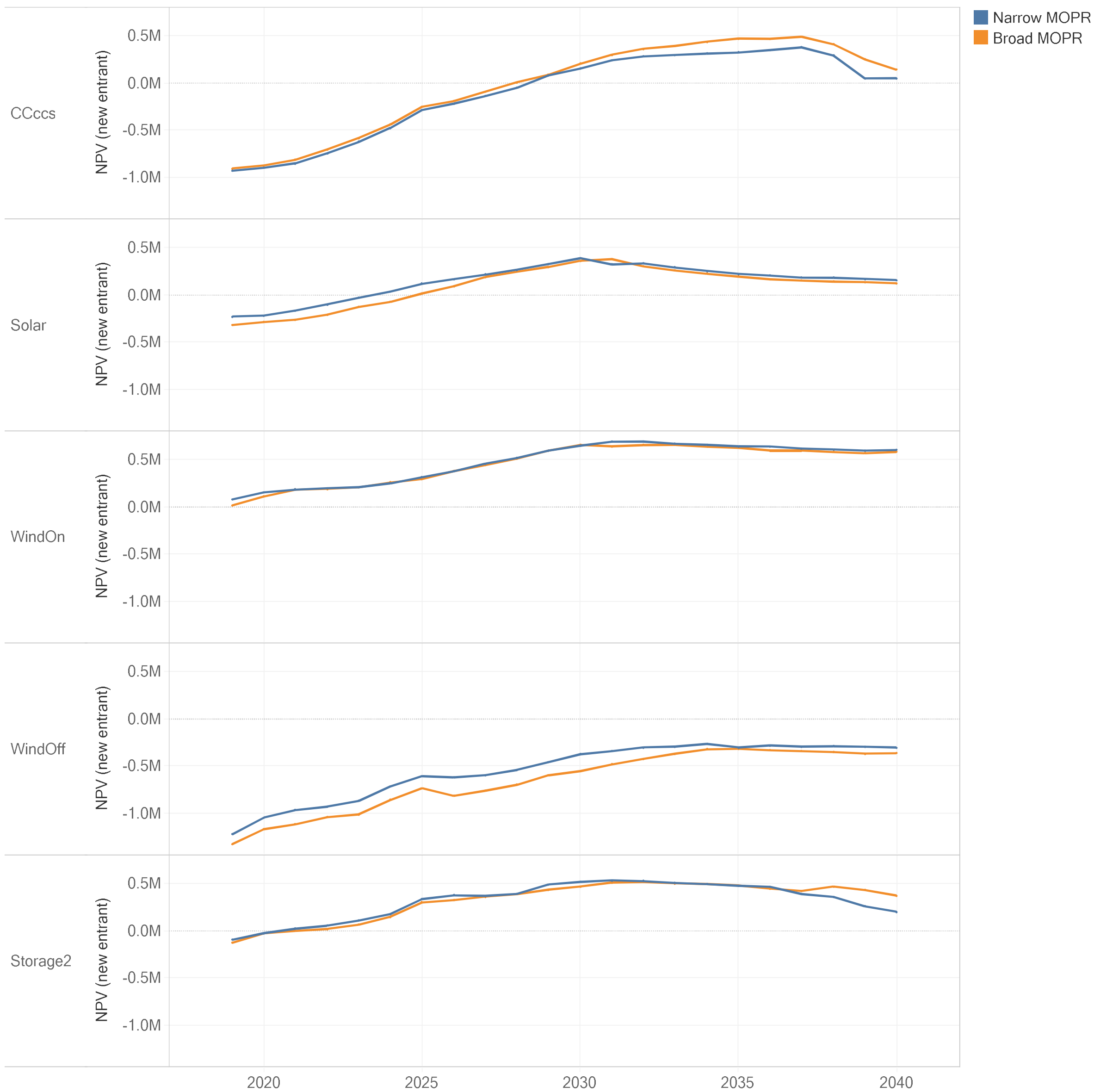


Figure 8.7: New entrant net present value by year, 2019-2040



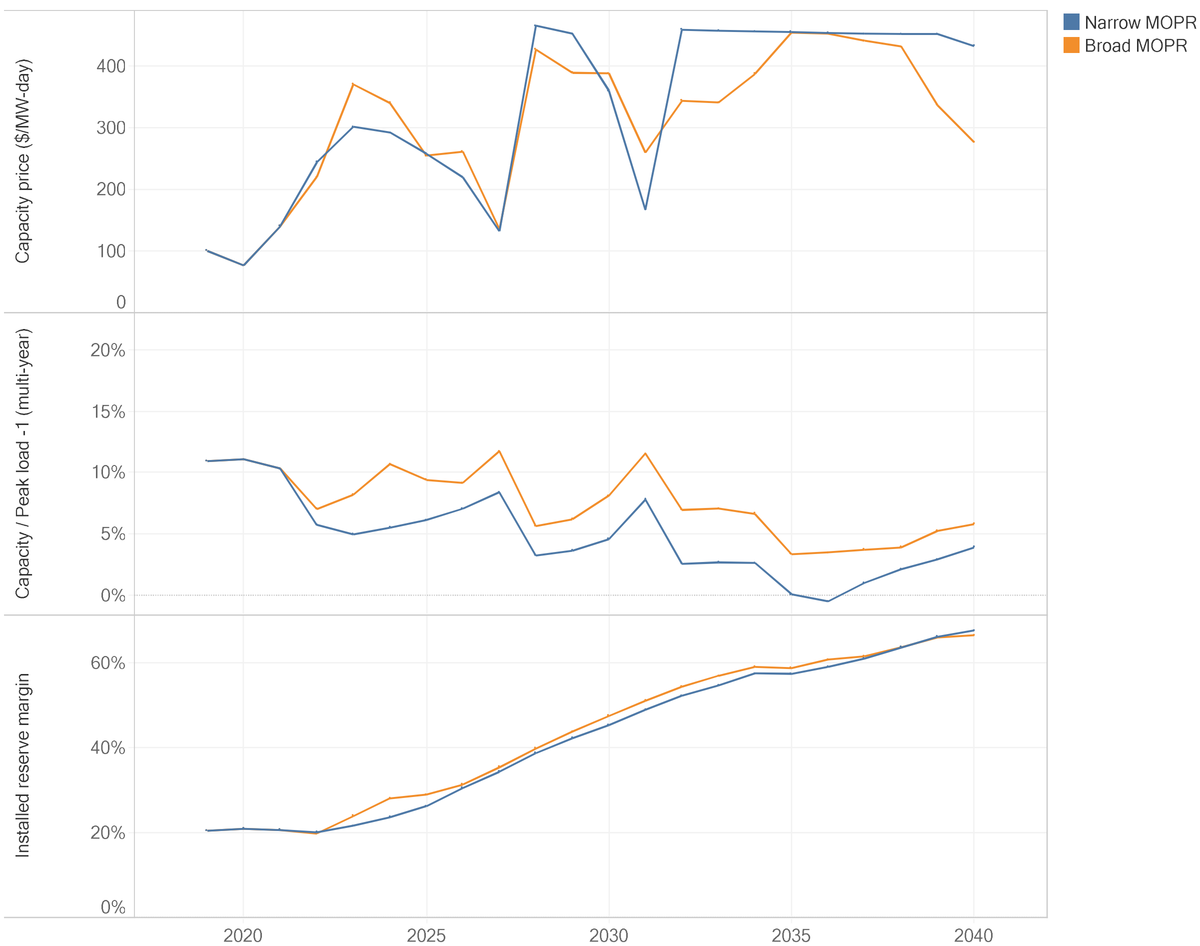
The introduction of the broad MOPR leaves net present values—and therefore entries, exits, and the evolution of the energy mix—largely unaffected. The broad MOPR lowers the net present value of onshore wind slightly in the first two years of the simulation and delays the date when storage and solar become economic by two years. These effects are small for two reasons. First, renewable resources and storage are already economic or near-economic at the beginning of the simulation (except for offshore wind). Thus, the broad MOPR is unlikely to bind, and once the resource clears, the offer floor drops from Net-CONE to Net-ACR, which is much lower. Second, the capacity value of renewable resources drops over time as more renewable resources enter the market. Thus, capacity payments become less critical for renewables. Figures 8.9 and 8.10 below show the profit components by technology and year and the performance and capacity values by technology and decade. Renewables rely less on capacity payments than traditional generation. Instead, the broad MOPR benefits combined cycle with carbon capture and sequestration. In the broad MOPR scenario, 1.5GW enters the market each year between 2021 and 2040 instead of 1.1GW in the narrow MOPR case.

Perhaps surprisingly, the broad MOPR lowers the net present value of traditional technologies slightly. The broad MOPR reduces renewables' penetration, especially of onshore wind, but increases entry of combined cycle with carbon capture and sequestration, which is a much closer substitute. Also, as discussed below, the broad MOPR does not affect capacity prices significantly but leads to more resources in the market. Thus, traditional resources' energy profits suffer, offsetting any potential revenue increase in the capacity market. This result highlights a well-known tradeoff. While the capacity market ensures that the system operator achieves the target capacity level, it can lead to too many resources in the market, shifting money away from the spot market into the capacity market and exacerbating the missing money problem (Joskow, 2008, and Hogan, 2013).

