Atlantic Council
GLOBAL ENERGY CENTER

Preparing for Decarbonization: Reforming US Power Markets for the Energy Transition

By Ben Hertz-Shargel

Atlantic Council
GLOBAL ENERGY CENTER

The Global Energy Center promotes energy security by working alongside government, industry, civil society, and public stakeholders to devise pragmatic solutions to the geopolitical, sustainability, and economic challenges of the changing global energy landscape.

Cover: Transmission lines in the fog, in Washington state. Unsplash/Peter Geo (@petergeokent)

Atlantic Council
1030 15th Street NW, 12th Floor
Washington, DC 20005

For more information, please visit
www.AtlanticCouncil.org.

ISBN-13: 978-1-61977-180-2

May 2021

This report was designed by Donald Partyka and Nikita Kataev.

The Atlantic Council wishes to thank the MacArthur Foundation. This report would not have been possible without its support.

Atlantic Council
GLOBAL ENERGY CENTER

Preparing for Decarbonization: Reforming US Power Markets for the Energy Transition

By Ben Hertz-Shargel

Table of contents

Executive summary

Organized electricity markets in the United States today face growing regulatory, policy, and commercial pressures stemming from the decarbonization of the power and transportation systems. State clean energy standards and the falling cost of renewables, particularly utility-scale wind and solar, have hastened the penetration of these resources into the generation fleet. Not only are these resources less dependable for grid balancing than the coal-fired and—increasingly—natural gas-fired generators that they displace, but their low marginal cost to operate depresses energy prices to the point that, under current market designs, even otherwise economic conventional units can be threatened. At the same time, the Federal Energy Regulatory Commission (FERC) has ordered markets to update their legacy participation models in order to accommodate new technologies, such as battery storage and virtual power plants, which are often interconnected off the bulk transmission grid and even behind customers' meters. Electric vehicles, set to be the power system's first mobile energy resource—and likely its largest storage resource overall—have yet to come under federal regulatory focus. But, with the accelerating changeover of the vehicle fleet, that day is not far away.

Ensuring energy security with a fleet relying significantly on intermittent renewables and smart inverter-based resources—none of which contribute to the rotational inertia on which the bulk power system has historically depended—is unchartered territory for system operators.[1] This new mix of resources poses market design challenges as well, creating more financial interdependencies with retail energy markets, and physical interdependencies with distribution systems. Wholesale markets must address these challenges against a backdrop of conflicting state mandates and priorities, most notably around resource adequacy and carbon pricing. With the expansion of the Western Energy Imbalance Market (EIM) and the Regional Greenhouse Gas Initiative (RGGI) in the northeast, four of the nation's seven real-time energy markets include states with and without a cap-and-trade system for generator carbon emissions. Europe faces similar challenges with respect to its neighbors, who are not parties to its emissions trading system (ETS) and represent a source of carbon leakage. These markets are, thus, thrust into the position of deciding whether and, if so, how to incorporate carbon emissions into market clearing and prices.

The challenges brought on by the energy transition do not arrive in a vacuum. They join, and have exacerbated, long-standing deficiencies in organized markets related to price formation, alignment between real-time and forward markets, and out-of-market operator actions. Comprehensive reform is required to simultaneously address these challenges, old and new, in order to ensure a successful energy transition. This reform will be multifaceted, regional, and incremental, and has begun playing out across the United States' seven organized markets through various regulatory initiatives.

Five key questions underly this reform, whose resolution will determine the viability of markets and their role in the coming decarbonized power system.

1 What role, if any, should organized capacity markets play?

Centralized capacity markets have a poor track record of delivering resource adequacy. To remain viable, they must leverage, rather than discriminate against, renewables and demand-side resources, and evolve the product they procure into one that meets the flexibility needs of real-time energy markets.

2 What is the role of organized spot markets in accounting for generator emissions?

Independent system operators (ISOs) and regional transmission organizations (RTOs) have no mandate to incorporate state policy objectives into markets, and have three radically different options with respect to individual state actions to subsidize clean resources or otherwise price carbon: accommodate such actions passively; counteract them, in order to undo their economic effect on participants; or directly integrate them into operations.

1 The US Energy Information Administration's "Annual Energy Outlook 2020" forecasts that renewables will represent 38 percent of electricity generation in the United States by 2050 under current laws and regulations. The continued adoption of state renewable portfolio standards (RPS) and clean energy standards (CES), a federal RPS or CES, or a federal price on carbon could significantly increase that share. Rotational inertia is a feature of conventional generation, including coal, natural gas, and nuclear, in which a large rotating mass within the facility helps the grid withstand sudden changes in load or generation.

3 What price formation enhancements are capable of solving the "missing money" problem and reducing operator dependence on out-of-market actions?

Two mutually reinforcing challenges plague ISOs and RTOs today. Competitive resources earn insufficient revenue in energy markets to meet their costs and to make investments that are most effective for society. This "missing money" causes, and is reinforced by, system-operator reliance on actions taken outside of economic dispatch. Solving these two issues requires revisiting the coordination of forward and spot markets, as well as price formation within markets.

4 How should distributed energy resources (DERs) and flexible load resources participate in markets?

New products and participation models must be developed for DERs and load resources, such as electric vehicles, to effectively participate in markets. Aggregation models present novel complexities, however, and enhancements to demand-side models may be required to incent load resource owners to participate as flexible demand, rather than supply. FERC's landmark Order 2222 formally opened markets to DER aggregations, but many key questions remain.

5 How can markets ensure energy security when supply is intermittent?

The intermittence of renewables renders the bulk power system vulnerable to both momentary fluctuations in power and prolonged supply-demand imbalances. Maintaining system stability will require innovative—and massively scaled—use of complementary resources, such as CCS, nuclear, geothermal, gas-fired generators, as well as renewables themselves, which, while limited by the sun and wind, have proven surgically dispatchable.

This report makes the following policy recommendations, which will help energy markets resolve these questions in a manner most adaptive to the exigencies of the energy transition.

- Consider decentralizing resource adequacy, returning responsibility to load-serving entities and greater authority to states, and changing the focus of resource adequacy from installed capacity to flexible capacity.[2] Resource adequacy must become defined in terms of flexible capacity and procured through products that satisfy real-time essential reliability needs. Advanced technologies, including distributed resources, must be eligible to provide it.

- Integrate carbon prices imposed by states into real-time markets, including through regional cap-and-trade programs such as RGGI. This is necessary to prevent emissions leakage and indiscriminate cost increases across consumers, which arise when markets passively accommodate or counteract carbon prices.

- Take a multi-pronged approach to shore up scarcity pricing in real-time energy markets. One priority is to base operating reserve demand curves on customers' value of lost load, rather than on assumed resource costs. Another is to ensure that real-time prices are effective for a majority zero marginal cost fleet, which may require harmonizing long- and short-term markets such that the financing costs that displace fuel costs are reflected in these prices.

- Recognize renewables as dispatchable resources and leverage the vast flexible capacity they offer. Markets should develop technology-neutral participation models for variable energy resources that enable both ramping and fast-response ancillary services, while accounting for the opportunity cost of real-time energy.

- Do not impose unreasonable limitations on DER aggregation formation or ability to provide market services. Aggregations should be permitted to span load service entities as well as transmission nodes, provided that distribution factors can be produced. They should not be required to provide metering and telemetry below the transmission node level. ISO/RTO requirements to the contrary should be closely scrutinized by FERC on the basis of technical necessity.

- Embrace statistical estimation-based telemetry methodologies, such as the New York Independent System Operator (NYISO) alternative telemetry option, in order to facilitate residential DER participation in real-time markets. These

2 Flexible capacity refers to generation capacity (measured in megawatts) that can be dispatched up or down by the system operator, in order to maintain grid stability.

methodologies permit smaller resources to share status information with the system operator based on occasional measurements and statistical estimation, rather than costly high-frequency measurements. Such methodologies should be permitted to incorporate DER and smart device data, and should be approved solely on the basis of demonstrated accuracy.

- Permit dual participation in retail and wholesale markets whenever the unbundled services for which DERs earn revenue in each market are non-overlapping. This is the case, in particular, when the purpose of the retail program is distribution system value. New York's Value Stack methodology serves as a reference.

- Study whether flexible load resources would provide greater value in markets as price-sensitive demand, compared to supply-side demand response, as they largely participate today. This includes exploring variants of PJM's Price Responsive Demand participation model, in which resources would be compensated for price sensitivity across all price levels, not simply emergency capacity.

- Increase regionalization through direct ISO/RTO expansion or regional organized markets, such as the California Independent System Operator's (CAISO's) Western Energy Imbalance Market. Vertical utilities across the west should consider joining the CAISO or Southwest Power Pool (SPP) regional markets, and the southeastern states should consider deregulating and either joining or forming a new RTO.

- Evaluate productizing inertial response as an ancillary service. Absence of grid inertia is perhaps the greatest energy security risk of a decarbonized fleet. Investigations should build off that concluded by the Australian Energy Market Commission in 2018, and compare the efficacy of a market product for inertial response to requirements established in interconnection agreements.

Solar panels by Lincoln Electric Systems, in Lincoln, Nebraska.
Unsplash/American Public Power Association (@publicpowerorg)

Introduction

Organized power markets today face numerous challenges stemming from policy, regulatory, and commercial pressures. Since 2015, sixteen US states and territories, and more than two hundred cities and counties in the United States, have committed to achieving 100 percent clean energy, through either legislation or executive action, acting on the realization that decarbonization of the electricity sector is critical to achieving global emissions reductions goals.[3] A recent study suggests that it is feasible for 90 percent of US electricity generation to be zero emission by 2035, with 70 percent of power being produced by wind, solar, and batteries during normal demand and generation conditions.[4] In particular, the Energy Futures Initiative estimates that it is feasible for California to meet its aggressive renewables portfolio standard (RPS) of 60 percent renewable generation by 2030, required by Senate Bill 100.[5]

Solar photovoltaic (PV) and wind resources, which would make up a vast majority of that renewable capacity, are not only unpredictable in their output, but bid into markets at near-zero marginal cost, potentially pricing out the predictable fossil-based units on which the grid has historically depended. These conventional generators are responsible for the power system's inertia—in the form of massive, rotating magnets turned by water and steam—which protects against sudden imbalances in supply and demand. As these generators exit the market, they take with them a crucial source of system stability.

3 "Emissions Gap Report," UN Environment Programme, 2019; Kelly Trumbull, et al., "Progress Toward 100% Clean Energy In Cities and States Across the U.S.," UCLA Luskin Center for Innovation, November, 2019.

4 Amol Phadke, et al., "The 2035 Report," Goldman School of Public Policy, June 2020. The authors estimate that nuclear and hydropower could contribute 20 percent during normal conditions. During atypical hours of very high demand or very low renewable output, natural gas would play an increased role, though accounting for only 10 percent of annual energy production.

5 "Optionality, Flexibility & Innovation. Pathways for Deep Decarbonation in California," Energy Futures Initiative, May 2019.

This study focuses on power markets in the United States, which vary enormously in their market design, relationships with the states in which they operate, and approaches to market reform. The challenges they face are not unique, however, and comparisons with Europe and Australia are often germane. Moreover, not all challenges relate to energy security. State and local subsidies for clean resources have been accused of undermining the non-discriminatory model of US markets, whose operators must decide whether to account for the effects of these subsidies on resource economics. Varying state policies on carbon pricing pose similar questions, with resources on different sides of a border facing unequal compliance costs, altering their competitiveness. Narrowly scoped decisions by FERC—which have encouraged an oppositional stance from market operators to state policies—have polarized stakeholder factions, and only raised the stakes of market policy decisions.

Less controversial has been federal regulatory pressure on power markets to modernize their legacy resource participation models, in order to accommodate advanced technologies demanding market access. FERC's Order 841, issued on February 15, 2018, requires market operators to offer participation models for electric storage technologies across energy, capacity, and ancillary service markets. A successor order in 2020 recognizes aggregations of DERs and the entities that aggregate them as first-class market participants, afforded the same opportunities as traditional resources, but with a dedicated participation model that respects their unique characteristics.[6] DERs are interconnected to low-voltage distribution systems managed by utilities, rather than the high-voltage transmission systems managed by system operators (RTOs and ISOs). Their participation in wholesale markets raises a host of challenges, from un-envisioned reverse power flow from the distribution system to the transmission system, to the distinction between retail and wholesale compensation.[7] The distributed nature of DER aggregations has added to the complexity, prompting debates over their required localization, metering, telemetry, and interconnection procedures. Electric vehicle (EV) fleets are a notably important—and complicated—example of DER aggregations, owing to the mobility of their constituents and the alternative interpretation of their battery charging as retail energy sale, wholesale energy sale, or transportation fueling.

The challenges to power markets posed by renewables and advanced technologies did not arise in a vacuum. Markets have been plagued by inefficiencies and barriers to entry in organized capacity markets, misalignment between forward and real-time energy markets, and numerous failings of energy- and reserve-price formation. The price formation failings deal primarily with two interrelated phenomena: the inability of generators in times of scarcity to recover sufficient revenue to support efficient long-term investment—the so-called "missing money" problem—and the increasing reliance on out-of-market actions by system operators, in the form of resource-uplift payments, reliability-must-run (RMR) contracts, and reliability unit commitments.

While not universally held, the predominant view among experts is that wholesale market reform is required to resolve both legacy challenges and those introduced by the energy transition.[8] Failure to act may see states pull out of capacity markets, deny market access to key technologies, and lead markets to become overburdened by varying state policies.[9] What that reform should look like has been debated among the usual stakeholder coalitions, including clean energy advocates, fossil fuel interests, and utilities. Proposals have been put forward, ranging from high-level

6 Electric Storage Participation in Markets Operated by Regional Transmission Organizations and Independent System Operators, 162 FERC, paragraph 61, 127. Order 841 requires ISOs and RTOs to dismantle barriers to electric-storage resources in capacity, energy, and ancillary service markets; Participation of Distributed Energy Resource Aggregations in Markets Operated by Regional Transmission Organizations and Independent System Operators, 172 FERC, paragraph 61, 247.

7 Traditionally, power is produced by large-scale generators, travels regionally via high-voltage transmission, and is delivered locally via the distribution system. Interconnection refers to a generator electrically connecting to either of these networks.

8 For a countervailing view, see Fritz Kahrl, "Electricity Markets Don't Need an Overhaul," *UtilityDive*, September 4, 2019, https://www.utilitydive.com/news/electricity-markets-dont-need-an-overhaul/562150/.

9 Dana Hanson, et al., "In an Accelerated Energy Transition, Can US Utilities Fast-Track Transformation?" GridWise Alliance and EY, December 2019, https://gridwise.org/wp-content/uploads/2019/12/Perspectives-on-a-Future-Distribution-System.pdf; "Illinois Governor Wants Clean Energy Legislation, Could Push State Out of PJM Power Grid," Reuters, January 30, 2020, https://www.reuters.com/article/us-usa-illinois-pjm-nuclearpower/illinois-governor-wants-clean-energy-legislation-could-push-state-out-of-pjm-power-grid-idUSKBN1ZT259; "Wholesale Market Barriers to Advanced Energy—and How to Remove Them," Advanced Energy Economy, May, 2019, https://info.aee.net/wholesale-market-barriers-to-advanced-energy; Rob Gramlich and Michael Goggin, "Too Much of the Wrong Thing: Capacity Market Replacement or Reform," Grid Strategies, November, 2019, https://gridprogress.files.wordpress.com/2019/11/too-much-of-the-wrong-thing-the-need-for-capacity-market-replacement-or-reform.pdf; David Farnsworth, "What I Wish I'd Said to California's Chief Air Regulator about Clean Transportation," *UtilityDive*, December 5, 2019, https://www.utilitydive.com/news/what-i-wish-id-said-to-californias-chief-air-regulator-about-clean-transp/568292/.

principles to concrete market designs, some of which represent significant departures from the status quo.[10]

Whether or not these proposals ultimately influence reforms made by FERC and RTOs/ISOs, what remains critical is that stakeholders throughout the energy sector are aware of the key questions that underlie the reform process. These questions are not readily apparent even to industry practitioners, buried within regulatory filings and abstruse white papers issued by markets and their independent monitors. The purpose of the present work is to identify the five most consequential questions facing markets as they prepare for the energy transition, and to discuss them in a self-contained manner, for greatest accessibility.

1 What role, if any, should organized capacity markets play?

Capacity markets have a poor track record of delivering resource adequacy, procuring more capacity than is required, and paying too much for it, most notably by discriminating against state-subsidized renewables and advanced technologies. The forward product they procure, moreover, is poorly suited to the needs of real-time energy markets, an issue that will become only more acute as renewable penetration increases. To offer a viable alternative to state management of resource adequacy, capacity markets must undo their discriminatory practices and tighten their reserve margins, ensuring that they are acquiring appropriate capacity at the lowest cost. They must also refine the products that they procure, ensuring that these products fully leverage the capabilities of all resource types and are designed to meet real-time market needs, particularly around resource flexibility. Where capacity markets fail to win the confidence of all stakeholders, they can be demoted to an optional venue for load-serving entities to acquire forward capacity, alongside a bilateral market.

The main issue is who is responsible for procuring power. When procurement responsibility is unclear, RTOs believe they are compelled to step in and procure power on society's behalf for peak system needs. An alternative approach is to push procurement responsibility back down to wholesale customers. Wholesale customers can be utilities (investor owned, municipal, or cooperative), competitive retail suppliers, and community aggregators. If these entities hold sufficient procurement responsibility, RTOs need not fill in the void.

2 What is the role of organized spot markets in accounting for generator emissions?

Created to facilitate open and non-discriminatory access to the transmission system, ISOs/RTOs have no mandate to incorporate state policy objectives in markets. They have three options with respect to individual state actions to subsidize clean resources, or otherwise price carbon: passively accommodate such actions, counteract them (to undo their economic effect on participants), or directly incorporate them. For example, organized markets have accommodated sulfur dioxide (SO_2) and nitrogen oxides (NO_x) prices for many years by allowing them to operate in the background, and by simply considering the prices of SO_2 and NO_x permits to be among the many legitimate costs that make up generators' bids. By contrast, CAISO's approach of directly integrating a price on carbon into market pricing and dispatch, which minimizes emissions leakage and the effect on other EIM states, is a compelling example of the third option. CAISO and NYISO's carbon price integration proposal serve as models for PJM and ISO New England (ISO-NE), whose stakeholder states include a growing number of RGGI members, and who would be best served by an integrated approach. Alternatively, the CLEAN Futures Act proposed by Democrats on the House Energy and Commerce Committee—informed by the Brattle Group's proposed Forward Clean Energy Market—represents an "accommodate" approach, in which environmental policies operate in the background. It establishes auctions and a bilateral market for new clean energy credits used to implement a national clean energy standard, but does not have ISO/RTOs directly incorporate policies into wholesale markets.

3 What price formation enhancements are capable of solving the "missing money" problem and reducing operator dependence on out-of-market actions?

Two mutually reinforcing challenges plague ISOs and RTOs today. Competitive resources earn insufficient revenue in certain energy markets to meet their costs and to make investments that are most effective for society, from both an economic and reliability perspective. This "missing money" causes, and is reinforced by, resource dispatch and payments made administratively by system operators to individual resources, outside of economic dispatch and pricing, in order to ensure

[10] "Wholesale Electricity Market Design For Rapid Decarbonization," *Energy Innovation*, June 25, 2019, https://energyinnovation.org/publication/wholesale-electricity-market-design-for-rapid-decarbonization/; Michael Goggin, et al., "Customer-Focused and Clean. Power Markets for the Future," Grid Strategies, November 2018, https://gridprogress.files.wordpress.com/2019/03/power-markets-for-the-future-full-report.pdf.

system reliability. While preventing contingencies in the short term, this practice mutes valuable scarcity price signals and reduces market efficiency. ISOs and RTOs have attempted a variety of tacks to address the price formation defects that lead to "missing money," including better aligning real-time and day-ahead products, enabling inflexible and fast-start resources to set price through extended locational marginal pricing, and modifying administrative demand curves to improve scarcity pricing. Accurate scarcity price signals, in particular, are critical to ensure that supply is present when and where it is urgently needed, reducing operator dependence on out-of-market payments.

4 How should DERs and flexible load resources participate in markets?

Numerous challenges confront loads, generators, and storage resources on the distribution system seeking to participate in wholesale markets. They must be aggregated to achieve minimal resource size, which has raised hotly debated questions as to the allowable location, interconnection, capabilities, and compensation of aggregated resources. Flexible loads have historically bid into markets as demand response, a supply-side resource, rather than as price-sensitive demand, but new demand-side participation models are making the latter route more compelling. Demand response and other legacy market products have proved insufficient for leveraging the unique capabilities of storage and other flexible load technologies. Moreover, they have prompted investigations into newer products, such as the "Shift" demand response in California, in which resources are paid to shift load from high-demand to low-demand periods, rather than eliminating that demand. Such market product innovation must continue for the full value of flexible distributed resources to be realized. Electric vehicles, in particular, could be game changing as a flexible storage resource, but their involvement in grid services hinges on basic questions of mobile-inverter interconnection and metering, and what distinguishes the sale of retail energy from the sale of wholesale energy, or vehicle-fueling energy.

