About the Authors

Chiara Lo Prete is Associate Professor of Energy Economics in the John and Willie Leone Family Department of Energy and Mineral Engineering at The Pennsylvania State University. Her research centers on the economics of energy markets, with a focus on the areas of competition and design of electricity markets, natural gas market design to enhance grid reliability, geopolitics and energy security, and the impacts of environmental regulations on electric power generation. Recent work has centered on the development of mathematical models and application of empirical methods to study electricity market structures for resource adequacy and wind energy integration, interdependent natural gas and electric power systems, the weaponization of electricity trade, emission leakage and cross-product manipulation.

Karen Palmer is a senior fellow at Resources for the Future (RFF), director of the Electric Power Program, and an expert on the economics of environmental, climate and public utility regulation of the electric power sector. Her work seeks to improve the design of environmental and technology regulations in the sector and the development of new institutions to help guide the ongoing transition of the electricity sector. To these ends, she explores climate policy design, analyzes efficient ways to promote use of renewable and other clean sources of electricity, and investigates new market designs, new approaches to electricity pricing and regulatory reforms to pave the way for long-term decarbonization of electricity supply and electrification of the energy economy.

Molly Robertson is a senior research associate at RFF working on topics related to the electric power sector, including grid decarbonization, electrification, and electricity market design. She has also contributed to RFF’s growing work on equitable community transition and environmental justice. She holds a master’s in public policy from the University of Michigan’s Ford School.

Acknowledgements

Karen Palmer and Molly Robertson wish to acknowledge funding from the RFF Electric Power Program. Chiara Lo Prete would like to acknowledge funding from the National Science Foundation (Grant 1943992). We thank Peter Cramton, Steven Corneli, Natalia Fabra, Eric Gimon, Malcolm Keay, David Nelson, Brendan Pierpont, Kathleen Spees, Leigh Tesfatsion, and Frank Wolak for helpful feedback. We also thank participants in the Electric Power Innovation in a Carbon Free Society (EPICS) Thrust 3 workshop for helpful comments.
About RFF

Resources for the Future (RFF) is an independent, nonprofit research institution in Washington, DC. Its mission is to improve environmental, energy, and natural resource decisions through impartial economic research and policy engagement. RFF is committed to being the most widely trusted source of research insights and policy solutions leading to a healthy environment and a thriving economy.

The views expressed here are those of the individual authors and may differ from those of other RFF experts, its officers, or its directors.

Sharing Our Work

Our work is available for sharing and adaptation under an Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) license. You can copy and redistribute our material in any medium or format; you must give appropriate credit, provide a link to the license, and indicate if changes were made, and you may not apply additional restrictions. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use. You may not use the material for commercial purposes. If you remix, transform, or build upon the material, you may not distribute the modified material. For more information, visit https://creativecommons.org/licenses/by-nc-nd/4.0/.
Abstract

Existing wholesale electricity market designs are poorly suited to address challenges associated with the evolving resource mix. For example, recent scarcity events in the United States show that reliability challenges in renewable- and gas-dominated electric power systems arise not from the lack of generation capacity to serve peak customer demand, but from the lack of available capacity to provide the requisite energy at times of need. We review 11 proposed electricity market designs for the clean energy transition and compare them based on 10 criteria. Enhancing reliability in electric power systems with a significant amount of variable renewable energy requires incentivizing resource flexibility, both in investment and in operation. Electricity market structures should allow resources needed for reliability to earn adequate revenues to recover their variable and fixed costs. Good market designs also enable low-cost financing to support investments in capital-intensive resources that are instrumental in meeting decarbonization objectives. An additional property of well-designed markets is promoting short-run efficiency by reducing incentives to exercise market power and supporting efficient renewable curtailment outcomes. Besides achieving reliability, long-run efficiency, and short-run efficiency, some proposals in our review seek to achieve energy affordability objectives and integration with clean energy goals. Our evaluation highlights several open questions and directions for future research: the determination of mandatory purchase obligations of load-serving entities and associated enforcement mechanisms; the interplay between long-term hedging requirements and incentives for demand participation in real time; and the compatibility between long-term contract design and efficient operations in short-term energy markets.
Contents

1. Introduction 1

2. Traditional Approaches to Resource Adequacy and Revenue Sufficiency 5
   2.1. Scarcity Pricing 5
   2.2. Capacity Remuneration Mechanisms 6
       2.2.1. Capacity Payments 6
       2.2.2. Capacity Requirements 6
       2.2.3. Capacity Markets 7
       2.2.4. Reliability Options 7
       2.2.5. Strategic Reserves 8
       2.2.6. Incentivizing Capacity Performance 9

3. Evolving Market Approaches to Support Reliability, Flexibility, and Price Certainty 10
   3.1. Flexible Capacity Requirements 10
   3.2. Seasonal Forward Capacity Markets 11
   3.3. Basic Generation Service 12

4. Proposed Approaches to Meet the Challenges of the Clean Energy Transition 14
   4.1. An Integrated Clean Capacity Market 14
   4.2. Standardized Fixed-Price Forward Contracts 16
   4.3. A Forward Energy Market with Flow Trading 18
   4.4. A Two-Part Market for Energy 21
   4.5. A Market for Long-Term Energy Contracts 22
   4.6. Long-Term Contracts Based on the Attributes of Generation Technologies 24
   4.7. Market-Based Integrated Resource Planning 26
   4.8. An Organized Long-Term Market 28
   4.9. Dual Short-Term Market Designs 29
   4.10. Swing Contracts 32

5. Discussion 36

6. Conclusions 41

References 43
1. Introduction

Independent system operators (ISOs), regional transmission organizations (RTOs), and electric utilities serving as balancing authorities must ensure there are sufficient resources to meet the electricity demand of end-use customers at all times, accounting for scheduled and reasonably expected unscheduled outages of system elements and subject to standard industry targets for loss of load, such as the “one day in ten year” loss of load expectation.\(^1\) This concept is referred to as resource adequacy (NARUC, 2023; NERC, 2013).

To achieve resource adequacy, both traditionally regulated and competitive wholesale electricity markets have adopted frameworks to address two well-known problems: the missing money problem and the missing market problem (Newbery, 2016; Wolak, 2022). The missing money problem arises because revenues from the energy and ancillary services markets are not sufficient to encourage adequate investment in new generation capacity (Cramton and Ockenfels, 2012). For example, during scarcity conditions, short-term energy prices are unlikely to rise to politically unacceptable levels reflecting the estimated amount that customers would be willing to pay to avoid a power disruption (Roques and Finon, 2017). On the other hand, the missing market problem results from the retailers’ incentive to delay their electricity purchases to the short-term energy market. This arises for two reasons. First, price caps in the short-term energy market limit the cost of failing to procure adequate energy to meet expected demand before delivery. Second, during scarcity conditions, retailers are randomly curtailed regardless of the amount of energy procured in the forward market, creating an incentive to under-procure expected needs in the forward market (Wolak, 2022). The gap between maturities of long-term contracts (that generally do not extend

\(^1\) In April 1996, the Federal Energy Regulatory Agency (FERC) issued Orders 888 and 889, where it introduced the concept of an ISO to “operate the transmission systems of public utilities in a manner that is independent of any business interest in sales or purchases of electric power by those utilities” and to facilitate nondiscriminatory access to the electric transmission grid. FERC Order 2000, issued in December 1999, encouraged transmission-owning public utilities to voluntarily form and participate in an RTO and established a set of 12 technical requirements to obtain RTO status. While some differences exist between ISOs and RTOs (for example, ISOs are formed at the direction or recommendation of the FERC and do not petition the FERC for approval and thus may not meet the requirements needed to hold the designation of RTO), their basic functions are the same and include operating the electric grid within a defined region, ensuring nondiscriminatory access to transmission, optimally dispatching generation assets to keep demand and supply in balance, administering the region’s wholesale electricity markets, and ensuring grid reliability. In the rest of the paper, we will use RTOs to refer to both types of organizations.
beyond a few years; Keppler et al., 2022) and time frames for investing in generation capacity increases the cost of capital and distorts the energy mix toward less capital-intensive and less risky generation technologies (Neuhoff and De Vries, 2004).\(^2\) Resource adequacy frameworks attempt to address these problems in various ways. Vertically integrated utilities operating in traditionally regulated markets adopt a centralized procurement framework. In decentralized procurement regions like the California Independent System Operator (CAISO) and Southwest Power Pool (SPP), load-serving entities (LSEs) have an obligation to individually procure the capacity they need to meet their customers’ demand. Obligations are subject to guidance and standards set by a central authority, but there is no central bidding process. The Electric Reliability Council of Texas (ERCOT) reformed its energy market so that peak prices better reflect the scarcity of capacity, while other regions like ISO New England (ISO-NE), the Midcontinent Independent System Operator (MISO), the New York Independent System Operator (NYISO), and PJM created mechanisms that provide additional remuneration for capacity on top of energy sales through centralized forward capacity markets.

Despite many differences among these approaches, all traditional resource adequacy mechanisms focus on procuring sufficient generation capacity to serve peak demand plus a reserve margin.\(^3\) While such approaches have been successful at maintaining resource adequacy in thermal generation–dominated systems under most contingencies, they are poorly suited to address challenges associated with the evolving resource mix. For example, evidence from recent scarcity events in the United States shows that reliability challenges in renewable- and gas-dominated electric power systems arise not from the lack of generation capacity to serve peak customer demand, but from the lack of available capacity to provide the requisite energy at times of need. In other words, supply shortfalls are at higher risk of occurring because of inadequate energy to serve demand during net demand peaks, rather than because of inadequate generation capacity to satisfy system demand peaks (NERC, 2020; ESIG, 2021).

High penetration of renewable energy sources also exacerbates the long-standing missing money problem because increased entry of variable energy resources (VERs) with near-zero marginal costs decreases the amount of thermal generation cleared from the merit-order stack and induces lower capacity factors for thermal generators. In addition, energy prices tend to be lower (on average) and more volatile as a result of increasing VER participation (Mallapragada et al., 2023). Between 2014 and 2020,

\(^2\) Long-term contracts are traded on organized exchanges like Intercontinental Exchange (ICE, 2024), but liquidity is concentrated in shorter periods, with only small amounts traded more than 2 or 3 years out. Contracts for more than 5 years are rare and usually traded over the counter. As a point of reference, construction times range from 2 years for a natural gas combined cycle plant to 3-4 years for a coal plant to 10 years for a nuclear plant, and a generation unit’s lifetime may be 20 or more years after the final investment decision.

\(^3\) In some cases, capacity commitments have significantly exceeded the target reserve margins, resulting in chronic over-procurement and negative cost impacts for the consumers (Wilson et al., 2020).
generation resources in some regions of the United States have observed diminishing energy market earnings and a greater contribution of capacity market revenues to the recovery of investment costs (Figure 1). This trend is problematic because capacity markets do not provide the necessary incentives to invest in an adequate amount and mix of flexible generation to meet evolving system needs (PJM Interconnection, 2017b). Additionally, price volatility creates greater revenue uncertainty for variable resources, which in turn raises financing costs for capital-intensive generators such as renewables (Pierpont and Nelson, 2017).

Figure 1. Components of PJM Price

Some researchers and analysts argue that deep decarbonization of electricity systems will require a shift away from the textbook electricity market design in which long-term investment decisions are driven by short-term energy prices. For example, Roques and Finon (2017) suggest an evolution of market design toward a hybrid regime where “competition for the market” via long-term contracts to support investment would be followed by “competition in the market” for short-term dispatch and balancing operations as at present. An example of hybrid markets combines mandates for LSEs to enter into long-term bilateral power purchase agreements (PPAs) with renewable power generators and storage (which are often mediated through competitive procurement programs; Joskow, 2022) and dispatch of the new generation capacity in short-run energy markets. Keppler et al. (2022) further elaborate on the motivation for hybrid markets.

The increase in energy market earnings observed in 2021 and 2022 was largely driven by the upward trend in natural gas prices (EIA, 2024).
This paper presents a review of electricity market designs to support the clean energy transition. Meeting system reliability challenges associated with increasing levels of renewable generation requires incentivizing resource flexibility, both in investment and in operation. Electricity market structures should be designed to allow resources needed for reliability to earn sufficient revenues to recover their fixed and variable costs. Good market designs should also enable low-cost financing to support investments in capital-intensive resources that are instrumental in meeting decarbonization objectives. An additional property of well-designed markets is promoting short-run efficiency by reducing incentives to exercise market power and supporting efficient renewable curtailment outcomes. Besides reliability, long-run efficiency, and short-run efficiency, electricity markets may seek to achieve equity objectives (e.g., lower costs of renewable generation must be passed through to retail customers) and in some cases may be integrated with clean energy goals.