Capacity prices and reserve margin

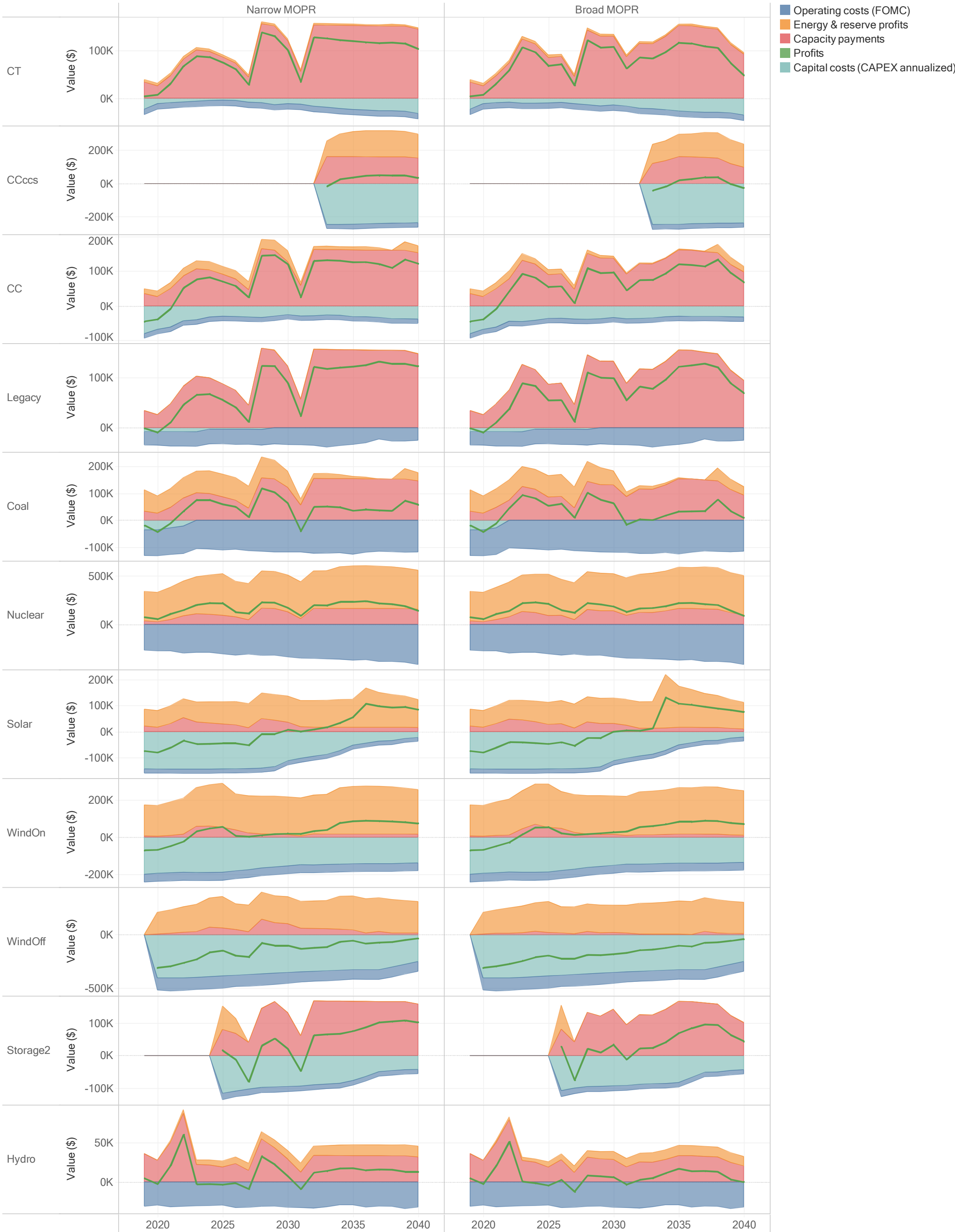
Figure 8.8 shows capacity prices and reserve margins by year. We report both the installed reserve margin (the ratio of nameplate capacity to peak load) and the capacity-to-peak load ratio (capacity being defined as nameplate times the capacity value).

Figure 8.8: Capacity price and reserve margin by year, 2019-2040



For 2019, 2020, and 2021 capacity prices are exogenous and set using the capacity prices from PJM's Base Residual Auction for those years. For 2022 and subsequent years, the capacity price clears the market, ensuring that the system operator achieves the target capacity level. Fossil resources earn little energy profits, especially legacy gas, combustion turbines, and combined cycle plants. These units account for over 50% of nameplate capacity in 2019 and are necessary to achieve the procurement target. Thus, the capacity price increases providing these resources with enough revenues to remain economic. Figure 8.9 shows profits by technology type and year together with revenues and costs. Capacity payments are the most significant contributor to profits for legacy gas, combustion turbines, and combined cycle plants.

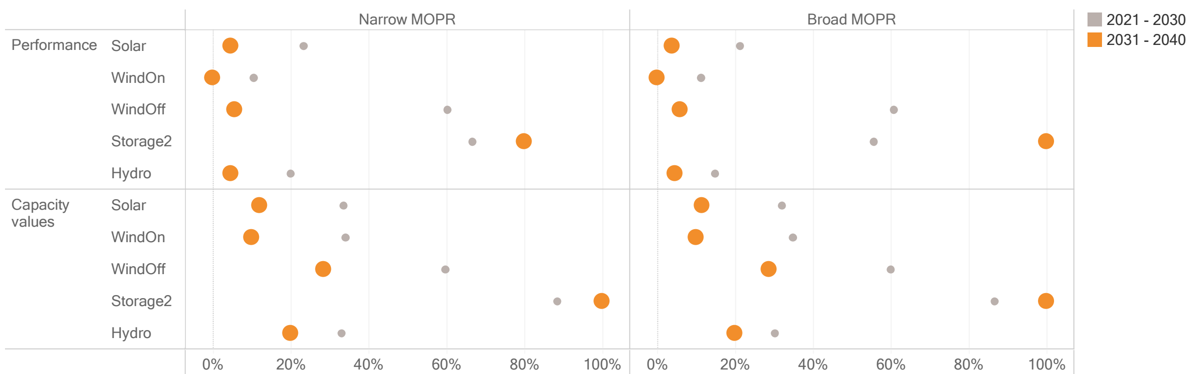
Figure 8.9: Profit breakdown by year, 2019-2040



In subsequent years the capacity price remains high. Figure 8.10 displays the performance and capacity values for renewables and storage by decade. Renewable resources have low capacity values and cannot enter the market in large enough quantities to replace traditional technologies. As seen in Figure 8.5, entry is at the cap throughout the simulation period. Capacity prices have to remain high to keep conventional resources from exiting the market.

Additionally, capacity values decline over time as more renewables enter the market and their performance decreases. Battery storage mitigates the reliability issues from the intermittency of renewables; its performance and capacity value rise over time. Interestingly, the performance of pump-storage declines as renewables enter the market because these resources are scheduled day-ahead and not optimized in real-time. In the last eight years of the simulation, when renewable performance is lowest, combined cycle plants with carbon capture replace renewables in the entry queue, easing the procurement of adequate resources.

Figure 8.10: Performance and capacity values by decade



The capacity-to-peak load ratio drops gradually over time. The drop happens for two reasons. First, the market has many resources in 2019. The capacity-to-peak load ratio is 11%, and the installed reserve margin is 20%. Therefore, we should expect a reduction in the equilibrium reserve margin. Second, procuring capacity is expensive because thermal resources' energy offsets are low. As the reserve margin declines, combined cycle energy profits recover. Still, thermal resources continue to earn relatively little money in the energy market, requiring large capacity payments.

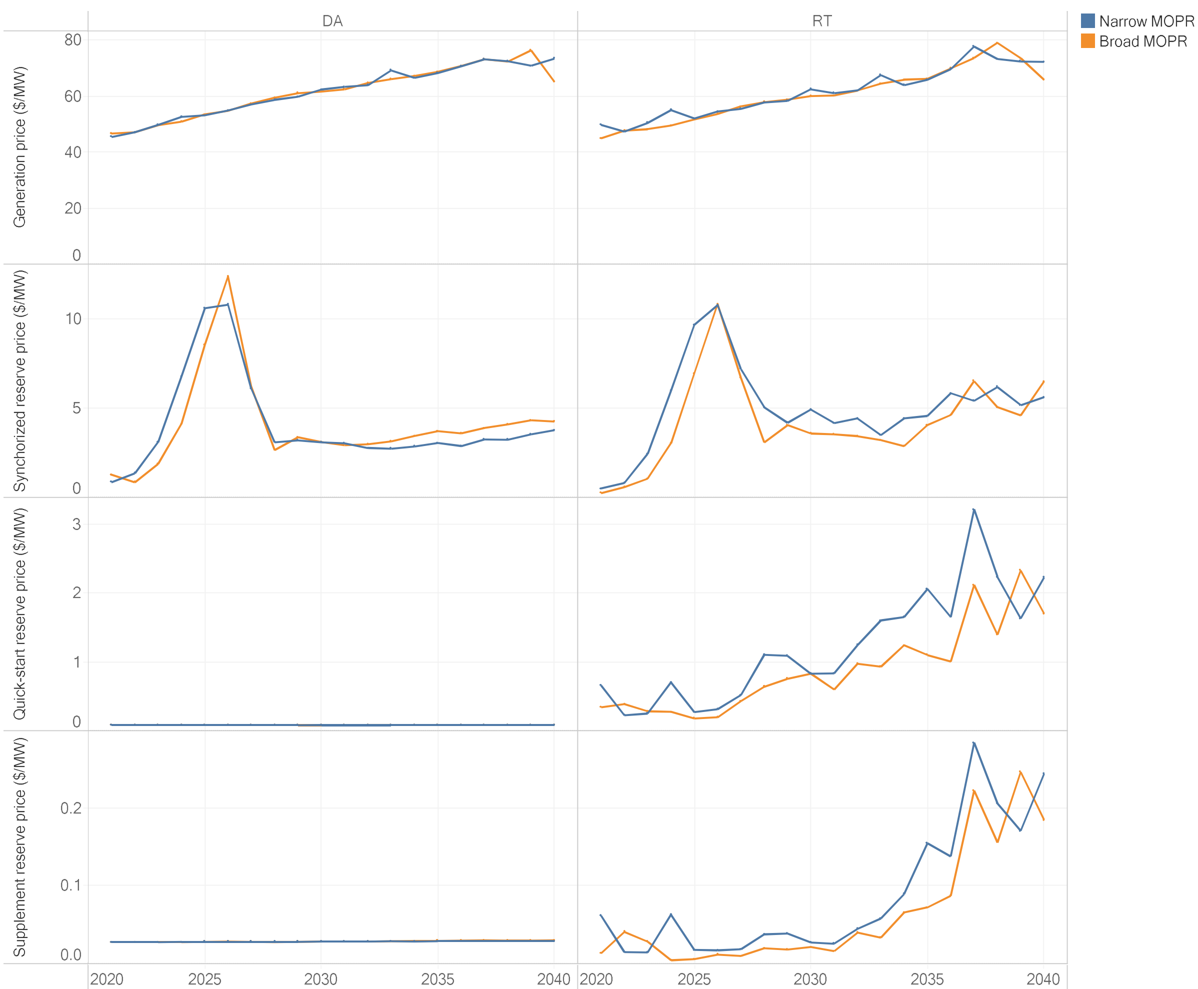
While the capacity-to-peak load ratio decreases over time, the installed reserve margin rises continuously as renewable resources enter the market, reaching 66% in 2040. The divergence between the capacity-to-peak load ratio and the installed reserve margins reveals the latter's limits as a reliability metric. While thermal resources have similar capacity values irrespective of technology, renewables' contribution to capacity varies significantly across technologies. Also, it changes over time depending on the correlation with load, total nameplate, and storage availability. Correctly measuring performance is crucial to design a capacity market suited for the energy transition.