5 How can markets ensure energy security when supply is intermittent?

The conversion of the generator fleet from predominantly dispatchable thermal units to variable energy resources has profound implications for energy security and system stability. With the North American Electric Reliability Corporation (NERC) calling for additional flexible capacity, some operators have proposed novel ancillary service products to address system load ramping and day-ahead forecast deviations. Others have leaned more heavily on transmission, using vast regional markets to average out variable renewable production and to find distant consumers for excess generation. Advanced energy technologists have not given up on renewables as inflexible, however, proving in utility-scale demonstrations that while limited by the sun and wind, these resources are capable of fast regulation services, offering a tantalizing tradeoff between energy and ancillary service capacity.

For regulators, system operators, and other stakeholders involved in wholesale market reform, these five questions must be tackled head on, rather than continuing to be addressed on an ad hoc and implicit basis through debates on individual reforms. Their resolution would represent a strategy for market reform, a marked improvement over a process currently characterized by issue-by-issue tactics and flashpoint regulatory rulings. For those outside of stakeholder meetings and commission hearings, these questions are the lens through which the technical, and, at times, abstruse market reform debates can be understood. On that basis, they should be used as a rubric when commissions issue public documents and decisions. With greater clarity around the issues at stake, new experts and stakeholders can join the debate, offering their view on questions that will affect not only the evolution of organized electricity markets, but the success of the energy transition as a whole.

These five questions are by no means independent of each other, but individually they offer a valuable perspective by which to consider candidate market reforms. They are discussed in the sections that follow, including the market challenges that prompted them, recent developments in which they are reflected, and the options available to regulators and market operators.

Commonly used acronyms

AMI — Advanced metering infrastructure

BSM — Buyer-side mitigation

BYOT — Bring your own thermostat

CAISO — California Independent System Operator

CCA — Community choice aggregation

CONE — Cost of new entry

CPUC — California Public Utility Commission

DER — Distributed energy resource

DERMS — Distributed energy resource management system

EIM — Energy Imbalance Market

ELMP — Extended locational marginal price

ERCOT — Electric Reliability Council of Texas

ESDER — Energy Storage and Distributed Energy Resources

ETS — Emissions trading scheme

FERC — Federal Energy Regulatory Commission

FRR — Fixed Resource Requirement

ISO — Independent system operator

ISO-NE — Independent System Operator New England

LMP/LBMP — Locational marginal price

LRV — Load Reduction Value

LSE — Load serving entity

LSRV — Locational System Relief Value

MISO — Midcontinent Independent System Operator

MOPR — Minimum Offer Price Rule

NERC — North American Electric Reliability Council

NOPR — Notice of proposed rulemaking

NREL — National Renewable Energy Laboratory

NYISO — New York Independent System Operator

NYPSC — New York Public Service Commission

OATT — Open access transmission tariff

ORDC — Operating reserve demand curve

PPA — Power purchase agreement

PRD — Price Responsive Demand

RA — Resource adequacy

RGGI — Regional Greenhouse Gas Initiative

RMR — Reliability-must-run

RTO — Regional transmission operator

RUC — Reliability unit commitment

SCC — Social cost of carbon

SPP — Southwest Power Pool

V2G — Vehicle-to-grid

VDER — Value of Distributed Energy Resources

VER — Variable energy resource

VGI — Vehicle-grid integration

VOLL — Value of lost load

1 What role, if any, should capacity markets play?

The great capacity market debate

Capacity markets compensate resources for their commitment to offer energy in the future. This revenue is additional to the revenue earned in energy and ancillary service markets for the actual provision of energy and services, as an incentive mechanism to ensure future resource adequacy (RA). Capacity markets are an alternative to the cost-based approach to RA used in vertically-integrated markets, in which the monopoly utility earns a fixed rate of return on capital investments required for reliability, and the energy-only approach to RA employed in Europe, Australia, and Texas. The energy-only approach relies on elevated prices during times of supply scarcity, known as scarcity pricing, as a sufficient incentive for resources to be available and to provide at times of greatest need.

With the exception of the Electric Reliability Council of Texas (ERCOT), the primary wholesale market in Texas, all system operators in the United States utilize a capacity market. In the west, the SPP, Midcontinent ISO (MISO), and CAISO impose RA requirements on load-serving entities (LSEs), such as utilities and community choice aggregations, requiring a minimum reserve margin in megawatts (MW) with respect to their expected peak demand. In this decentralized RA model, LSEs can self-provide this capacity, or procure it through bilateral contracts or, in MISO's case, an optional centralized capacity auction. In the northeast, PJM, NYISO, and NE-ISO take on RA responsibility themselves, procuring capacity they deem necessary and passing the cost on to LSEs based on their contribution to system peak load.[11] That is, the system operator is the buyer of capacity, rather than LSEs. Procurement happens through a centralized reverse auction, in which the system operator administratively sets a demand curve based on load forecasts and risk tolerance, and competitive resources bid to supply capacity.

The other organized markets in North America reflect the diversity of US markets. In Canada, the Alberta Electric System Operator employs an energy-only market, but has migrated to a centralized RA model, as has Ontario's Independent Electricity System Operator. Meanwhile, through its 2014 Energy Reform, Mexico has undergone significant market liberalization, including the establishment of a decentralized RA model.[12]

RA programs originated from a public good market failure in restructured markets, which persists today. These markets fail to sufficiently compensate generation during supply shortages, as energy prices rarely reflect scarcity. In efficient markets, it is during scarcity periods that suppliers recoup their capital costs, when prices are permitted to rise well above marginal production costs. Insufficient scarcity revenue stifles generation investment, pushing markets toward unreliability in the longer term, including blackouts, with societal costs that outweigh the cost of the investment necessary to preempt them.[13] RA programs, however they are implemented, seek to correct this market failure by preventing under-builds of capacity and, therefore, societally suboptimal outcomes.

While the necessity of resource adequacy is not in dispute, the use of capacity markets to achieve it very much is. The manner in which these markets compensate for insufficient scarcity revenue further severs the relationship between generator economics and real-time grid conditions.[14] In particular, there is a significant disparity between the product procured by capacity markets and the needs of energy markets, where resource adequacy is ultimately determined. The precise location, timing, and service needed by the market

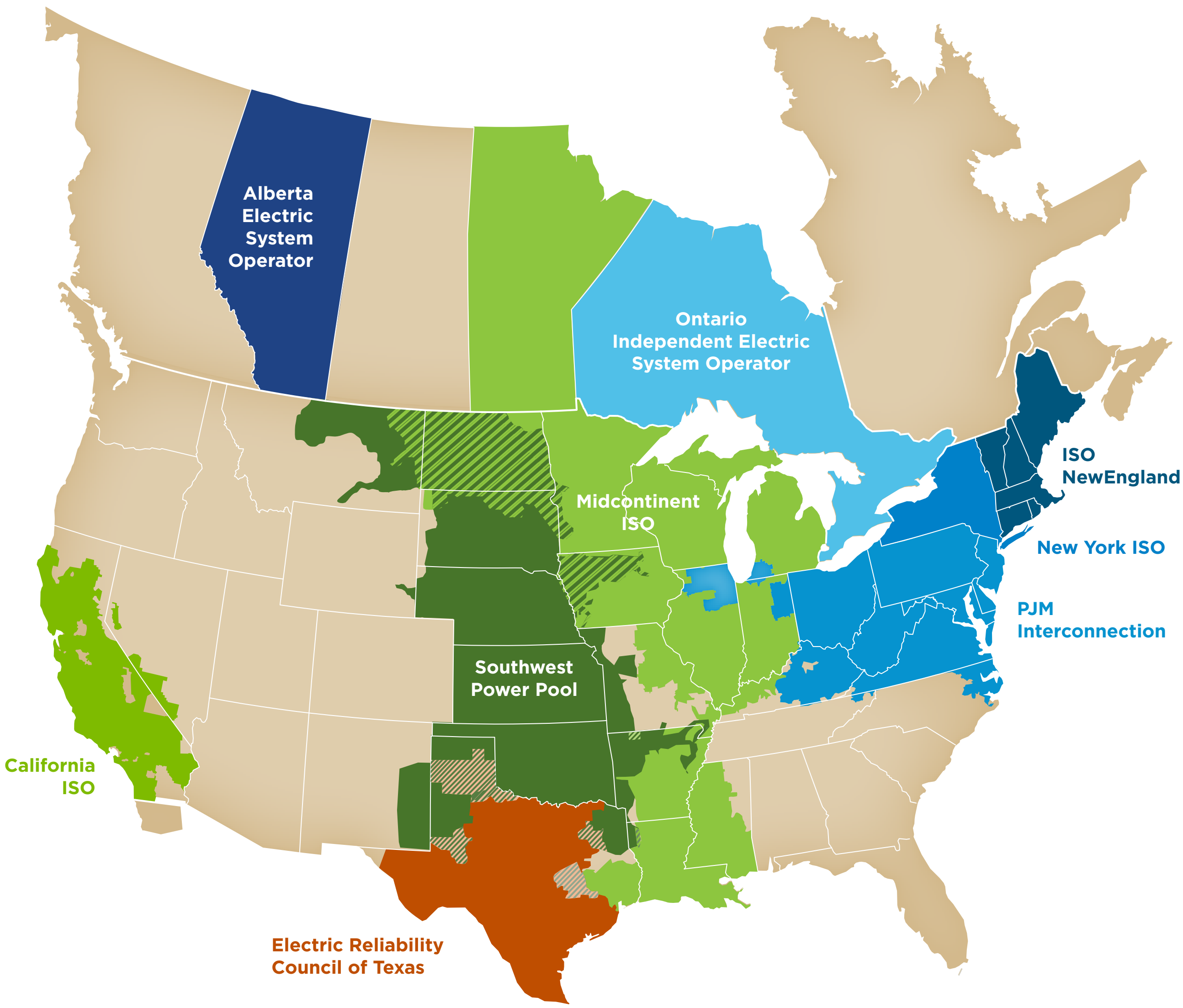
11 In the northeast, as in Texas, retail electricity markets are deregulated, meaning retail energy providers compete to sell power to mass-market customers. They, therefore, qualify as LSEs, like regulated utilities.

12 Duncan Wood, "Mexico's New Energy Reform," Wilson Center, October 2018, https://www.wilsoncenter.org/sites/default/files/media/documents/publication/mexicos_new_energy_reform.pdf.

13 Peter C. Cramton and Steven Stoft, "The Convergence of Market Designs for Adequate Generating Capacity with Special Attention to the CAISO's Resource Adequacy Problem," Harvard Electricity Policy Group, 2006, https://hepg.hks.harvard.edu/publications/convergence-market-designs-adequate-generating-capacity.

14 William W. Hogan, "Electricity Scarcity Pricing Through Operating Reserves," International Association for Energy Economics, 2013, http://www.lmpmarketdesign.com/papers/Hogan_ORDC_042513.pdf.

Figure 1: Resource adequacy (RA) paradigms in North America



Centralized RA
The system operator is the buyer of capacity, centrally managing RA

Energy-only
No capacity market. RA is a byproduct of the real-time energy market and in some cases credit-worthiness requirements on LSEs

Decentralized RA
LSEs are the buyers of capacity, with RA obligations

Cost-based RA
Legacy monopolistic model in which utility earns fixed rate of return for reliability-based investments

is not known at the time that capacity is contracted, nor is there a guarantee that the resource will fulfill its obligation to schedule or bid into the market consistent with those needs. Capacity products "have nothing to do with the responsiveness of resources," and have been accused of failing to increase reliability.[15]

The imperfection of capacity markets as a means to achieve RA is best captured by the European Union's tepid embrace. In its recast Electricity Regulation, part of the "Clean Energy for all Europeans" package passed in May of 2019, the European Union (EU) explicitly affirmed the legality of capacity markets, but only as a temporary last resort for RA, citing such markets' "distortive" effects.[16] Before implementing capacity markets, member states must evaluate and mitigate their distortive effects, consult with neighboring member states, and limit their longevity to a maximum of ten years.[17]

Capacity markets provide value outside of RA. A key benefit that they offer to renewables is upfront revenue, which mitigates project investor risks and reduces financing costs, particularly in low-energy price environments. However, the same benefit is achieved by forward contracting energy through power purchase agreements (PPAs), a core component of the energy-only model and one that is deemed more critical to renewable financing.[18]

The need for reform

Whether or not capacity markets are a beneficial construct in power markets, the mandatory centralized markets deployed in the northeast have a dismal track record. FERC Commissioner Richard Glick has warned that their current approach to RA, including the disregard for state policy objectives, risks imperiling the regional transmission operator model itself.[19] From an efficiency standpoint, these markets are wasteful, imposing excessive reserve margins at ratepayers' expense. PJM's margin of excess power in 2018 was 32.8 percent, more than twice the target reserve margin of 16.1 percent, estimated by NERC.[20] This is consistent with NERC's 2020 "Long-Term Reliability Assessment," which found PJM's 2025 anticipated reserve margin to be 41.1 percent, close to three times the target of 14.8 percent.[21] Former Governor Martin O'Malley of Maryland has accused PJM and other RTOs of "imposing ever-expanding price supports" for conventional generation, and operating "like a cartel, with suppliers on committees voting on rules.[22] A cost estimate for the excess reserve margin across PJM, NYISO, and ISO-NE is approximately $1.4 billion per year.[23] Further evidence of market distortion at work is that, even with excess reserve margins, entry into the capacity market continues, and an estimated 18 gigawatts (GW) of coal-fired units in PJM would be uneconomical without capacity payments.[24] The ISO-NE Market Monitor observed in 2018 that capacity prices and revenues in the most recent auction did not "facilitate efficient entry and exit decisions."[25]

The primary cause of the inefficiencies in the northeast markets is their administratively set demand curves, which determine the quantity and price of capacity procured. Different methodologies are employed, but each is generally based on the net cost of new entry (Net CONE) of a new reference generator to the market, equal to its levelized costs minus estimated energy and ancillary service market revenues. PJM's reference unit is a combustion turbine, a legacy technology that dramatically inflates the height of its demand

15 James Bushnell, Michaela Flagg, and Erin Mansur, "Capacity Markets at a Crossroads," Energy Institute at Haas, working paper 278, April 2017; Peter Cramton, "Electricity Market Design: The Good the Bad and the Ugly," Proceedings of the 36th Annual Hawaii International Conference on System Sciences, Track 2-Volume 2, 2003.

16 Regulation (EU) 2019/943 of the European Parliament and of the Council of 5 June 2019 on the internal market for electricity, http://data.europa.eu/eli/reg/2019/943/oj.

17 Ibid.

18 Rob Gramlich and Frank Lacey, "Who's the Buyer? Retail Electric Market Structure Reforms in Support of Resource Adequacy and Clean Energy Deployment," Grid Strategies, March 2020, https://windsolaralliance.org/wp-content/uploads/2020/03/WSA-Retail-Structure-Contracting-FINAL.pdf.

19 Jasmin Melvin, "FERC's Glick Urges Broader Commission to Look at Capacity Markets, States' Rights," *S&P Global Market Intelligence*, February 5, 2020, https://www.spglobal.com/marketintelligence/en/news-insights/latest-news-headlines/ferc-s-glick-urges-broader-commission-look-at-capacity-markets-states-rights-56969569.

20 Stephanie Tsao and Richard Martin, "Overpowered: PJM Market Rules Drive an Era of Oversupply," *S&P Global Market Intelligence*, December 3, 2019, https://www.spglobal.com/marketintelligence/en/news-insights/latest-news-headlines/overpowered-pjm-market-rules-drive-an-era-of-oversupply-54111666.

21 "Long-Term Reliability Assessment," North American Electric Reliability Corporation, 2019.

22 Tsao and Martin, "Overpowered: PJM Market Rules Drive an Era of Oversupply;" Martin O'Malley, "Ex-Maryland Gov O'Malley: States Must Reassert Authority on Clean Energy Policy," *UtilityDive*, March 28, 2019, https://www.utilitydive.com/news/ex-maryland-gov-omalley-states-must-reassert-authority-on-clean-energy-po/551461/.

23 Gramlich and Goggin, "Too Much of the Wrong Thing: Capacity Market Replacement or Reform."

24 Ibid.

25 "2018 Annual Markets Report," ISO-NE Internal Market Monitor.

curve. Combined-cycle plants have been the dominant new entry technology in PJM for more than fifteen years, with a Net CONE 44–76 percent lower than the assumed Net CONE.[26] Commissioner Glick has noted that 31 GW of resource have cleared PJM at prices, on average, 60 percent below the inflated Net CONE, which has, at the same time, encouraged the entry of new resources that are not needed.[27] Additional factors that make the demand curve fatter—biased toward greater capacity and higher prices—are over-forecasted load and an inflated value of lost load (VOLL), representing the demand-side value of capacity, which can be greater than ten times conventional estimates.[28]

Many of the findings above hold true for the United Kingdom (UK) capacity market as well, including prices that clear below the administratively set CONE and excessive capacity procurement.[29] As with the United States, this can be attributed to centralized procurement decisions managed by cautious regulators, system operators, and politicians, rather than the emergent determination of markets.

Beyond demand curve development in centralized RA procurement, the capacity product itself requires reform if capacity markets are to remain viable. The obligations of a capacity resource consist only of self-scheduling or economically bidding into day-ahead or real-time energy markets, a minimal availability commitment. NERC has identified fast frequency response, ramping and balancing, and voltage support as key capabilities for reliability under high renewables penetration, none of which are obtained through such commitments.[30] Some capacity products, in fact, discriminate against flexible technologies possessing these very capabilities, such as battery energy storage and HVAC (heating, ventilation, and air conditioning) demand response. Neither can support PJM's requirement of ten- to fifteen-hour emergency events, and air conditioner-driven demand response does not exist half of the year, violating PJM's year-round capacity requirement.

Such discriminatory practices disrupted the UK's capacity market in 2018, when the General Court of the European Union ruled that the European Commission had failed to investigate its compliance with EU state aid rules.[31] The UK market offers contract terms of up to fifteen years to new or refurbished generators, whereas demand response resources are eligible for one-year contracts only. Also, as in PJM, demand response resources must commit to long, open-ended events, which is often infeasible. Demand response service provider Tempus Energy argued that these constitutes discriminatory treatment of demand-side resources, a claim that was later rejected by the court, but which forced the UK to operate the market without payments for a year. The capacity product recognized by the European Commission, like its US counterpart, commits a resource only to availability in the real-time market, rather than energy or flexibility.[32]

It is imperative that capacity markets, like energy markets, be nondiscriminatory, evaluating resources solely by offer and by cost. Moreover, capacity product reform must introduce new products, better aligned with the flexibility needs of real-time markets and the operating profiles of advanced technologies capable of providing them.

Approaches to reform

Given the flaws in present-day capacity markets, most notably in centralized markets, system operators and their regulators have several options. The first is to do away with capacity markets altogether and adopt an energy-only RA framework, similar to Texas and international markets. This is consistent with the view that capacity markets sever the relationship between resource investment decisions and energy market incentives, violating a foundational principle of power markets.[33] Evidence of the sufficiency of the leaner energy-only model is strong. ERCOT entered 2019

26 Samuel A. Newell, et al., "Fourth Review of PJM's Variable Resource Requirement Curve," Brattle Group, 2018, https://www.pjm.com/-/media/library/reports-notices/reliability-pricing-model/20180425-pjm-2018-variable-resource-requirement-curve-study.ashx.

27 Commissioner Richard Glick Dissent Regarding PJM Interconnection, L.L.C., tariff, Docker No. ER19-105-001, April 15, 2019. https://www.ferc.gov/news-events/news/commissioner-richard-glick-dissent-regarding-pjm-interconnection-llc-tariff.

28 Gramlich and Goggin, "Too Much of the Wrong Thing: Capacity Market Replacement or Reform."

29 David Newbery, "Missing Money and Missing Markets: Reliability, Capacity Auctions and Interconnectors," Cambridge Working Paper in Economics 1513, February 2015, https://www.eprg.group.cam.ac.uk/wp-content/uploads/2015/03/1508_updated-July-20151.pdf.

30 NERC, "Essential Reliability Services. Whitepaper on Sufficiency Guidelines," 2016.

31 Dave Keating, "EU Accused of Subsidizing Fossil Fuels Through Capacity Markets," *Forbes*, October 26, 2019, https://www.forbes.com/sites/davekeating/2019/10/26/eu-accused-of-subsidizing-fossil-fuels-through-capacity-markets/?sh=1ca082db774c.

32 Regulation (EU) 2019/943 of the European Parliament and of the Council of 5 June 2019 on the internal market for electricity.

33 Hogan, "Electricity Scarcity Pricing Through Operating Reserves."

with only an 8.6 percent reserve margin, but, despite above-normal load growth, new generation is expected to strengthen it to upward of 12 percent through 2024.[34] Moreover, much of that new capacity is in renewables and flexible gas plants, showing that the energy-only model can succeed, even in a high demand and high renewables environment.[35] Critical to the energy-only option, however, is to ensure that there are robust buyers of long-term power. If competitive retail suppliers only buy power six months at a time, for example, then there will likely be under-procurement and high generation financing costs for merchant generators that have no one to whom they can sell a PPA.