Our work builds on three recent reviews of electricity market design. Bhagwat, De Vries, et al. (2016) survey US capacity market experts with the purpose of drawing lessons for the EU and conclude that US capacity markets have achieved reliability goals in an economically inefficient manner because they tend to lead to excess generation capacity. As a result, consumers have not benefited from the implementation of capacity markets in terms of lower overall costs. Bublitz et al. (2019) discuss capacity remuneration mechanisms from across the globe and review experiences with them. They note that “as the adequate design depends on a variety of factors such as the existing capacity mix and demand characteristics, no general advantageousness of single (capacity remuneration) mechanisms could be determined so far. ... Nevertheless, it can be concluded that market-based mechanisms, e.g., a forward capacity market, are usually advantageous compared to interventionist mechanisms such as capacity payments” (2019, 1074). Duggan (2020) reviews the literature that evaluates the relative performance of various capacity mechanisms, including the “energy-only” market approach. He discusses recent developments in different RTO markets, which he concludes collectively indicate that experimentation on capacity remuneration for resource adequacy support is an ongoing process. We contribute to this literature by broadening the focus to include proposed electricity market designs, and by providing a comprehensive review and comparison of these options based on key metrics.

The rest of the paper proceeds as follows. The next section reviews traditional approaches to ensure resource adequacy and revenue sufficiency, including scarcity pricing and a range of capacity mechanisms. Section 3 discusses evolving US market approaches to support reliability, incentivize flexibility, and provide forward price certainty to meet load requirements of eligible customers. The paper then turns in Section 4 to a discussion of 11 proposed designs that offer potential solutions to the challenges faced by electricity markets in the United States and abroad. We compare the designs based on their potential ability to support the development of reliable mixes of generation resources, promote efficient long- and short-run resource allocation, and help achieve energy affordability and decarbonization objectives. Section 5 highlights open questions and directions for future research. Section 6 provides concluding remarks.
2. Traditional Approaches to Resource Adequacy and Revenue Sufficiency

2.1. Scarcity Pricing

All RTOs in the United States use some form of scarcity pricing to incentivize availability during times of need (Mehrtash et al., 2023). Scarcity pricing allows energy or ancillary services prices to rise above the average variable cost of all operating plants when the system is capacity-constrained. In such instances, units earn a short-run profit, known as scarcity rent. A sufficient number of scarcity hours would increase the cumulative scarcity rents earned by operating units, and firms would invest in new capacity as long as the expected cumulative scarcity rents exceeded the costs of building new capacity (Bublitz et al., 2019). However, the existence of price caps and lack of short-term price-responsive demand pose practical challenges in the implementation of scarcity pricing in real-world electricity markets. To address these challenges, markets have amended their pricing mechanisms to generate high prices when operating reserves are scarce. For example, in ERCOT, the only grid operator in the United States that relies on an “energy-only” market, prices for both energy and ancillary services rise above the offer prices of generation units when reserve margins are low. Prices in times of scarcity are based on an operating reserve demand curve that reflects the system operator’s demand for reserves (Hogan, 2005). This mechanism provides incentives for the performance of generation capacity under most operating conditions (Potomac Economics, 2021), although after the February 2021 outages in Texas, critics argued that the energy-only paradigm fails to ensure adequate levels of installed capacity and maintain reliability under stress conditions (Borreson, 2022).

As part of the redesign of the state’s energy-only market in the wake of Winter Storm Uri, in January 2023 the Public Utility Commission of Texas (PUCT) endorsed the performance credit mechanism (PCM) to incentivize generators to be available during times of high demand. The new market mechanism awards credits to generators at the end of a compliance period, based on their availability at times of grid stress. Utilities have an obligation to purchase credits based on their share of system load during high reliability risk hours. In February 2024, ERCOT filed a memorandum laying out options for design parameters and a proposed evaluation methodology to select the final parameters (PUCT, 2024). The PUCT is expected to adopt rules regarding PCM in the first half of 2025.

5 ERCOT does not currently co-optimize provision of energy and operating reserves in real time, although co-optimization is expected to be implemented by 2025 (Mehrtash et al., 2023). As a result, the reserve price is determined ex post and added to the real-time energy price to provide additional revenues to recover fixed and variable costs.

6 It is unclear whether any alternative market structures would have provided better incentives for power suppliers to weatherize against the extreme cold-weather conditions observed in Texas, relative to an “energy-only” market (Busby et al., 2021; Palmer and Cleary, 2021).
2.2. Capacity Remuneration Mechanisms

In addition to scarcity pricing, resource adequacy concerns may also be addressed through capacity remuneration mechanisms whereby generation units earn revenue streams for capacity in addition to revenues earned through the sale of energy and ancillary services. A wide range of options includes capacity payments, capacity requirements, capacity markets, reliability options, and strategic reserves.

2.2.1. Capacity Payments

Capacity payments award generators an administratively determined price for capacity made available to the system. These mechanisms have been in place for years in Spain, Italy, Greece, Chile, Colombia, Peru, and South Korea, and they may be fixed or variable, as well as targeted (i.e., paid exclusively to new generation capacity) or market-wide (i.e., paid to both new and existing generation capacity). Absent a competitive process, the major challenge is the identification of the correct level of the per-unit capacity payment. Capacity payments are therefore likely to be the least efficient type of capacity mechanism (European Commission, 2016).

2.2.2. Capacity Requirements

In some regions of the United States (CAISO and SPP), capacity requirements mandate that LSEs procure an administratively determined quantity of capacity from generators (often through bilateral contracts) to meet planning requirements. Typically, capacity resources procured using such mechanisms have a requirement to offer into wholesale energy markets (NARUC, 2023). Bushnell et al. (2017) note that “capacity prices in these areas are less transparent than in areas with centralized capacity markets, due to the lack of a standardized auction or clearing-house for capacity.” Over time, California has expanded the range of capacity requirements for LSEs to deal with different categories of needs. Currently each LSE must acquire three types of reserves: system reserves to meet the reserve planning margin determined by the California Public Utility Commission, local reserves to buttress against extreme weather and relevant outage contingencies in a local area, and flexible reserves that can vary their output rapidly to meet expected ramping needs (see Section 3.1 for more details on flexible reserves). Beginning in 2025, LSEs will be subject to a more detailed slice-of-day approach to acquiring system-wide reserves that moves away from focusing on peak-load hours and instead considers hourly generation profiles throughout several representative days. This is intended to provide a straightforward resource adequacy framework for LSEs to balance resources with correlated outages and complementary supply profiles, such as solar and energy storage (Robertson et al., 2023).
2.2.3. Capacity Markets

Alternatively, the value of capacity may be set through centralized capacity markets, as in ISO-NE, MISO, NYISO, PJM, and the UK. In this case, a generation capacity target to provide an appropriate level of system reliability (commonly referred to as capacity requirement) is established, and system operators organize its procurement through auctions in which generators bid existing and new capacity. A market-clearing price for capacity is determined based on the supplier bids and a demand curve that reflects the reserve margin, and generators receive capacity payments that complement revenues from energy and ancillary services markets. Capacity markets in the United States differ with regard to timing, shape of capacity demand, reserve margin, resource contribution to the reserve margin, and performance incentives. For example, ISO-NE and PJM operate forward capacity markets that procure capacity three years in advance of when the capacity must be in operation and produce energy. Generators that clear the capacity market auction make a commitment to supply capacity in three years’ time and receive payments in the delivery year based on the capacity market-clearing prices and quantities (Monitoring Analytics, 2023b; PJM Interconnection, 2022). On the other hand, NYISO relies on capacity market auctions that occur six months and one month ahead of the performance period. In some jurisdictions, capacity markets are supplementary to other resource adequacy practices. For example, LSEs in MISO may procure capacity using bilateral contracts or through a voluntary centralized planning resource auction in addition to the MISO capacity market.

The capacity value of resources across these frameworks is moderated by accreditation processes, which calculate the probability that the resource will be available in times of supply shortfall. Most jurisdictions are shifting from an average accreditation value, which calculates a resource’s average availability factor, to a marginal accreditation value, which calculates a resource’s marginal contribution to improved reliability. In the marginal framework, resources with correlated outages are valued less, particularly as they make up a greater share of the generation mix. While marginal accreditation is considered the most efficient way to identify the necessary generation mix, it may mean that some resources, particularly variable ones, can expect very low compensation in the capacity market.

2.2.4. Reliability Options

Reliability options are one-way contracts-for-differences (CfDs) that entitle the buyer (system operator) to receive from the seller (generator) any positive difference between the spot price of electricity and the contract strike prices for each megawatt (MW) procured under the contract (Vazquez et al., 2002; Bidwell, 2005; Oren, 2005; Cramton and Stoft, 2008). In exchange, the seller receives a premium. From the generator’s perspective, the option provides a fixed payment that stabilizes its income, thus increasing revenue certainty. From the system operator’s perspective, the option reduces the risk of generation inadequacy by incentivizing resources to be available during scarcity conditions. Additionally, consumers are protected against price spikes, as the system operator has the right to buy electricity at the strike price, which is typically set above the marginal cost of the most expensive unit on the system (Bidwell,
When the option is called (i.e., when the spot price exceeds the strike price), nonperforming generators not only lose out on potential income from the spot market (at the strike price level) but also must pay a penalty equal to the difference between the spot price and the strike price, times the amount of electricity contracted in the reliability option but not produced. Reliability options are implemented in Ireland, Italy, and Colombia. In the United States, ISO-NE’s centralized capacity market includes reliability options backed by physical generation assets.

### 2.2.5. Strategic Reserves

Finally, strategic reserves are out-of-market supplemental generation resources that are contracted by the system operator and deployed during exceptional situations when existing commercial generation capacity in the market cannot meet short-term electricity demand. Capacity for strategic reserves is commonly procured via competitive auctions taking place before the capacity is available for offering strategic reserves (ACER, 2013). The winning resources are paid with a market-clearing price for the option to deliver 1 megawatt-hour (MWh) of energy for each MW of capacity when a reserve shortage is announced. Strategic reserves have been implemented in hydro-dominated countries like Finland and Sweden, and in Germany, where capacity reserves typically consist of units on the verge of retirement that are kept on standby in case of output shortfalls (Lindboe et al., 2016; Neuhoff et al., 2016). Since a capacity policy based on strategic reserves would procure a smaller amount of capacity, any errors would be less consequential and costly than in centralized capacity markets (Neuhoff et al., 2016).

However, as Cramton et al. (2024) suggest, there are some limitations on the use of strategic reserves. When energy prices are high for an extended period, RTOs may use reserves to reduce energy prices, breaking the commitment that strategic reserves be used only during a shortage. Such “preshortage” use harms the market, since it would be hard for reserve providers to anticipate the circumstances under which the strategic reserves would be used. Furthermore, preshortage use creates missing money for the generators because it reduces market prices. Bublitz et al. (2015), Bhagwat, Richstein, et al. (2016b), and Bhagwat et al. (2017) argue that a capacity market can be more effective than strategic reserves for ensuring energy reliability, especially with a high share of energy coming from renewable sources and in the presence of high levels of demand uncertainty.

---

7 A strategic reserve is smaller because the procurement is only for the capacity held in reserve and not for capacity that is expected to be operating most of the time, all of which is eligible for payments in capacity markets.
2.2.6. Incentivizing Capacity Performance

While capacity mechanisms incentivize installed capacity with the potential to provide reliable generation supply, energy shortfalls may occur if capacity is not available to operate when it is needed (Bushnell et al., 2017). Centralized capacity markets attempt to tackle this problem by financially penalizing nonperformance of contracted capacity during stress events (Mastropietro et al., 2015). For example, the capacity market in the UK penalizes capacity providers that fail to meet their resource obligations during system stress events. The penalty payment is the product of a fixed penalty rate and the resource's undelivered output during the stress event and is capped at 200 percent of the generating resource's monthly capacity payments; to put this into perspective, a resource that fails to deliver any capacity during a stress event may incur penalties that reach the penalty cap in four hours of non-delivery (BEIS, 2021).

In 2012, ISO-NE was the first grid operator in the United States to propose pay-for-performance incentives to improve generator response when called on to operate during scarcity events. The Federal Energy Regulatory Commission approved ISO-NE's proposal in 2014 (FERC, 2014), and pay-for-performance requirements were first introduced in the capacity market auction held in February 2015 for the 2018–19 capacity commitment period (Coffman Smith, 2018). PJM also developed a similar pay-for-performance program after the 2014 polar vortex event, receiving FERC approval in June 2015 (FERC, 2015). PJM’s capacity performance mechanism went into effect in June 2016.

Under a pay-for-performance scheme, during capacity scarcity conditions, generators participating in the capacity market are rewarded or penalized for their performance relative to a baseline (Gillespie, 2018). Generators receive monthly capacity payments in the form of a base payment and a performance payment. The base payment is fixed and determined by the generator’s capacity supply obligation, while the performance payment depends on system conditions and the generator’s performance relative to a baseline during scarcity. Overperformers are rewarded with performance credits, while underperformers are penalized. These changes to the market design appear to have helped improve performance in subsequent cold weather events.\(^8\) Capacity performance provides an incentive to meet capacity obligations, with a marginal incentive for generators that is similar to real-time scarcity pricing. However, at the current penalty rates,\(^9\) and with infrequent performance evaluation periods, penalties may not provide sufficient incentives for generator compliance. Performance would be more effectively incentivized if it were measured more often than just during infrequent shortage events (e.g., not just during high-reliability risk hours) (Bowring, 2023). Another option may be to measure performance not only when emergency actions are declared by the grid operator but also during additional intervals with tight supply cushions so that a certain minimum number of triggers per year is met.