The broad MOPR has little impact on capacity prices. By contrast, it results in a higher reserve margin—about two percentage points higher in most years. The main implication of broad MOPR is not to raise the capacity price. Instead, broad MOPR excludes non-clearing units from the resource adequacy calculation and brings in extra resources without letting the capacity price decrease. The additional resources might be desirable if they improved reliability. However, as shown below, reliability is the same with the broad MOPR as with the narrow MOPR. This result should be expected. In the narrow MOPR case, the capacity market already guarantees that the system operator procures adequate resources. The broad MOPR forces the system operator to exclude some units when assessing resource adequacy and procure more resources than needed.

Energy and reserves prices and reliability

Figure 8.11 shows the average annual prices for energy and reserves by year in the day-ahead and real-time markets. The generation price almost doubles between 2019 and 2040, mainly due to the rising carbon price. The slope of the generation price curve in the real-time market is essentially the same as in the day-ahead market, indicating that the increase is mainly due to the carbon price. However, we observe a slightly steeper trend for synchronized reserves in the real-time market relative to the day-ahead market, suggesting that increasing net-load volatility due to renewables also plays a role.

Figure 8.11: Generation and reserves prices by year, 2019-2040 (day-ahead and real-time markets)



The price for synchronized reserves starts rising in 2024 and then reverts to trend. The initial increase is due to the entry of renewable resources and the exit of plants providing ramping capability, e.g., combined cycle (see Figure 8.5). Then, in 2026 batteries begin entering the market, lowering the price for this reserve product.

The zero prices for quickstart and supplemental reserves in the day-ahead markets are expected. Units providing these services do not incur costs, and supply exceeds the maximum ORDC quantities for those products—we do not model the opportunity cost of providing reserves. Instead, in the real-time market, prices gradually rise, again due to the increasing net-load volatility related to renewable entries. The carbon price should play no role in these markets because units proving quickstart and supplemental reserves are offline and do not burn fuel (except for some supplemental reserves provided by online resources). Finally, the price for supplemental reserves in the real-time market sharply increases after 2032. This sharp rise is due to combined cycle units replacing coal units in the merit order as coal resources exit the market (see Figures 4.3 and 8.5). Combined cycle units have higher and steeper bid curves than coal, even in the last few years of the simulation when the carbon price is highest.

Figure 8.12 displays duration curves for generation, reserve prices, and shortages by decade. Prices are below \$150/MWh about 99.8% of the time. The market preserves the long-term reliability of service to load despite the seemingly low capacity-to-peak load ratio (see Figure 8.8). Figure 8.13 zooms into the right tail of the duration curves to better highlight these facts, showing 99.5% and above. Shortage never occurs day-ahead and is extremely rare in real-time, occurring once every 10,000 hours.

Figure 8.12: Duration curves for generation price and shortage by decade

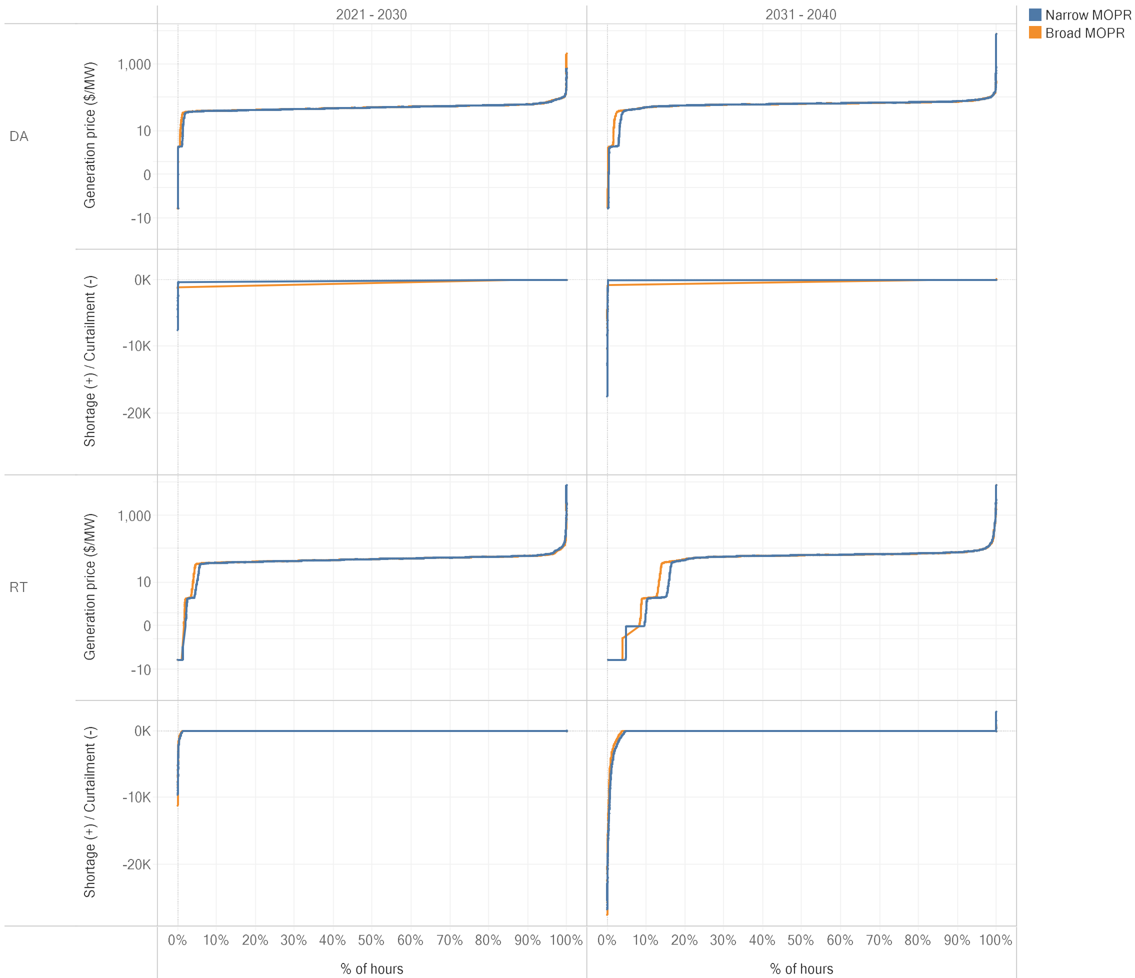


Figure 8.13: Right tail duration curves for generation price and shortage by decade

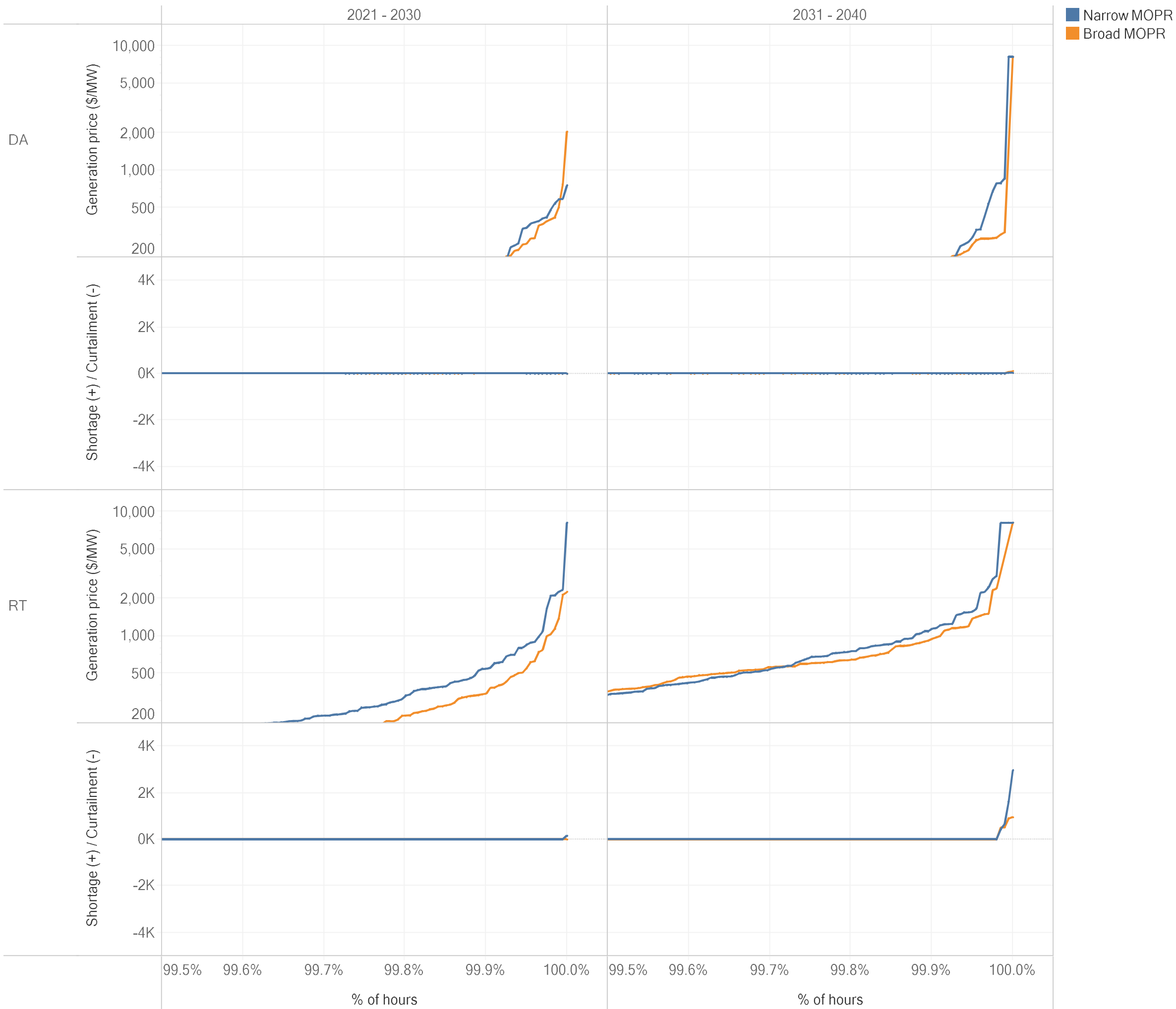


Figure 8.14 shows the frequency of price spikes—five-minute intervals with prices above $100/MWh. As the share of renewables grows, price spikes become more frequent. As expected, real-time price spikes are more frequent and more extreme than day-ahead price spikes. Figure 8.15 zooms in to see the frequency of prices above $400/MWh. The conclusion is the same. Price spikes primarily occur in real-time and grow in frequency with higher renewable penetration. The frequency of extreme price spikes above $1600/MWh does not increase as renewable resources enter the market. Batteries step in to provide the flexibility required with more renewable resources. In the last four years of the simulation, the frequency of prices between $100/MWh and $200/MWh grows substantially as combined cycle plants replace exiting coal units in the merit order.