Short of abolishing capacity markets altogether, there are various gradations that move markets in the energy-only direction. The first incremental step would be to maintain centralized RA—with the system operator still the single buyer for the system—but permit LSEs to opt out and procure their own capacity. This is the case with PJM's Fixed Resource Requirement (FRR), a carve-out originally designed for vertical utilities, which allows them to opt out of the capacity market for a period of five years, provided that they meet PJM's reserve capacity requirements.[36] This option has been given a hard look recently by unexpected utilities, such as New Jersey's PSE&G, eager to buy capacity from state-sponsored resources that may fail to clear the capacity market due to FERC's draconian Minimum Offer Price Rule (MOPR; see section 4).[37] PJM's FRR exemptions are highly regulated, however, forcing LSE's to be all in or all out of the capacity market. A reformed market's FRR should be designed flexibly for LSEs of all stripes. In its 2019 Expanded MOPR Order, FERC declined to require PJM to implement a more flexible, resource-level FRR alternative.[38]

The next incremental step toward the energy-only model is to entrust RA responsibility to states, rather than system operators. LSEs are then empowered to secure their own RA, with financial penalties for failing to secure their required reserve margin, based on historic peak load, and centralized capacity markets become optional venues for procuring capacity. This is the case for MISO today. A benefit of this approach is that states can imprint their own policy objectives on RA markets, including support for clean resources.

If capacity markets are a priority, but today's centralized and decentralized variants are deemed inadequate, a final option is to fundamentally redesign them. One example is ISO-NE's contemplated Seasonal Forward Market, a component of the operator's Energy Security Improvements initiative and a potential successor to today's centralized Forward Capacity Market, but which was deemed a significant undertaking and ultimately not submitted to FERC in ISO-NE's Energy Security Improvements filing.[39] The products envisioned for this market are forward contracts for the new ancillary services that ISO-NE proposed as part of its Energy Security Improvements, enabling the same two-settlement relationship between the forward and real-time markets that real-time markets currently enjoy with day-ahead markets.[40] Not only would this align energy and capacity markets in ways that do not currently exist, but it would reform the capacity product into the kind flexible product that the real-time market needs. FERC rejected the Energy Security Improvements filing, in significant part because it focused entirely on day-ahead products, rather than longer-term ones—such as seasonal forward products—that would better align generator revenue with upfront fuel reservation.[41]

A more ambitious set of proposals envisions capacity markets as procuring not simply flexible capacity, but a portfolio of resources that complements the precise mix of variable energy resources (VERs) anticipated in the future.[42] It is premised on the fact that the stability of real-time markets, and their ability to provide

---

34 Report on the Capacity, Demand, and Reserves (CDR) in the ERCOT Region, 2020-2029. ERCOT. December 5, 2019. http://www.ercot.com/content/wcm/lists/167023/CapacityDemandandReserveReport-Dec2019.pdf

35 Ibid.; This report does not delve into the winter storm events in Texas of 2021, which occurred after writing. However, it should be noted that the ERCOT reserve margin was not a contributor to the outages that occurred. The cause of the outages was the failure of committed resources to perform, and the total number of resources committed by ERCOT.

36 "Security Resources Through the Fixed Resource Requirements," PJM, 2020, https://www.pjm.com/-/media/about-pjm/newsroom/fact-sheets/securing-resources-through-fixed-resource-requirement-fact-sheet.ashx.

37 Molly Christian, "PSEG Considering Exit from PJM's Capacity Market Following FERC Order," *S&P Global Market Intelligence*, February 26, 2020, https://www.spglobal.com/marketintelligence/en/news-insights/latest-news-headlines/pseg-considering-exit-from-pjm-capacity-market-following-ferc-order-57289960.

38 "Understanding FERC's 'Minimum Offer Price Rule' Order," Advanced Energy Economy, January 2020, https://info.aee.net/hubfs/Federal%20Policy%20(2018-2020)/PJM%20MOPR%20Explainer%2001_20.pdf.

39 Energy Security Improvements, ISO Discussion Paper, 2019; ISO New England Inc.: Compliance Filing of Energy Security Improvements, Docket No. ER20-1567 (April 15, 2020).

40 Ibid.

41 Order Rejecting Proposed Tariff Revisions, 173 FERC, paragraph 61,106.

42 "Wholesale Electricity Market Design for Rapid Decarbonization."

sufficient revenue for all resources—including renewables—depends on the right mix of resources competing in those markets: solar and wind from different regions, for example, so that one resource may be producing when another is not; and complementary load and storage resources to increase net load when VER production is high, and reduce it when production is low.[43] Arranging this balanced mix of resources in the short run requires proper long-run incentives, given the timescales of resource entry and exit. This can be achieved through a carefully designed centralized forward market, in which either system planners or market software select for complementary portfolios, using resource and system modeling tools.[44] The key insight of these proposals is the role of the forward market as a puzzle solver, taking full account of the shape of each resource and its offer as it fits them together to meet future system needs. This is in contrast to the current practice of lumping all resource offers into the same featureless capacity product, losing any guarantee of complementarity.

[43] Ibid.

[44] Ibid.

2 What is the role of organized spot markets in accounting for generator carbon emissions?

An unclear role for market operators

RTOs and ISOs are tasked with ensuring safety, reliability, and open access to transmission systems, and for independently operating wholesale power markets. Even if their territory resides entirely within a single state, as long as the interconnection to which their network belongs crosses state lines, they are regulated by FERC, not by state public utility commissions. RTOs and ISOs therefore have no mandate with respect to the policies of states in which they operate, and states do not have direct control over them. In light of this, market operators are put in a difficult position when states have differing policies with respect to generator carbon emissions, which can impact the dynamics of the markets in which these resources participate.

A prominent example of this dilemma stems from subsidies for renewable or nuclear generators offered by a subset of PJM states, including Illinois and New Jersey, which serve to support clean energy targets.[45] These subsidies enable the resources to bid into PJM's capacity market at lower cost than would otherwise be possible, improving their competitiveness against unsubsidized resources. Another example is carbon pricing, a policy with a strong empirical basis for decarbonizing energy systems.[46] A number of states in the northeast are part of the RGGI, a cap-and-trade system that imposes compliance costs on generators: they must purchase an allowance for each ton of carbon dioxide (CO_2) they emit, raising their variable production costs and, therefore, their bids into energy markets. In both cases, states are altering the economics of resources on the basis of their carbon emissions, thereby altering the dynamics of system operators' markets.

The issue is not confined to RTOs, which extend across multiple states. CAISO, an ISO confined to California, operates the Western EIM, which includes utilities across western states. California operates a cap-and-trade system similar to RGGI, which puts its emitting resources at a disadvantage compared to those in other states. NYISO in New York, moreover, is considering putting a price on carbon based on the gross social cost of carbon (SCC). While all New York resources face the same rules for increased costs, loads and generators outside of the state—which participate in NYISO through imports and exports—do not. Both CAISO and NYISO must, therefore, be concerned not only with the economic impact of their state's policies on resources both inside and outside their borders, but with the likelihood of emissions leakage if these impacts are not managed properly.[47]

The role for market operators would be clarified considerably if a carbon policy, such as a price on carbon, were to be established at the federal level, either by statute or rule. As federally regulated entities, this policy would dictate their market operations, including their accommodation of individual state policies. FERC's role would be equally clear, as the enforcer of this policy. Whether FERC itself has the authority under the Federal Power Act to put a price on carbon is unclear, however. FERC acting as an economic regulator might require the finding that the absence of a carbon price is a market failure, implying unjust or unreasonable electricity rates. A more convincing case can be made for the Environmental Protection Agency's authority to price carbon under the Independent Offices Appropriation Act, which grants it the authority to charge for the disposal of polluting gases as a user fee (technically distinct from a tax).[48]

In the absence of federal leadership on carbon pricing, however, ISOs and RTOs have three fundamental options for handling differing state emissions policies: treat them as exogenous to the market and accommodate them, much as a new technology that reduces cost does not warrant special treatment; treat them as endogenous biasing factors and counteract them,

[45] Kathryne Cleary, "What the Minimum Offer Price Rule (MOPR) Means for Clean Energy in PJM," *Resources*, January 21, 2020, https://www.resourcesmag.org/common-resources/what-minimum-offer-price-rule-mopr-means-clean-energy-pjm/.

[46] Jonas Meckling, Thomas Sterner, and Gernot Wagner, "Policy Sequencing Toward Decarbonization," *Nature Energy 2*, 12, 2017.

[47] Emissions leakage is the displacement of emissions from resources that are covered by a carbon policy to those that are not, negating the impact of the policy.

[48] E. Donald Elliott, "EPA's Existing Authority to Impose a Carbon 'Tax,'" *Environmental Law Reporter*, October 2019, https://digitalcommons.law.yale.edu/cgi/viewcontent.cgi?article=6379&context=fss_papers.

in order to return market economics to an "unbiased" state; or treat them as endogenous but legitimate factors, and directly integrate them into market processes. These options and examples from recent market policies are discussed in the sections that follow.

Accommodate

The approach of passively accommodating state policies regarding carbon emissions is the most natural one for system operators, given that market designs predated concerns about carbon. Two market-based solutions for carbon emissions—one existing, one proposed—best capture the extent and the limitations of this approach.

RGGI AND EU ETS: CARBON MARKETS NOT INTEGRATED INTO POWER MARKETS

Initiated in 2009, RGGI is the first carbon cap-and-trade system in the United States. It is based on an agreement signed by governors, and currently includes eleven member states in New England and the mid-Atlantic. RGGI only covers the carbon emissions of power plants, with carveouts for units below 25 MW and those that consume more than 10 percent of the power they produce, such as refineries. Plants covered by the regime, known as covered entities, must acquire a carbon allowance for every ton of CO_2 they emit through power generation, and surrender them at the conclusion of three-year compliance periods. Allowances can be purchased through quarterly auctions, or bilaterally through a secondary market, and up to 3.3 percent of an entity's obligation can be satisfied by carbon offsets. Allowance prices rose from 2017 through the end of 2020, following a tightening of the emissions cap in 2014, reaching $7.41 per short ton CO_2 in the December 2020 auction.[49] Prices are likely to increase further, becoming more material to producers, as the cap continues its scheduled reduction of 2.5 percent per year through 2030.[50]

The allowance-compliance costs that generators face affect their variable costs, and therefore increase their offers into power markets. The nature of these costs is opaque to the market, however, which simply sees the elevated offers. To the extent that these resources fail to clear the market, the supply balance will shift toward similar resources in other states, causing emissions leakage. When covered entities are on the margin, by contrast, they raise the price of energy for consumers, including those outside of RGGI states.

Similar dynamics have played out in the EU under its emission trading scheme (EU ETS). Energy prices generally reflect the cost of EU ETS allowances, and the lack of ETS integration with energy markets has resulted in emissions leakage to neighboring countries. Thirty-three terawatt hours (TWh) of energy were imported to the EU in 2019, a majority of them coal-based energy from producers in Russia, Ukraine, and the Western Balkans.[51] These importers are at a competitive advantage compared to EU producers, benefiting from higher prices and no obligation to purchase allowances for their carbon-intensive output. Fifty-seven GW of coal-fired capacity is planned in countries that are connected to, or soon will be connected to, the EU, which may augment the existing leakage.[52]

Emissions leakage and indiscriminate cost increases represent downsides of state carbon pricing not integrated into markets. Another feature of non-integration is that allowances are with respect to a generic short ton of CO_2, without regard to the time, location, or context in which emission takes place. Power markets are adept at taking such factors into account when computing prices. This is not necessarily a defect, however. Whereas the cost of serving electricity is highly dependent on the attributes above, the geologic time and spatial scales over which emissions affect the Earth's climate render precise time and location irrelevant in distinguishing emissions events. Nevertheless, the distinction of a clean resource generating on the margin and actively displacing the production of a dirtier resource, versus not having this effect, is ignored in the non-integration model.

Aside from New York, which is developing its own carbon pricing scheme, no plans exist today to mitigate emissions leakage from RGGI states. The independence of RTOs from states, and their lack of alignment in market objectives, creates infertile ground for such a solution to develop. The European Commission, on the other hand, plans to take action to mitigate emissions leakage by petitioning the World Trade Organization (WTO) to include power imports in Europe's planned carbon border adjustment mechanism, a component of the commission's proposed European Green

49 "Allowance Prices and Volumes," Regional Greenhouse Gas Initiative, accessed December 26, 2020, https://www.rggi.org/auctions/auction-results/prices-volumes.

50 "The Regional Greenhouse Gas Initiative: Background, Impact, and Selected Issues," Congressional Research Service, July 16, 2019, https://fas.org/sgp/crs/misc/R41836.pdf.

51 Chris Rosslowe, "The Path of Least Resistance. How Electricity from Coal is Leaking into the EU," Sandbag, January 2020, https://ember-climate.org/wp-content/uploads/2020/01/2020-SB-Path-of-least-resistance-1.2b_DIGI.pdf.

52 Ibid.

Offshore wind near Rhode Island, United States.
Unsplash/Shaun Dakin (@dakinshaun)

Deal.[53] A carbon price would be applied to imports, which would, in a yet-to-be-defined way, reflect the price of ETS allowances. Significantly, application of the carbon border adjustment to electricity is supported by Eurelectric, a 3500-company association that represents the common interests of the electricity industry in Europe, including generation, markets, distribution, and consumers.[54] Eurelectric offers that revenue captured by the adjustment could be earmarked for investment in clean energy in external countries, offsetting the increased energy costs.[55]

THE CLEAN FUTURES ACT AND A FORWARD CLEAN ENERGY MARKET

The CLEAN Futures Act proposed by Democrats on the House Energy and Commerce Committee in January 2020 represents another solution for managing generator emissions, which does not integrate with power markets.[56] It has the potential for supplanting current or future state policies, but explicitly welcomes additive measures that states may take to support clean energy.

The act imposes a nationwide clean energy standard, setting a target of 100 percent clean energy by 2050. It achieves this in a manner very different from RGGI. Rather than cap the emissions of generators, it sets a floor on the percentage of clean energy that LSEs provide to end customers. Just as RGGI's cap is ratcheted down over time, this floor is ratcheted up, rising from an LSE-specific baseline value in 2022 to 100 percent in 2050, in constant increments. Clean energy is denominated by clean energy credits (CECs), which are awarded to generators annually based on their carbon intensity. Zero-emission resources, such as solar, wind, tidal, and nuclear, are awarded one credit per megawatt hour (MWh) of energy produced, and emitting units are awarded a fraction of a credit, and the fraction declines with carbon intensity. LSEs can acquire CECs from generators through federally managed auctions, bilateral trading, and other contract mechanisms, and third parties are permitted to participate and transact in credits as well.

The CLEAN Futures Act was apparently influenced by Brattle's proposal of a forward clean energy market (FCEM), which contemplates clean energy attribute credits (CEACs) as a tradeable unit of the "clean" value of power generation, as unbundled from its

53 Siobhan Hall, "EC to Include Power in EU Carbon Border Import Tax Plans," *S&P Global Platts*, March 5, 2020, https://www.spglobal.com/platts/en/market-insights/latest-news/metals/030520-ec-to-include-power-in-eu-carbon-border-import-tax-plans.

54 "Carbon Border Adjustment: Opportunities to Complement Efforts under the Green Deal," Eurelectric, March 2020, https://cdn.eurelectric.org/media/4271/eurelectric_position_on_carbon_border_adjustment-2020-030-0170-01-e-h-D754C926.pdf.

55 Ibid.

56 Discussion draft of CLEAN Energy Act.

energy and capacity values. A centralized auction, operated by a state or regional authority, such as an ISO/RTO, would occur three years in advance of a delivery period. States, cities, corporate entities, and others would place demand bids for CEACs, with suppliers placing competitive offers for CEACs based on operating costs and expected availability through the delivery period. The aggregate demand curve would be expected to contain flat (price-insensitive) segments, based on fixed renewable portfolio standards and other clean energy targets, followed by price-sensitive segments based on municipal and corporate discretionary targets. The price of the demand curve's initial segments would reflect either the SCC or the Net CONE of a clean resource.

The FCEM is designed to complement, but not interact with, existing wholesale markets for energy and capacity. As these three markets transact in distinct attributes of power generation, there is an argument that the exchange and price formation of these attributes should occur independently. At the same time, the authors make the claim that by bringing clean energy subsidies "in-market," in the form of CEAC revenue, the benefiting resources would be exempt from crippling MOPRs in capacity markets. This claim is dubious, given FERC's explicit identification of state procurement of an attribute of power generation as a state subsidy, which falls afoul of PJM's MOPR (see the discussion in the following section). Additionally, the commission has determined that state actions that *increase the costs* of certain resources, such as RGGI, do not trigger the MOPR, whereas those that *increase revenues*, such as CEACs, do.[57] Regardless, as a regional construct, the FCEM proposal does not resolve the fundamental question of how ISOs/RTOs should accommodate differing state policies. Unless all states within the region join the CEAC market—a scenario that begs the question—resources from states that join will benefit from an additional revenue stream, resulting in the same ISO/RTO conditions that exist today because of state subsidies.

Counteract

Rather than passively accommodate state policies that affect their markets, system operators or their regulators can choose to counteract them. This approach is based on the somewhat radical view that the role of market operators is to insulate market competition not only from the market power of buyers and sellers, but also the market power of states. In particular, it restricts the ability of states to address externalities in markets, such as carbon emissions, that affect their constituents and bear on their policy objectives.

The interpretation of state subsidies in FERC's 2018 PJM MOPR Order, for example, is that they suppress prices by enabling resources to "offer below a competitive price."[58] In other words, subsidies for resource attributes not valued by the capacity market, legitimate or not, render the bids of those resources uncompetitive. For avoidance of doubt, FERC's successor 2019 MOPR Expansion Order explicitly includes "the procurement of . . . an attribute of the generation process," such as its lack of carbon intensity, as an example of a state subsidy.[59] FERC insists that its order, which may prevent state-supported resources from clearing the market, and therefore receiving capacity revenue, "acknowledges the rights of states to pursue legitimate policy interests," and "does not prevent states from making decisions about preferred resources," begging a narrow view of prevention.[60] FERC's expansion order, and a follow-up decision in April 2020 to deny rehearing requests, was met with widespread dismay across the energy sector, and was viewed as a major setback for clean energy in the United States.[61] Diverse groups of stakeholders, ranging from states to environmental groups to public and municipal power associations, have sued in federal court.[62]

Another example of the counteract approach to state policies are NYISO's buyer-side mitigation (BSM) rules, accepted in part by FERC in February 2020, which subject capacity offers in congested load zones near New York City to tests for price suppression from LSEs.[63] These buyers have an incentive to offer capacity into the market below cost, in order to reduce the clearing price and, therefore, the price they are charged for capacity. NYISO's BSM rules, however, apply indiscriminately

57 Glick dissenting, 171 FERC, paragraph 61,034 at P 7 (footnote 24).

58 June 2018 Order, 163 FERC, 158, paragraph 61,236.

59 169 FERC, paragraph 61,239.

60 2011 MOPR Order, 135 FERC, 143, paragraph 61,022; 169 FERC, 6, paragraph 61,239.

61 Order on Rehearing and Clarification, 171 FERC, paragraph 61,034; Jeff St. John, "FERC Denies Rehearings on PJM Capacity Orders, in a Blow to States' Renewables Plans," *Greentech Media*, April 16, 2020, https://www.greentechmedia.com/articles/read/ferc-denies-rehearings-on-its-pjm-capacity-rulings-opening-door-for-legal-challenges.

62 Catherine Morehouse, "Broad Array of Groups Sue FERC over PJM MOPR Decision as Chatterjee Rejects Cost, Renewable Concerns," *UtilityDive*, April 22, 2020, https://www.utilitydive.com/news/broad-array-of-groups-sue-ferc-over-pjm-mopr-decision-as-chatterjee-rejects/576478/.

63 170 FERC, paragraph 61,121.

to all resource offers, including demand-side ones, not simply those for which there is a basis for market power concern. Like PJM's MOPR, these rules discriminate against state-sponsored renewables, applying administrative price tests based on Net CONE, which is higher for renewables owing to their capital intensity relative to natural gas-based resources.[64] Battery energy storage, another resource with high upfront costs, is even more likely to be impacted, as New York has targeted significant capacity precisely in the congested load zones subject to the BSM rules, in order to mitigate the congestion there.[65]

In light of the disproportionate harm it causes clean energy resources and advanced technologies, as well as its reliance on pervasive and controversial administrative rules that invite litigation, the counteract approach to state policies would not be effective in supporting the energy transition.[66] The president of ISO-NE acknowledged as much when he conceded that the ISO's new Competitive Auctions with Sponsored Policy Resources (CASPR) program "is a second-best solution," and that the ISO has "long advocated that the region instead adopt a carbon price, which is a simple and easily-implemented mechanism for reducing (or eliminating) carbon and sparking a clean energy transition."[67]

CASPR represents an extension to ISO-NE's forward-capacity auction, in which clean resources that fail to clear the market due to a MOPR targeting subsidized resources are able to participate in a second auction. In this substitution auction, clean resources bid to take over the capacity obligations from retiring, uneconomic (generally fossil) resources that won them in the initial auction, earning a lower price. While providing a recourse to the MOPR, "forc[ing] state-sponsored clean energy to wait for fossil fuel generators to retire before these clean resources can enter the capacity market," and offering them reduced revenues when they do, again represents a restriction on states' abilities to affect a meaningful externality in markets.[68]

Integrate

In contrast to the limitations and regressive qualities, respectively, of the "accommodate" and "counteract" market approaches to state policies on carbon emissions, there is a growing consensus in the power sector to take an integration-based approach. A group of US stakeholders from across traditional fault lines of the supply side of the industry, including the American Council on Renewable Energy, the Electric Power Supply Association, and the Natural Gas Supply Association, petitioned FERC in April 2020 for a technical conference on integrating state, regional, and federal carbon pricing into wholesale markets.[69] As noted in the petition, "The unique features of organized wholesale electricity markets create an opportunity for integrating policies that directly price carbon emissions into energy market operations."[70] This conference was held in late September 2020.