---

\(^8\) For example, in PJM, generator outages were halved during a severe cold spell in January 2019 relative to the 2014 event (Chen et al., 2020).

\(^9\) For example, the penalty rate in ISO-NE is $5,455/MWh for the 2024–25 capacity commitment period and $9,337/MWh for all subsequent periods (ISO-NE, 2024).
3. Evolving Market Approaches to Support Reliability, Flexibility, and Price Certainty

Overall, there is a sense that traditional methods for ensuring resource adequacy are both insufficient and inefficient, meaning that payments to ensure sufficient resources are high, and they have not done enough to guarantee that supply will be available in times of system need. In this section, we discuss three market approaches in the U.S. that aim to address these and related concerns, including incentivizing flexibility and providing forward price certainty, without wholly redesigning market structures. Non-market resource adequacy approaches like the Western Resource Adequacy Program (WRAP) are outside the scope of our review. 10

3.1. Flexible Capacity Requirements

Although California does not have a capacity market, the state does impose reserve procurement requirements on LSEs. In 2014, the growth of variable energy resources in the state prompted concerns over the level of rampable capacity with the ability to meet swings in variable supply. This concern led to the creation of a new category of reserves that LSEs are required to procure (Ela et al., 2014). The Flexible Resource Adequacy Criteria and Must Offer Obligation (FRAC-MOO) initiative defines resources by their attributes of flexibility and requires LSEs to have a certain amount of flexible capacity to meet the RTO’s operational needs based on their historical contribution to the maximum three-hour net load ramp at a system level (CAISO, 2023). LSEs must submit year-ahead and month-ahead resource adequacy plans to the CAISO, listing the capacity committed to serve flexible capacity requirements (Bushnell et al., 2017). If there is a shortage in the aggregate supply of flexible capacity to the system, the RTO procures backstop flexible capacity through a competitive offer process on a one-year forward basis.

10 The WRAP is a joint planning, resource-sharing, and compliance approach that was established by utilities participating in the Western Power Pool (WPP), formerly known as the Northwest Power Pool. It includes 22 investor-owned and public utilities from within the WPP (see https://www.westernpowerpool.org/news/wrap-area-map). This agreement sets regional targets for planning reserve margins and determines each member’s share of the region-wide requirement. The agreement requires a “forward showing,” nine months in advance of both summer and winter peaks, that the utility has sufficient resources to meet its contribution. The agreement also includes an operational program that allows participants to adjust to deviations from planning assumptions in real time and indicate either needs for additional resources or the presence of excess supplies to other WRAP participants to facilitate resource sharing. More information is available at https://www.westernpowerpool.org/news/what-is-the-wrap.
Flexible capacity providers offering into the FRAC-MOO initiative must submit economic energy bids into the RTO's day-ahead and real-time energy markets. Self-scheduling is not allowed to meet must-offer obligations, as it implies that resources are not available for dispatch by the RTO without adjusting the self-schedule (and therefore are not flexible). Ramps of different magnitude require different types of flexible capacity resources. Therefore, flexible capacity falls into three categories with different offer obligations. Base flexibility resources have to submit economic bids for 17 hours a day and must be available seven days a week. Peak flexibility resources have to submit economic bids for 5 hours a day (determined by season) and must be available seven days a week. Super-peak flexibility resources have to submit economic bids for 5 hours a day, only on weekdays and nonholidays (CAISO, 2023). All types of flexible capacity supply can be sold or committed in these categories, provided that they meet availability requirements and must-offer obligations. It is not clear, however, whether demand resources could also be committed in these categories. Further, CAISO assesses the adequacy of each LSE's flexible capacity showings.

In the years following the adoption of FRAC-MOO, reliability experts and CAISO began to express concern that the product, as designed, failed to adequately cover flexibility needs prompted by unpredictable ramping needs. Critics felt that too many resources were credited as flexible even though they were slow-starting or had onerous minimum operating capacities (Holman, 2018). In 2018, CAISO proposed an update to FRAC-MOO to develop three flexible resource adequacy (RA) products: five-minute flexible RA, fifteen-minute flexible RA, and day-ahead shaping RA (CAISO, 2018). This proposal also suggests other changes, such as limiting must-offer obligations to the day-ahead market, but it has not been adopted.

### 3.2. Seasonal Forward Capacity Markets

With the changing resource mix and electrification shifting the profile of net load, many jurisdictions have considered, and some have adopted, resource adequacy frameworks with multiple capacity markets. Seasonal capacity markets have been discussed in most RTOs that have existing capacity markets because seasonal characteristics affect the relationship between peak load and resource supply. For example, in the summer, peak demand associated with increased demand for air conditioning may be effectively served by high solar availability. However, peaks in the winter driven by increasingly electrified home heating demand may require a different resource mix to meet demand. Seasonal markets are thought to meet resource adequacy needs more cost-effectively by matching seasonally variable demand with seasonally variable resources.

NYISO and MISO currently operate seasonal forward capacity markets, where resources may bid different quantities and prices based on their projected availability. In NYISO, the resource adequacy requirement does not change seasonally, but rather still reflects the annual peak demand. In MISO, the resource adequacy requirement reflects seasonal peak demand and is currently lower in the winter than in the summer market. However, accreditation can differ by season in both RTOs. Other jurisdictions, like ISO-NE and PJM, have expressed interest in the seasonal forward market framework but have yet to implement one (Graf, 2022).
3.3. Basic Generation Service

Centralized forward markets for energy have played an important role as a mechanism for determining energy prices for residential and small commercial customers who reside in states with retail choice but choose not to purchase energy from a competitive retail provider. The service available to these customers, which goes by different names in different states, is typically priced using a competitive auction where generators bid to supply all the requirements (energy plus ancillary services and capacity) needed to serve a predefined share of load for a future time period.

For example, New Jersey local distribution utilities work with the state to design an auction each year that identifies the suppliers that will provide energy and other services to some portion (or tranche) of load served by each local distribution utility that does not choose to purchase power from a competitive retailer. In New Jersey, this type of service is known as basic generation service (BGS). Participants in the BGS auction are essentially competing for the opportunity to be the LSE for a portion of New Jersey load of particular types.

Since its inception in 2002, the BGS auction has used a multi-round descending clock auction in which the announced price is lowered between rounds until the amount of energy plus related services that bidders are willing to provide at the declared price equals the total number of tranches of future load up for bid. The process of designing the auction and qualifying and registering bidders starts roughly 11 months before the first delivery period, with the actual auction occurring four months before initial delivery. Each supplier participating in the auction is bidding to supply all the energy necessary to serve a given portion of a particular type of load, called full requirements service, at each point in time throughout the life of the contract, which is typically three years. One-third of the total BGS load is procured each year, and the price to ultimate consumers of BGS service at any point in time is based on a rolling average of the three contract prices in effect. The winners in the auction also take on

---

11 Industrial and large customers typically face prices that are more reflective of nearer-term wholesale energy prices.

12 The service provided to customers who decide not to choose a provider goes by different names in different states, including, for example, default service in New Hampshire, last resort service in Rhode Island, and standard offer service in the District of Columbia. When retail markets were first introduced, the expectation was that most customers would migrate away from these default services to purchase needed energy from retail providers, and the demand for these types of default services would diminish over time. However, retail competition has not taken off in most places, and default services still capture a large portion of the market in most restructured states.


14 The timing of the steps leading up to the BGS auction and the auction itself are described at https://www.bgs-auction.com/bgs.calendar.asp.

15 Winning suppliers are also responsible for meeting the renewable portfolio standard requirements for their portion of realized load.
the requirement to pay for the cost of capacity and ancillary services procured by PJM in other relevant auctions, much the way any LSE in PJM takes on those costs from PJM and passes them on to load.16

This type of full requirements contract imposes substantial amounts of risk on the suppliers bidding in BGS actions because the total amount of MWh in any given tranche at any point in time is uncertain in advance, as is the precise amount of ancillary services or capacity needed to serve customers (Hogan, 2010). There is also risk in terms of the total size of the BGS market, as consumers may float in and out of BGS service as competitive suppliers leave and enter the market. Suppliers of BGS service, which may own generators or may instead contract with generators, face energy price uncertainty in real time, and they bear the risks associated with transmission congestion, as the winning prices in BGS auctions are typically set at the distribution utility level.

BGS offers an interesting point for comparison to some of the market designs described in the next section. Unlike most of the proposed forward energy markets, the BGS auction does not occur very far in advance of ultimate delivery of energy (only about four months); hence the potential to provide a strong signal for future investment is small. The obligation is for a predetermined share of customer load, but it is not for a fixed quantity of energy. While BGS or related contracts range between one and three years in length (depending on the types of customers being supplied), some of the proposals for new markets described in Section 4 call for longer terms, although typically not that much longer. The BGS auction is a mandatory market for physical power delivery, while some of the proposals explored in the next section focus on voluntary financial markets to help create incentives for investment with no requirements for LSEs or generators to enter into forward contracts. Finally, while the full requirements nature of the BGS contract includes provision of clean or renewable energy certificates, it does not include explicit consideration of greenhouse gas emissions limits, envisioned in some of the proposals.

16 The apportionment of capacity costs to load-serving entities is described in PJM Interconnection (2023a), and the assignment of ancillary services costs to load-serving entities is described in PJM Interconnection (2023b).
4. Proposed Approaches to Meet the Challenges of the Clean Energy Transition

In this section, we discuss 11 proposed designs that seek to support the development of reliable mixes of generation resources, promote long- and short-run efficiency, and help achieve energy affordability and decarbonization objectives. These ideas are in varying stages of development and have yet to be tested in the real world. We begin our discussion with a design that proposes the combined procurement of forward capacity and clean energy attribute credits. This approach represents a variation on existing capacity markets and is the least radical departure from current market designs. Next, we consider proposals that feature long-term contract auctions and energy markets, centralized resource planning that may be coupled with bidding to uncover inexpensive solutions to resource needs, or separate short-term markets for different generation technologies. We conclude with a design that proposes linked swing-contract markets to ensure the availability of net load balancing services for real-time operations. The swing-contract market design represents the most fundamental shift in electricity market operation relative to other proposals considered in this review, and it would entail a complete overhaul of the market design.

4.1. An Integrated Clean Capacity Market

One proposal for adapting existing resource adequacy mechanisms to incorporate explicit consideration of the changes in resource mix necessary to meet climate policy goals is an integrated clean capacity market (ICCM). This approach, proposed by Spees (2020) would combine procurement of forward capacity by the regional system operator with procurement of clean energy attribute credits (CEACs) used to demonstrate compliance with clean energy procurement goals or requirements. Note that renewable energy credits (RECs) are a specific kind of CEACs. RECs are associated with renewable power generation and typically satisfy the specific requirements of a state renewable portfolio standard (RPS) law, while CEACs include other sources of carbon-free energy such as nuclear power. Because the ICCM would procure both resource adequacy and environmental policy needs, it would replace the existing capacity market construct.

The ICCM consists of a combination of two simultaneous bidding processes and a single bid evaluation algorithm that jointly optimizes procurement of CEACs and of

---

17 This section was informed by and benefited from comments and suggestions from Peter Cramton, Steven Corneli, Natalia Fabra, Eric Gimon, Malcolm Keay, David Nelson, Brendan Pierpont, Kathleen Spees, Leigh Tesfatsion, and Frank Wolak. Any errors are the authors’ sole responsibility. Note that Section 4.9 includes two proposed designs.

18 The creation of a broad regional market for procurement of clean energy attribute credits was first proposed by Spees et al. (2019).
accredited capacity to meet RTO reserve requirements (New Jersey Board of Public Utilities, 2021). Suppliers that qualify as clean energy sources (renewables and other non-emitting sources) can offer to supply separate quantities of capacity and CEACs at a single combined price, while generators that burn fossil fuels would offer into the capacity auction only. Demand curves for capacity would be set by the RTO and could vary by season and location. Demand curves for CEACs would be based on an aggregation of state requirements across constituent states within a multi-state RTO and any additional demand for CEACs that voluntary agents, such as corporate entities or municipalities with more ambitious clean energy targets than the state, want to bring to the auction. The joint market would be run by the RTO with the same forward-looking timeline as existing capacity markets (for example, three years in the case of PJM and ISO-NE). CEAC procurement would also be three years in advance, but the proposal includes a provision to fix the price for CEACs provided by new resources for a longer time horizon, proposed to be between 7 and 12 years and to be determined in the future. This provision would provide more revenue certainty to new renewable resources and would be a closer replication of the terms of existing procurement mechanisms used in different states, such as New York’s periodic REC procurements, which can reward contracts to winning bidders for up to 20 years into the future.

By offering clean resources the opportunity to receive compensation for two revenue streams within a single auction, the ICCM approach is expected to shift down the supply curve both for capacity and for clean energy credits relative to what suppliers would bid in separate auctions for the two products. It would thus both lower the cost of procuring capacity and reduce the carbon intensity of the resources on the system.

Recently, the Massachusetts Department of Energy Resources (2023) proposed a variation on this proposal in which a regional market focuses solely on the procurement of clean energy attribute credits, known as a forward clean energy market (FCEM). Different vintages of the proposal have different parameters, but the basic structure remains the same. The FCEM market proposed by Massachusetts would be a voluntary market that could be either administered by the states and entirely separate from the RTO or administered by ISO-NE subject to FERC jurisdiction. The FCEM auction would occur before the forward capacity auction and thus would inform generator bidding strategies in the capacity market. The Organization of PJM States is also considering an FCEM concept.