The broad MOPR does not affect the average generation price. It slightly reduces reserve prices and the incidence of price spikes at the cost of procuring extra capacity. Importantly, the frequency of shortages remains unchanged. Thus, the additional resources procured in the capacity market relative to the narrow MOPR case do not enhance reliability.

Figure 8.14: Price spikes in day-ahead and real-time markets by year, 2019-2040

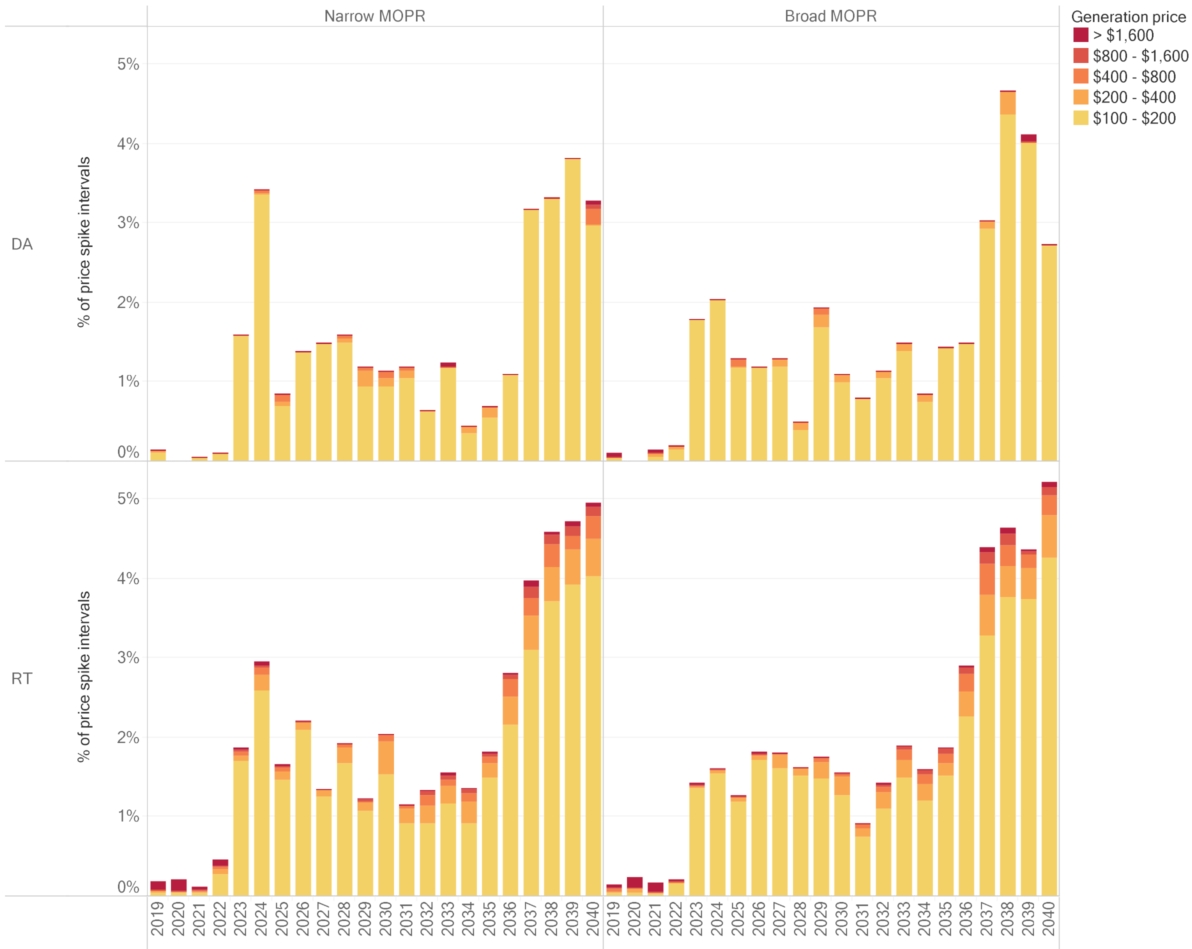
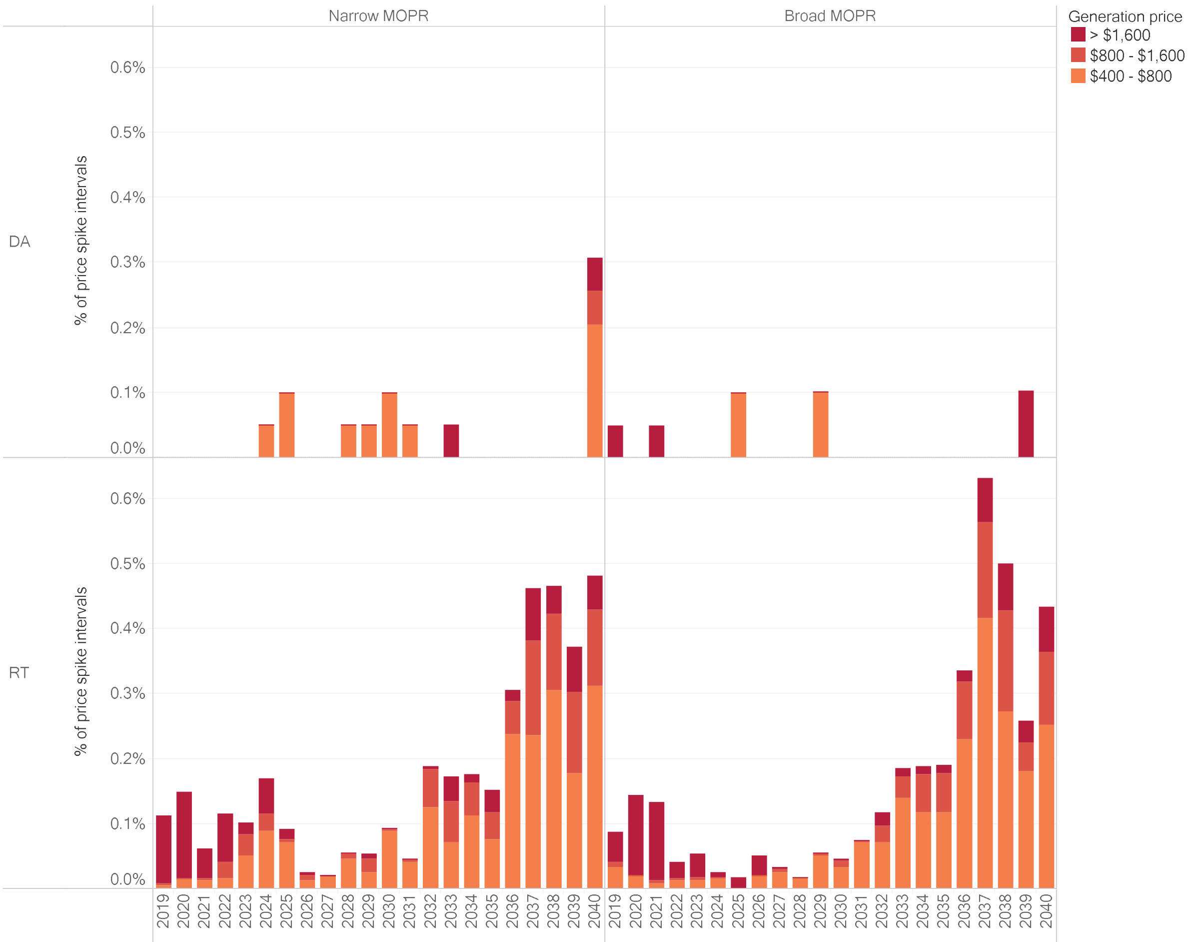


Figure 8.15: High price spikes in day-ahead and real-time markets by year, 2019-2040



Components of the electricity price

Figure 8.16 shows the all-in electricity price and its components by year. The all-in price includes all payments made to generation and storage units for energy, reserves, and capacity. We split energy into base energy when the energy price is below $200/MWh and peak energy when the energy price is above $200/MWh. As expected, the main price components are capacity and base energy. Uplift, reserves, and peak energy are much less significant. The share of capacity payments is 13% in 2019 and increases over time, reaching 33% in 2040. These payments are needed along the energy transition to sustain existing resources while renewables, batteries, and later combined cycle with carbon capture and sequestration gradually enter the market. On average, capacity payments account for 28% of generation and storage revenues over the simulation period.

Figure 8.16: Electricity price components by year, 2019-2040

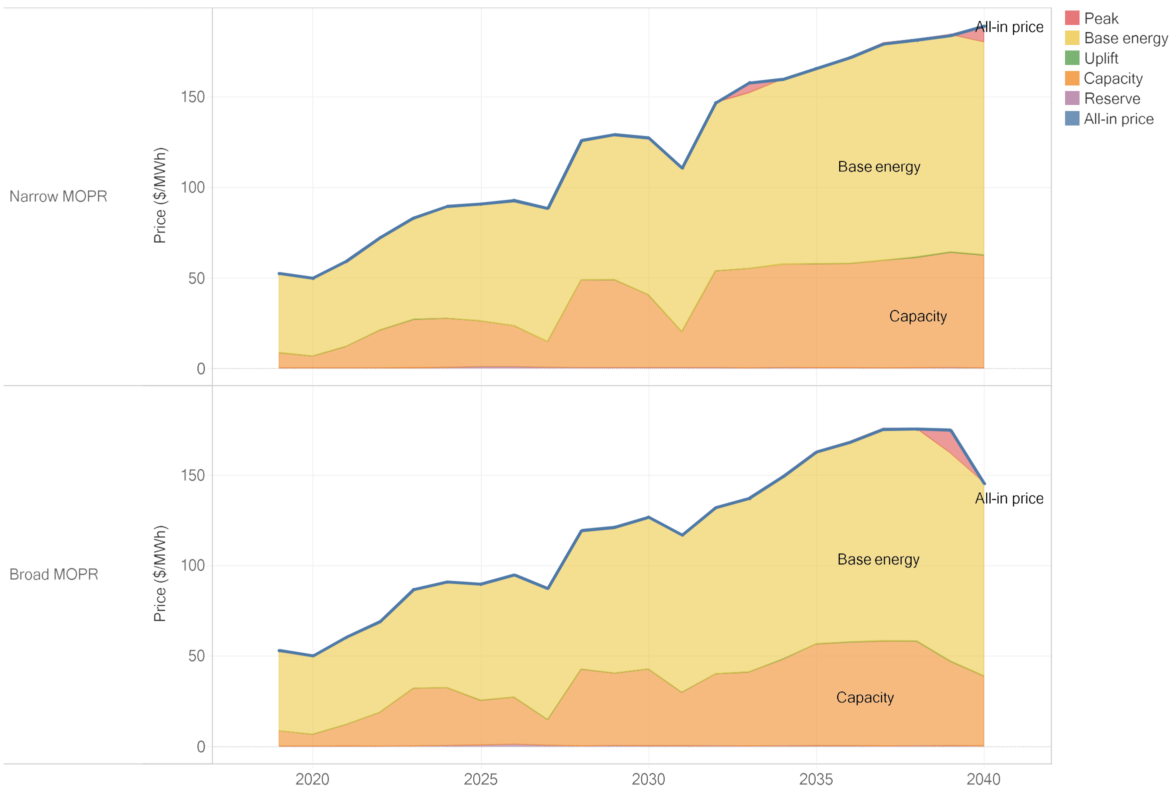


Figure 8.16 offers an alternative breakdown of the all-in electricity price into costs and profits. The all-in price is the amount paid by consumers for reliable electricity as the market transitions to a lower emission energy mix, not accounting for carbon rebates. It is $52.5/MWh in 2019 and climbs to $189.3/MWh in 2040. The increase reflects several trends.