There are various ways such integration could take place. For markets in which a cap-and-trade system exists, markets could force suppliers to include their carbon allowance costs into their market offers, reflecting the true, increased cost of their production. Alternatively, markets could impose a carbon tax on generator emissions. In both cases, imports and exports must be carefully handled to prevent emissions leakage and to minimize the impact of carbon pricing on jurisdictions that do not support it. California and New York, respectively, offer examples of the two approaches.

64 Kathryne Cleary and Karen Palmer, "Buyer-Side Mitigation in the NYISO: Another MOPR?" *Resources,* March 6, 2020, https://www.resourcesmag.org/common-resources/buyer-side-mitigation-nyiso-another-mopr/.

65 K. Mongird, et al., "Energy Storage Technology and Cost Characterization Report," Hydrowires, US Department of Energy, July 2019, https://www.energy.gov/sites/prod/files/2019/07/f65/Storage%20Cost%20and%20Performance%20Characterization%20Report_Final.pdf.

66 Brooksany Barrowes, Robert S. Fleishman, and Nicholas Gladd, "Federal-State Capacity Market Tensions Shift into New York," Kirkland & Ellis, February 28, 2020, https://www.kirkland.com/publications/blog-post/2020/02/federal-state-capacity-market.

67 Gordon van Welie, "Open Letter," ISO New England, November 21, 2019, https://www.iso-ne.com/static-assets/documents/2019/11/combined_iso_us_senate_nov_18_and_22_letters.pdf

68 Ibid.

69 Request for technical conference or workshop of advanced energy economy, American Council on Renewable Energy, American Wind Energy Association, Brookfield Renewable, Calpine Corporation, Competitive Power Ventures, inc., Electric Power Supply Association, Independent Power Producers of New York, Inc., LS Power Associates, I.P., Natural Gas Supply Association, NextEra Energy, Inc., PJM Power Providers Group, R Street Institute, and Vistra Energy Corp, Docket No. AD20-14-000, April 13, 2020. https://www.rstreet.org/wp-content/uploads/2020/04/Carbon-Pricing_Request-for-Tech-Conf-or-Workshop.pdf.

70 Ibid.

CAISO: THE ONLY US MARKET WITH CURRENT INTEGRATED EMISSION PRICING

California operates the only cap-and-trade system in the United States outside of RGGI. Resources that serve loads within California, including those that import through the EIM's real-time market, must surrender allowances covering emissions produced in the course of serving that load. Resources within California include a greenhouse gas (GHG) compliance cost component to their energy bids, ensuring their emissions are taken into account by CAISO's economic dispatch. Higher-emitting resources are less likely to clear the market, and when they are on the margin, their compliance cost increases energy prices for all California customers. Resources external to California face higher costs in serving California load and are, therefore, allowed to determine how many MW of their energy bid may be considered for import into California, on an hourly basis, along with a declared compliance cost for this energy.

Key to CAISO's model is the allocation of supply from each generator in the network to each load. This allocation is only loosely related to the underlying physics of the grid, which dictates which way electrons actually flow, and is instead based carefully on price. Supply that is allocated to load in California, regardless of where the supply is located, is assigned a cost equal to the sum of its energy cost and its GHG compliance cost. Supply that is allocated to external loads is assigned only its energy cost, as no allowances will need to be surrendered. Importantly, different allocations of supply to load will result in different overall costs for the system, which the EIM market engine seeks to minimize in its optimization.

The EIM was designed to insulate external resources from California's emissions compliance costs. It does this by adding a compliance cost component to the locational marginal price (LMP) at all nodes outside of California.[71] This marginal carbon price is paid to external resources for each MWh they are allocated as an import into California, to offset the cost of the allowances that they must purchase. This revenue does not precisely match their compliance costs, but instead equals the difference between the EIM-wide marginal cost to serve California load and the marginal cost to serve external load. The introduction of a marginal carbon price within a real-time power market, which is predicated on a price on carbon outside of the market, is novel to CAISO's emissions integration model.

While CAISO's model successfully insulates external resources from California's emissions regime, there was concern from California's Air Resource Board (CARB) after the EIM's launch regarding emissions leakage. Based on the considerations above, the real-time market engine is incentivized to maximize the allocation of clean energy imports into California and high-emissions exports out of the state, in order to reduce the compliance costs—and, therefore, the overall costs—of the system. This has the perverse effect of administratively shifting carbon emissions from California to its neighbors, without counting the emissions toward California's cap or requiring allowances for it. After investigation of several solutions to the problem, CAISO settled on a simple one: prevent energy already accounted for in resource base schedules—committed energy around which the imbalance market is "balancing," which makes up a vast majority of real-time power flows—from being allocated as an import or export.[72] Despite acknowledgment that spurious import/export allocations remain possible, FERC accepted the proposal in October 2018.[73]

CAISO's model is a compelling one for the integration of a cap-and-trade system into a real-time wholesale market. It accommodates the participation of resources and LSEs from states outside the carbon regime, insulating them from compliance costs while minimizing emissions leakage to their territory. It could apply to PJM and ISO-NE, RTOs that partially overlap with RGGI states. ISO-NE's President Gordon van Welie has acknowledged that "pricing carbon could be implemented through state or federal policy including through the existing Regional Greenhouse Gas Initiative structure."[74]

NYISO'S PROPOSED CARBON PRICE INTEGRATION

NYISO has proposed a very different carbon pricing scheme to California, which also presents a potential model for PJM and ISO-NE. Rather than integrate a carbon price based on the market price of RGGI allowances, which will reflect the scarcity value of carbon under RGGI's cap, it would instead impose a carbon

71 "Energy Imbalance Market: Draft Final Proposal," CAISO, September 23, 2013, http://www.caiso.com/Documents/EnergyImbalanceMarket-DraftFinalProposal092313.pdf.

72 "California Independent System Operator Corporation Energy Imbalance Bid Adder," Federal Energy Regulatory Commission, August 29, 2018; "EIM GHG Enhancement: Draft Final Proposal," CAISO, May 24, 2017, http://www.caiso.com/Documents/DraftFinalProposal-EnergyImbalanceMarketGreenhouseGasEnhancement.pdf.

73 165 FERC, paragraph 61,050.

74 Michael Kuser, "ISO-NE: States Must Lead on Carbon Pricing," *RTO Insider*, March 1, 2020, https://rtoinsider.com/iso-ne-states-must-lead-on-carbon-pricing-156431/.

Figure 2: A comparison of the EIM's carbon price integration, and the integration currently proposed by NYISO

	CAISO EIM	NYISO proposal
Basis of carbon price	CA ETS allowance price	Gross social cost of carbon (SCC)
Cost to generators in carbon regime	Carbon allowance procurement, outside of power markets	SCC charge, applied during real-time market settlement
Costs to loads in carbon regime	Elevated prices due to generator supply offers increased by allowance costs	Elevated prices due to generator supply offers increased by SCC costs
Allocation of revenues from carbon charges	CA Greenhouse Gas Reduction Fund (progressive emphasis on environmentally disadvantaged and low-income communities)	Residual revenue from SCC charges allocated pro rata to NY loads
Emissions leakage prevention	Non-CA loads pay LMP absent allowance cost inflation. Non-CA generators purchase allowances but are credited marginal emissions price. Clean non-CA supply, scheduled prior to real-time, cannot be attributed as an import to CA, leaking emissions.	NY generators exporting power pay SCC charge based on actual emissions, but are credited marginal emissions price. Non-NY generators importing power to NY receive LMP elevated by carbon impact but are debited marginal emissions price.
Real-time marginal emissions price	True marginal cost. Zero if CA is net exporting, otherwise the difference in marginal cost of non-CA generation serving CA load vs. serving non-CA load.	Ex post approximation, termed LBMPc. Gross SCC times heat rate of the marginal unit, estimated by LMP, fuel prices, and assumed generator technology.

Source: Susan F. Tierney and Paul J. Hibbard, "Clean Energy in New York State: The Role and Impacts of a Carbon Price in NYISO's Electricity Markets", Analysis Group, October 3, 2019; Center for Climate and Energy Solutions website; CAISO, August 29 Filing.

charge on emitters equal to the gross SCC.[75] The SCC would be determined by the state, and is expected to remain more than five times the price of RGGI allowances through 2030, the year that many aggressive clean energy targets from the state's Climate Leadership and Community Protection Act must be hit.[76] In fact, the effect of RGGI would be effectively negated by the policy; resources covered by RGGI would see the current allowance value subtracted from their carbon charge, meaning a just-in-time procurement strategy of allowances would make the resource indifferent to its price. The resource would only be exposed to the gross SCC.

New York is aiming not only for impact—with a carbon price that eclipses $60 per ton by 2027 (more than eight times the current price of RGGI allowances)—but simplicity, as this price is not subject to market forces. As emitting resources pad their bids to account for the assessed carbon charge, clean resources would enjoy a significant advantage in clearing the market, and would benefit from higher prices when emitting resources or storage resources bidding carbon charge-influenced opportunity costs are on the margin.[77] Moreover, the revenue side of the scheme is progressive; residual carbon revenue obtained from suppliers would be allocated proportionally to New York loads to mitigate the impact of higher prices.

[75] "IPPTF Carbon Pricing Proposal Prepared for the Integrating Public Policy Task Force," New York Independent System Operator, December 7, 2018, https://www.nyiso.com/documents/20142/2244202/IPPTF-Carbon-Pricing-Proposal.pdf/60889852-2eaf-6157-796f-0b73333847e8.

[76] Warren Myers, "Recommended CO2 Value to Use in IPPTF Analysis," New York Department of Public Service, April 23, 2018, https://www.nyiso.com/documents/20142/1393516/IPPTF%20CO2%20Value%204%2023%202018%20final%20%20pd.pdf/9b8ad8e6-8766-368e-43cd-171b55391a1d; Susan F. Tierney and Paul J. Hibbard, "Clean Energy in New York State: The Role and Economic Impacts of a Carbon Price in NYISO's Wholesale Electricity Markets," Analysis Group, October 2019, https://www.analysisgroup.com/Insights/publishing/clean-energy-in-new-york-state-the-role-and-economic-impacts-of-a-carbon-price-in-nyisos-wholesale-electricity-markets/.

[77] "IPPTF Carbon Pricing Proposal Prepared for the Integrating Public Policy Task Force."

The principal weakness of the proposal is its approach to calculating the marginal effect of the carbon charge on location-based marginal prices (LBMP—New York's term for LMP). This effect is termed LBMPc, and represents the shadow price of carbon compliance, which is added to the LBMP alongside the usual shadow prices of energy, transmission congestion, and transmission loss. Whenever an emitting resource is on the margin, LBMPc would be positive, reflecting the increase in the resource's cost due to the carbon charge, reflected in increased LBMP. This price would be credited to energy exports and debited from energy imports, each of which settle at the full LBMP, so that the ultimate price does not reflect the presence of the carbon charge. This allows external resources to compete on a level playing field with internal ones, without affecting their economics. After several investigations, the calculation that NYISO has arrived at for computing LBMPc is an ex post facto one, which does not flow directly from the real-time market solution. Instead, it relies on numerous administrative approximations and assumptions regarding the marginal unit and the manner in which the carbon charge affects LBMP through its bid. The calculation is simple and transparent for participants to forecast and incorporate into their bids, but does not inspire confidence that it will fully shield external resources from carbon pricing in all scenarios. While at a much earlier stage of carbon price investigation, PJM is considering a border adjustment mechanism similar to NYISO's, in order to prevent emissions leakage.[78]

Comparing New York's carbon price integration approach to California's, if enacted it is likely to have a more significant impact over a ten-year horizon, being anchored to an elevated SCC. (Auction prices for California ETS allowances have increased from 2018 through 2020, but remain below $18 per ton.[79]) At the same time, New York's carbon price would be fixed, and not reflect the scarcity value of emissions as emissions caps are reduced. If RGGI allowance prices were to eclipse the administratively determined SCC, the scheme might need to be reworked. New York's carbon price is also not as deeply integrated into the real-time market as California's, with its marginal effect being estimated ex post facto, rather than incorporated into the market engine's pricing and dispatch, as it is in the EIM.

Nonetheless, as integration-based approaches, both New York and California's models leverage markets to send price signals regarding carbon emissions in such a way that the impact on external resources and the risk of emissions leakage are minimized.[80] These approaches, therefore, enjoy distinct advantages compared to market approaches that accommodate or counteract state carbon policies.

78 "PJM Study of Carbon Prices and Potential Leakage Mitigation Mechanisms," PJM, January 14, 2020, https://www.pjm.com/-/media/committees-groups/task-forces/cpstf/2020/20200114/20200114-item-03-pjm-study-of-carbon-pricing-and-potential-leakage-mitigation-mechanisms.ashx.

79 "California and Quebec Carbon Allowance Prices," California Air Resources Board, December 4, 2020, https://ww2.arb.ca.gov/sites/default/files/2020-09/carbonallowanceprices_0.pdf.

80 Tierney and Hibbard, "Clean Energy in New York State: The Role and Economic Impacts of a Carbon Price in NYISO's Wholesale Electricity Markets."

3 What price formation enhancements are capable of solving the "missing money" problem, and reducing operator dependence on out-of-market actions?

The challenges

Two interdependent challenges face ISOs and RTOs today. If prices are set by generators' marginal operating costs only, at the exclusion of occasional scarcity prices, competitive resources earn insufficient revenue in energy markets to meet their fixed costs and make investments that are most effective for society, both from an economic and reliability standpoint. This "missing money" both causes and is reinforced by the manual dispatch by system operators of individual, often uneconomic resources for the sake of reliability, outside of the economic dispatch process. Out-of-market dispatch suppresses prices, prevents the natural retirement of uneconomic resources, and muffles price signals necessary to drive resource investment. Both of these challenges preceded the widespread deployment of renewables, but are exacerbated by it, and both must be addressed if both energy security and economic efficiency are to survive grid decarbonization.

THE "MISSING MONEY" PROBLEM

The "missing money" in electricity markets refers to insufficiency of real-time energy prices to drive the efficient entry and exit of resources in the long run.[81] The root cause is depressed scarcity prices, which do not adequately incentivize or reward resources for their availability during shortage times, when society values them most.[82] That is, the price assigned to energy at those times is less than the value that consumers assign to avoided curtailment, known as the value of lost load (VOLL). It is during scarcity periods that resources typically recover their capital costs, as, at other times, ample supply ensures that they earn only a modest margin on their variable costs.

"Missing money" has been a concern in organized markets since their inception, but has been exacerbated in recent years by declining natural gas prices and the growth of solar and wind resources.[83] These resources push higher-margin resources further down the merit curve, lowering the clearing price that all resources—these efficient ones included—receive. "Missing money" is a global phenomenon, not limited to the United States. Aggressive renewables targets, as well as the early failure of the EU ETS to maintain an "adequate, durable, and credible" carbon price, contributed to an early "missing money" phenomenon in Europe.[84]

The intermittence of renewables means that the higher-margin resources that survive in the depressed price environment that they create must generally be flexible, fast-ramping ones with low startup costs, capable of reacting to the rapid swings in renewable output. Recent retirements in California, which leads the United States in solar deployment, have primarily been inflexible oil and gas steam units, historical providers of baseload and mid-merit production.[85]

---

81 Roy Shanker, "Comments on Standard Market Design: Resource Adequacy Requirement," Federal Energy Regulatory Commission, 2003, http://elibrary.ferc.gov/idmws/common/opennat.asp?fileID=9619272.

82 Hogan, "Electricity Scarcity Pricing Through Operating Reserves."

83 Paul L. Joskow, "Challenges for Wholesale Electricity Markets with Intermittent Renewable Generation at Scale: The U.S. Experience," MIT Center for Energy and Environmental Policy Research, January 2019, https://economics.mit.edu/files/16650.

84 David Newbery, "Missing Money and Missing Markets: Reliability, Capacity Auctions and Interconnectors," University of Cambridge Working Paper in Economics 1513, February, 2015, https://www.eprg.group.cam.ac.uk/wp-content/uploads/2015/03/1508_updated-July-20151.pdf.

85 Joskow, "Challenges for Wholesale Electricity Markets with Intermittent Renewable Generation at Scale: The U.S. Experience;" Austin Perea, et al., "U.S Solar Market Insight. 2019 Year in Review," Wood Mackenzie and Solar Energy Industries Association, March 2020.

A market that today suffers from "missing money," but not as a direct cause of renewables and cheap natural gas, is ISO-NE, which operates in a region that is natural gas transmission constrained. ISO-NE has identified a misaligned incentives problem, in which the value that customers place on a generator making an upfront investment in liquefied natural gas (LNG) fuel reserves may be much greater than the market price that the generator can expect by making it.[86] The problem is one of insufficient market coordination. Resources that invest in fuel for the same anticipated scarcity period, and successfully deliver during it, prevent the very scarcity for which they were preparing. In doing so, they reduce the real-time price, eliminating the payoff of their investment.[87]

More generally, the "missing money" problem has roots in the design and operation of real-time markets. Negligible price-sensitive demand participation in markets has ceded the role of setting scarcity prices to system operators, who do this through an administrative operating reserve demand curve (ORDC). As its name suggests, this curve sets the price that the market is willing to pay, on behalf of loads, to procure a specific type of operating reserve capacity at various MW levels. Certain RTOs, including PJM, set the maximum price of system-level demand curves—representing the value of the minimal reserve capacity necessary to avoid load shedding—based on estimated resource cost to operate during scarcity periods, rather than the much higher VOLL.[88] This effect, along with offer and other price caps designed to mitigate sell-side market power, significantly depresses scarcity prices. Such tight price caps can interfere with economic dispatch during shortage situations, resulting in spurious scarcity events and out-of-market dispatch.[89]

OUT-OF-MARKET DISPATCH

Another feature of market operations that contributes to the reduction of scarcity revenue for generators—and, therefore, "missing money"—are actions taken by system operators that bypass the economic pricing and dispatch of the real-time market process. Such out-of-market actions are taken for the sake of system reliability, to prevent or prepare for a contingency, but have the effect of muting scarcity price signals and reducing market efficiency.[90] Operator reliance on them indicates an incomplete market design.[91]

A common form of out-of-market action is a reliability unit commitment (RUC)—the manual commitment of a generation resource by the system operator after the day-ahead market has closed—anticipating a reliability risk that was not addressed by the market. That risk is typically a transmission constraint, such as a faulted transmission line, requiring additional local capacity downstream of the line.[92] By acting outside of the market, the operator prevents a need from developing in real time, which would, in turn, send a short-run scarcity price signal to competitive resources to contribute, and a long-run price signal to invest in the high-risk region. Not only is the committed resource uneconomic (it did not clear the day-ahead market), but it did not participate in price formation at all. Its perceived need in real time is not represented in the day-ahead LMP, i.e., a form of price suppression.

Operator-initiated commitments can also be made at the outset of day-ahead schedule formation. The units committed address contingencies that may arise in real time, but their role is already baked into the schedule before the market optimization is run. They participate in price formation, unlike RUCs, but they do so by displacing a unit that would otherwise be committed through the economic process.[93] They do so effectively as a price taker, suppressing the day-ahead clearing price, and often require uplift payments in real time to cover their costs. Uplift payments are made administratively by system operators to individual resources, to incent them to follow market dispatch instructions when their costs (startup, variable, opportunity, and other) and the prevailing LMP may otherwise compel them not to. Uplift paid to manually committed resources is what enables operators to include them out of merit order; it ensures that they are made whole when prices are lower than they would otherwise tolerate, while providing no such support to other resources impacted by their price-suppressive effect.

86 "Energy Security Improvements. Discussion Paper," ISO New England, April 2019.

87 Ibid.

88 William W. Hogan and Susan L. Pope, "Priorities for the Evolution of an Energy-Only Market Design in Texas," FTI Consulting, May 2017, https://hepg.hks.harvard.edu/files/hepg/files/hogan_pope_ercot_050917.pdf?m=1523367673.

89 "Enhanced Price Formation In Reserve Markets of PJM Interconnection, L.L.C.," PJM Interconnection, March 29, 2019.

90 Ibid.

91 Matthew White, et al., "Energy Security Improvements: Creating Energy Options for New England," ISO New England, April 15, 2020, https://www.iso-ne.com/static-assets/documents/2020/04/esi-white-paper-final-with-cover-page-04152020.pdf.

92 Hogan and Pope, "Priorities for the Evolution of an Energy-Only Market Design in Texas."

93 Emma Nicholson, "Operator-initiated Commitments in RTO and ISO Markets," Federal Energy Regulatory Commission, December, 2014, https://www.ferc.gov/sites/default/files/2020-05/AD14-14-operator-actions.pdf.