19 In addition, the RTO could define some requirements related to flexibility for a portion of the capacity needs (Spees, 2020).

20 The Massachusetts proposal includes a number of different types of potential environmental attributes in the regional FCEM market beyond CEACs, including clean capacity credits (used to demonstrate compliance with a clean capacity requirement, should one be established by a state or other jurisdiction in the future), renewable energy credits, and greenhouse gas marginal abatement certificates (which indicate the quantity of emissions displaced by the additional renewable energy). These different types of certificates are intended to accommodate different approaches to clean resource requirements across the states.
The ICCM approach seeks to maximize benefits from joint procurements of clean energy requirements of state and other energy buyers together with regional resource adequacy needs by allowing for simultaneous identification of the least-cost mix of resources to meet both goals. Procuring capacity and clean energy credits separately could lead to failure to acknowledge the impact of clean energy goals (and associated policy incentives) on the best approach to meet resource adequacy needs and could raise the cost of meeting both goals. This integrated approach avoids some of the pitfalls of existing capacity markets that tend to favor fossil generators. However, the approach would need modifications, such as an additional requirement for a minimum amount of flexible capacity or improved capacity accreditation procedures to address the disconnects between capacity markets and energy adequacy in nonpeak hours to deal with variability of generation from renewable energy sources. It would also face some of the same challenges as existing capacity markets when it comes to aligning payments for capacity with generator performance in real time (see Section 2.2).

4.2. Standardized Fixed-Price Forward Contracts

In Wolak (2022)’s long-term resource adequacy approach electricity retailers must hold standardized fixed-price forward contracts (SFPFCs) that cover significant shares of system energy demand far enough in advance to allow new entrants to compete to supply this energy. The obligation increases linearly, e.g., from 85 percent of forecast demand four years before delivery to 100 percent of realized system demand in the current year (Wolak, 2022). The system operator procures SFPFCs in one-sided auctions for a compliance period (defined as the length of time covered by the contract) and allocates the obligation to the SFPFC energy to retailers based on their share of system demand. SFPFC auctions clear for the entire compliance period (i.e., not on an hourly basis) at the system level (i.e., yielding a single forward price, not a price that varies by load zone or node). Physical constraints of the network are not considered in these auctions.

On the supply side, power generators bid to supply energy (up to their firm capacity value21) at a fixed price in the compliance period. To strengthen incentives to bid efficiently in the spot market, the mechanism proposes that 100 percent of the realized electricity demand be covered by SFPFCs. To this end, “true-up” auctions may take place after energy demand is realized in each compliance period. In these auctions, unused SFPFCs are bought back (if actual demand is lower than the contract quantities cleared in the forward auctions), or additional SFPFCs are purchased (if actual demand is higher than the contract quantities cleared in the forward auctions).

A core element of the proposed design is the retroactive load shape adjustment. SFPFC obligations for a compliance period are allocated to power generators and retailers across hours based on hourly shares of system demand. For example, SFPFCs sold by a power supplier in a given compliance period are allocated across hours based on the hourly shares of system demand in that period. Similarly, the obligation of each

---

21 Firm capacity values would be determined through resource accreditation mechanisms like the ones currently in use in RTOs.
A retailer is based on the share of total demand served by each retailer in the compliance period. Hourly load shapes from the previous year may be used to do an initial settlement in the compliance auctions, while actual hourly load shapes may be used for the final settlement in the true-up auction.

To ensure that the allocation of SFPFC energy follows the actual hourly pattern of demand as closely as possible, Wolak (2021a) argues in favor of compliance periods for SFPFCs that are no longer than quarters of the year. If the compliance period is quarterly, up-front compliance auctions would be run every year, with quarterly products for delivery two, three, and four years in the future. Additionally, true-up auctions would be run at the end of each quarter, based on the realized energy demand for the compliance period (Wolak, 2021a).

For each MWh cleared in the long-term market auctions, suppliers receive (and retailers pay) the market-clearing price in the forward auction and pay (and retailers receive) the demand-weighted average short-term price in the compliance period. The sale of energy slated to be delivered years in the future provides suppliers with a revenue stream that helps recover the costs of new investment in generation capacity, although, as noted in Section 1, generation units’ lifetimes are typically in the 20-plus-year range.

An important feature of the proposed mechanism is that retailers are shielded from price spikes in the short-term energy market because fixed-price forward contracts must cover 100 percent of realized energy demand in the system. Adjusting the contract quantities to cover the realized demand over the compliance period incentivizes power generators to have sufficient capacity available to produce in real time and to offer that capacity at marginal cost in the short-term energy market, removing the incentive to exercise unilateral market power in the energy market (Thurber et al., 2022). The reason is that generators are uncertain about their final SFPFC obligation when they bid in the short-term energy market. Withholding capacity or submitting high offer prices in that market risks reducing a supplier’s generated quantity and increasing the price, and any shortfall in output relative to the final allocation of SFPFCs must be purchased at that higher price. Since retailers are shielded from price spikes, the offer cap in the short-term energy market may be raised to provide generation units with stronger incentives to be available and produce during times of scarcity. Higher price caps may provide stronger incentives for consumers to participate in real-time markets (Borenstein, 2007), and retailers serving these customers would benefit by taking advantage of opportunities to sell unused contracted energy in real time.

On the other hand, while SFPFCs would need to cover a substantial fraction of forecast demand years before delivery, forecasts tend to be imprecise in that time frame, potentially leading to inaccurate procurements that would then require large

Note that retailers don't need to hedge locational risk completely, because they pay average prices. On the other hand, when an entity has a PPA with a project, that entity (usually a retailer or a utility) bears all the locational price risk from transmission congestion.
quantities to be traded in the true-up auction. A related complication stems from the fact that retailers serving residential customers typically sign contracts for up to two years and may resist hedging for demand they haven't signed yet. In addition, forward contracting yields less volatile total procurement costs and lower bankruptcy risk for the retailers; as a result, retailers may have little incentive to encourage demand participation in real time. We return to these points in Section 5.

The SFPFC approach is not currently implemented in any electricity market, but it would create similar incentives for supplier behavior as existing long-run resource adequacy mechanisms in Chile and Peru (Wolak, 2022). Shu and Mays (2023) compare centralized capacity markets and SFPFCs using stochastic equilibrium models that focus on efficient risk sharing. Their analysis does not explicitly model the various market stages of each design. They represent capacity markets as operating like a call option with robust penalties for nonperformance, which differentiates their setting from the reality in most US capacity markets, which include low nonperformance penalties with associated weak incentives for performance when called. Their analysis allows for long-term contracts in support of low-carbon technologies and considers the interactions between these types of contracts and traditional capacity contracts with a focus on risk-averse investors. Risk-averse variable resources prefer offtake agreements, which match the shape of their production (e.g., generation power purchase agreements), to capacity obligations, which typically do not have such matching. This observation suggests that if capacity markets continue to exist, it would be more efficient to remove variable resources from the market and to reduce capacity demand commensurately. As an alternative to current capacity markets, SFPFC approaches result in higher consumer surplus, lower costs for a given level of reliability, and lower price volatility for consumers. One potential downside is that an SFPFC approach could lead to greater consolidation across generators and a need for an insurance mechanism to help generators hedge fuel supply risks.

4.3. A Forward Energy Market with Flow Trading

Similar to the approach of Wolak (2022), Cramton et al. (2024) propose requiring electricity retailers to hold forward contracts that cover significant shares of system energy demand before delivery. The obligation increases from 0 to 100 percent of realized real-time demand as the market moves from 48 months ahead to day ahead. Forward energy contracts are traded in two-sided auctions that clear on an hourly basis and by load zone. Monthly forward energy is traded 48 to 1 month ahead by hour, day type (weekday and weekend), and load zone (yielding 24 * 2 * 48 = 2,304 products for each load zone). Hourly forward energy is traded 720 to 8 hours ahead by load zone (yielding 24 * 30 = 720 products by load zone). In addition to forward energy, market participants may purchase energy options, forward reserves (in the month ahead of power delivery), and renewable energy certificates.

23 Energy options deliver energy when the real-time energy price is $1,000/MWh or more and provide price coverage for unanticipated demand.
Frequent batch auctions are implemented with flow trading (Budish et al., 2023), which allows market participants to trade linear combinations of products. Orders for such portfolios are expressed as downward-sloping piecewise linear net demand curves that are represented by a set of price-quantity pairs with quantity as a flow—that is, the rate of trade over a one-hour window in MWh/hour or MW. Supply quantities are expressed as negative net demand. The amount of energy a supplier can sell may be more or less than its capacity value, as determined by resource accreditation mechanisms currently in use in RTOs.

The design proposed by Cramton and colleagues shares some similarities with Wolak’s SFPFC approach. For example, retailers are shielded from price spikes in the short-term energy market because contracts must cover 100 percent of realized demand on the day before delivery, resulting in a strong hedge for consumers. Further, similar to Wolak’s design, that of Cramton and colleagues does not explicitly account for transmission constraints and resource characteristics (such as ramp rates, minimum up and down times, and upper limits) in the forward energy market. Transmission congestion is represented by arbitrage—that is, market participants adjust their orders to anticipate the congestion that will ultimately arise day ahead. Finally, as in Wolak’s design, there are no incentives for direct demand-side participation of retail customers in the forward markets. In Cramton and colleagues’ proposal, retail choice is an essential component of having the forward market foster innovation on the demand side. Forward energy prices are expected to encourage retailers to offer dynamic rates that expose consumers to the real-time price on the margin but hedge them from downside risk.

Despite these similarities, Cramton and colleagues’ proposal differs from Wolak’s in several ways. First, whereas Wolak proposes forward energy procurement through annual or quarterly auctions that clear at the system level, Cramton and colleagues propose frequent batch auctions that occur every hour (Budish et al., 2015; Graf et al., 2024). Forward prices are determined at a delivery point (load zone), and settlement is updated every hour for each market participant. The second difference relates to the schedule of retailer obligations. In Wolak’s proposal, LSEs may be required to purchase, for example, 85 percent of forecast demand four years before delivery to 100 percent of realized system demand in the current year. In Cramton and colleagues’ proposal, the schedule is more gradual, going from 0 percent 48 months ahead and increasing linearly to 100 percent of realized demand day ahead. Finer granularity of discrete auctions and a more gradual schedule of retailer obligations enable market participants to trade smaller positions over time and are expected to enhance liquidity and reduce adverse price impact (Kyle, 1985). Given the large number of traded products, and based on historical market performance (Weber, 2010), liquidity may pose practical

24 For example, on December 20, 2025, a supplier could bid hourly P-Q of weekday and weekend energy by month for the next 48 months (January 2026 to December 2029) and hourly P-Q of energy by day for the next 30 days (December 21, 2025, to January 19, 2026).

25 In markets without retail choice, the regulator may define two default plans, a fixed-rate plan and a dynamic plan with a hedge, and let retailers offer these plans to their customers.
challenges in the implementation of this proposal. However, Cramton and colleagues suggest that a forward energy market run by the grid operator would have higher transparency and lower transaction costs than futures trading on private exchanges like ICE and CME.

The two proposals also differ with regard to the type of market power that is a point of focus. Wolak’s discussion centers on supplier market power in the energy market. The requirement that forward contracts cover 100 percent of realized system demand creates a strong incentive for a supplier to offer at least as much energy at its marginal cost as it expects its final SFPFC allocation to be for that hour of the compliance period. This is because any shortfalls relative to that final allocation must be purchased at a short-term energy price that may be higher than the supplier’s marginal cost. Although the discussion in the proposal focuses on market power in the energy market, the design also provides incentives for a supplier to offer capacity at marginal cost in the forward energy market. This is because the supplier’s final SFPFC obligation for each hour of the compliance period depends on the realized system demand and increases as hourly system demand increases. Therefore, the supplier has an incentive to offer its capacity in the forward energy market to ensure that its final SFPFC obligation for any hour in the compliance period does not exceed its hourly production; this would require purchasing any shortfalls between the supplier’s final SFPFC obligation and its production at a short-term energy price that may be higher than marginal cost. As a result, Wolak’s design also reduces incentives to exercise supplier market power in the forward energy market.26

Cramton and colleagues argue that market power in capacity markets is a primary concern.27 In their proposed design, the ability to trade forward small quantities would provide little incentive or opportunity for market participants to exercise market power. To address supplier market power in the real-time energy market, they require that each retailer’s day-ahead position be at least as high as its realized real-time demand, with a linear penalty for shortfalls. To meet this requirement, retailers would purchase more than their realized demand most of the time, using financial energy options for price coverage.

Finally, the two proposals differ with respect to the role of resource accreditation. In Wolak’s design, the firm capacity value from existing capacity-based resource adequacy approaches would be used to limit the amount of SFPFC energy a supplier can sell in the forward energy market. In contrast, accreditation of capacity value plays a limited role in Cramton and colleagues’ proposal. As discussed, the quantity of energy a supplier can sell forward is not tied to its firm capacity value and may even

26 Wolak (2022) also describes bidding incentives in the “true-up” auction that would take place after energy demand is realized in each compliance period, noting that “suppliers are extremely unlikely to offer to supply this energy below the demand-weighted average short-term price over the compliance period because their overall profits would decline.”