Between 2019 and 2040, load rises by about 13%, while generation costs grow by 33%, from $24/MWh to $32/MWh. The carbon price props up the entry of renewables that have zero marginal cost but increases the generation costs of coal and gas-powered plants. The latter effect dominates.

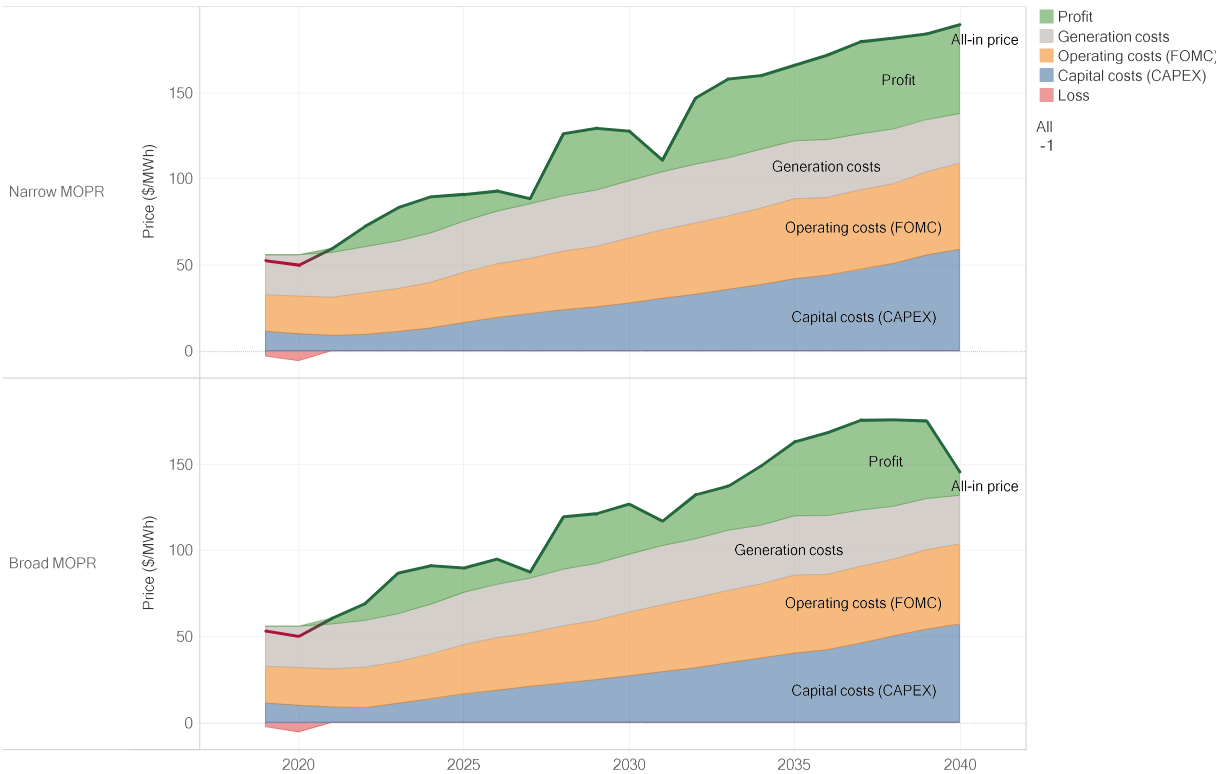
Second, operating costs increase from $21/MWh in 2019 to $50/MWh in 2040. Coal and nuclear have rising operating costs due to technical obsolescence but are kept in the market to ensure reliability and satisfy states' preferences. In addition, renewables, batteries, and combined cycle with carbon capture and sequestration have much higher operating costs per unit of nameplate capacity than traditional technologies. Renewables also have lower capacity values, meaning that more nameplate capacity needs installing for each conventional retiring unit.

Third, capital costs increase from $11/MWh to $59/ MWh as investments are made to transition to a lower emission energy mix. Again, the lower capacity values of renewables mean that these investments are large.

Finally, profits are negative or nearly zero in the first three years of the simulation because capacity prices are exogenously low. Then capacity prices rise to keep existing resources viable and shore up reliability while the capacity-to-peak load gradually decreases. As a result, capacity and energy payments increase over time. In 2040 profits amount to $52/MWh.

As expected, the broad MOPR does not significantly affect the all-in price and its components since it leaves the energy mix essentially unchanged. The capacity price drop in 2040 reduces capacity payments and the all-in price. This reduction is transitory, as also seen in 2027 and 2031.

Figure 8.17: Electricity cost components by year, 2019-2040



9 Conclusion

A transition to renewable generation is essential to achieve climate goals. Decarbonization of electricity is a requirement of decarbonization in other sectors such as transport. Our model enables us to evaluate the impact of market design and climate policy on the evolution of the market over several decades. Although initially developed to answer questions in the PJM market, our methods can be applied to study electricity markets worldwide.

References

Allcott, Hunt (2012) "Real-time Pricing and Electricity Market Design," Working Paper, New York University. Accessed 2 Jan 2021.

American Clean Power Association (2020) *Renewable Energy and Infrastructure Policy Scenario Analysis*, Wood Mackenzie. Accessed 2 Jan 2021.

Anderson, Robert and Neal Wolkoff (2019) "Report of the Independent Consultants on the Greenhat Default," PJM Interconnection. Accessed 2 Jan 2021.

Ausubel, Lawrence M, Peter Cramton, Marek Pycia, Marzena Rostek, and Marek Weretka (2014) "Demand Reduction and Inefficiency in Multi-Unit Auctions," *Review of Economic Studies*, 81:4, 1366-1400.

Bartlett, Jay (2019) "Reducing Risk in Merchant Wind and Solar Projects through Financial Hedges," Working Paper, Resources for the Future. Accessed 2 Jan 2021.

Bloom, Aaron (2018) "Interconnections Seam Study," NREL, presented at TransGrid-X Symposium, Ames, Iowa, July 26. Accessed 2 Jan 2021.

Borenstein, Severin (2002) "The Trouble with Electricity Markets: Understanding California's Restructuring Disaster," *Journal of Economic Perspectives*, 16:1, 191-211.

Burger, Scott P, Christopher R Knittel, and Ignacio J Perez-Arriaga (2020) "Quantifying the Distributional Impacts of Rooftop Solar PV Adoption Under Net Energy Metering," Working Paper, MIT. Accessed 26 Jan 2021.

Burger, Scott P, Christopher R Knittel, Ignacio J Perez-Arriaga, Ian Schneider, and Frederik vom Scheidt (2020) "The Efficiency and Distributional Effects of Alternative Residential Electricity Rate Designs," *Energy Journal*, 41. Accessed 26 Jan 2021.

Carbon Pricing Leadership Coalition (2017) *Report of the High-Level Commission on Carbon Prices*. Washington, DC: World Bank. Accessed 2 Jan 2021.

Caspary, Jay, Michael Goggin, Rob Gramlich, Jesse Schneider (2021) "Disconnected: The Need for a New Generator Interconnection Policy," *Americans for a Clean Energy Grid*. Accessed 12 Jan 2021.

CDP (2017) *Putting a Price on Carbon: Integrating Climate Risk into Business Planning*. Accessed 19 Jan 2021.

Chen, Cong, Lang Tong, Ye Guo (2021) "Pricing Energy Storage in Real-time Market," *arXiv*:2101.10151. Accessed 30 Jan 2021.

Clack, Christopher TM, Aditya Choukulkar, Brianna Cote, and Sarah A McKee (2020) "WIS:dom-P v9.0: Weather-Informed Energy Systems for Design, Operations, and Markets," Vibrant Clean Energy LLC. Accessed 30 Jan 2021.

Corneli, Steven (2020) "A PRISM-based Configuration Market for Rapid, Low Cost, and Reliable Electric Sector Decarbonization," WRI/RFF Workshop on Market Design for the Clean Energy Transition. Accessed 2 Jan 2021.

Crampes, Claude and Jean-Michel Trochet (2019) "Economics of Stationary Electricity Storage with Various Charge and Discharge Durations," *Journal of Energy Storage*, 24, 100746.

Cramton, Peter (2004) "Competitive Bidding Behavior in Uniform-Price Auction Markets," *Proceedings of the Hawaii International Conference on System Sciences*, January. Accessed 2 Jan 2021.

Cramton, Peter (2017) "Electricity Market Design," *Oxford Review of Economic Policy*, 33:4, 589–612. Accessed 2 Jan 2021.

Cramton, Peter (2021a) "A Forward Energy Market for a Reliable Grid," Working Paper, University of Cologne. Accessed 26 Jun 2021.

Cramton, Peter (2021b) "Affidavit of Peter Cramton on the PJM MOPR Proposal," submitted by PJM Interconnection LLC before the Federal Energy Regulatory Commission. Accessed 26 Jun 2021.

Cramton, Peter, Axel Ockenfels, and Steven Stoft (2013) "Capacity Market Fundamentals," *Economics of Energy & Environmental Policy*, 2:2, September. Accessed 2 Jan 2021.

Cramton, Peter, David JC MacKay, Axel Ockenfels and Steven Stoft (2017) *Global Carbon Pricing—The Path to Climate Cooperation*, MIT Press. Accessed 2 Jan 2021.

Daniel, Joseph (2018) "The Coal Bailout Nobody is Talking About," Union of Concerned Scientists, September 24. Accessed 2 Jan 2021.

Daniel, Joseph (2019) "The Billion-Dollar Coal Bailout Nobody is Talking About: Self-Committing in Power Markets," Union of Concerned Scientists, June 3. Accessed 2 Jan 2021.