A third form of out-of-market action taken by RTOs and ISOs is the long-term contracting of reliability-must-run (RMR) resources. These are units that are uneconomic and selected for retirement but are compelled to remain in operation by the RTO/ISO in order to address a perceived reliability need. In all markets outside of CAISO, LSEs (and, therefore, end customers) bear the cost of RMR units, which are compensated according to cost-based formulas, including a fixed rate of return. California has historically assigned these costs to transmission operators, but it has recently petitioned FERC to migrate these costs to LSEs going forward.[94]

CAISO sees elevated risk of retirement of natural gas-based resources, which are increasingly uncompetitive, necessitating RMR designations.[95] In its petition to FERC, it sought and was granted additional authority to designate a unit as RMR based on any reliability concern, rather than strictly a local one, a subtle distinction, but one that may dramatically amplify the practice and set a precedent for markets across the country.[96] It suggests a more expansive usage of these resources, and a heavier reliance on them, at the expense of market-based solutions.

RMRs act as an expensive form of insurance against real-time reliability needs. Like RUCs and other operator-initiated unit commitments, they address contingencies before they arise, rather than allowing a scarcity price signal to form and the market to act on it through resource investment. This locks in higher costs for customers and "missing money" for competitive suppliers, and blocks innovation; the incentives necessary for the system to evolve to a resource mix that addresses energy security needs at lowest cost are dropped. This harms advanced technology in particular, such as battery energy storage, flexible loads, and other DERs.

Approaches

All US wholesale markets face challenges with "missing money" and operator reliance on out-of-market actions. Each has put forward proposals to mitigate them, either tackling them directly or indirectly through adjacent market issues. Many of these proposals, as well as those put forward by outside economists, fall under three broad categories: improving the coordination of markets operating on different time horizons, improving scarcity price formation, and rethinking locational marginal pricing.

BETTER COORDINATION ACROSS MARKET HORIZONS

Effective operation of wholesale markets depends on the coordination of markets operating on different time horizons, each of which is responsible for setting up its successors for efficient outcomes. Today, those markets include long-term forward markets that operate months or years in advance, day-ahead markets, hour-ahead markets (in some cases), and real-time spot markets. By improving this coordination, markets would be able to improve operating plans going into real time, making better use of resources and mitigating potential contingencies. Not only would this reduce operator reliance on out-of-market dispatch, but it would enable a higher proportion of revenue to be captured by competitive resources in market, reducing "missing money."

One approach that markets are pursuing to improve this coordination is to extend the day-ahead market to a multiday-ahead one. The SPP's Market Monitoring Unit has recommended an extension to two days in order to make better use of long lead-time resources, which may not be available if committed a day ahead, therefore requiring manual commitment.[97] Advanced market commitment would also reduce the unnecessary shutdown and startup of resources in between operating days. NE-ISO has similarly contemplated an extension to multiday ahead as part of its Energy Security Improvements initiative, in order to better handle fuel shortages during prolonged cold snaps.[98] Compensating resources multiple days in advance of the delivery period, when their role in a multiday plan is already known, supports their investment in fuel or other preparations they may require.

Another approach for improved market coordination is to better align day-ahead products with real-time ones. The real-time spot market is the basis of price formation, so when day-ahead products do not settle against real-time analogs, there are likely to be inefficiencies in real time. Such was the case until recently for PJM, which procured a thirty-minute reserve product in

94 Order Accepting Tariff Revisions, 168 FERC, paragraph 61,199, September 27, 2019.

95 "2019 Draft Three-Year Policy Initiatives Roadmap and Annual Plan," California ISO, September 11, 2018.

96 Order Accepting Tariff Revisions, 168 FERC, paragraph 61,199, September 27, 2019.

97 "Self-Committing in SPP Markets: Overview, Impacts, and Recommendations," Southwest Power Pool Market Monitoring Unit, December 2019.

98 Matthew White and Christopher Parent, "Energy Security Improvements: Market Solutions for New England," FERC public meeting, July 15, 2019. https://www.iso-ne.com/static-assets/documents/2019/07/07_12_2019_ferc_white_final_web.pdf.

the day-ahead market, but a ten-minute reserve product in the real-time market.[99] Since resources capable of ramping to a given capacity in thirty minutes may not be able to do so in ten minutes, this misalignment put the system at risk of inadequate reserves in real time—and, at the very least—indicated that suboptimal resources may have been committed a day ahead. More subtly, if different resources are enlisted in real time compared to the day-ahead schedule, transmission flows will be different, risking unnecessary congestion and transmission losses. PJM successfully petitioned FERC last year to correct this asymmetry by procuring the same products both a day ahead and in real time: two ten-minute products (Synchronized and Non-Synchronized Reserve) and a thirty-minute Secondary Reserve.[100]

The new Synchronized Reserve product, moreover, is the consolidation of two former ten-minute products, one of which was deeply problematic as it provided close to no compensation, incurred no penalties for performance, and yet counted against PJM's Synchronized Reserve requirement. The new product is market priced (co-optimized with the energy price) and penalizes nonperformance, providing operators greater confidence in the level of reserve (reducing need for out-of-market commitment), while compensating resources more appropriately for their reserve capacity in real time.

A key feature of PJM's proposed alignment between day-ahead and real-time products is that day-ahead ancillary services should settle against the value of their real-time counterparts. That is, the day-ahead version is simply a forward procurement. This aligns with the market design of all other RTOs/ISOs today (indeed, that was one of PJM's justifications). ISO-NE has proposed an alternative: to price day-ahead ancillary services as call options on real-time energy.[101] More precisely, the supplier would sell to the market a call option that pays out the real-time LMP minus a strike price, if that difference is positive. The resource would, therefore, be selling the market insurance against the risk of a high real-time price, which the resource can cover simply by delivering energy in real time.

ISO-NE justifies this unorthodox design by claiming that exposing generators to real-time energy prices, which can be far greater than real-time reserve prices, will increase their incentive to fulfill reserve obligations in real time. Selling this option without the ability to deliver energy is equivalent to selling a naked call option on a stock, a trade with significant downside risk. At the same time, when the strike price is set "at the money" (equal to the day-ahead forecasted real-time LMP), the product would be priced higher than traditional day-ahead reserve products, better compensating suppliers and mitigating the "missing money" problem.

FERC rejected ISO-NE's proposal, not because it fails to properly align the day-ahead and real-time markets, but because it fails to align the day-time market with the forward market for LNG.[102] This market operates far in advance of the cold season, during which fuel (and, therefore, electricity) shortages occur, and offers generators the opportunity to secure fuel for future scarcity periods. Generators are often unwilling to make this upfront investment, resulting in recurring energy security concerns in New England. ISO-NE claimed that the revenue opportunity from its new day-ahead construct would incentivize generators to make the investment, and that the exposure to real-time prices would ensure they deliver on their obligations. The commission disagreed, however, noting that nothing compels generators to invest in fuel and provide the day-ahead service. Moreover, ISO-NE expects generators to incur a significant cost in the forward fuel market for uncertain revenue in the day-ahead market, a gamble that suppliers have already proven they are loath to make. In other words, ISO-NE has failed to align the two markets, such that the forward fuel market sets up the day-ahead market for success.

IMPROVE SCARCITY PRICE FORMATION

There is no more direct way to tackle the "missing money" problem than by addressing scarcity price formation. It is through scarcity revenue that resources recover fixed costs and earn profit in an efficient market. In the absence of meaningful price-sensitive demand participation, scarcity price formation is largely administrative. An ORDC sets the price that the market assigns to reserve capacity, particularly in the vicinity of minimum reserve margins, while energy prices and offer caps limit revenue for the sake of market power mitigation.

A basic reform that markets can make is to ensure that its ORDC is based on VOLL, and not assumed operating costs. By taking the latter approach, the market is explicitly denying suppliers compensation for the full value of avoided curtailment perceived by consumers.

[99] A ten-minute product indicates the resource must be capable of ramping to its committed MW capacity in ten minutes, and analogously for thirty-minute products.

[100] "Enhanced Price Formation In Reserve Markets of PJM Interconnection, L.L.C.;" 171 FERC, paragraph 61,153.

[101] White, et al., "Energy Security Improvements: Creating Energy Options for New England."

[102] Order Rejecting Proposed Tariff Revisions, 173 FERC, paragraph 61,106.

PJM takes this tack, for example, setting the peak price of operating reserves to $850/MWh, which it nevertheless has proposed increasing to $2,000/MWh to better reflect resource opportunity costs.[103] A value of $2,000 implies, for instance, that customers would be unwilling to pay $2 to run their air conditioner for twenty minutes during a blackout, even during a heat wave.[104] MISO takes the same tack, setting its peak price based on natural gas fuel costs, reaching close to $3,400/MWh in August 2019.[105] NE-ISO and NYISO's peak prices fall below these. ERCOT is a notable exception, setting its peak price to a VOLL estimated at $9,000/MWh. This is not coincidental; as the sole energy-only market in the United States, ERCOT has a particular need to support scarcity pricing.

Another important feature of ERCOT's ORDC, which PJM has proposed to adopt, is that past the initial segment of the curve, which quantifies the market's willingness to meet minimum reserve margins (based on the largest single possible generator contingency), the curve is downward sloping. Unlike PJM and SPP's current ORDCs, this reflects the continuous dependence of the probability of load shedding—and, therefore, customer value—on the current reserve margin. That is, each MW of reserve capacity past the minimum reserve margin decreases the probability of system capacity falling below that margin in real time, and therefore decreases its value.

A shortcoming of ERCOT's model, however, is that reserve requirements—as well as their corresponding ORDC—are not defined locally, but at the system level. PJM, NYISO, MISO, and NE-ISO each define ORDCs at the zonal level, ensuring sufficient capacity is procured adjacent to where it will be needed, minimizing transmission risk. ERCOT also does not co-optimize operating reserve price with real-time energy price, instead treating the former as an ex post facto price adder. These are valuable reforms as well, if secondary to the height and shape of the ORDC.

When wading through the technical details of ORDCs, it is critical for markets and their stakeholders to critically question the very existence of these administrative constructs. As customer loads become increasingly cloud-connected and flexible, and DERs enable customers to participate more directly in power markets, there is an opportunity for operating reserve demand to be set by those who actually bear the cost of system failure: end customers. This must be a regulated activity, to ensure societally reasonable outcomes, but would otherwise empower end consumers of electricity to determine the price that they are willing to pay for energy security, and the price beyond which they are willing to forego it. This becomes a meaningful tradeoff when customers are able to invest in distributed generation and battery storage, reducing their exposure to short-term supply deficiencies. Enlisting greater price-sensitive demand participation in wholesale markets is largely the responsibility of LSEs, but markets themselves have a role to play through the continued development of participation models that are simple, flexible, and widely accessible, and that strike the right balance between market exposure and risk mitigation.

RETHINKING MARGINAL COST PRICING

Location marginal pricing is one of the cornerstones of modern power markets. It prescribes that the price of power should be based on the marginal cost of providing it at the particular time and location on the grid.[106] In practice, this means that the LMP at a given node of the transmission grid during a real-time market interval equals the increment in *total system cost* if an additional MWh of load were to appear there.

While conceptually elegant and backed by significant economic theory, this pricing model has basic flaws when applied to real wholesale markets. The first is that it does not take into account the startup or no-load costs of generators, or the economic minimum ("ecomin") levels of certain resources.

Consider a simple scenario of two co-located generators: one with a high marginal cost and no startup cost (Generator A), and another with a low marginal cost but a high startup cost (Generator B). For demand below Generator A's maximum capacity, it is dispatched solely to supply all demand, avoiding Generator B's high upfront cost. But, for demand greater than that value, the optimal dispatch is very different: *Generator B* should be dispatched on its own. Once it must be committed, and its startup cost incurred, it is advantageous from this point on to incur its lower marginal cost in place of Generator A's. As seen in Figure 1, the result is that total system cost as a function of total demand has a bump in it. It increases at Generator A's steep marginal

[103] William W. Hogan and Susan L. Pope, "PJM Reserve Markets: Operating Reserve Demand Curve Enhancements," Harvard University, March 21, 2019, https://scholar.harvard.edu/whogan/files/hogan_pope_pjm_report_032119.pdf.

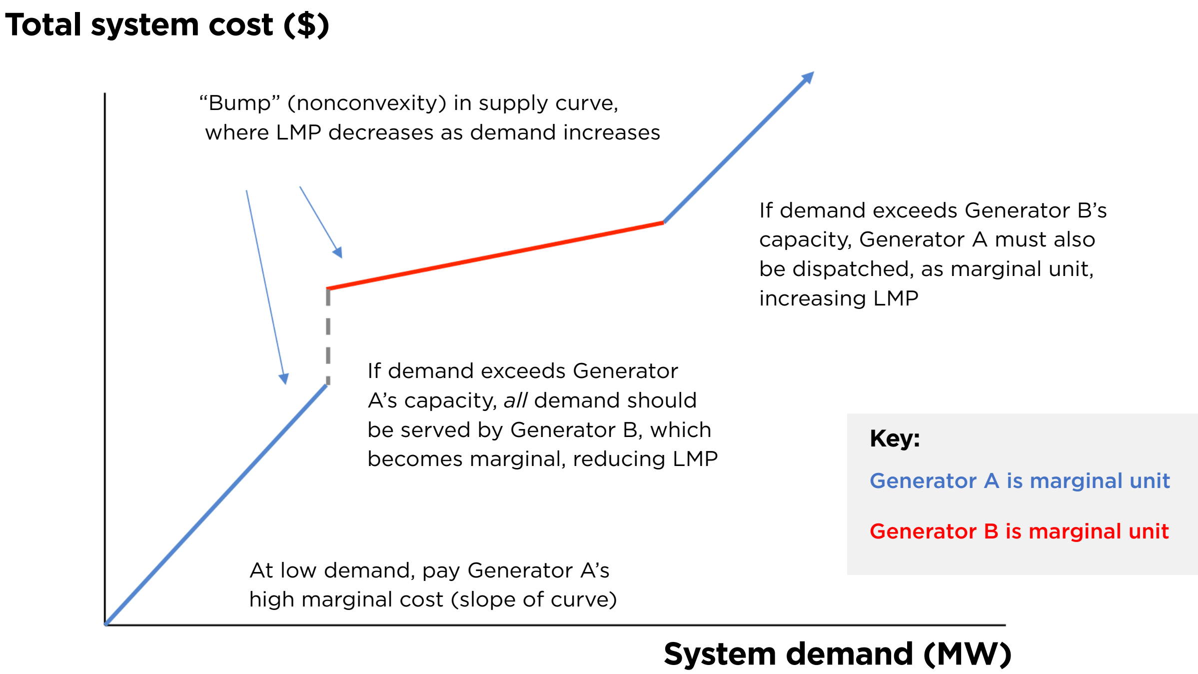
[104] This assumes a three-kilowatt central air conditioner, operating at a cost of $2/kWh, equivalent to $2,000/MWh. It does not account for distribution charges or LSE profit margin.

[105] "Energy and Operating Reserves Pricing and Market Reports August 2019 Demand Curve," MISO, July 23, 2019, https://www.misoenergy.org/markets-and-operations/notifications-overview/energy-and-operating-reserve-pricing-notifications/august-2019-demand-curve/.

[106] Fred C. Schweppe, et al., *Spot Pricing of Electricity* (Boston: Kluwer, 1988).

Figure 3: Two generator example of a nonconvex supply curve

Generator B's startup cost and lower marginal cost create an upward bump in the curve. While total system cost increases with demand, price decreases, requiring out-of-market uplift payments to make Generator B whole.



price until the point that Generator B must be committed, at which point it jumps up to Generator B's all-in cost at that value, after which it increases more shallowly at Generator B's marginal price. The result is that the LMP is *reduced* as demand on the system increases, compensating the units less per unit for doing more.

While this phenomenon can be explained as an economy of scale of sorts, in fact, Generator B will not be made whole when paid at its marginal cost only; it requires an out-of-market uplift payment to cover its startup cost. Absent this payment, it would not follow the dispatch instruction. In particular, the reduced LMP ignores Generator B's significant startup cost, rather than, for example, amortizing it over each MWh the unit produces. The price set by the marginal unit is paid to all units below it in the supply stack, so the startup and no-load costs missing from that one resource's variable cost can have enormous financial implications for the entire fleet.

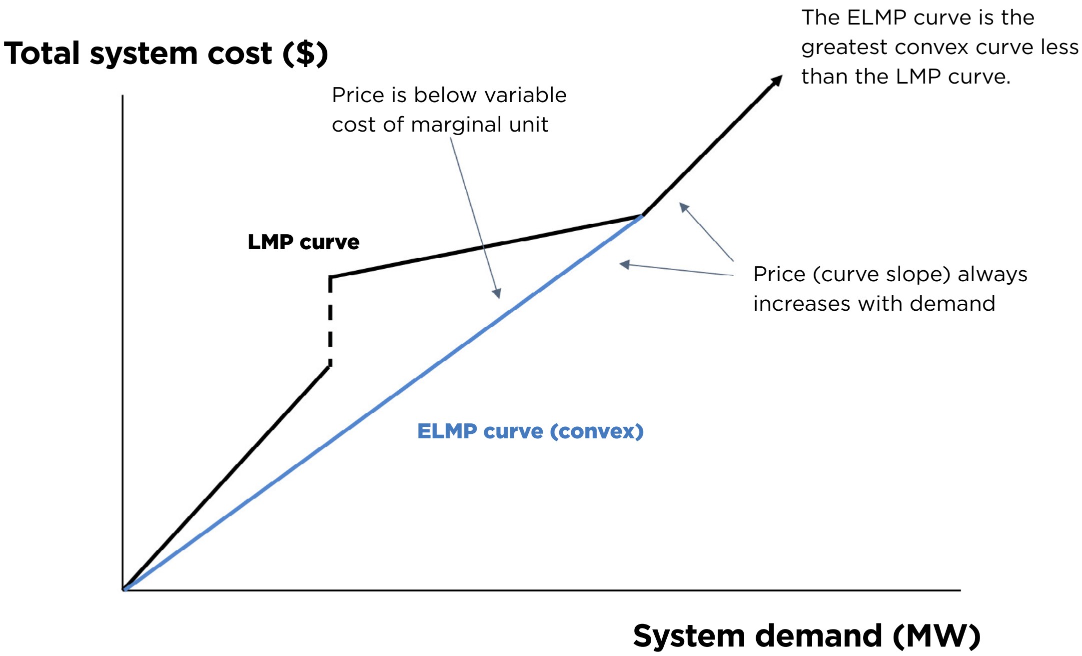
A related flaw is that certain resources, notably block-loaded ones (those that can only operate at a single output level), are unable to set the marginal price. In the block-loaded case, this is because the resource cannot be dispatched up 1 MWh, and therefore, by definition, cannot be treated as marginal. This is a significant concern during the ramp up to peak load periods, when block-loaded peaker plants are predominantly the ones dispatched and should, in principle, be setting price.

A final flaw in locational marginal pricing is that it may not make sense in a decarbonized power sector, in which most generators, including renewables and battery storage, have near-zero marginal cost. At worst, it may dramatically worsen the "missing money" problem. When renewables meet all but a fraction of system load, the marginal price will be low, forcing them to recover their capital costs at infrequent times of scarcity.[107] But, storage and flexible load resources would be expected to respond at precisely those times, preventing scarcity from occurring and prices from rising. This is an economically efficient outcome—indeed, the very goal of the energy transition—and yet it is wholly unsustainable for renewables, which would be unable to recover their upfront costs.

[107] "Wholesale Electricity Market Design For Rapid Decarbonization."

Figure 4: Extended LMP (ELMP) versus LMP

Extended LMP pricing ensures that energy price increases with demand and minimizes (but cannot eliminate) uplift payments. The downside is a more challenging optimization problem, and no simple interpretation of price as the variable cost of the marginal unit.



A common approach applied by markets today to the first two flaws is a technical one, known as integer relaxation. The LMPs that are implied by the economic dispatch optimization are ignored, and a second "pricing run" is performed, in which troublesome constraints are relaxed. The minimum operating level of block-loaded units is artificially reduced, for example, and startup costs can be amortized over MWh produced. A more sophisticated approach, pursued most aggressively by MISO, is known as extended locational marginal pricing (ELMP), or "convex hull" pricing.[108]

The mathematical defect with generator upfront costs and constraints is that they make the resource's offer curve non-convex (not upward-curving). Convex offer curves correspond to each MWh of generation being at least as expensive to produce as the previous one. Fundamental mathematical theorems imply that least-cost optimization over non-convex curves may result not only in a suboptimal solution—greater total system cost than is actually achievable—but in marginal prices that do not incent all resources to follow dispatch instructions, instead requiring uplift. This can be prevented if one replaces the aggregate supply curve (indicating the system cost associated with each level of demand) with the greatest convex curve less than it: its convex hull. This curve is equal to the original one outside of the upward bumps, implying the same LMP, but will have an LMP within the bumps equal to a weighted average of the slopes of its sides. This blended LMP "takes the long view" of price within this range of demand.

The key property of the convex hull supply curve is that its slope is always increasing, and the price of power can only increase with demand, i.e., the expectation for supply curves. A second key benefit of ELMPs is that they provably minimize the required uplift, and, in particular, reduce uplift compared to prices resulting from today's integer relaxation.[109] The cost of these two benefits, however, is that one gives up the interpretation of prices as marginal costs. Computing ELMPs is also much more intensive than other approaches, requiring mathematical and computational advancement if this pricing model is to become viable in markets.