27 Aagaard and Kleit (2022) discuss how inelastic demand and supply can lead to market power in capacity markets and the difficulties each of the three RTOs have faced in their efforts to mitigate the effects of market power on capacity market-clearing prices.
exceed that value. The rationale is that capacity values do not provide assurance of physical performance. In contrast, the ability of generators to select forward hours to sell energy based on their predicted availability can lead to forward periods of high prices that provide an indication of shortage, and the costs of failing to perform would drive generator performance in real time. As an additional means to make up for any potential short-term energy shortfalls, the system operator may procure emergency reserves as reliability backstop.

4.4. A Two-Part Market for Energy

Pierpont and Nelson (2017) focus on the challenges that non-dispatchable resources like renewables face in terms of ability to adapt to price variations in wholesale energy markets and potential for greater price variability in those markets going forward. To address these challenges, they propose a two-market approach that combines a long-term market for a single energy commodity that does not vary by location or time of day and a short-term market for energy delivery. The energy market seeks to minimize the financing cost of renewable power generation while allowing for capacity expansion. Procurement targets in the energy market are informed by a robust planning process. On the other hand, the delivery market concentrates incentives for flexible behavior created by time-varying prices on those resources best equipped to respond, including flexible generators, storage assets, and aggregated flexible loads responding to time-varying price signals.

The energy market operates through rolling annual auctions for long-term contracts. These auctions happen 1 to 3 years before delivery, and contracts are 15 or more years in length. Purchases are based on forecast demand from either the system operator or a collection of LSEs, and all load is required to be covered by contracts. As contracts expire and load forecasts evolve, energy needs procured in future auctions are adjusted. Offers that generators submit to these auctions are expected to reflect long-run levelized costs of generation technologies. These offers are for energy and are independent of time and location. Further, the price from the energy market represents a “notional wholesale price for electricity used in system planning and investment decisions” (Pierpont and Nelson, 2017).

The delivery market replaces existing day-ahead and real-time markets as a mechanism to meet the difference between expected generation from the energy market and real-time electricity demand at different locations. Thus, the bids into this market are increments and decrements to those emerging in the energy market, and the prices from the delivery market reflect the relative value of electricity at each point in time in a particular location. The short-term market does not include scarcity price caps and has many similarities to the ISO-operated energy markets today, except that it does not include supply from inflexible resources. Capacity mechanisms may be needed to support dispatchable resources that participate in the delivery market.

Buyers and sellers can interact with the delivery market in many different ways. Consumers can buy from the delivery market operator and pay the sum of energy and delivery charges for each MWh consumed. They can also transact directly with
aggregators that can help control flexible resources and the LSE demand in a way that reduces delivery charges. This type of aggregator could be similar to the role that some competitive retailers play in deregulated electricity markets. Suppliers in the delivery market include flexible resources that can deliver additional energy at the last minute; generators that can back off when supply is too abundant (as evidenced by lower prices); storage resources that can absorb excess generation and make it available at times of shortage; and resources that can adjust their consumption, thereby contributing to day-ahead and real-time supply-demand balance in the markets. Generators that sell in the delivery market receive a payment for energy supplied in each hour that equals the sum of the energy and delivery prices (even though they did not lock in an energy price through participation in the long-term energy market). Hydro and fossil generators may choose to participate in the energy market with some of their output, but they may be more attracted to the delivery market, where their ability to contribute generation at times of system shortage is directly rewarded.

Pierpont and Nelson describe their proposal as largely a concept for further discussion and recognize that several aspects would require additional refinement. One is the potential need for mechanisms to address resource adequacy. In particular, the prospect of making money in the delivery market may not be sufficiently certain to provide an incentive for investment in batteries or other capital-intensive flexible technologies. Therefore, some type of long-term contract for flexibility or resource adequacy services may be needed to support the reliable operation of a delivery market. A second aspect for further development is to more explicitly recognize the interplay between the energy market and the flexibility needs of the system, the latter of which they anticipate will evolve over time as the mix of resources and the shape of electricity demand change (e.g., a shift from summer to winter peaks).

4.5. A Market for Long-Term Energy Contracts

Pierpont (2020) proposes a market for standardized long-term energy contracts that adds some specifics to the energy market component of Pierpont and Nelson (2017). The mechanism is intended primarily to provide revenue certainty to carbon-free resources that are needed to achieve decarbonization goals (like wind and solar) and support lower costs of capital from investors. While the proposed mechanism is intended to value energy production at the times it is most needed, it does not explicitly seek to ensure long-term resource adequacy; additional mechanisms may be necessary to ensure sufficient resources are built to cover resource adequacy requirements.

Sellers (i.e., project developers of carbon-free resources) specify a fixed price per MWh of contracted energy across all hours, the total annual MWh of production, and an expected hourly and seasonal energy production profile they are willing to commit to. This profile is not binding with regard to how the resource operates but serves as a basis for financial settlement.
Pierpont envisions two possible ways for buyers of energy (load-serving entities) to participate. The first approach involves submitting a quantity bid for the amount of MWhs they wish to purchase for the relevant aggregate time period. Each supplier's fixed price bid would be evaluated relative to an hourly forward price projection developed by the grid operator. The value per MWh is established for each shape bid based on how the hourly availability of the supplier's offer matches the high-valued hours in the forward price projection. Thus, because projects differ based on the shape of energy suppliers expect to be able to deliver over time, the value per MWh differs for each shape bid. After establishing the value per MWh for each shape bid, the market operator would determine the bid's net value per MWh, calculated as the difference between the estimated value and the fixed price bid by the supplier. Next, bids would be ranked from highest to lowest net value up to the amount of demand from load-serving entities. Cleared resources would receive their as-bid price, rather than a uniform price. Further, winning bids would be allocated proportionally to the LSEs, based on the total they have asked to be served through the market. While this approach allows for the market-clearing entity to identify the collection of contracts with the highest value to buyers, it heavily relies on the construction of forward price forecasts against which to evaluate each bid. Additionally, this approach may produce knife-edge solutions in which only one type of resource (characterized by similar costs across projects bidding into the long-term market) is procured in a given year, rather than an efficient mix of resources.

In the second approach, buyers submit an hourly willingness to pay and a demand profile at the same level of granularity as the sellers' bids. In this case, the clearing mechanism would match those bids and allocate output from each of the sellers to the buyers with the highest value. While introducing greater complexity for the buyers, this alternative approach would have the advantage of relying on bidding parameters submitted by market participants, rather than market price forecasts. In addition, it may lead to a more efficient mix of resources and a better match for buyers' demand profiles.

The contract procured in the market is a financial contract based on a specific production shape. Buyers pay a fixed contract price that represents the weighted average of the as-bid prices offered by the suppliers (denoted as shape-weighted average contract price) and receive the real-time price applied to the contracted shape (which would offset their exposure to real-time prices for physical electricity purchases in the short-term market). On the other hand, sellers settle their long-term fixed price contract based on the real-time price applied to their contracted shape. Specifically, for each hour of their contracted energy production shape, suppliers receive their fixed as-bid price and pay the real-time energy price applied to that contracted shape (which would be offset by their own physical electricity sales in the short-term market). This insulates the seller from future wholesale market price risk by guaranteeing a fixed price while providing an incentive to deliver power to the physical electricity market to match the load shape (or perform better). Any physical deliveries in excess of expected production may be sold in the spot market. 28

Note, however, that the product traded in the long-term market is purely financial, and thus there is no requirement that a generation resource that sells in this market actually be available to produce. Nonetheless, participating resources would be incentivized to perform against their contracted shape, especially when prices are high and they could incur a large cost to buyers if they were unable to produce during those hours.
Capital-intensive renewable energy is more sensitive to the cost of capital than nonrenewable energy. Pierpont’s market mechanism is intended primarily to provide revenue certainty and thereby lower financing costs for carbon-free resources, rather than ensuring resource adequacy for the electricity system as a whole. Unlike most other market proposals evaluated in this paper, Pierpont (2020) explicitly recognizes the need to design long-term markets so as not to distort short-run market efficiency and price signals. More work remains to test and validate Pierpont’s concept, and several unresolved questions must be answered to make his design implementable. For example, it is not clear what entity (grid operator, government agency, or third party) would be best placed to operate the long-term market just outlined. Further, participation in the market is purely voluntary, although Pierpont envisions situations where regulators might require participation by LSEs for some portion of load.

4.6. Long-Term Contracts Based on the Attributes of Generation Technologies

As a part of discussions around reforms to electricity market design in the European Union, the proposal by Fabra (2023) facilitates efficient investments in generation capacity through long-term contracts signed between the regulator (acting on behalf of consumers) and the generators. Contracts are of two types, depending on the attributes of the generation technologies, and are allocated through competitive bidding mechanisms whenever possible. Competition for these contracts through auctions would allow efficiency gains to be passed on to final consumers.

The proposed reform of EU electricity market design places emphasis on private bilateral contracting through PPAs (European Commission, 2023). As noted in Section 1, PPAs have often been used to acquire renewable energy and storage through competitive procurement programs. However, these contracts have some drawbacks; for example, counterparty risk arises because buyers have an incentive to renege on the contracts if short-run electricity prices are low (Ambec et al., 2023). To address these shortcomings and help derisk investments (thus enabling access to lower financing costs), Fabra envisions a greater role for CfDs in her proposal. Contracts are allocated through technology-specific or technology-neutral auctions, and tailored to the specific attributes of the generation technologies. For example, two-way CfDs are suitable for intermittent

29 To mitigate these risks, the EU’s market design reform requires member states to provide state-backed guarantees for PPAs.

30 Fabra and Montero (2023) analyze the trade-offs between technology-specific and technology-neutral auctions. For example, technology-specific auctions would sacrifice cost efficiency and may procure suboptimal amounts from each technology, but they reduce inframarginal rents. In general, the balancing between different effects will determine, on a case-by-case basis, whether technology-specific or technology-neutral auctions should be used in each case.
renewables (e.g., solar photovoltaic, wind, and run-of-river hydro), as the primary goal is to reduce generators’ uncertainty over cost recovery, leading to higher financing costs.\textsuperscript{31}

These contracts may be procured via pay-as-bid auctions, whose design forces bidders to compete at the margin and thus is less easily manipulated than uniform price auctions. Flexibility contracts, a type of CfD that exposes generators to price variation by calculating the settlement based on difference between the strike price and the average market price over an extended period, incentivize electricity production by dispatchable renewables (e.g., solar thermal and biomass) at times when their output is most valuable to the system. Fabra also proposes flexibility contracts for hydropower and nuclear plants, whose contribution to security of supply is critical. Finally, the system may rely on reliability options to promote investments in firm capacity (e.g., CCGTs, coal plants, and peakers). This type of one-way CfD provides a secure source of revenues for generators that can offer firm energy in exchange for becoming subject to price caps and penalties for not being available when they are most needed.

Fabra advocates for long-term remuneration for capacity, which may be determined based on administratively set prices or through a competitive process, to support flexible resources that are capable of shifting demand and supply across time and locations (e.g., energy storage and demand response). This option would ensure that investors receive an amount contributing to cost recovery, while preserving exposure to short-run price variations.

While Fabra’s proposal relies on CfDs to address some of the drawbacks of PPAs, CfDs may pose challenges to short-term operations. For example, two-way CfDs may lead to dispatch distortions because they remunerate generators based on actual electricity production (Crampes et al., 2023; Newbery 2021, 2023a). As an illustration, an electricity generator signing a two-way CfD will earn the strike price regardless of realized market prices. Thus, the generator has an incentive to produce if the strike price exceeds its production cost. To this end, the supplier may bid the lowest possible price to be sure it will be called on to produce, and it may be dispatched even though it is not the cheapest. Conversely, a generator whose production cost is higher than the strike price would bid high enough not to be called, leading to situations where its output may not be dispatched even though its production cost lies below the market price. Some of the potential dispatch distortions that CfDs may create can be eliminated with good contract design (e.g., a zero-price floor, as in Spain). Another option is to rely on financial CfDs (Schlecht et al., 2024).

\textsuperscript{31} Under a two-way CfD, generators sell their electricity in the market, and if the market price is above the strike price, they pay the difference between the two for each unit of output. Conversely, if the market price is below the strike price, generators receive the difference between the two for their metered output. Thus, two-way CfDs differ from the one-way CfDs discussed in Section 2, where payments may be payouts or clawbacks, but not both (European University Institute, 2023). For example, a reliability option is a one-way CfD because generators pay back any positive difference between the reference price and the strike price times the committed quantity. However, they don’t receive the difference. In contrast, with two-way CfDs, generators pay or receive the difference between the strike price and the reference price times the actual quantity.
4.7. Market-Based Integrated Resource Planning

Integrated resource plans (IRPs) are developed by electric utilities to plan which resources to procure or develop to meet demand over a 10- to 20-year horizon. As described in Cleary and Ratz (2021), the process typically begins with a review of the utility’s internal goals (e.g., its projected load growth) and state-mandated IRP requirements over the planning horizon. To help identify combinations of existing and other potential resources that could meet future needs, the utility relies on capacity expansion models that optimize around key goals (e.g., minimize cost, meet carbon emission reduction targets, and maintain reliability) and determine candidate portfolios of resources. These models are typically run for up to 30 years into the future, with hourly granularity for representative days of the year (e.g., a peak day and an off-peak day per calendar month, for a total of 24 calendar days per year), and group wind and solar resources by zone. Further, models typically rely on estimated technology and fuel costs as inputs. Candidate portfolios are then evaluated under various demand and weather scenarios using production cost models or stochastic models that quantify the risk associated with each portfolio. Selection of the optimal portfolio of resources considers model results as well as qualitative criteria and possible impacts such as job creation.