Dyson, Mark, Alexander Engel, and Jamil Farbes (2018) *The Economics of Clean Energy Portfolios*, Rocky Mountain Institute. Accessed 2 Jan 2021.

Ela, Erik, Michael Milligan, Aaron Bloom, Audun Botterud, Aaron Townsend, Todd Levin, and Bethany A. Frew (2016) "Wholesale Electricity Market Design with Increasing Levels of Renewable Generation: Incentivizing Flexibility in System Operations," *Electricity Journal*, 29, 51-60.

Energy Information Administration (2020a) *Annual Energy Outlook 2020*, Washington, DC: U.S. Department of Energy. Accessed 2 Jan 2021.

Energy Information Administration (2020b) *Capital Cost and Performance Characteristic Estimates for Utility Scale Electric Power Generating Technologies*, Washington, DC: U.S. Department of Energy. Accessed 2 Jan 2021.

Energy Information Administration (2020c) *Assumptions to the Annual Energy Outlook 2020: Electricity Market Module*. Washington, DC: U.S. Department of Energy. Accessed 2 Jan 2021.

Environmental Protection Agency (2018) *Emission Factors for Greenhouse Gas Inventories*. Accessed 2 Jan 2021.

Environmental Protection Agency (2020) *Inventory of U.S. Greenhouse Gas Emissions and Sinks*. Accessed 20 Jan 2021.

Federal Energy Regulatory Commission (2002) "Working Paper on Standardized Transmission Service and Wholesale Electric Market Design," March.

Federal Energy Regulatory Commission (2019) "Order Establishing Just and Reasonable Rates," Dockets EL16-49-000 and EL18-178-000.

G30 Working Group on Climate Change and Finance (2020) "Mainstreaming the Transition to a Net-Zero Economy," Washington, DC: Group of Thirty. Accessed 25 Jan 2021.

Gates, Bill (2021) *How to Avoid A Climate Disaster*, Penguin Random House.

Gimon, Eric (2020) "Let's Get Organized! Long-Term Market Design for a High Penetration Grid," WRI/RFF Workshop on Market Design for the Clean Energy Transition. Accessed 2 Jan 2021.

Gorman, Will, Andrew Mills, Mark Bolinger, Ryan Wiser, Nikita G Singhal, Erik Ela, Eric O'Shaughnessy (2020) "Motivations and Options for Deploying Hybrid Generator-plus-battery Projects within the Bulk Power System," *Electricity Journal*, 33, 106739. Accessed 2 Jan 2021.

Gramlich, Rob and Jay Caspary (2021) "Planning for the Future: FERC's Opportunity to Spur More Cost-Effective Transmission Infrastructure," Americans for a Clean Energy Grid. Accessed 30 Jan 2021.

Gramlich, Rob and Michael Goggin (2019) "Too Much of the Wrong Thing: The Need for Capacity Market Replacement or Reform," Grid Strategies LLC. Accessed 2 Jan 2021.

Griffith, Saul (2020) *Rewiring America*. Accessed 2 Jan 2021.

Hobbs, Benjamin F, Ming-Chi Hu, Javier G Inon, Steven E Stoft, and Murty P Bhavaraju (2007) "A Dynamic Analysis of a Demand Curve-Based Capacity Market Proposal: The PJM Reliability Pricing Model," *IEEE Transactions on Power Systems*, 22, 3-14.

Hogan, William W (2013) “Electricity Scarcity Pricing Through Operating Reserves,” *Economics of Energy and Environmental Policy*, 2:2, 65-86.

Holmberg, Par and Robert A Ritz (2020) “Optimal Capacity Mechanisms for Competitive Electricity Markets,” *Energy Journal*, 41, 33-66.

Intergovernmental Panel on Climate Change (2014) *Fifth Assessment Report*. Accessed 2 Jan 2021.

Intergovernmental Panel on Climate Change (2018) *Global Warming of 1.5°C.* Accessed 2 Jan 2021.

International Energy Agency (2020) *Energy Technology Perspectives 2020*. Accessed 2 Jan 2021.

Inzunza, Andres and Christopher R Knittel (2020) “Distributional Effects of Net Metering Policies and Residential Solar Plus Behind-the-Meter Storage Adoption,” Working Paper, MIT. Accessed 26 Jan 2021.

IRENA (2020) *Renewable Power Generation Costs in 2019*, International Renewable Energy Agency, Abu Dhabi. Accessed 2 Jan 2021.

Jenkins, Jesse D, Max Luke, and Samuel Thernstrom (2018) “Getting to Zero Carbon Emissions in the Electric Power Sector,” *Joule*, 2, 2498-2510.

Jha, Akshaya, and Frank A Wolak (2020) “Can Forward Commodity Markets Improve Spot Market Performance? Evidence from Wholesale Electricity,” Working Paper, Stanford University. Accessed 2 Jan 2021.

Joskow, Paul (2008) “Capacity payments in imperfect electricity markets: Need and design.” *Utilities Policy*, 16, 159-170.

Junge, Cristian, Dharik Mallapragada, and Richard Schmalensee (2021) “Energy Storage Investment and Operation in Efficient Electric Power Systems,” Working Paper, MIT. Accessed 26 Jan 2021.

Larson, E, C Greig, J Jenkins, E Mayfield, A Pascale, C Zhang, J Drossman, R Williams, S Pacala, R Socolow, EJ Baik, R Birdsey, R Duke, R Jones, B Haley, E Leslie, K Paustian, and A Swan (2020) “Net-Zero America: Potential Pathways, Infrastructure, and Impacts,” Interim Report, Princeton University. Accessed 19 Jan 2021.

Lombardi, Francesco, Bryn Pickering, Emanuela Colombo, and Stefan Pfenniger (2020) “Policy Decision Support for Renewables Deployment through Spatially Explicit Practically Optimal Alternatives,” *Joule*, 4, 1-23.

MacKay, David JC, Peter Cramton, Axel Ockenfels and Steven Stoft (2015) “Price Carbon—I will if you will,” *Nature*, 526, 315-316. Accessed 2 Jan 2021.

Mays, Jacob (2021) “Missing Incentives for Flexibility in Wholesale Electricity Markets,” *Energy Policy*, forthcoming. Accessed 26 Jan 2021.

Mays, Jacob, David P. Morton, and Richard P O’Neill (2019) “Asymmetric Risk and Fuel Neutrality in Electricity Capacity Markets,” *Nature Energy*, 4, 948-956.

McPherson, Madeleine, Brendan McBennett, Devon Sigler, and Paul Denholm (2020) “Impacts of Storage Dispatch on Revenue in Electricity Markets,” *Journal of Energy Storage*, 31, 101573. Accessed 2 Jan 2021.

Milligan, Michael, Bethany A Frew, Aaron Bloom, Erik Ela, Audun Botterud, Aaron Townsend, and Todd Levin (2016) “Wholesale Electricity Market Design with Increasing Levels of Renewable Generation: Revenue Sufficiency and Long-term Reliability,” *Electricity Journal*, 29, 26-38.

Mills, Andrew D, Joachim Seel, Dev Millstein, James Hyungkwan Kim, Mark Bolinger, Will Gorman, Yuhan Wang, Seongeun Jeong, and Ryan Wiser (2021) “Solar-to-Grid: Trends in System Impacts, Reliability, and Market Value in the United States,” Lawrence Berkeley National Laboratory. Accessed 7 Feb 2021.

Miotti, Marco and Jessika E Trancik (2021) *Carboncounter.com*. Accessed 20 Jan 2021.

Monitoring Analytics (2020) *2019 State of the Market Report for PJM*. Accessed 2 Jan 2021.

Muratori, Matteo and Trieu Mai (2021) “The Shape of Electrified Transportation,” *Environmental Research Letters*, 16. Accessed 26 Jan 2021.

Murphy, Caitlin, Trieu Mai, Yinong Sun, Paige Jadun, Matteo Muratori, Brent Nelson, and Ryan Jones (2021) "Electrification Futures Study: Scenarios of Power System Evolution and Infrastructure Development for the United States," Golden, CO: National Renewable Energy Laboratory. Accessed 25 Jan 2021.

Musial, Walt, Donna Heimiller, Phillip Beiter, George Scott, and Caroline Draxl (2016) "Offshore Wind Energy Resource Assessment for the United States," NREL. Accessed 2 Jan 2021.

National Academies of Sciences, Engineering, and Medicine (2021) *The Future of Electric Power in the United States,* Washington, DC: The National Academies Press. Accessed 25 Feb 2021.

National Renewable Energy Laboratory (2020a) *2020 Standard Scenarios Report: A U.S. Electricity Sector Outlook*, Golden, CO. Accessed 2 Jan 2021.

National Renewable Energy Laboratory (2020b) *2020 Annual Technology Baseline*, Golden, CO. Accessed 2 Jan 2021.

NuScale (2020). Portland, OR. Accessed 23 Mar 2020.

Pechman, Carl (2021) "Whither the FERC? Overcoming the Existential Threat to Its Magic Pricing Formula through Prudent Regulation," National Regulatory Research Institute. Accessed 22 Jan 2021.

Phadke, Amol, Umed Paliwal, Nikit Abyuankar, Taylor McNair, Ben Paulos, David Wooley, Ric O'Connell (2020) *2035 Report*, Goldman School of Public Policy, University of California Berkeley. Accessed 2 Jan 2021.

Pierpont, Brendan (2020) "A Market Mechanism for Long-Term Energy Contracts to Support Electricity System Decarbonization," WRI/RFF Workshop on Market Design for the Clean Energy Transition. Accessed 2 Jan 2021.

PJM Interconnection (2016) *PJM Load Forecast Report*, January. Accessed 2 Jan 2021.

PJM Interconnection (2017) *PJM Load Forecast Report*, January. Accessed 2 Jan 2021.

PJM Interconnection (2018) *PJM Load Forecast Report*, January. Accessed 2 Jan 2021.

PJM Interconnection (2019) *PJM States See Big Possibilities in Wind*. Accessed 2 Jan 2021.

PJM Interconnection (2020a) *PJM Load Forecast Report*, January. Accessed 2 Jan 2021.

PJM Interconnection (2020b) *Compliance filing concerning the minimum offer price rule*. Accessed 2 Jan 2021.

PJM Interconnection (2020c) *Enhanced price formation in reserve markets of PJM interconnection*, Docket No. EL19-58-000. Accessed 4 Feb 2021.

PJM Interconnection (2020d) *2020 Load Forecast Supplement*, Resource Adequacy Planning Department. Accessed 2 Jan 2021.

PJM Interconnection (2021a) "Summary of the Updated PJM MOPR Proposal."

PJM Interconnection (2021b) "Flow Diagrams to Support the PJM MOPR Proposal."