[108] "ELMP III White Paper I. R&D Report and Design Recommendation on Short-Term Enhancements," MISO, January 31, 2019, https://cdn.misoenergy.org/20190117%20MSC%20Item%2005%20ELMP%20III%20Whitepaper315878.pdf.

[109] Paul R. Gribik, William W. Hogan, and Susan L. Pope, "Market-Clearing Electricity Prices and Energy Uplift," 2007, http://www.lmpmarketdesign.com/papers/Gribik_Hogan_Pope_Price_Uplift_123107.pdf.

The third flaw of locational marginal pricing—that it relates price to variable production cost when that cost will be near zero for a significant fraction of the decarbonized fleet—can be overcome in a number of ways. The first, and simplest, way is to ensure that power procurement occurs through long-term arrangements, such as PPAs, rather than spot markets. This reduces the exposure of renewable generators to LMPs, while leaving these prices to be influenced to a greater degree by other resources, such as gas-fired plants, which are, in fact, driven by marginal costs. The key question remains, however, regarding who are the buyers of long-term power, and how these entities are governed and regulated, particularly if they are centralized authorities.[110]

Another approach is to prevent overbuilds of renewables or complementary flexible resources, either of which would be financially unsustainable.[111] With the right proportion of each resource set, loads and storage would consume enough power during peak renewable generation to offer those generators a healthy margin, and not shed or discharge enough power during peak load periods to deny the generators a scarcity premium. To achieve this resource balance in the short run, forward markets would need to set the proper long-run price signals.

A more radical approach would tackle the zero marginal nature of the fleet head on, rather than cleverly work around locational marginal pricing's flaws. Marginal cost in electricity pricing is a proxy for scarcity, under the assumption that there is meaningful marginal cost variation among resources, and the highest cost resources are dispatched only in scarcity situations. For a fleet dominated by zero-margin renewables and battery storage, which are capable of meeting system demand at all times (for the sake of argument), this relationship does not hold. LMPs will be fixed at zero in this case. While one may argue that this price is technically correct, as there is no scarcity and real-time power is effectively free—producible at zero marginal cost—there is a defect to the argument. PPAs, through which long-term power is procured, are priced based on futures prices for electricity, and futures prices converge to spot prices due to the availability of arbitrage strategies: arbitrageurs can buy or sell a short-term futures contract and fulfill it by selling or buying in the spot market. More intuitively, no power purchaser will contract in long-term markets if they can buy free power in real-time markets.

Healthy long-term power markets, therefore, require healthy real-time energy prices, even if all the resources have zero marginal cost. Unless a meaningful proportion of system capacity comes from non-zero marginal resources, a basis for real-time prices other than marginal cost may be necessary. What that basis is, and whether it can match the historical efficiency of LMPs while supporting healthy long-term investment, remains an open question.

110 Rob Gramlich and Frank Lacey, "Who's the Buyer? Retail Electric Market Structure Reforms in Support of Resource Adequacy and Clean Energy Deployment," Grid Strategies, March 2020, https://windsolaralliance.org/wp-content/uploads/2020/03/WSA-Retail-Structure-Contracting-FINAL.pdf.

111 "Wholesale Electricity Market Design for Rapid Decarbonization."

4. How should DERs and flexible load resources participate in markets?

DERs and flexible demand resources are playing an increasingly prominent role in power markets today, a role that will only increase as the electric power system both decarbonizes and becomes more decentralized. Market analyses forecast compound annual growth rates of more than 10 percent from 2020 through 2030 for several classes of DERs in the United States, including microgrids and residential PV, with a rate as high as 40 percent for residential battery systems.[112] This growth is already being felt within markets, with MISO reporting an 80 percent year-over-year growth of DERs in its footprint from 2017 to 2018, led by a staggering 170 percent growth in rooftop PV. 2019 continued the trend of significant coal retirements across the United States, most notably in MISO, PJM, and regions outside of competitive markets, with nearly all of the new capacity consisting of wind, solar, and natural gas plants.[113] In Europe, electric and hybrid heat pumps are projected to grow fivefold from 2020 through 2030, a significant increase in flexible winter demand capacity.[114]

Not only are DERs and flexible load resources making up a greater share of the US energy mix, but they are increasingly winning on price. A shot across the bow was fired in February 2019, when solar retailer SunRun cleared 20 MW of solar-plus-storage capacity in ISO-NE's 13th Forward Capacity Market, clearing despite a capacity price of $3.80/kW month, the lowest in six years.[115] This achievement was notable not only for the technology type, but the fact that the resource was not a utility-scale solar farm but a virtual power plant, composed of five thousand residential solar and storage units. The capability of residential resources to band together and outcompete dedicated centralized generation will be the driver of the distributed resource movement.

The growing advantage of DERs and flexible loads extends to reliability, as well as price. Conventional fossil generators are susceptible to fuel shortages, contingencies that often arise when their production is needed (and rewarded) most. Such was the case during the polar vortex of 2014, when coal piles froze and natural gas was scarce, due to competing space- and water-heating demand.[116] In the polar vortex of January 2019, mechanical failure was the primary factor forcing generators offline, with coal and natural gas plants again the primary casualties.[117] Australia, a country on the front lines of climate change, has seen these resources wilt in the heat as well. The thermal generation cycle requires a cooling step, in which the hot steam used to drive the turbine is condensed back into a liquid, evacuating latent heat to the environment surrounding the plant. This step can fail in extreme temperatures, causing the plant to trip offline, dropping hundreds of MW of load within minutes. Australia has been plagued by such incidents in recent years, with coal and gas plants—old and new—tripping systematically, in some cases multiple times per month.[118]

Flexible loads are the best equipped to perform in such conditions, having no fuel to run out on and no generator to fail, and they have the added benefit of being located downstream of the transmission system, where failures often occur. During a rare and

[112] Hanson, et al., "In an Accelerated Energy Transition, Can US Utilities Fast-Track Transformation?"

[113] FERC, "State of the Markets 2019," March 19, 2020.

[114] Sander van Ginkel, et al., "Flex and Balances. Unlocking Value from Demand-Side Flexibility in the European Power System," Accenture, 2018, https://www.accenture.com/_acnmedia/accenture/conversion-assets/dotcom/documents/global/pdf/dualpub_26/accenture_flex_balances_pov.pdf.

[115] Garrett Hering, "At 'Tipping Point,' Battery-Backed Solar Homes Gain Foothold on New England Grid," *S&P Global Market Intelligence*, February 8, 2019, https://www.spglobal.com/marketintelligence/en/news-insights/trending/PXvIDLhX2Wx9-Xht6Iw4Bg2; "Results of the Annual Forward Capacity Auctions," ISO New England, accessed May 11, 2020, https://www.iso-ne.com/about/key-stats/markets#fcaresults.

[116] "Polar Vortex Review," NERC, September 29, 2014.

[117] Emma Foehringer Merchant, "Surviving the Polar Vortex: A Look at How the Electricity System Fared," *Greentech Media*, February 6, 2019, https://www.greentechmedia.com/articles/read/polar-vortex-electricity-system-fared - gs.dyyQeQ8W.

[118] Ebony Bennett, "Solar Energy Shines as Heatwaves Switch Off Gas and Coal Plants," *Sydney Morning Herald*, January 12, 2018, https://www.smh.com.au/opinion/solar-energy-shines-as-heatwaves-switch-off-gas-and-coal-plants-20180112-h0hfkj.html.

much-analyzed PJM Maximum Generation Emergency/Load Management event on October 2, 2019, in which elevated fall temperatures drove increased system load and an important transmission line had to be taken out of service, demand response resources delivered—incredibly—more than twenty-two times their committed emergency capacity.[119] In dramatic counterpoint, a significant fraction of conventional generators were offline during this event due to scheduled maintenance.[120]

Analyses show DERs providing a host of benefits to states and end customers across the United States, a sample of which are listed in the figure below. This makes it critical that they are empowered, and incentivized, to offer their full physical capabilities to power markets. Despite this urgency, basic questions remain over how DERs and flexible demand can, and should, participate in markets. Some of these questions reflect basic challenges, such as how to reconcile these resources' retail and wholesale market participations, and when they should bid into markets as demand versus supply. Other questions are focused on how these resources can overcome unnecessary barriers to entry as they exist today, posed by regulation and market rules that were designed for the generator fleet of a generation ago. Dismantling these barriers, and facilitating market participation that harnesses the unique capabilities of DERs and flexible demand, should be among the chief priorities of power market reform.

Evolving market products toward advanced technologies

The generator fleet of the 1990s has left its mark on power markets, in subtle and not-so-subtle ways. Spinning and non-spinning reserves—ancillary services in which generators commit to ramping up to full capacity within ten and thirty minutes of a contingency, respectively—are so-called because the large, rotating mass in a conventional synchronous generator must already be spinning at the frequency of the grid in order for the unit to be fully loaded within ten minutes. Market rules and analyses are replete with references to implied heat rates and fuel costs, factors that do not relate to modern renewable, storage, and load resources.[121] These artifacts are not simply a historical curiosity, however, as they reflect market structures that are not properly adapted to modern technologies.

The Federal Power Act requires RTO/ISO tariffs that are "just and reasonable" and "not unduly discriminatory or preferential."[122] In light of this standard, FERC has tackled the discriminatory treatment of energy storage and load resources through a pair of regulatory orders. Order 745, issued in 2011, requires load resources to be compensated equivalently to generation, via LMP, and Order 841, issued in 2018, requires ISOs and RTOs to offer participation models for energy storage resources based on their physical characteristics, which allow them to provide all energy products and ancillary services of which they are capable.[123] Market participation models are ISO/RTO tariff provisions for resources whose physical characteristics warrant dedicated treatment.

Markets continue to work toward compliance with these rules today. All markets currently compensate demand response resources by LMP, but only as recently as June 1, 2018, did ISO-NE begin dispatching these resources economically, rather than solely based on reliability need.[124] Order 841's requirements are less prescriptive, however. All markets offer some form of participation model for energy storage resources that respects their bidirectional power flow capabilities and finite storage capacity, but whether those models permit these resources to offer all products and services of

[119] "Review of October 2, 2019 CP Event Performance Assessment Interval Non-performance," PJM Market Implementation Committee, February 5, 2020; "PJM Studying Grid Performance During Challenging Autumn Heat," PJM Inside Lines, October 21, 2019, https://insidelines.pjm.com/pjm-studying-grid-performance-during-challenging-autumn-heat/.

[120] "Review of October 2, 2019 CP Event Performance Assessment Interval Non-performance," PJM Market Implementation Committee, February 5, 2020, https://www.pjm.com/-/media/committees-groups/committees/mic/2020/20200205/20200205-item-08-review-of-20191002-cp-event-performance-assessment-interval-non-performance.ashx; "PJM Studying Grid Performance During Challenging Autumn Heat," PJM Inside Lines, October 21, 2019, https://insidelines.pjm.com/pjm-studying-grid-performance-during-challenging-autumn-heat/.

[121] "State of the Market 2018," Southwest Power Pool Market Monitoring Unit, May 15, 2019, https://spp.org/documents/59861/2018%20annual%20state%20of%20the%20market%20report.pdf.

[122] Federal Power Act, 16 U.S.C. § 824d.

[123] Demand Response Compensation in Organized Wholesale Energy Markets, 134 FERC, paragraph 61,187, March 15, 2011; Electric Storage Participation in Markets Operated by Regional Transmission Organizations and Independent System Operators, 162 FERC, paragraph 61,127.

[124] ISO New England Inc. and New England Power Pool, Docket No. ER17-2164, July 27, 2017.

Figure 5: Projected growth in the US DER ecosystem

	2020	2030	10-year Compound Annual Growth Rate
Generation from distributed solar PV	63 TWh	165 TWh	10.1%
Installed residential battery storage capacity	2 GWh	58 GWh	40.0%
Number of EVs and hybrids	1.9m	18.7m	25.7%
Electricity demand from EVs	4 TWh	60 TWh	31.1%
Microgrid capacity (North America)	11.2 GW	29.1 GW	10.0%

Source: EY analysis on EIA, IDC, IEA, Navigant Research, IEI/EEI data

Reproduction: Hanson et. al., "In an accelerated energy transition, can US utilities fast-track transformation?", GridWise Alliance and EY, December 2019.

which they are capable in a non-discriminatory manner remains contentious.[125]

The compliance plans that PJM and SPP submitted to FERC, for example, do not specify minimum run-time requirements for storage resources to provide RA and capacity, i.e., the time duration that a resource must commit to delivering energy when called by the operator.[126] PJM's forward capacity product, Capacity Performance, requires a lengthy ten-hour runtime, and federal investigators have opened an inquiry into whether this is implicitly discriminatory toward battery storage. Batteries deployed by project developers to participate in wholesale markets face a strict cost tradeoff between power capacity (in MW), which determines maximum charge and discharge rates, and energy capacity (in MWh), which determines how long they are capable of delivering power. These systems are often optimized for the former, focused on lucrative frequency regulation service and energy arbitrage, and a ten-hour minimum runtime would effectively disqualify them from capturing capacity revenue.[127] The other northeast operators, NYISO and NE-ISO, require a four-hour minimum runtime, by comparison, indicating long runtimes are not essential to reliability.[128]

[125] Iulia Gheorghiu, "Tesla, Others Question Storage Hourly Requirements, Charges in FERC Order 841 Compliance Plans," *UtilityDive*, February 13, 2019, https://www.utilitydive.com/news/tesla-others-question-storage-hourly-requirements-charges-in-ferc-order-8/548315/.

[126] "FERC Approves First Compliance Filings on Landmark Storage Rule," Federal Energy Regulatory Commission, October 17, 2019, https://www.ferc.gov/news-events/news/ferc-approves-first-compliance-filings-landmark-storage-rule.

[127] Frequency regulation is an ancillary service in which a resource must respond to a four-second Automatic Generation Control signal from the system operator to increase and/or decrease production by a precise amount (regulation "up" and "down," respectively).

[128] Comments of Advanced Energy Economy, Docket No. ER19-469-000, February 7, 2019.

Figure 6: DER benefit assessments

State/ISO	Benefits from DERs
Massachusetts	Up to $2.3 billion in ratepayer savings from using advanced energy storage to improve overall grid utilization and economics, and $250 million in regional benefits across ISO New England from optimizing storage.
New York	Using DERs to flatten peak demand would result in avoided capacity and energy savings of $1.2 billion to $1.7 billion per year, while using DERs to improve overall system efficiency would reduce line loss costs by approximately $200-$400 million per year.
Missouri	Demand-side resources, including energy efficiency and distributed resources, would create $1 billion in economy-wide benefits, including avoided retail and wholesale costs, and represent a "'no regrets' resource investment for customers"
Michigan	Offsetting Michigan's peak demand growth by 2,000 MW through a combination of demand reduction strategies could save the state as much as $1 billion over a ten-year period.
Pennsylvania	The Public Utility Commission found that use of demand response and energy efficiency produced $4.2 billion in benefits for consumers at a cost of $1.8 billion.
California, Maryland, and Pennsylvania	According to the joint brief submitted by the Public Utility Commissions of California, Maryland, and Pennsylvania in the FERC v. EPSA Supreme Court proceeding, "Demand response provides a critical competitive presence in FERC's wholesale markets, limiting market power and lowering prices to end use customers by billions of dollars."

Reproduction: Answer of Advanced Energy Economy to Supplemental Comments of Arkansas Public Service Commission, Docket No. RM18-9-000, March 29, 2019.

Even if a long-duration capacity product provides greater value than a shorter-duration one, the latter indisputably provides value, and Order 841 requires that storage resources capable of providing only the latter should be permitted to do so. As discussed in Section 3, increasing the number and breadth of capacity products to meet the flexibility needs of real-time markets should, in fact, be a key priority of capacity market reform. There should be no false choice between a ten-hour and four-hour market product; the breadth of technologies and scales of today's resources demand an equal breadth of market products they can provide.

Forward capacity is just one domain in which storage assets, renewables, and DER aggregations are unable to offer their full capabilities. MISO, for example, treats wind resources as dispatchable, but solar as non-dispatchable, and, therefore, price taking and inflexible from the perspective of the operator. MISO's market subcommittee is advocating to petition FERC to treat solar as dispatchable as well, given the three-fold increase in solar deployment that was expected in MISO's territory from 2020 to 2021.[129] NYISO is similarly considering upgrading solar to a dispatchable resource through its Large Scale Solar on Dispatch initiative, but has made no commitments.

Basic dispatchability is the tip of the iceberg with respect to utility-scale solar and wind's capabilities. The National Renewable Energy Laboratory (NREL) and CAISO, along with industry partners, have demonstrated the ability of these plants to provide an entire suite of essential reliability services, including frequency regulation (both up and down), voltage regulation, and frequency response.[130] A separate NREL and First Solar partnership study indicates that at 30

[129] "Solar as a Dispatchable Intermittent Resource (DIR)," Midcontinent Independent System Operator Market Subcommittee, September 12, 2019, https://cdn.misoenergy.org/20190912 MSC Item 07 Solar DIR Proposal381080.pdf.

[130] Frequency response refers to the ability to respond near instantaneously to changes in grid alternating current frequency, providing an inertial response to deviations away from the target frequency. Clyde Loutan, et. al., "Demonstration of Essential Reliability Services by a 300-MW Solar Photovoltaic Power Plant," National Renewable Energy Laboratory, 2017, https://www.nrel.gov/docs/fy17osti/67799.pdf; Clyde Loutan, et. al., "Avangrid Renewables Tule Wind Farm. Demonstration of Capability to Provide Essential Grid Services," California Independent System Operator, March 11, 2020, https://www.esig.energy/resources/avangrid-renewables-tule-wind-farm-demonstration-of-capability-to-provide-essential-grid-services/.

percent annual solar penetration, flexible solar could save the western region at least $268 million per year.[131] However, no energy market today permits renewables to provide these services, denying them the opportunity to mitigate the challenges posed by their variability. It is an underappreciated fact that while renewable generators cannot control their maximum output capacity at any given moment, due to its weather dependence, they have near-surgical control over their production below that level. This fact should guide ancillary service market reforms to leverage that capability, taking into account the opportunity cost when a MW of energy production is withheld as a MW of ancillary service capacity.

Solar and wind are not the only clean generation resources capable of providing flexibility. A recent study has shown that flexible nuclear would be expensive to develop, but advanced reactors capable of cycling to meet load could result in nuclear being the largest or second-largest form of in capacity in several major regions, including New England, California, and Florida.[132] This scenario is predicated on nuclear reaching costs of $50/MWh in 2006 dollars, however, with the potential of nuclear being phased out if prices remain above $76/MWh.[133] While its future is uncertain, advanced nuclear must remain a part of the conversation around deep decarbonization planning.

For their part, load and storage resources have a unique ability that is not recognized by most markets today: that of shifting load from one time period to another. Energy markets entertain bids to buy and offers to sell energy, but the manner in which these offers are cleared does not allow them to be linked. A recent study conducted by NREL and collaborating institutions, on behalf of the California Public Utilities Commission, has identified an acute need for a Shift demand response product in California, to mitigate its "duck curve" load profile.[134] CAISO's grid becomes dangerously underloaded during the day, when solar generation peaks, and then quickly ramps up to a challenging load peak in the evening as that generation vanishes and residential demand increases. Shifting load from the evening to the afternoon would reduce both solar curtailment in the afternoon and reliance on expensive, high-emitting peaker plants in the evening.

California introduced a Proxy Demand Response-Load Shift Resource (PDR-LSR) for batteries in 2019, during the third phase of its Energy Storage and DER (ESDER) initiative, but the NREL study concludes that a technology-neutral product would enable demand-side resources to displace capital-intensive battery storage.[135] The study authors estimate that a latent 53 GWh of load shift capacity exists today, primarily in commercial and industrial loads, but that the demand for load shift will quickly eclipse this capacity, calling for greater residential participation.[136] Technology to harness residential air conditioning through smart thermostats is widely deployed today through "bring your own thermostat" (BYOT) utility demand response programs, and utilities such as Arizona Public Service have created residential load shift programs as well, leveraging grid-interactive water heaters and residential battery storage and DER management systems (DERMS).[137] No market outside of CAISO supports a load shift product of any kind. As the penetration of solar PV brings the duck curve phenomenon to territories across the United States, this creates a market product gap in need of reform.[138]

A final example of the inadequacy of current market participation models and products to leverage advanced technologies regards DER aggregations. In January 2020, NYISO received FERC approval for a new Aggregation Participation Model, to enable dispersed and potentially smaller-scale DERs to participate in

[131] Steven Dahlke, Mahesh Morjaria, Vahan Gevorgian, and Barry Mather, "The Economics of Flexible Solar for Electricity Market in Transition," First Solar, May 1, 2020, https://www.firstsolar.com/es-CSA/-/media/First-Solar/Documents/Grid-Evolution/The_Economics_of_Flexible_Solar_for_Electricity_Markets_in_Transition.ashx?la=en.

[132] Karen D. Tapia-Ahumada, et al., "Deep Decarbonization of the U.S. Electricity Sector: Is There a Role for Nuclear Power," MIT Joint Program on the Science and Polity of Global Change, Report 338, September 2019, https://globalchange.mit.edu/sites/default/files/MITJPSPGC_Rpt338.pdf.