While IRPs have been used for decades, their processes and tools are evolving to accommodate a changing resource mix and state goals. In particular, some states are using all-source procurement to solicit competitive proposals from companies to meet resource needs (Henderson, 2018). A request for proposal (RfP) outlining procurement demand and evaluation criteria is typically the first step of the process. In response to the RfP, companies submit a range of parameters for their proposed technology (e.g., fuel type, heat rate, capacity factor, expected generation), which are converted into a bid by the utility soliciting the RfP. IRP planning that integrates bid data into portfolio selection is referred to as a market-based IRP (Bade, 2018). The key difference between traditional IRP planning and market-based IRP is that the former uses estimated costs as model inputs, whereas the latter takes bids as inputs. Backgrounders explaining the use of utility all-source procurements in Hawaii and Colorado are covered in Wilson et al. (2020) and Cleary and Ratz (2021).

Corneli (2020) proposes building on market-based IRPs but applying this approach to a much larger region instead of focusing on a single utility. Corneli also suggests that bids from various types of clean energy projects be evaluated by precise renewable integrating system models (PRISMs), which represent a new type of electricity system planning model with highly granular regional data of wind and solar availability across many hours and years and under a variety of weather conditions. A PRISM takes bids as inputs to the optimization process and conducts the analysis for a broad market region. For example, each project bids its annual levelized cost along with its expected energy production, and performance parameters such as the time profile of availability

32 Capacity expansion models being used for market-based IRPs (e.g., the model used by Xcel to evaluate competitive bids as part of Colorado’s integrated grid planning process) do not have the level of spatial and temporal granularity in wind and solar data availability proposed by Corneli (2020).
for solar and wind resources. Given bids and publicly available information as inputs, the model selects the combination of new and existing projects that minimizes the total system costs of meeting load under various weather and demand conditions, while satisfying the carbon budget during the next several years. The winning generators in this selection process enter into a type of contract known as a tolling agreement, in which generators essentially agree to be available to supply power whenever the underlying prime mover is available.

The long-term resource procurement is conducted every three years. As a possible start-up mode, participation in the proposed configuration market may be voluntary for new resources. Existing resources may participate in the market if expected revenues from the short-term energy market would be insufficient to support their operation costs (in this case, they would bid their going-forward costs in the configuration market).

Selected projects are offered a long-term hedging contract with load for an extended period. The contract is structured as a financial swap—that is, the project receives a performance-adjusted fixed revenue stream based on its as-bid levelized fixed cost and compliance with operator dispatch instructions. Failure to perform when called at a time when the underlying prime mover is available results in a penalty (hence the revenue stream is performance-adjusted). On the other side of the swap, the market operator takes over the plant dispatch, pays for its variable costs, and retains any energy market revenues. Thus, the LSE receives a floating revenue stream based on the short-term market price. If the fixed payments due to generation resources (reflecting their as-bid costs) are higher than the spot market revenues received by the LSEs, there is a net cost to the LSEs, which is allocated to the customers.

To the extent that the PRISM models used to evaluate bids account for the potential for correlated outages across similar resource types and enable the identification of complementary generators at low cost on a regional basis, the approach should lead to relatively efficient capacity investment decisions. Moving toward practical implementation would require a deeper consideration of several features and further exposition and development, as Corneli discusses in his proposal.

However, based on the information put forth so far, the practical implementation of this idea faces several challenges. First, several states coming to agreement on regional climate goals within an RTO footprint would likely be a challenge, as evidenced by the wide range of climate policies, from ambitious to no policy at all, across the states within some existing RTOs. Absent convergence between the RTO footprint and climate goals, it is not clear how an RTO could run a long-term market with a carbon budget where that budget applies to only part of its footprint. Second, while the configuration market builds on notional long-term planning exercises that are already used by some RTOs to identify implications of demand growth expectations and state clean energy policies for future resource needs (PJM Interconnection, 2017a; MISO, 2023), and these exercises

---

33 Capacity expansion models being used for market-based IRPs (e.g., the model used by Xcel to evaluate competitive bids as part of Colorado’s integrated grid planning process) do not have the level of spatial and temporal granularity in wind and solar data availability proposed by Corneli (2020).
have been proposed for broader use in the future by Joseph (2022, 2023), the proposal takes the role of the model to a whole new level of identifying investments and retirements. Implementing a configuration market would require substantial buy-in by states and FERC to rely on a highly complex model to govern investment choices that would ultimately allocate costs of the resource procurement across states.

Coming to consensus in the near term around such a big change in how investments are selected seems unlikely, especially given recent difficulties in getting support for a proposal that the Massachusetts Department of Energy Resources put forth for a less complex regional forward clean energy market, as described in Section 4.1 (Massachusetts Executive Office of Energy and Environmental Affairs, 2023). Nonetheless, although PRISM models may not become the engine for a long-term market clearing approach, they could play a central role in advisory RTO planning efforts to help inform regional assessments of resource needs and regional transmission expansion plans, both of which play a role in future strategies to support resource adequacy.

4.8. An Organized Long-Term Market

Gimon (2022) proposes the outlines of a long-term market for electricity that he labels an organized long-term market (OLTM). The OLTM helps inform investment and portfolio optimization decisions that will shape the mix of resources participating in day-ahead and real-time energy markets and thus how those markets operate. The OLTM design is built on three principles that Gimon believes should apply to forward markets in the electricity space: all advanced markets should be derivative of the MWh of energy traded in the day-ahead and real-time markets so as not to distort incentives or behavior in those markets; forward markets should be voluntary; and forward markets should be nondiscriminatory, transparent, and liquid, with a focus on enabling greater participation by the demand side than in current forward markets.

Gimon’s conceptual framework merges aspects of Corneli’s configuration market approach with a schedule-based forward energy product similar to Pierpont’s proposal. Like Corneli’s approach, the OLTM uses detailed planning and dispatch models combined with supplier bids of long-run marginal costs, and other parameters to identify the least-cost mix of suppliers to meet demand from among those that bid into the market. The OLTM then combines the energy output of the selected resources to create schedules for a specified profile of supply across the hours of the day (perhaps differentiated between weekdays and weekends), similar to the schedules traded in Pierpont’s proposal, albeit with a very different market-clearing approach.34 Consumers pay a common price for a particular schedule based on the average cost identified in

34 Gimon envisions that the schedules could be either as produced by the resource or for some fixed schedule designed to meet load (WRI/RFF, 2020).
the model solution, and suppliers are paid their as-bid price. Gimon suggests that the products traded in the OLTM could be either purely financial hedges or contracts for physical delivery. He indicates that the markets could operate annually and start many years in advance of delivery, much like renewable procurements run by states and utilities today, although other frequencies could be used.

He envisions three main types of actors in this market: producers, consumers, and facilitators. The producers include all utility-scale generators as well as some distributed generation. The expectation is that generators with little flexibility in generation would rely heavily on OLTM market participation for revenues, while more flexible resources could either help shape energy supply schedules in the OLTM or operate primarily in the shorter-term markets to help manage forecast errors or other sources of volatility. The consumers are largely LSEs but could also include other electricity resellers. Facilitators, expected to play a greater role as the market matures, include entities that balance demand and supply and are defined broadly to include storage facilities, transmission assets, flexible demand, virtual bidders, wholesale traders, and even insurance providers.

The link between the OLTM and resource adequacy comes (as in Wolak's and Cramton and colleagues' proposals) from an expectation that the introduction of an OLTM will lead to greater reliance on shortage pricing in the real-time market, which will help support resource adequacy. If the OLTM reduces LSE exposure to revenue fluctuation, and if having a forward energy market weakens incentives for generators to exercise market power in real time (Wolak, 2000), this arguably reduces the need for price caps in the real-time market. If price caps are raised or removed, increased reliance on shortage pricing could increase the incentives for suppliers to perform in real time by elevating the cost of making up shortfalls in their performance in the real-time market to make good on forward commitments. Additionally, greater reliance on shortage pricing in real-time markets will increase the appeal to LSEs of covering their loads with contracts purchased in the OLTM market to avoid exposure to price risk.

4.9. Dual Short-Term Market Designs

In response to the UK government’s 2021 objective of decarbonizing the electricity system by 2035 and unprecedented rises in gas prices in early 2022, the Department for Business, Energy and Industrial Strategy launched a review of electricity market arrangements (REMA) in July 2022. This public consultation considers a range of options for prospective reforms to the wholesale electricity market in Great Britain. In March 2023, the Department for Energy Security and Net Zero published a summary of responses to the REMA consultation, ruling out only six of the more than thirty options for electricity market reform. Options for capacity adequacy discussed in REMA include an optimized capacity market, capacity payments, reliability options,
targeted tenders, strategic reserves, and an equivalent firm power auction. As of June 2024, no timeline has been proposed to implement any of the options that advocate for fundamental changes to the operation of wholesale electricity markets.

Some analysts propose a split between two electricity markets for renewable versus nonrenewable resources based on the following considerations. Marginal cost pricing ensures both short- and long-run economic efficiency by operating plants with the lowest short-run marginal costs as much as possible, with more expensive ones called on only as needed. However, the design assumes dispatchable plants with varying short-run marginal costs. This assumption breaks down in the presence of intermittent resources characterized by zero or low short-run operating costs. In such circumstances, the electricity merit order structure is split between plants that cost little to run (such as renewables and nuclear) and gas- and oil-fired plants that operate as a residual to balance the system but drive wholesale electricity prices feeding through to electricity consumption. The gap between wholesale electricity prices and average electricity generation costs raised concerns during the 2022 European energy crisis (Grubb, 2022). In addition, as described in Section 1, the presence of increasing shares of renewable generation induces a decrease in energy prices, on average, leading to increased reliance on capacity payments and other mechanisms to support thermal power plants. Electricity systems are also split in terms of generation investments because low-marginal-cost generators (including existing nuclear and renewables) are mostly financed through mechanisms that provide long-term price security, such as feed-in tariffs and contracts-for-differences, rather than those based on signals from the electricity market. Hence, proponents of dual market structures argue that the theory of marginal cost pricing is inadequate to support efficient operation and investment in power systems and that electricity markets must be fundamentally reformed.

In the “two market” approach of Keay and Robinson (2017), intermittent plants with low and zero marginal costs operate as available, participate in an “as available” market, and are paid a price reflecting their current and future levelized costs of electricity. In contrast, dispatchable plants are called based on merit order, participate in a separate “on demand” market that is similar to current wholesale electricity markets, and are paid on the basis of short-run marginal costs. Some generators (e.g., nuclear, biomass, and storage) could choose the short-term market they participate in. The grid operator would first rely on the “as available” market to meet demand and then purchase dispatchable energy from the “on demand” wholesale market when demand exceeds the available renewable energy.

Consumer choice is central to Keay and Robinson’s proposal. Consumers may choose between different types of supply (“as available,” at low and stable prices, and “on demand,” at higher and more volatile prices) or combinations thereof, via separate metering. Therefore, consumers who express their preferences for each type of power would receive an “as available” price (reflecting the purchase price of electricity from the “as available” wholesale market) as long as sufficient power of that type was being generated and a more expensive “on demand” price (reflecting the costs of supply in the “on demand” market, including any congestion and carbon costs) to the extent
that their supplier does not have sufficient “as available” supply to meet demand. Consumers who do not have the right type of metering or equipment would pay the “on demand” price.

In the long term, investments in renewable and thermal power plants would be remunerated solely from the market. In the “as available” market, prices would cover both investment and operating costs of low-carbon resources. However, many existing renewable power plants currently operate with feed-in tariffs or other forms of support, and it may be necessary to continue to provide such support for these plants in the short and medium term. In the “on demand” market, price caps could be raised to allow scarcity pricing, given the ability of consumers to reveal their preferences for different types of power. Capacity payments would not be necessary in the long term, but they may be retained in the early stages of implementation of the dual market structure if there is reluctance to invest in generators participating in the “on demand” market.