Richstein, Jörn C, Emile J Chappin, and Laurens J de Vries (2014) "Cross-Border Electricity Market Effects Due to Price Caps in an Emission Trading System: An Agent-Based Approach, *Energy Policy*, 71, 139-158.

S&P Global Market Intelligence (2020) "Green Competition and Asset Valuation," Webinar on 15 Dec 2020. Accessed 2 Jan 2021.

Sepulveda, Nestor A, Jesse D Jenkins, Fernando J de Sisternes, and Richard K Lester (2018) "The Role of Firm Low-Carbon Electricity Resources in Deep Decarbonization of Power Generation," *Joule*, 2, 2403-2420.

Spees, Kathleen, Samuel A Newell, Walter Graf, and Emily Shorin (2019) "How States, Cities, and Customers Can Harness Competitive Markets to Meet Ambitious Carbon Goals," Brattle Report for NRG. Accessed 2 Jan 2021.

Tanenblatt, Eric, Crawford Schneider, Michael Malterer, Henry Litong Chen, Michael Ledgett, Kurt B. Gerstner, Ben Allen, Marcell Szőnyi, Şahin Ardıyok, and Karol Laskowski (2021) *Global Guide to Autonomous Vehicles 2021*, Dentons. Accessed 30 Jan 2021.

Tesla (2020). "Powerpacks, overall system specs". Accessed 10 Jan 2021.

Tierney, Susan F (2020) "Wholesale Power Market Design in a Future Low-Carbon Electricity System," WRI/RFF Workshop on Market Design for the Clean Energy Transition. Accessed 2 Jan 2021.

U.S. Department of Energy (2020) *Energy Storage Grand Challenge: Energy Storage Market Report*. Accessed 2 Jan 2021.

Williams, James H, Ryan A Jones, Ben Haley, Gabe Kwok, Jeremy Hargreaves, Jamil Farbes, and Margaret S Torn (2021) "Carbon-Neutral Pathways for the United States," *AGU Advances*, 2. Accessed 7 Feb 2021.

Wilson, Robert (2002) "Architecture of Power Markets," *Econometrica*, 70:4, 1299-1340.

Wolak, Frank A (2020) "Market Design in a Zero Marginal Cost Intermittent Renewable Future," Working Paper, Stanford University, October 15. Accessed 2 Jan 2021.

Wolak, Frank A (2021a) "Market Design in an Intermittent Renewable Future," *IEEE Power and Energy Magazine,* January-February, 29-40.

Wolak, Frank A (2021b) "Wholesale Electricity Market Design," in Glachant JM, Joskow P, Pollitt M., eds., *Handbook on the Economics of Electricity*, Northhampton: Edward Elgar Publishing. Accessed 2 Jan 2021.

Appendix: Multi-year simulation

Let $M_t = \left\{ S_t, \left\{ P^C_{t+k} \right\}_{k=0}^{h}, H_t \right\}$ be the electricity market state in delivery year *t*, where $H_t = \{ J_{t+1}, P_t^F, P_t^C, P_t^O, \bar{P}_t, D_t, PRD_t, B_{t+1} \}$. Denote with $\pi^{j,a}\left(M_t, g_t\right)$ the profits per MW of nameplate capacity earned in *t* by a unit that operates vintage *v* of technology *j*:

$$\pi^{j,v}\left(M_t, g_t\right) = EP^{j,v}\left(M_t\right) + 365 \times g \times cv_t^j \times P_t^C$$

Profits are equal to energy profits plus capacity payments. Capacity is a daily product, hence the multiplication by 365; *g* is a dummy that equals one if the unit cleared in the capacity auction in year *t-p* and cv_t^j is the capacity value measuring the technology *j* capacity contribution.

The marginal cost is assumed to be constant and equal to:

$$mgc_{t,m}^{j,v} = c^{j,m}\left[VOMC^{j,v} + HeatRate^{j,v} \times \left(P_t^{F,j} + EmissionFactor^j \times P_t^O \right) \right]$$

Where $c^{j,m}$ is the technology and month-specific calibration coefficient matching the energy offer curve to the data—see the subsection Energy offer curves. The marginal cost is an argument passed to the energy market model to construct offer curves.

We use τ_j^E, τ_j^D, τ_j^P for the plant construction time, service life, and repayment period, $\phi^{j,v}$ for the scrap value per MW of nameplate capacity, δ for the discount factor, and $C_t^{j,v,a} = capex_t^{j,v} + FOMC^{j,v,a}$ for the annual payment made by the investor to cover the entry fees and fixed operation and maintenance costs. Fixed operation and maintenance costs depend on the age of the plant, *a*. A new unit pays the entry cost in constant installments, beginning with construction, and goes online after *h* periods. Then, annual entry fees are equal to the capital recovery factor times CAPEX if $t < v + h - \tau_j^E + \tau_j^P$ and zero otherwise. The capital recovery factor is routinely defined as:

$$\frac{\delta(1+\delta)^{\tau_j^P}}{(1+\delta)^{\tau_j^P} - 1}$$

A plant remaining technical life in year *t* is $l = v + h + \tau_{j(i)}^{D} - t$.

Value function and exit and entry decisions

The system operator runs the capacity market at the beginning of each year. The market state is M_{t-1} and investors know the frontier technology for the current year J_t and the capacity market parameters B_t.

Exit and entry decisions are announced in the capacity market and are irreversible. An existing plant deciding to retire stays in the energy market for *h* periods and then exits. A new resource goes online after *h* periods. The value function per MW of nameplate capacity of an existing plant *i* entering the capacity market solves the system of Bellman equations:

$$\text{for } l \le h: X^{j,v,a,l}\left(M_{t-1}, G_{i,t-1}\right) = E_{t-1}\left[\sum_{k=0}^{l-1} \delta^k \left[\pi^{j,v}(M_{t+k}, g_{i,t+k}) - C_t^{j,v,a+k}\right] + \delta^l \phi^{j,v,a+l}\right]$$

$$\text{for } l > h: V^{j,v,a,l}\left(M_{t-1}, G_{i,t-1}\right) = \max\Big\langle X^{j,v,a,h}\left(M_{t-1}, G_{i,t-1}\right),$$
$$E_{t-1}\left[\left[\pi^{j,v}(M_t, g_{i,t}) - C_t^{j,v,a}\right] + \delta V^{j,v,a+1,l-1}\left(M_t, G_{i,t}\right)\right]\Big\rangle$$

A plant decides to exit if the first branch (the “exit branch”) in the second equation is higher than the second branch (the “continuation branch”). In this case, the plant’s residual life is set to *h*, and the plant goes offline in year *t+h*.

The net present value of a new project is:

$$E^{j,v}\left(M_{t-1}\right) = -\frac{capex^{j,v}}{\tau_j^E} \delta^{h-\tau_j^E} \frac{1-\delta^{\tau_j^E}}{1-\delta} + \delta^h E_{t-1}\left[V^{j,v,0,\tau_j^D}\left(M_{t+h} \cup i; g_{t+h}\right)\right]$$

The investor considers the effect of her entry decision on the plant’s net present value. The investor commits to building the plant it has a positive net present value.

We reset seeds in each iteration of the capacity market so that a) incumbents are selected in the same order when rerunning the exit; b) the same size is drawn for a new project when rerunning the exit and entry subroutines.

Simplifying assumptions and value functions, revisited

We consider a perfect foresight equilibrium and assume that the complete list of resources S can be replaced with a small set of near-sufficient statistics, the resource mix $\tilde{S}$, for predicting energy profits and performance. Then we substitute $\tilde{M}_t = \left\{\tilde{S}_t, \left\{P_{t+k}^C\right\}_{k=0}^h, H_t\right\}$ in the value functions for incumbents and new entrants.

Expectation formation

Finally, we need to specify how investors form expectations for capacity prices and the resource mix. Since exit and entry decisions are taken *h* periods in advance, the list of energy resources S_{t-1}, including existing and planned units, encodes all the resources that will be active at time *t,…, t+h-1*. In addition, exit and entry decisions are consistent with the capacity price announced in *t* for delivery year *t+h*. Thus, in equilibrium, investors know at time *t* what resources will be active at time *t+h*. For *t+h+1* and beyond, we use the sequence updated at the end of the previous run of the model, $\left\{\tilde{S}_{t+h}, \tilde{S}_{t+h+1}, \ldots\right\} = \left\{\tilde{S}_{t+h}^{r-1}, \tilde{S}_{t+h+1}^{r-1}, \ldots\right\}$

. Similarly, capacity prices are known (or provisionally known) for *t*,…, *t+h*, and we set expectations for the following years using the sequence from the previous run.

Given expectations for future resource mix and capacity prices, investors can form expectations for energy market outcomes at each future date. Thus, they can estimate future energy offsets and performances and update their expectations for energy offsets, capacity values, and minimum offer price floors. Finally, investors can forecast their capacity payments and compute the net present value of future profits.

When we reach the end of the simulation horizon, we update expectations for the resource mix and capacity prices by simple exponential smoothing using the realization from the latest iteration, $\{\tilde{S}_t, P_t^C\}_{t=0}^T$:

$$\begin{aligned}\tilde{S}_t^r &= (1-\eta)\tilde{S}_t + \eta\tilde{S}_t^{r-1},\ t=0,1,\ldots \\ P_t^{C,r} &= (1-\eta)P_t^C + \eta P_t^{C,r-1},\ t=0,1,\ldots\end{aligned}$$

The sequence for the resource mix is initialized by setting expectations equal to the initial resource mix adjusted for demand changes over time:

$$\tilde{S}_k^{-1} = \hat{S}_0 \frac{D_k}{D_0},\ k=0,1,\ldots$$

The sequence for the capacity price is initialized at a constant value equal to the capacity price historical average. We run the simulation repeteadly until the distance between $\tilde{S}_k$ and $\tilde{S}_k^{r-1}$, and between P_k^C and $\tilde{P}_k^{C,r-1}$ falls below a critical value.