[133] Ibid.

[134] Brian F. Gerke, et. al., "The California Demand Response Potential Study, Phase 3: Draft Final Report on the Shift Resource through 2030," Lawrence Berkeley National Laboratory, February 19, 2020, https://eta-publications.lbl.gov/sites/default/files/ca_dr_potential_study_-_phase_3_-_shift_-_final_report.pdf.

[135] California Independent System Operator Tariff Amendment to Implement Demand Response Enhancements, Docket No. ER19-2733, September 23, 2019.

[136] Gerke et. al., "The California Demand Response Potential Study, Phase 3."

[137] Robert Walton, "APS Rolls Out 3 New Customer-Sited Storage and Efficiency Programs," *UtilityDive*, September 12, 2018, https://www.utilitydive.com/news/aps-rolls-out-3-new-customer-sited-storage-and-efficiency-programs/532189/; "The Future of Utility 'Bring Your Own Thermostat' Programs," Peak Load Management Alliance, March 2018, https://www.peakload.org/assets/Groupsdocs/PractitionerPerspectives-UtilityBYOTPrograms-March2018.pdf.

[138] Julian Spector, "Massachusetts Is Staring Down a Duck Curve of Its Own. Storage Could Help," *Greentech Media*, April 23, 2018, https://www.greentechmedia.com/articles/read/massachusetts-is-staring-down-a-duck-curve-of-its-own-storage-could-help.

NYISO installed capacity (ICAP), energy, and ancillary service markets.[139] A subtle, but key, limitation of this participation model is that an aggregation is only permitted to provide an ancillary service if each individual member of the aggregation is capable of providing it. This defeats one of the chief purposes of aggregations, which is to manage the power flow of individual resources in concert to achieve increased flexibility, as well as capacity. For example, a virtual power plant composed of smart thermostat-enabled air conditioners may be capable of providing ten-minute operating reserves, even if only a fraction of the air conditioners can curtail load within that ramping window. The aggregation would not bid the sum of the nameplate capacities of all the appliances, but rather the number of MW that can be shed within ten minutes and maintained throughout the service duration. When called on by the market, intelligent dispatch algorithms would then command individual thermostats based on real-time cycling conditions to meet the ramping and sustained capacity obligations. This capability is widely deployed in markets today. DER aggregation participation models must not deny providers the opportunity to use such software smarts to create flexible grid assets from populations of less flexible ones. A stronger argument can be made, in fact, that the maintenance of resource-specific participation models is doomed to perpetual technology catch-up, and that a suitably generic "universal" participation model might accommodate all resources.[140]

The aggregation participation model

While somewhat of a niche concept within power markets today, the aggregation participation model is gaining importance through several trends. The first is the proliferation of DER aggregator entities—sometimes focused on particular technologies, such as solar and storage, EV charging, and smart thermostats—whose business model relies on monetizing end-user DER assets in markets. The second is community solar, a shared ownership model in which a community invests in one or more commercial- to utility-scale solar projects to benefit from economies of scale. Early ventures have focused on community self-consumption of the solar output, but as opportunities for solar aggregations in markets increase, it is to be expected that future projects will seek additional revenue streams in markets.

A trend that is closely aligned with both DER aggregations and community solar is community choice aggregation (CCA), a model that has burgeoned particularly in New York and California.[141] CCA enables a set of residents within a given municipality to purchase power as a bloc, like a small cooperative. The aggregation is large enough to shield individuals from the complexity and risk of the wholesale market, but small enough to ensure that community preferences for clean or local energy can be prioritized. As with community solar, increased opportunities for aggregations within markets may incent CCAs to harness the load flexibility and DER capacity of their members to offer energy and services into markets.

In light of the barriers to market participation that exist for DER owners and aggregators, FERC issued a landmark order, known as Order 2222, in September 2020.[142] The order adopted a majority of the proposals that the commission put forward in a notice of proposed rulemaking (NOPR) in 2016—the same one that led to Order 841—but with significant modifications, most notably regarding dual participation in retail and wholesale markets.[143] The primary effect of the order is to recognize DER aggregations and the aggregators who bring them to market as first-class participants in wholesale markets, to be afforded the same opportunities and face the same limitations as existing participants, but with dedicated participation models that respect their unique characteristics.

Citing regional and market-specific complexities in carrying out the reforms, the commission provided considerable discretion to RTOs and ISOs, which are tasked with updating their tariffs to comply with the ruling's broad, and often subjective, criteria. For example, RTO/ISO limitations on the geographic dispersion of assets in an aggregation are required to be "as geographically broad as technically feasible," a criterion that hinges on what is "feasible" in a wholesale market, upon which few stakeholders will agree.[144] The discretion left to system operators means that few of the reforms can be taken at face value, as they may be eviscerated by tariff amendments that hew to the letter, but not the spirit, of the order. Nevertheless, Order 2222 represents a significant bipartisan achievement of the commission's

139 170 FERC, paragraph 61,033.

140 Mark Ahlstrom, "The Universal Market Participation Model," Energy Systems Integration Group, April 5, 2018, https://www.esig.energy/blog-the-universal-market-participation-model/.

141 Deanne Barrow, "Community Choice Aggregators and Community Solar," Norton Rose Fulbright, April 10, 2018, https://www.projectfinance.law/publications/community-choice-aggregators-and-community-solar.

142 172 FERC, paragraph 61,247.

143 Docket Nos. RM-16-23-000 and AD16-20-000; 157 FERC, paragraph 61,121.

144 172 FERC, paragraph 61,247, 161.

often-acrimonious Democratic and Republican factions, and will bring meaningful change to the market participation of DER aggregations in a range of areas.

INTERCONNECTION AND LOCALIZATION

Interconnection is an area in which the commission's declination to exercise its authority was as significant as in any area in which it asserted it. While an arcane engineering subject, interconnection is central to DER and renewable deployment. Any equipment capable of injecting power into the grid, such as a PV array or residential battery, requires permission to operate by the utility or other governing authority. An interconnection application must be filed, which requires a study process by a technician. This process typically begins with a sequence of simple technical screens, seeking a quick pass-or-fail decision, but can also result in a detailed study to determine whether the local grid infrastructure can support the equipment's expected behavior, and whether an upgrade is required.

As demand for DERs and other renewables has increased, interconnection applications have piled up in ISO/RTO queues. PJM reports that applications have increased steadily from 2015 through 2019, and their makeup has transitioned from primarily natural gas generation to solar and wind, with an uptick in storage as well.[145] In early 2019, with 288 generation and storage projects languishing in its interconnection queues, CAISO's board of governors was forced to take action, approving improvements to streamline the application review process.[146]

A key question for DER aggregations is whether each DER in the aggregation requires its own interconnection application—and, therefore, its own spot in the queue—or whether the aggregation can be reviewed at once, holistically. NYISO, for example, enters Generator Interconnection Agreements with facilities, rather than individual assets within that facility (such as batteries or generators) or aggregations of those assets. There is a legitimate engineering argument to consider DERs behind different points of interconnection separately, but combining these processes would respect the integrity of the aggregation as a single resource, and could accelerate DER deployment. In Order 2222, FERC declined to create universal standards for the interconnection of DERs intending to participate in a wholesale aggregation, ceding these policies to system operators, as well as state and local authorities.[147]

A more controversial question is which power authority has jurisdiction over the interconnection application process itself. Resources that participate in wholesale markets fall under FERC jurisdiction through the ISO/RTO's open access transmission tariff (OATT), while others fall under either state or distribution utility jurisdiction. DERs complicate this logic when they begin participating in markets after obtaining a non-ISO/RTO interconnection agreement, or do not participate in markets but interconnect to distribution equipment through which other DERs participate. In September 2019, during its evidence gathering on DER aggregations, FERC issued a letter to RTOs and ISOs—the so-called "September letter"—with a set of questions designed to elicit how the operators handle situations such as these. For example, NYISO affirmed in its response to the September letter that, in the latter scenario, the non-participating asset would, in fact, fall under its OATT.[148] This follows a convention known as the "first use" test, according to which the first DER that interconnects to a piece of distribution equipment for the purpose of wholesale activity converts that equipment to commission jurisdiction for all future interconnections, regardless of their intent. In other words, FERC jurisdiction acts like an infection, spreading to electrically neighboring assets through the distribution equipment to which they mutually connect, a questionable policy that reduces state influence over interconnection. While FERC did not question the validity of the first-use test in Order 2222, it created the meaningful exception that DERs interconnecting to distribution equipment solely for the purpose of participating within an aggregation would not trigger conversion to commission jurisdiction.

The dispersion of DERs in an aggregation poses important questions outside of interconnection. Several ISOs/RTOs argued in comments to the 2016 NOPR that aggregations should be confined within a single node of the transmission network, the electrical location to which distribution equipment connects, and which represents the most granular logical component of the bulk transmission system.[149] This would enable aggregations to be truly treated as a single, localized resource. A single LMP would apply to the aggregation,

---

145 Onyinye Caven, "PJM Interconnection Queue Status Update," PJM, November 14, 2019, .

146 CAISO, "California ISO Board Takes Action to Improve Interconnection Process," CAISO, press release, February 7, 2019, http://www.caiso.com/Documents/CaliforniaISOBoardTakesActiontoImproveInterconnectionProcess.pdf.

147 172 FERC, paragraph 61,247, 82.

148 New York Independent System Operator, Inc., Response to September 5, 2019, Letter in Docket No. RM18-9-000, October 7, 2019.

149 Emily Fisher, Erika Myers, and Brenda Chew, "DER Aggregations in Wholesale Markets," Smart Electric Power Alliance and Edison Electric Institute, September 2017.

Neil Chatterjee, then-chairman of FERC, answers questions during a Reuters interview in Washington, DC, in November 2017. REUTERS/Jim Bourg

and its physical dispersion would be transparent to the market's power flow optimization. Advanced Energy Economy and the Advanced Energy Management Alliance pushed back on this, however, arguing that aggregations should be permitted to span transmission nodes, up to the zonal level, roughly corresponding to distribution utility territory. This would enable aggregators to recruit participants freely across utility territories, rather than needing to microtarget within specific neighborhoods, while still enabling system operators to dispatch granularly at the nodal level. For example, CAISO already supports aggregation across transmission nodes, and both PJM and CAISO demonstrated during FERC's technical conference following the NOPR that potential solutions exist for any ensuing reliability issues.[150]

In its ESDER phase three filing, CAISO also received permission to relax the requirement that constituents of an aggregation must lie within the same LSE, citing market participants' difficulty in meeting the one 100 kW minimum aggregation capacity.[151] Given the proliferation of CCAs and competitive retail energy providers across California and other regions, setting boundaries on aggregations based on financial relationships would impose needless challenges and limit the ability of customers to choose their retail energy and DER aggregation providers freely.

FERC adopted a moderate position on the localization of aggregations in Order 2222. It declined to require RTOs and ISOs to restrict aggregations to single nodes, or to support them across multiple nodes. As referenced above, operators are required to permit aggregations to be as broad as is technically feasible. Given both CAISO and PJM's demonstration of multi-node aggregation feasibility, the commission's determination is a curious one, and may be designed to offer other operators, such as ISO-NE, an out from supporting it.

A chief concern with multi-node aggregations is that they hide the electrical attributes of their constituents from the system operator, who might not have sufficient information to dispatch them properly. To resolve this concern, FERC ruled that aggregators must share during resource registration distribution factors for

[150] Post-Technical Conference Comments by Advanced Energy Economy, Docket No. RM18-9-000, June 26, 2016.

[151] California Independent System Operator Tariff Amendment to Implement Demand Response Enhancements, Docket No. ER19-2733, September 23, 2019.

the subset of the aggregation behind each pricing node under which the aggregation has a presence. Distribution factors are mathematical values that capture a resource's effect on transmission lines throughout the network. By providing these values at a nodal level, the aggregator would provide comprehensive information to the operator. Whether it is feasible, and reasonable, to expect aggregators to possess the technical capabilities and detailed grid information in order to produce accurate distribution factors remains an important question, and was a point of concern for ISO-NE.[152]

METERING AND TELEMETRY

The distributed nature of aggregations represents a technical challenge for system operators, insofar as operators must be capable of representing and optimizing these unconventional resources in market software. Aggregations face their own technical challenges, however, involving power metering. Participation in energy markets carries strict requirements in the granularity and accuracy with which energy is metered for settlement, as well as the frequency with which data must be transmitted to the market for real-time telemetry. Energy readings must be captured on at least a five minute basis, for instance—the granularity of real-time market dispatch—and telemetered data must be provided at a frequency on the order of seconds.[153] Reporting requirements are even more strict for resources providing ancillary services.

In almost all cases, these requirements cannot be met by retail advanced metering infrastructure ("AMI," also known as smart meters), which are only capable of fifteen-minute or hourly energy readings. Dedicated submeters and communication gateways are, therefore, frequently the only option for participation. Such equipment can be prohibitively expensive—comparable to the cost of the customer's DER resource itself—compromising the economics of market participation.

DER vendors point out that it is the aggregation that should be metered, not its constituents, as it is the resource participating in the market.[154] As long as statistical techniques can be used to estimate the net load of its constituents behind each transmission node to the accuracy required by the market, no additional requirement should be imposed on the individual DERs.[155] NYISO has adopted this approach, proposing the option of an "alternative telemetry requirement" to its standard, six-second direct metering requirement.[156] DER aggregation providers would be permitted to propose a custom methodology for estimating six-second readings in between five-minute physical readings, subject to NYISO approval. These methodologies could make use of data reported by smart-home devices to cloud platforms, and to DERMS, data that do not meet rigid market requirements but, in practice, enable real-time situation awareness. In comparison to NYISO, some ISOs/RTOs have taken a harder line on metering and telemetry, with PJM insisting that behind-the-meter DERs capable of injecting power be subject to the same requirements as comparably sized merchant power plants.[157]

Like aggregation localization, metering and telemetry is an area where FERC declined to set universal standards, and instead gave considerable latitude to RTOs and ISOs. Operators are empowered to set their own rules, including requiring physical telemetry from individual constituents of an aggregation, but must justify why their requirements are necessary and do not pose an unnecessary burden.

Should operators opt to require individual metering, metering cost will remain one of the most significant participation hurdles for DERs and flexible demand resources. Fortunately, it is a problem with a limited horizon, as AMI technology will surely advance to the point that today's commercial-grade meters—which are capable of market requirements—become cost-effective in the retail setting. 5G cellular technology may be required to meet the bandwidth challenge of orders of magnitude more assets sending telemetered data in real time. Notably, utilities are uniquely positioned to support 5G's close-range requirements, given their ubiquitous pole-top infrastructure. In the meantime, NYISO's alternative telemetry policy represents a useful testbed to solicit and evaluate estimation-based metering methodologies.

DUAL PARTICIPATION IN RETAIL MARKETS

In its 2016 NOPR, FERC proposed excluding from wholesale markets resources that receive any form of

[152] 172 FERC, paragraph 61,247, 170.

[153] See e.g., "ISO New England Operating Procedure No. 18. Metering and Telemetry," ISO New England, March 5, 2020.

[154] Fisher, et al., "DER Aggregations in Wholesale Markets."

[155] "Wholesale Market Barriers to Advanced Energy—And How to Remove Them," Advanced Energy Economy, May 2019.

[156] Michael Lavillotti and Zachary T Smith, "DER Energy and Capacity Market Design," Business Issues Committee, NYISO, April 17, 2019, https://www.nyiso.com/documents/20142/6006612/BIC DER Market Design Presentation.pdf/9cdc8700-ab90-d741-c28d-0c29b3468807.

[157] Comments of PJM Interconnection, L.L.C., Docket Nos. RM16-23-000 and AD16-20-000, February 13, 2017.

retail compensation, effectively requiring resources to choose one market or the other.[158] This was premised on a concern about "duplication of compensation," whereby resources would be compensated twice for the same service. A highlighted example was that of a DER, such as a rooftop solar system, on a net metering rate (paid the retail price of energy for every MWh injected into the grid) that is seeking to participate in a wholesale market aggregation.

While the duplication concern is a valid one, it is best approached with a chisel, rather than a hammer. As explained by the New York Public Service Commission (NYPSC) and New York State Energy Research and Development Authority (NYSERDA) in their comments to the NOPR, "DERs may provide distinct benefits to the wholesale and retail markets, and their participation in each market should be supported to maximize the potential deployment of these resources."[159]

To unpack the first point, retail value is generally different from wholesale value. The former benefits the distribution system, including the distribution utility and its ratepayers, and the latter benefits the bulk transmission system and wholesale buyers. Much as a single MWh of production can be assigned distinct energy (commodity), capacity, and environmental value, so can a single MW simultaneously provide distribution and wholesale value. The NYPSC has systematized this valuation logic in its Value of DER (VDER, or Value Stack) methodology, which has replaced net metering for solar PV resources.[160] In addition to energy, capacity, and environmental value—the first two of which represent wholesale values—the commission recognizes two distinct distribution values: Load Reduction Value (LRV) and Locational System Relief Value (LSRV). LRV captures the distribution costs avoided as a result of an injection, such as the upgrade of a substation or replacement of a transformer, and LSRV is similar, but only exists in specific locations with acute distribution needs that can be met by DERs. These distribution values are significant: New York utilities have avoided significant capital projects through the deployment of DERs and load resources, which are known as non-wire alternative (NWA) procurements.[161]

In Order 2222, FERC acknowledged these arguments and reversed its NOPR position. It ruled that DERs should be permitted to participate in retail, as well as wholesale, markets, so long as they are not doubly compensated for any market services.[162] This correctly rules out real-time wholesale energy revenue for assets such as rooftop solar or batteries that are on a net-metering retail rate, as this would represent a double payment for the commodity.

A key scenario identified by the commission as one in which dual participation *should* be restricted is a DER "included in a retail program to reduce a utility's or other load-serving entity's obligations to purchase services from the RTO/ISO market."[163] This is problematic, as most retail demand response and load-shifting programs inherently reduce wholesale energy needs, even though their purpose is to provide distribution value, such as congestion relief or voltage support. The condition above suggests that a single DER cannot provide distribution value through a retail program, as well as market value in a wholesale aggregation—the central purpose of dual participation—because the retail program indirectly reduces utility energy needs. If the principal market service being provided is energy, the restriction is debatable, and depends on the materiality of the utility demand reduction being counted a second time as wholesale supply, perhaps in dollar relation to the distribution value provided. If the market service is an ancillary service or capacity product, however, the restriction is unjust and unreasonable.

To address this problem, RTOs and ISOs should amend their tariffs such that: dual participation is permitted whenever the services for which DERs earn revenue in retail programs and wholesale aggregations do not overlap; and that this is the case, in particular, when the purpose of the retail program is distribution system value.

Participating as demand vs. supply

A major unintended feature of power markets today is that flexible demand resources most frequently participate as supply, rather than demand. While energy market designers envisioned prices being set by the intersection of an upward-sloping supply curve and a downward-sloping demand curve, in practice the demand curve is nearly vertical, reflecting

[158] Docket Nos. RM-16-23-000 and AD16-20-000; 157 FERC, paragraph 61,121.

[159] Comments of the New York State Public Service Commission and New York State Energy Research and Development Authority, Docket Nos. RM-16-23-000 and AD16-20-000, February 13, 2017.

[160] Order Regarding Value Stack Compensation, Case 15-0751, State of New York Public Service Commission, April 18, 2019.

[161] See "Non-Wires Solutions," ConEdison, https://www.coned.com/en/business-partners/business-opportunities/non-wires-solutions/ and links therein for past and present NWA procurements by ConEdison.

[162] 172 FERC paragraph 61,247, 129.

[163] Ibid, 131.

price insensitivity. The supply participation of flexible demand resources, known as demand response, involves reducing the amount of energy that they otherwise would have consumed, and being paid for this difference. This load reduction is theoretically equivalent to the resource behaving as normal and injecting the same number of MWh, a fact that led FERC to require it to be paid the LMP in Order 745, over ISO/RTO protests.

DEMAND RESPONSE: DEMAND ACTING AS SUPPLY

Demand response is much more complicated than simply consuming power in a price-sensitive manner. It requires the market to compute a baselining load curve for the resource, providing the counterfactual of what the resource would have consumed absent dispatch, and requires a costly and complex two-way market integration on the part of the resource, including real-time dispatch and telemetry. The reasons that resource owners nevertheless opt for this participation model are twofold. First, the resource is compensated for not consuming power when it is expensive to do so. The previous equivalence of demand response to generation begs the significant assumption that a load would behave as normal, consuming its baseline power, when the price of energy rises above its energy offer. The offer price is precisely the value at which the resource is indifferent between behaving normally and being paid not to do so. For prices above this level, it is conceivable that the resource would elect to consume less than its baseline offer; therefore, it is being compensated at a premium.[164]

The second reason that resource owners opt for demand response is that it offers the opportunity for a capacity payment in addition to an energy payment. Demand response is permitted to offer into capacity markets, which compensate for the forward commitment to curtail—technically at present, to commit to economically *offer* to curtail—in real time. Conventional demand participation does not provide capacity revenue, since the resource makes no commitment in advance as to its real-time behavior.