An alternative dual market structure proposed by Grubb et al. (2022) aims at separating electricity prices from gas prices and connecting consumers more directly to the cost of renewable generation. Their approach is based on a green power pool (GPP), which they define as “a combined volume of electricity from many renewable generators, made available to consumers directly rather than through the current wholesale market.” The pool operator serves as an aggregator and buys electricity from the wholesale electricity market when pool generation cannot satisfy its customers’ needs, while selling to the wholesale market when pool generation exceeds its customers’ needs. In the latter case, some pool generators may be directed or paid to cease generation. Consumers contract with the pool operator and pay an “assured” price reflecting the average cost of renewable generation for their electricity consumption when pool generation exceeds pool demand. When pool generation is insufficient to meet demand, consumers pay a combination of the cost of renewable generation for their proportionate share of pool generation and the marginal cost of generation from thermal resources for demand exceeding that share. In other words, the short-run operating cost of thermal generators is added to the consumer bill for consumption only above the volume of “as available” renewable generation. More complex two-tier pricing structures that preserve the incentives provided by marginal pricing (e.g., with proportionate consumption paid at the assured price and additional consumption charged at the price from the “on demand” wholesale market) may be designed to enhance consumer engagement and flexibility.

Dual market designs discussed in this section differ from that of Pierpont and Nelson (2017) because Keay and Robinson (2017) and Grubb et al. (2022) focus on two short-run energy markets for operations, whereas Pierpont and Nelson couple a long-term energy auction with a short-term market. A split electricity market represents a radical departure from existing designs that may reduce short-run market efficiency (Pollitt et al., 2022). Near-term implementation appears difficult because of debates around the structure of long-term contracts, insufficient volumes of renewables to match a substantial fraction of electricity demand, and limited storage options to guarantee firm power to the pool’s customers other than through access to “as required” sources in the wholesale market.
An important element of the design to be further developed is which generators would be directed or paid to cease generation, and on which terms, if the electricity system has an overall surplus of renewable generation. Identification of customer groups that have priority access, implications for consumers who would not be eligible to receive electricity from a GPP (at least in its initial years), and interactions with the market for power purchase agreements would also have to be carefully assessed. In the early stages, design of a GPP could leverage existing government underwriting of fixed-price long-term contracts. For example in Great Britain, where a government-backed counterparty to fixed-price CfDs exists, the Low Carbon Contracts Company, a GPP may bring together electricity from renewables already operating on CfDs. This design would not require significant changes to the financial terms of existing CfDs and may serve as a proof of concept for future developments. Additionally, existing government underwriting may help determine how the economic benefits of a GPP could be fed through to consumers, since several options for consumer access to the GPP are available (e.g., priority access may be given to vulnerable groups of consumers such as fuel-poor households).  

4.10. Swing Contracts

In 2011, Tesfatsion and collaborators at Iowa State University and Sandia National Laboratories began working on a new approach to wholesale electricity market design centered around swing contracts. A linked swing-contract market design was proposed in 2013 (Tesfatsion et al., 2013) and refined over the years (Heo and Tesfatsion, 2015; Li and Tesfatsion, 2018; Tesfatsion, 2021, 2022). In Tesfatsion’s view, all wholesale electricity markets must necessarily be forward markets because of the speed of real-time grid operations. A problematic aspect of current RTO-managed wholesale electricity markets is their artificial distinction between the delivery of energy based on an optimal dispatch schedule and the provision of reserve to handle any real-time deviations from that schedule. Instead, the only product that can be offered in wholesale electricity markets is reserve, which ensures availability of power production capabilities for dispatch in future operating periods. As a result, Tesfatsion’s proposed design consists of a linked collection of ISO-managed forward reserve markets for future periods $T$.

Core types of participants in each reserve market include LSEs, VERs whose generation is not fully firmed by storage, and dispatchable power resources (DPRs), such as dispatchable generators and storage facilities. LSEs submit reserve bids that

36 Newbery (2023b) argues that a “two market” solution is not necessary for existing renewable generators selling on CfDs because the Low Carbon Contracts Company already receives the difference between the wholesale market price and the (fixed) strike price of the contracts, and this difference is passed on to consumers.
represent fixed or price-sensitive demands for power path delivery during period $T$.37
A power path for period $T$ refers to a sequence of power injections or withdrawals at a grid location during $T$. VERs submit forecasts for their non-dispatched power injections at specific grid locations during $T$ without accompanying price information that signals required compensation. Finally, DPRs submit reserve offers by means of swing contracts, defined as “two-part pricing contracts that offer a collection of dispatchable power-paths for possible RTO dispatch during $T$” (Tesfatsion, 2021). A reserve offer consists of four elements: an offer price to cover ex ante any avoidable fixed cost that the DPR must incur to ensure the availability of its offered reserve (e.g., start-up cost); a set of contract exercise times available to the RTO; a production possibility set that conveys the degree of swing (flexibility) in the DPR's offered reserve in terms of attributes such as delivery location, start time, and ramp rate limits; and a performance payment method to recover ex post any variable cost (e.g., fuel cost) incurred for verified delivery of a power path in response to RTO dispatch instructions. Thus, reserve offers permit DPRs participating in a swing-contract market to ensure their revenue sufficiency—that is, the revenues they earn will cover all the avoidable costs (i.e., avoidable fixed costs plus variable costs) incurred from market participation.38

The costs incurred to ensure balancing of net must-service load during period $T$ are allocated across market participants in accordance with their contributions to these costs.39 Specifically, cost allocations are determined in accordance with the three-part ISO Cost Allocation Rule of Tesfatsion (2021). For example, VERs with non-dispatched power injections during $T$ receive payments (negative charges) to the extent that their injections reduce the RTO's need to dispatch cleared DPR reserve offers. Conversely, VERs whose non-dispatched power injections during $T$ increase the anticipated volatility of net must-service load are charged a corresponding portion of the ISO's reserve acquisition costs for $T$. VERs thus face revenue insufficiency risk because their period $T$ reserve acquisition charges can exceed the payments they receive for period $T$ reductions in variable cost. Consequently, VERs have an incentive to firm their non-dispatchable power with storage, thereby enabling their participation in reserve markets as DPRs.

---

37 A fixed reserve bid specifies a must-service demand for delivery of a designated power path at a given location in period $T$, with no accompanying price information indicating willingness to pay. In contrast, a price-sensitive bid consists of a set of power paths for period $T$ along with information conveying willingness to pay for each power path. Traditional downward-sloping demand functions linking price and power usage at a given time represent a special case of price-sensitive reserve bids, as shown by Tesfatsion (2021).

38 Reserve offers may be submitted as firm contracts or option contracts. Firm contracts obligate the holder (buyer) to take delivery of energy at the exercise time, whereas option contracts give the buyer the right, but not the obligation, to procure services from the seller. Tesfatsion (2021) recommends firm contracts for handling normal net load balancing requirements and option contracts for handling contingencies.

39 Net must-service load consists of non-dispatched power withdrawals (plus inadvertent power losses) minus non-dispatched power injections.
The RTO undertakes a contract-clearing optimization that maximizes expected total net benefit (i.e., total benefit minus total avoidable cost) to market participants, conditional on initial state conditions and subject to power balance, transmission, and other system constraints. Further, the three-part ISO Cost Allocation Rule ensures a direct correspondence between the two-part pricing reserve offers submitted by dispatchable power resources and the resulting charges paid by customers for both price-sensitive power usage and fixed power usage. By design, cleared seller bids cover all the offer price and performance payment costs, and the Cost Allocation Rule ensures that these costs are charged to market participants in an equitable manner.

As noted, Tesfatsion does not envision a single swing-contract market, but rather a collection of markets with look-ahead periods ranging from multiple years to minutes. Linkages among these markets are established through the contracts: reserve bids and offers cleared in earlier markets are carried forward as contracts that can be updated in subsequent markets. Forward reserve markets pertaining to a distant future operating period $T$ typically have to include a long-term forward market (LTFM) with a one-year or longer look-ahead horizon to encourage appropriate entry of new DPRs. Reserve bids and offers cleared in the LTFM can be adjusted in subsequent markets, such as the day-ahead market, intraday, and real-time markets, before the start of the operating period. However, credible specification of system constraints in the LTFM may be difficult because of deep uncertainty arising from its long look-ahead horizon (Tesfatsion, 2021). Thus, the ISO might need to supplement any LSE-submitted reserve bids with reserve bids based on its own long-term assessment of reserve needs for period $T$. Tesfatsion (2021) also notes the challenges associated with the representation of customer preferences in LTFM reserve bids, given uncertain power usage needs in the distant future.

On the other hand, reserve offers into the LTFM are submitted by entities that already own power facilities or are contemplating investment in such assets. For example, a swing contract for a new generation unit would define an offer price covering its expected capital costs and a performance payment method offering compensation for services contingent on market conditions at the time of performance (e.g., fuel prices). The offer price payment received by cleared participants in long-term swing-contract markets functions as a capacity payment, while the performance payment received by cleared participants is similar to the day-ahead/real-time energy market payments received by cleared capacity market participants. There are, however, some key differences. Capacity payments are determined by the intersection of demand and supply curves for generation capacity, and generation capacity cleared in the capacity market must be offered as fixed quantity obligations in the energy markets. As a result, there is no guarantee that cleared capacity market participants fully cover their avoidable fixed costs and actual performance costs. In contrast, the two-part pricing form of the reserve offers that DPRs are permitted to submit into the LTFM ensures revenue sufficiency through appropriate specification of the offer prices and performance payment methods. Also, since all performance payments are paid ex post contingent on verified performance, a DPR can make its required performance payments contingent on actual period $T$ market conditions (e.g., fuel prices) tailored to its own unique operational attributes.
The linked swing-contract market design may enhance resource adequacy because of its emphasis on reserve as the primary transacted product. Further, the design seeks to ameliorate revenue insufficiency concerns for dispatchable resources through its distinctive two-part pricing form for reserve offers, and it provides incentives for non-dispatchable VERs to reduce their revenue insufficiency risk by firming up their non-dispatchable power with storage. In particular, the long-term forward market would encourage efficient investment in new dispatchable resources, as well as efficient retention and exit of existing dispatchable units. We note that swing contracts could be designed to allow for availability and performance cost compensation of various attributes of a power path, including reliability services such as frequency regulation and primary frequency response. However, to date, work on the linked swing-contract design has not addressed the provision of these important types of ancillary services.

The swing-contract proposal represents the largest departure from current practice of all the proposals considered in this review. As such, it raises a number of questions with respect to some of the concerns that other market designs seek to address. For example, with its focus on resource adequacy, the swing-contract market design does not address financing uncertainties for variable resources that are not explicitly firmed by storage. Such generators are not allowed to bid into the market but instead are rewarded or charged based on their anticipated contribution to the system’s need for balancing resources. To qualify to bid into this market, variable resources must be coupled with a firming resource such as energy storage that enables it to become dispatchable. Exactly how much storage is needed for a resource to transition from one category to another is uncertain and would need to be specified before such a market is implemented. Implementation does appear to be challenging, given the complexities of power flow bidding, which involves articulating numerous potential flow offers to include in a bid that the system operator can account for in finding the most beneficial solution. Since the swing-contract model represents a big change from current market mechanisms, making this change would require careful testing, planning, and a clear articulation of a transition strategy. Another uncertainty is the extent to which the complexity of the market creates opportunities for some participants to manipulate the market in ways that benefit them.
5. Discussion

We divide the 11 market design proposals explored in Section 4 into five groups, as shown in Table 1. Each row of the table represents a proposal, and each group of proposals is highlighted in a different color. The color shade within a group indicates the stage of development (light: nascent/conceptual; darker: some specifics; darkest: formally modeled or draft rules for implementation, or both). The first group (in gray) includes the forward market for capacity and clean energy attributes proposed by Spees (2020). The second category of proposals (in yellow) considers long-term contract auctions and energy markets: Wolak (2022), Cramton et al. (2024), Pierpont and Nelson (2017), Pierpont (2020), and Fabra (2023). The third group (in blue) encompasses proposals that rely on system-planning models that may be informed by bids to select investments: Corneli (2020) and Gimon (2022). Keay and Robinson (2017) and Grubb et al. (2022) consider separate short-term markets for different generation technologies and are highlighted in green. Finally, the swing-contract wholesale market design of Tesfatsion (2021), which would entail fundamental changes in product definition, contract design, and settlement rules for current electricity markets, is shown in orange.

We compare the proposals based on 10 criteria. Each column of Table 1 represents a criterion. Based on our assessment of the proposals, we assign a full circle to a table cell if a criterion is a primary focus of the proposal, a half circle if the criterion is discussed to some extent in the proposal, and a blank circle when the proposal does not provide enough information to assess the criterion. An empty cell indicates that the criterion is not explicitly considered in the proposal or is outside the scope. We have not conducted modeling exercises or empirical analyses to compare the proposed designs with each other or with existing approaches such as capacity markets. As discussed in the next section, such assessments would be valuable to more thoroughly assess and compare the proposals’ ability to deliver on key performance metrics.