Energy offsets and capacity values

Energy offsets, EO^j, and capacity values, cv^j, for technology *j* are updated by Holt-Winters double exponential smoothing using the average energy profits EP^j, and average performance $perf^j$, of units in that class for the latest available year. The updating rule for energy offsets is defined as:

$$\begin{aligned}s_t^{EO,j} &= EO_h^j, \text{ for } t \le h \\ b_0^{EO,j} &= 0, \text{ for } t \le h \\ s_t^{EO,j} &= \alpha^{EO} EP_{t-h-1}^j + (1-\alpha^{EO})(s_{t-1}^{EO,j} + b_{t-1}^{EO,j}), \text{ for } t > h \\ b_t^{EO,j} &= \beta^{EO}(s_t^{EO,j} - s_{t-1}^{EO,j}) + (1-\beta^{EO})b_{t-1}^{EO,j}, \text{ for } t > h \\ EO_t^j &= s_t^{EO,j} + h \times b_t^{EO,j}\end{aligned}$$

The updating rule for capacity values is similar, except we need to restrict the range. Capacity values should not be negative and cannot exceed one. In the model, we allow for strictly positive technology-specific floors to capacity values $\underline{cv}^j$. Thus we impose two additional restrictions:

$$s_t^{cv,j} = \min\left\langle \max\left\langle \underline{cv}^j, \alpha^{cv} perf_{t-h-1}^j + (1-\alpha^{cv})(s_{t-1}^{cv,j} + b_{t-1}^{cv,j})\right\rangle, 1\right\rangle, \text{ for } t > h$$

$$b_t^{EO,j} = \min\left\langle \max\left\langle \beta^{cv}(s_t^{cv,j} - s_{t-1}^{cv,j}) + (1-\beta^{cv})b_{t-1}^{cv,j}, \frac{\underline{cv}^j - s_t^{cv,j}}{h}\right\rangle, \frac{1-s_t^{cv,j}}{h}\right\rangle, \text{ for } t > h$$

MOPR

The minimum offer price floor for a unit operating technology j is equal to the cost of new entry net of energy offsets, EO^j, adjusted for the capacity value if the resource has never cleared in the capacity market:

$$\text{Net-CONE}_t^j = \frac{C_t^{j,t-h} - EO_t^j}{cv_t^j}$$

If the unit has cleared, it is equal to the avoidable cost rate net of energy offsets adjusted for the capacity value:

$$\text{Net-ACR}_t^j = \frac{FOMC^{j,t-h} - EO_t^j}{cv_t^j}$$

Appendix: Data

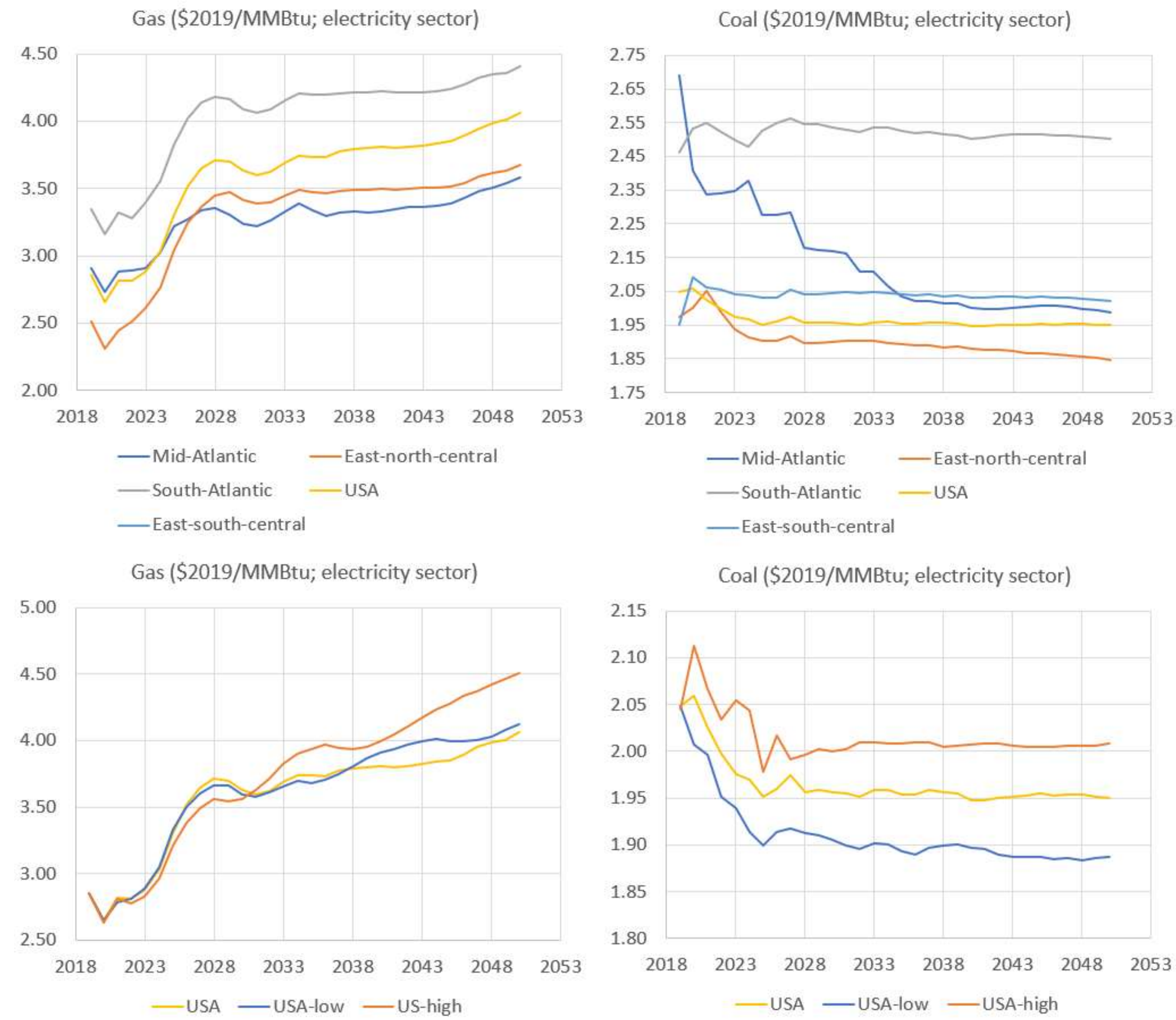
Fuel Prices, EIA (2020)

We consider three scenarios for the price of natural gas, coal and uranium. These scenarios are from EIA (2020a) and correspond to the “reference” scenario and to the “low oil price” and “high oil price” scenarios.

We use price series expressed in 2019 US dollars. For the reference scenario, series are available by sector of use and region. PJM footprint falls within the following four regions: Mid Atlantic (PA, NY, NJ; EIA, 2020a, Table 3.2), East North Central (OH, IN, MI, IL, WI; EIA, 2020a, Table 3.3), South Atlantic (DE, MD, WV, VA NC, SC, GA, FL; EIA, 2020a, Table 3.5) East South Central (AL, KY, MS, TN; EIA, 2020a, Table 3.6).[14] We select series for the power sector and average across these four regions using the fraction of nameplate capacity across states in PJM footprint as weights. Figure A.1, top panels, displays the input series for coal and uranium. The series for uranium does not vary across regions or scenarios and increases by approximately 0.2% per year – not reported. The figure for fuel prices in the main text displays the resulting series used in the simulation.

[14] Tables 3.2, 3.3, 3.5 and 3.6 are available at https://www.eia.gov/outlooks/aeo/tables_ref.php.

Figure A.1: Gas and coal prices different scenarios and regions



Sources: EIA (2020a).

For the low and high oil price scenarios the breakdown by region and sector of use is not available. We consider prices for the US power sector (EIA, 2020a, Table 3) and rescale those series using the ratio between the PJM series computed as above and the US series for the reference scenario: [15]

$$P_t^{PJM,c} = P_t^{USA,c} \frac{P_t^{PJM,ref}}{P_t^{USA,ref}}, \; c \in \{low,\ high\}$$

Figure A.1, bottom panels, displays the input series for coal and uranium for the US. The Figure in the main text displays the resulting series which we use in the simulation.

[15] Table 3 is available at https://www.eia.gov/outlooks/aeo/tables_side.php for the low and high oil price scenarios.

Technology, NREL (2020b)

We use series from NREL (2020b) for the entry cost (CAPEX), the fixed operation and maintenance cost (FOMC), the variable operation and maintenance cost (VOMC), heat rates for fossil resources and nuclear, and capacity factors for renewable technologies. We inflate prices from 2018 to 2019 US dollars using the consumer price index for the United States, 1.8% for 2019.

For coal, nuclear, combustion turbines, and combined cycle technologies, NREL series correspond to those underlying the annual energy outlook for 2020 (EIA, 2020a). The annual energy outlook makes assumptions for the degree of technological optimism and learning for each technology type, as detailed in EIA (2020c). Then, CAPEX and FOMC series are constructed based on production volumes simulated in the Electricity Market Module. Heat rates are essentially flat. VOMC is assumed to remain constant over time.

The representative technologies for thermal resources are described in EIA (2020b). We use series for ultra-supercritical coal (new coal), for combined cycle with and without carbon capture and storage, for combustion turbines without carbon capture and storage, and for nuclear.

For hydropower, photovoltaic, land-based wind, offshore wind, and battery storage, the series are internally developed by NREL (2020b). NREL considers a wide range of alternative scenarios described in NREL (2020a). In addition to the reference scenario, we also consider “high” and “low” cost series, corresponding to the cases of slow and fast technical progress for renewables, respectively. The description of the representative renewable technologies and their estimated potential across US regions are available on NREL website.

For hydropower, NREL reports series for powering non-powered dams (NPD), upgrading existing power dams, and greenfield developments (NSD). We select the series NSD-4, which correspond to the new development of large plants (MW>10 and dam head height>30ft). For solar we select the Kansas City series, which has an intermediate capacity factor (27% in 2019, as opposed to 22% and 35% for the Seattle and Daggett, CA series), and better approximates the average latitude of PJM’s footprint. For land-based wind we use the Class 7 series which correspond to sites with medium-to-low potential (CF 35% in 2019, as opposed to 17% and 52% for classes 10 and 1). Finally, based on recent initiatives in PJM (2019) footprint, and given the assessment of offshore wind potential across US regions in Musial, Beiter, Scott, Draxl (2016), we select the Class 1 series for offshore wind, corresponding to fixed bottom turbines installed at sites wind the highest wind potential.

Next generation nuclear

We rely on information from NuScale (2020). NuScale expects the first deployment to occur by mid-2020s. One reactor is expected to have a gross capacity of 60-MW and a typical plant to host 12 reactors. The construction cost is projected to be $3 billion for 684-MW of net capacity. Each reactor should have an output of 200-MW thermal per 60-MW electric. Accordingly, we set the expected time of commercial availability to 2030, the minimum efficient scale to 300-MW, the CAPEX to $4386/kW and the heat rate to 11.3 MMBtu/MWh. Next generation nuclear is characterized by enhanced modularity and a simplified design which are expected to lower operation costs. Accordingly, we set FOMC and VOMC at 20% below values for nuclear in NREL (2020b). We also consider a low-cost next generation nuclear scenario, where FOMC and VOMC are 50% below values for nuclear in NREL (2020b). Dispatch constraints are set equal to those computed from PJM energy offer data for nuclear.