The frequent mismatch between periods of high renewable resource availability and periods of high electricity demand, coupled with the variability and unpredictability of renewable energy supply, means that potential challenges to grid reliability extend beyond peak periods to encompass all hours of the day and increasingly occur during hours of peak net load instead of peak load. As a result, the focus is shifting from resource adequacy to energy adequacy. Proposals in our review support energy adequacy in various ways. For example, in Wolak’s design, the uncertainty in the final SFPFC obligation and higher price caps in the short-term energy market provide strong incentives for suppliers to be available and produce in real time. In contrast, Corneli’s approach supports energy adequacy by centrally selecting the optimal mix of investments based on highly granular regional data of renewable availability across many hours and years and under a variety of weather conditions. Tesfatsion offers yet another approach, whereby remuneration for ex post verified performance provides incentives for suppliers to be available and follow dispatch signals.
Ensuring energy adequacy under increasing shares of variable and unpredictable renewable energy supplies requires resources that meet evolving system needs, including flexible generation resources, demand response, and energy storage. About half of the proposals in our review discuss mechanisms to procure flexible generation resources, such as through long-term remuneration for capacity (Fabra) or forward reserves (Cramton et al.). Additionally, six proposals contain explicit provisions to promote energy storage participation in wholesale markets. For example, in Pierpont’s and Tesfatsion’s designs, suppliers are incentivized to combine renewable energy with storage technologies to manage risks associated with variability. Fabra proposes additional payments to support investments in energy storage, while long-term energy market designs seek to achieve this objective through higher short-term price spikes in the energy market. Since hourly forward prices years ahead may provide opportunities to engage in price arbitrage strategies that improve the economics of energy storage, Cramton and colleagues’ proposal receives a higher ranking on this criterion than other proposals. On the other hand, demand-side participation is typically not a focus of the proposed designs in our review, with the exception of Keay and Robinson (2017). In the long-term energy market proposals, consumers are expected to be more responsive in the short-term energy market, to the extent that offer caps are raised as a result of forward hedging and retailers offer innovative plans to their customers. We return to this point later in this section.

Growing penetration of variable energy resources brings with it greater fluctuations in the composition of energy supply across hours and seasons and accompanying large variations in the marginal cost of energy supply and energy prices. As a result, energy market revenues to generators could vary substantially depending on when generators are available and operating. The associated uncertainties surrounding energy revenues can raise financing costs for new renewable generators, with potential adverse impacts on energy prices paid by consumers. Some proposals in our review aim at ensuring revenue sufficiency for all generation technology types. For example, in Corneli’s design, new resources participating in the proposed configuration market would bid their annual levelized cost, while existing resources would bid their going-forward cost. In other proposals, the focus appears to be on specific resource types. For example, the two-part pricing feature of Tesfatsion’s design permits dispatchable resources to ensure full coverage of their avoidable costs, but it is not clear whether renewable energy sources would be revenue sufficient. Pierpont and Nelson (2017) and Pierpont (2020) do not discuss revenue sufficiency but tackle the related issue of enabling long-term, low-cost financing to support investments in capital-intensive resources.

Market power and energy affordability are considered in only two proposals. Section 4.3 discusses how Wolak’s and Cramton and colleagues’ designs differ with regard to the type of market power that is targeted, while energy affordability is a key objective of proposals conceived in Europe around the time of the 2022 energy crisis. Further, Spees and Corneli explicitly include clean energy objectives, while some of the other designs can be adapted to integrate the procurement of Renewable Energy Certificates. Finally, most proposals do not explicitly account for network constraints. Spees’s design could be segmented by region to account for interzonal transmission constraints, as in current capacity markets. Tesfatsion’s proposal considers system constraints in the shorter-term swing-contract markets but notes that “the ability
to incorporate such constraints with specificity will presumably be low for long-term forward markets, given their long look-ahead horizons” (2021).

Our evaluation highlights several open questions and directions for future research. First, the determination of mandatory energy purchase obligations of load-serving entities ex ante and enforcement of mandatory obligations deserve closer scrutiny. Australia’s Retailer Reliability Obligation offers an interesting point for comparison with some of the proposed designs. The mechanism obliges retailers to secure forward financial contracts for energy or capacity to support grid reliability if the Australian Energy Market Operator identifies a reliability gap three years and three months out (COAG Energy Council, 2019; AEMC, 2024). If the gap is still projected one year and three months out, liable entities must also disclose their contract positions. The amount of required qualifying contracts is set ex ante by the market operator based on its peak-load forecast. Since no reliability gap was ever identified one year and three months out, contracting obligations have never been disclosed, and no verification mechanism has been implemented to assess noncompliance (i.e., whether retailers procured sufficient contract positions to cover their customer peak load).

In the proposals we reviewed, forward retailer obligations are based on load forecasts at different look-ahead horizons. For example, in Wolak’s proposal, the mandatory schedule of purchase obligations is based on forecast demand up to four years in advance. Constructing these forecasts can be challenging as the United States and other regions anticipate historically high rates of electricity demand growth due to a combination of factors, including onshoring of manufacturing, growing demand from data centers, and electrification of buildings and transport (Wilson and Zimmerman, 2023). The ultimate impact of these factors on total load and load shapes is uncertain. As a result, forward auctions that procure a substantial fraction of energy years ahead may result in energy over- or under-procurement and may require significant adjustments ex post if 100 percent of realized electricity demand must be covered by forward contracts. In this regard, frequent forward auctions with a more gradual schedule of retailer obligations, coupled with energy options to purchase unexpected real-time quantities (as proposed by Cramton and colleagues), may present an advantage over big event auctions because they allow for continual load forecast adjustments.

A related question pertains to how forward markets affect the opportunity for demand to engage in real-time markets or to face time-varying prices that reflect changes in system conditions. Electrification of buildings and transport introduces new load sources, such as water heating and electric car batteries, for which electricity consumption can be separated in time from consumption of energy services. This separation provides an opportunity for shifting electricity consumption to take advantage of lower-priced hours and to help match load with periods of abundant renewable energy and away from times of shortage. To enable these opportunities, it is important that forward markets not mute or distort incentives for individual customers to engage in time-varying pricing regimes that enhance system efficiency by shifting load to meet supply.

In the proposals under review, retailers participate in the long-term markets on behalf of their customers, and forward quantities represent total forecast demand served
by the retailers. Hedging requirements reduce the need for price caps in short-term markets, leading to more robust shortage pricing and anticipated greater demand-side participation of retail customers in real-time markets. Further, since retailers could profit from selling unused contracted energy in real time, they would offer dynamic rates that expose the consumer to real-time prices on the margin.

An alternate view is that retailers may have little incentive to promote dynamic pricing in real time because forward contracts allow them to incur less volatile total procurement costs and face lower bankruptcy risks. Economists have long argued that greater exposure of customers to time-varying prices would enhance economic efficiency and save consumers money, but such approaches to retail prices have been slow to take hold, even as smart meters and other enabling technologies proliferate. More generally, the interplay between long-term hedging requirements and incentives for demand-side participation in real time remains an important topic for further investigation.

Our third consideration relates to the interactions between long- and short-term markets. Most proposed designs in our review provide little discussion of whether long-term markets and contracts would be compatible with efficient operations in short-term markets. Some proposals rely on contract designs that are not fully specified or may create short-run inefficiencies. For example, in the proposal of Grubb et al. (2022), if the electricity system has a surplus of renewables and must-run plants, some pool generators may be directed or paid to cease generation. However, which pool generators would be curtailed, and on which terms, would be determined by the specific structure of existing and future CfD contracts, and this structure is outside the scope of the proposal. More research is needed to ensure compatibility between long-term contract designs and efficient short-term operations, including efficient curtailment outcomes (Joskow, 2022).

Finally, most long-term market designs do not incorporate transmission and generator constraints. While modeling system constraints with specificity several years in the future is likely not possible because of frequent changes in network configuration, incorporating roughly anticipated constraints could limit unintended consequences associated with market design. For example, disregarding transmission constraints in the forward market may result in negative congestion surplus, triggering the need for additional mechanisms to make the grid operator whole (NEMO Committee, 2020). As the share of intermittent renewables increases, coordinating the physics of electric power systems operation with the design of long-term markets for energy and operating reserves becomes increasingly important to ensuring decarbonization and security of supply at the least economic cost (Wolak, 2021b).

Consider a two-node system with a low-cost generator at node A, a consumer and a high-cost generator at node B, and a limited transmission line between them. The low-cost generator at A and the consumer at B sign a forward contract with a quantity that exceeds the capacity of the line and a price that is below the marginal cost of the high-cost generator colocated with load. If the transmission constraint binds in real time, the more expensive generator is dispatched to serve load and sets the electricity price at node B. The system operator charges the consumer the nodal electricity price for withdrawing power and pays the generator the local electricity price for injecting power. The total amount paid by the consumer would be lower than the total amount received by the generators, yielding a negative congestion surplus.
Table 1. Comparison of Proposals

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Spees</th>
<th>Wolak</th>
<th>Cramton</th>
<th>Pierpont and Nelson</th>
<th>Pierpont</th>
<th>Fabra</th>
<th>Corneli</th>
<th>Gimon</th>
<th>Keay and Robinson</th>
<th>Grubb</th>
<th>Tesfatsion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ensures energy adequacy</td>
<td>●</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Procures generation resources that support flexibility and other essential grid needs</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Fosters demand-side participation</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Encourages energy storage participation</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Promotes revenue sufficiency</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Lowers financing cost for capital-intensive resources</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Weakens incentives to exercise market power</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Promotes energy affordability</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Includes clean energy goals</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Considers network restraints</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

Resources for the Future 40
6. Conclusions

The main goal of electric power systems is to provide reliable electricity at the lowest cost to consumers. Electricity markets are designed to achieve this goal by satisfying the objective of short-run efficiency (i.e., making the best use of existing resources) and long-run efficiency (i.e., providing adequate incentives for efficient long-run investment). Basic market design features that lead to good short-run performance include short-term (day-ahead and real-time) wholesale markets for energy and ancillary services, locational marginal pricing, and financial transmission rights for congestion hedging (Joskow, 2008). These are the features of the standard market design that has prevailed in the United States, despite some variations in regional rules (FERC, 2002). The objective of long-run efficiency has proven to be more challenging.

In response to concerns that energy markets do not provide adequate incentives for investment in sufficient amounts of generation capacity or a generation mix that satisfies acceptable reliability criteria, policymakers have implemented approaches that include investment coordination through capacity markets and improved scarcity pricing to provide stronger investment incentives. Existing market designs have been successful at managing system operations and maintaining resource adequacy in thermal-based electricity systems (Joskow, 2019; Hausker and Palmer, 2021). However, current designs are poorly suited to address challenges that were unknown, or less material, in the early designs of organized electricity markets, such as rapid growth in renewable energy generation and ambitious decarbonization objectives. This paper has provided an overview of electricity market designs that seek to support the development of reliable mixes of generation resources, promote long- and short-run efficiency, and help achieve energy affordability and decarbonization objectives.

The different stages of development and points of emphasis of the various designs reviewed in this paper make it difficult to identify specific recommendations for policy at this time. Clearly, more research is needed on several aspects, as discussed in Section 5. Our review does suggest that researchers providing greater support for one another’s efforts to develop new approaches is likely to have substantial value. It will be important to actively seek to share the results of ongoing research with market operators, federal regulators, market participants, and other stakeholder communities involved in market rules and governance, operation, and oversight. An important part of that sharing will be developing proposals that are accessible to these broader audiences and interacting with them at various points in the research. Such interactions will help researchers understand stakeholder concerns and potential points of confusion, enabling them to address these concerns as proposals undergo further development and refinement.

A factor that complicates comparison and evaluation of various design proposals is the evolving nature of our understanding of what resource adequacy means and how best to measure and support it in the future. While economists have long argued for using measures of the value of lost load to set resource adequacy and reliability standards, that value likely depends on a variety of features of the load loss experience, including length of the outage, geographic footprint of the outage, and outage frequency. Engineers have increasingly focused on different metrics of system reliability and
suggest that assessments of system performance should have multiple dimensions (ESIG, 2021). Coming to an understanding of what is required to support resource adequacy first requires a shared understanding among system engineers, economists, and regulators of how best to define society’s goals with respect to resource adequacy of the system.

For market operators in the electricity sector, engaging with researchers to further develop and test their proposals for new markets will help ensure that the proposals are relevant and implementable. One concrete way that this community could support researchers is by facilitating access to data that could be used to test some of these proposals in models or to explore the effectiveness of particular approaches to markets and contract structures that have been employed in the past. Ultimately, a fuller quantitative examination in a unified modeling framework would be valuable to more thoroughly assess and compare the proposals’ ability to deliver on the criteria we have identified. The success of a new market design will also depend on how well market participants understand the rules and associated rewards for particular bidding strategies. Market participant behavior is often not considered or empirically assessed in electricity market design. While laboratory experiments may be useful to this end, a fundamental challenge for research of this type is to design experimental markets that exhibit the same principles as the natural setting (Plott, 1987).

Market participants and other stakeholders could also support researchers by creating opportunities to explore proposal ideas with practitioners and thereby help identify aspects that need further development or should be abandoned, as well as new areas where research could be helpful. One challenge that market operators, regulators, and other practitioners face in this regard is that their attention is often consumed by near-term issues, leaving them with little time to focus on solutions that are in early stages of development. Bringing near-term aspects and longer-term market needs together can help enable the development of better practical long-term plans for future markets and of transition strategies that will guide today’s electricity markets to where they need to be to support the grid of the future.
References


Borreson, J. M. (2022). Houston, we have a market design problem: Why the legislative response to Winter Storm Uri does not yet develop a more efficient market mechanism to ensure reliability. *Oil & Gas, Natural Resources, and Energy Journal* 7 (4), 865-910.