A goal of power market reform should be a determination of whether the status quo is acceptable, or whether it would be societally beneficial for flexible demand resources to participate as intended, as demand. Indeed, real-time, price-sensitive demand has long been viewed as the optimal market-based solution to the intertwined reliability and "missing money" problems in markets.[165] If demand participation is deemed superior, a concerted effort should be made to encourage it, in a manner that is neither discriminatory nor preferential, and simply reflects the societal value of this approach.

Several aspects of supply-side participation should be taken into account. First, customer baseline methodologies are a burden both to markets, which must implement them, and market participants, who must advocate during their development. They also suffer from poor accuracy, particularly in the case of weather-sensitive loads, as evidenced during baseline development work during the second phase of California's ESDER initiative, to which the author contributed.[166] Second, the price elasticity implicit in a resource's willingness to participate in demand response is not necessarily present in the energy bid of the resource's LSE. For many utility residential and commercial demand response programs, for example, the customers' unhedged forecasted load is bid into the market as price insensitive, along with that of other utility customs, despite the fact that these customers stand ready to reduce consumption if called by the market. If this elasticity was present in the utility's bid, it could improve price formation, ensuring that the market-demand curve reflects the actual price sensitivity of consumers and, therefore, intersects the supply curve closer to the societally optimal point.

PJM'S PRICE RESPONSIVE DEMAND MODEL: DEMAND-SIDE PARTICIPATION WITH CAPACITY REVENUE

A final notable aspect of supply-side participation is that it is not, in fact, required for capacity revenue. PJM's Price Responsive Demand (PRD) participation model is a key counterexample. More than a decade ago, PRD was hailed by its creators as "a third generation of demand response," a successor to the foregoing

164 William W. Hogan, "Demand Response Pricing in Organized Wholesale Markets," IRC Comments, Demand Response Notice of Proposed Rulemaking, Federal Energy Regulatory Commission, Docket RM10-17-000, May 13, 2010; Robert King and Barry Huddleston, "The Debate About Demand Response and Wholesale Electricity Markets," the South-Central Partnership for Energy Efficiency as a Resource, October 2015, https://eepartnership.org/wp-content/uploads/2015/10/The-Debate-About-Demand-Response-Final-10.28.15.pdf.

165 Cramton and Stoft, "The Convergence of Market Designs for Adequate Generating Capacity with Special Attention to the CAISO's Resource Adequacy Problem."

166 "California ISO Baseline Accuracy Working Group Proposal," California ISO, June 6, 2017.

supply-side model.[167] Its purpose was to leverage smart-meter technology—capable of hourly power-consumption measurement—and dynamic retail rates to make customer loads price responsive. Price-responsive demand offers several values to markets: Less generation and transmission capacity must be procured to meet system peak load, much of which is ultimately wasted, and customers have the opportunity to save significantly on energy costs.

The model works as follows: A PRD provider, which could be an LSE or another entity, designates a portion of an LSE's load that is sensitive to real-time price as PRD.[168] The end customers responsible for this load must have smart meters installed and must be enrolled in dynamic retail rates, such as critical-peak pricing, that expose them directly or indirectly to real-time LMP.[169] Their price sensitivity is declared via a submitted PRD curve, which consists of price and MW pairs, indicating the energy consumption level that the consumers will remain below at a given LMP. Accompanying the curve must be evidence justifying it, in the form of methodologies, analyses, or pilot programs, which is subject to PJM rejection.

The MW value at the price corresponding to the generator offer cap—the height of the right-most point on the curve—indicates the level to which the PRD will reduce in the event of an emergency, and is known as its Firm Service Level. The PRD provider must demonstrate the ability to remotely reduce load to this level during an emergency, as a backstop to price sensitivity. The difference between the load's peak-load contribution—its usage during the system's peak hour from the previous year—and its Firm Service Level is termed its Nominal PRD Value. Significantly, committing to a Nominal PRD Value is equivalent to committing the same MW of capacity as a demand response resource. Therefore, PJM credits the PRD provider for that capacity at the cleared-capacity price in the load zone, subtracting that amount from the LSE's RA charge.

In this way, an LSE earns capacity revenue through demand-side participation, rather than demand response. This participation is simpler, and at lower cost, as it does not require dispatch integration to the market's energy management system, merely passive response to real-time prices. (This is true even for the backstop supervisory control during emergencies, which is dispatched by the PRD provider, not PJM.) It is a compelling market participation model not only for flexible loads, but for behind-the-meter DERs such as PV and battery storage. While there are no provisions to compensate injections into the grid (negative load), if the customer can consume them in a price-sensitive way—charging a battery from PV when prices are low, and powering loads from the battery when prices are high—their value can be captured through the demand-side participation of the premise.

While compelling as is, there is a significant opportunity to improve PRD. Currently, the LSE is compensated for its Nominal PRD Value, tied to its response in emergencies, but not the degree to which the PRD is actually sensitive to prices. While Nominal PRD Value amounts are offered into the market by the PRD provider at different reservation prices, mirroring capacity offers, the price-sensitivity PRD curve is declared as a fact, justified by data. This fails to value price-sensitivity outside of emergency settings, and the degree to which that sensitivity might be incentivized by market compensation. Indeed, both energy consumers and the system operator benefit when a buyer commits to more sharply curtail consumption as prices rise. This suggests that PRD compensation should be based not only on the final segment of the PRD curve, but on the shape of the entire curve.

How compensation should be tied to curve shape is an open question. A methodology could be developed to assign value based intrinsically on the shape; for instance, based on its average slope, perhaps with greater weight contributed by segments at higher prices. Revenue could be tied directly to that value, via an administrative process, or it could be determined by an auction mechanism, whereby PRD providers bid PRD curves at varying prices, and these bids are cleared based on curve value. Perhaps curves can be compared and valued without reducing them to their slope, instead unbundling the value of the curve segments, each a commitment with its own independent value. A comprehensive valuation of demand-price sensitivity via methods such as these could lead to improved demand-participation models and greater demand participation in markets overall.

Electric vehicles

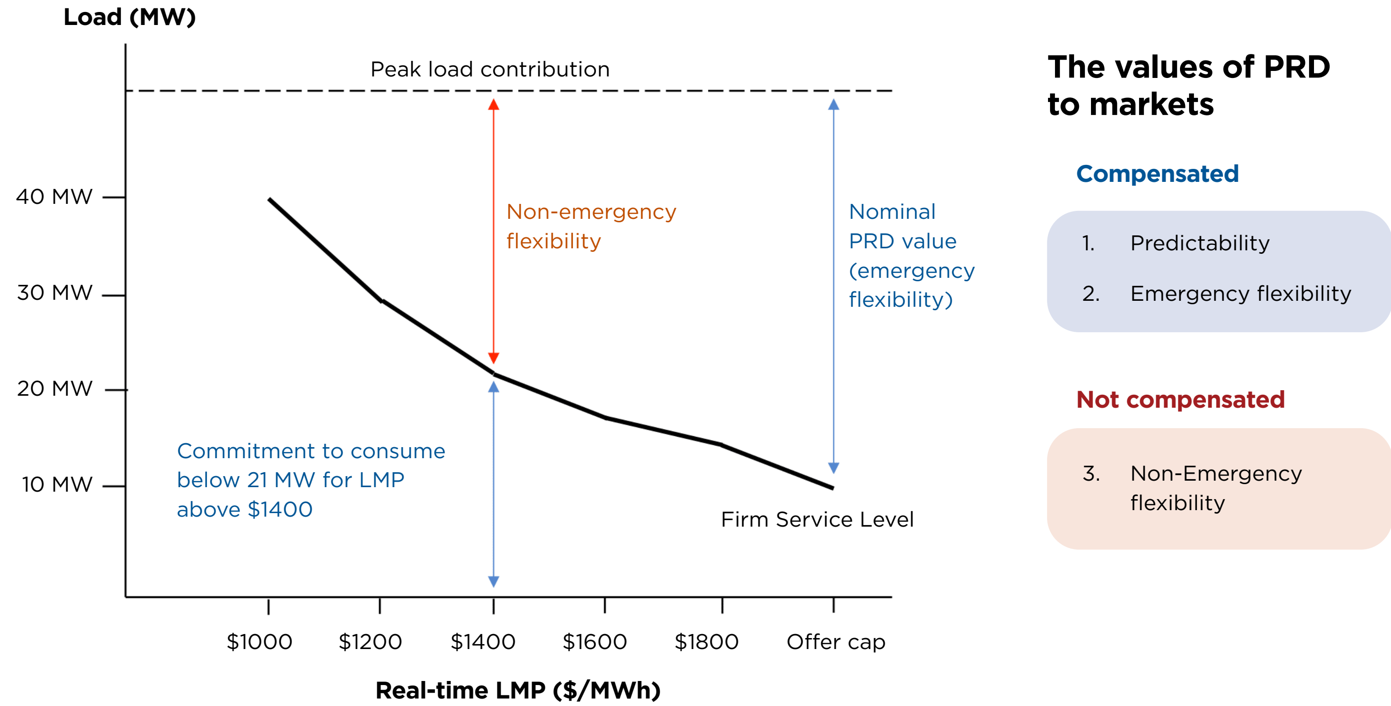
There exists no greater opportunity—or risk—for the success of DER participation in markets than electric vehicles. EVs are expected to grow in the United States from just more than one million in 2019 to almost nineteen million by 2030, representing 60 TWh of energy

[167] Paul Centolella and Andrew Ott, "The Integration of Price Responsive Demand into PJM Wholesale Power Markets and System Operations," *Energy* 35, 4, April 2020, 1568–1574.

[168] "PJM Manual 18: PJM Capacity Market," PJM, December 5, 2019.

[169] See e.g., "Recovery Act: Time Based Rate Programs," SmartGrid.gov, https://www.smartgrid.gov/recovery_act/time_based_rate_programs.html.

Figure 7: Price-responsive demand participation compensates load for emergency flexibility, but not for flexibility outside of those circumstances



annually.[170] EV sales in Europe are expected to grow 30 percent annually from 2020 to 2030, at which point EVs will represent 9 percent of the fleet.[171] If left unchecked, EV load could put perilous strain on both transmission and distribution systems: residential charging is concentrated during the evening, coincident with existing peak load and after the decline of solar generation. Time-of-use rates, an oft-cited strategy for preventing coincidence with systems peaks, would likely exacerbate the issue, with onboard charge scheduling capabilities enabling so-called "timer peaks" (mass coincidentally scheduled load) at the moment the off-peak rate begins.[172]

Despite the risks of inaction, EVs offer enormous value to energy markets and to the grid. The California Public Utility Commission's (CPUC's) Energy Division identifies three characteristics of EVs that are responsible for their grid service potential, including

- the capability to both consume and inject power;
- communication and actuation capabilities embedded by the manufacturer; and
- low capacity utilization (vehicles idle nearly 95 percent of the time).[173]

These characteristics, in addition to the abilities of EVs and smart chargers to modulate power flow near instantaneously, enable real-time coordination of EV charging for a variety of purposes. This includes renewables following, whereby charging rises and falls with solar and wind production—a practice that could eliminate thousands of GWh of wasted production, in the form of renewables curtailment—as well as staggered charging of vehicles on the same distribution circuit to eliminate price spikes and distribution overloads.[174] Enel X, an e-mobility technology provider, offers users

170 Hanson, et al., "In an Accelerated Energy Transition, Can US Utilities Fast-Track Transformation?"

171 Van Ginkel, et al., "Flex and Balances. Unlocking Value from Demand-Side Flexibility in the European Power System."

172 Erika Myers, et al., "Residential Electric Vehicle Rates That Work," Smart Electric Power Alliance, November, 2014; Dana Lowell, Brian Jones, and David Seamonds, "Electric Vehicle Cost-Benefit Analysis," M.J. Bradley and Associates, April 2017.

173 Adam Langston and Noel Crisostomo, "Vehicle-Grid Integration. A Vision for Zero-emission Transportation Interconnected Throughout California's Electricity System," Energy Division, California Public Utility Commission, March 2014.

174 David Farnsworth, et al., "Beneficial Electrification of Transportation," Regulatory Assistance Project, January 2019.

of its JuiceNet EV charging platform a service whereby charging is automatically shifted to periods of maximal renewable production. It charges a $50 software fee for this service, reflecting the assumption that customers value sustainability-driven consumption, despite the potential for inconvenience.[175]

California has undertaken a statewide effort around transportation electrification, involving the CPUC's Energy Division, the California Energy Commission, the California Air Resource Board, CAISO, and the governor's office. California's landmark Senate Bill 350 requires the state's investor-owned utilities to support transportation electrification under the guidance of the CPUC, which has produced a draft Transportation Electrification Framework for realizing a ten-year roadmap toward full electrification.[176] This framework offers a playbook for other states seeking to study and activate EVs within their own territory. While focused on distribution utilities and retail concerns, the framework also addresses numerous questions of importance to wholesale markets, of which ISOs and RTOs should take note.

One set of questions revolves around interconnection. EVs use an onboard inverter to convert direct-current electricity from the battery to the alternating-current electricity necessary to drive the electric motor. The same inverter can be used to discharge alternating current to the grid in order to provide grid services, functionality known as vehicle-to-grid (V2G). Like all assets that inject power into the grid, EVs utilizing V2G are subject to interconnection standards, although what exactly those standards should be is not clear in the case of EVs. Interconnection studies today examine prospective assets in the context of their electrical location—on a specific residential feeder, for instance—but the mobility of EVs makes this problematic. Interconnection authorities, including ISOs/RTOs, may require interconnection applications for EVs that are specific to an electrical location, permitting expansive grid services in some locations, and reduced or no service provision in others. Two of the three major investor-owned utilities in California, Southern California Edison and San Diego Gas & Electric, have even argued that vehicles that are capable of V2G but have this functionality deactivated still require an interconnection agreement.[177] If enacted broadly, such a policy could stymie the rollout of advanced inverter technologies in EVs.

Another set of questions with relevance to wholesale markets involves the grid services that EVs are capable of providing, and how service provision can be verified. As an energy storage resource, EVs are capable of providing most ancillary services procured in wholesale markets, including spinning and non-spinning operating reserves, contingency reserves, and ramping reserves. Realizing these physical capabilities will require robust communication between all entities in the EV charging ecosystem, including the vehicle, charging station, fleet aggregator, and local utility. The CPUC's Vehicle Grid Integration (VGI) Communication Protocol Working Group has identified a small number of protocol standards that meet all use cases, including ISO 15118 for vehicle-to-station communication and IEEE 2030.5 between all other entities.[178]

Owing to the sub-second response times of modern inverters, EVs are able to provide fast-response (and high-value) ancillary services as well, such as frequency regulation. This capability has been envisioned for quite some time, with a research partnership between the University of Delaware and PJM dating back to 2007, and proof-of-concept work sponsored by the California Air Resource Board dating back even further.[179] Frequency regulation is a critical service for maintaining supply-and-demand balance on the grid, in which a resource commits to increasing or decreasing its generation or load based on an Automatic Generation Control signal sent by the system operator every four to six seconds. Notably, EVs do not require V2G capabilities to provide this service, as it can be achieved by modulating charging power alone. Load-based frequency regulation has been demonstrated in MISO, for instance, using grid-interactive water heaters.[180]

175 "JuiceNet Green," Enel X, https://evcharging.enelx.com/products/juicenet-green.

176 "Transportation Electrification Framework. Energy Division Staff Proposal," California Public Utility Commission Energy Division, February 3, 2020.

177 "Rule 21 Working Group 3 In Person Meeting," Gridworks, March 6, 2019, https://gridworks.org/wp-content/uploads/2019/03/Rule-21-WG3-In-Person-Meeting-3.06-rev.pdf.

178 "VGI Communication Protocol Working Group. Energy Division Staff Report," California Public Utility Commission Energy Division, October 2018.

179 Willett Kempton et. al., "A Test of Vehicle-to-Grid (V2G) for Energy Storage and Frequency Regulation in the PJM System," University of Delaware, 2008, https://www1.udel.edu/V2G/resources/test-v2g-in-pjm-jan09.pdf; Alec N. Brooks, "Vehicle-to-Grid Demonstration Project: Grid Regulation Ancillary Service with a Battery Electric Vehicle," AC Propulsion Inc., December 10, 2002, http://www1.udel.edu/V2G/docs/V2G-Demo-Brooks-02-R5.pdf.

180 Douglas Danley, Dale Bradshaw, and Peter Muhoro, "Energy Storage—The Benefits of 'Behind-the-Meter' Storage. Adding Value with Ancillary Services," NRECA-DOE Smart Grid Demonstration Project, May 31, 2014, https://www.energy.gov/sites/prod/files/2016/10/f34/NRECA_DOE_Energy_Storage_May_2014.pdf.

An electric vehicle charging system. Unsplash/Chuttersnap (@Chuttersnap)

The key challenge to EV fleets providing frequency regulation in markets is not service provision itself, but verifying that provision to the market. EV onboard meters are capable of the 4-6 second energy readings that are necessary for performance reporting, but it is unlikely that system operators will accept settlement data from these sources. AMI meters are not a viable alternative, as the EV's usage cannot be disaggregated from that of other site loads, but submetering via EV charging stations is being evaluated in California.[181] Unlike dedicated submeters required of other DERs, Level 2 smart chargers are frequently purchased by EV and property owners for faster and more convenient charging in residential and commercial settings, and, therefore, do not represent an extraneous cost. The openness of wholesale markets to accept performance measurements originating either from onboard battery meters or EV chargers will be a significant factor in the emergence of EVs as wholesale resources.

In addition to frequency regulation, an important, if esoteric, ancillary service that EVs may be capable of providing is black-start service. Following a blackout, the bulk power system is fully de-energized. Black-start resources are those with a primary fuel source and which are not dependent on power from the grid to start up, which contribute to returning the grid to an energized state. California, which faces recurring blackout risks during wildfire season, is searching for alternatives to diesel generators, which are a reliable, but dirty and costly, provider of black-start service. National Grid, Britain's energy system operator, is researching both the technical and commercial challenges involved in procuring black-start services from DERs, and GE has

[181] Michael Sullivan, et al., "California Statewide PEV Submetering Pilot—Phase 2 Report," Nexant, April 26, 2019, https://www.ncei.noaa.gov/news/projected-ranks.

demonstrated this capability using stationary batteries.[182] The massive forecasted collective energy capacity of EV fleets makes them ideal candidates to provide this service, but it is not presently procured openly and competitively in all wholesale markets.[183] Modernization of the procurement process and support for DER aggregations such as EV fleets represent an opportunity of market reform, and a further value stream for these resources. It is important, however, that provision of this service be voluntary for EV owners, and that it does not jeopardize their autonomy or the lifespan of their vehicle.

An area in which EVs introduce greater complexity than other DERs is in the classification of energy transactions. In addition to the ambiguity between wholesale and retail transactions, delineated painstakingly for load and storage resources, power consumption by EVs can additionally be classified as a fuel purchase. The interpretation of the purchase of a kWh of energy by a charging station and end vehicle, both during and outside of wholesale market participation, affects the jurisdiction of that transaction. A key question is whether charging stations "resell" electricity, and should, therefore, be regulated as public utilities, or are merely selling vehicle fuel. A plurality of states, including California and Iowa, have exempted stations from burdensome utility regulation, whereas Texas has answered in the affirmative, applying prohibitive financial and operational requirements designed for competitive retail energy providers.[184] This question will become more complex in the context of wholesale market participation, when energy purchased from a charging station is re-sold to the market as electricity (wholesale energy) by the end vehicle. As noted by the Congressional Research Service, "The question of how to define sales of electricity to and from charging stations (including vehicle-to-grid transactions) may be subject to significant legal interpretation, and potentially represents the intersection of various federal and state statutes and regulations."[185] Resolving this question in a consistent manner across states will be crucial for the development of economically and politically efficient participation models for EVs in ISOs and RTOs.

[182] "Distributed ReStart. Project Brief," National Grid ESO, March, 2020, https://www.nationalgrideso.com/node/177691; Jeannine Anderson, "IID Demonstrates Battery's Emergency Black Start Capability," American Public Power Association, May 17, 2017, https://www.publicpower.org/periodical/article/iid-demonstrates-batterys-emergency-black-start-capability.

[183] ERCOT procures black-start service competitively, soliciting dollar-per-hour availability offers.

[184] Kendrick Vonderschmitt, "State Utilities Law and Electric Vehicle Charging Stations," Council of State Governments, October 9, 2013, http://knowledgecenter.csg.org/kc/sites/default/files/Electric%20Vehicle%20Charging%20Stations_0.pdf; A. Christopher Young, Marc D. Machlin, and Erica Hall Dressler, "Why Your Local Electric Vehicle Charging Station Doesn't (And Shouldn't) Look Like Your Local Gas Station," *Pratt's Energy Law Report* 16, 8, September 2016.

[185] Bill Canis, Molly F. Sherlock, and Corrie E. Clark, "Vehicle Electrification: Federal and State Issues Affecting Deployment," Congressional Research Service, June 3, 2019, https://fas.org/sgp/crs/misc/R45747.pdf.

5 How can markets ensure energy security when supply is intermittent?*

Energy security is deeply connected to the challenges previously discussed in the report. An energy-only approach to resource adequacy may yield different reserve margins in real time compared to a centralized capacity-market approach, for example, and "missing money" in energy markets risks the insolvency of otherwise-economic resources that are required for reliability. However, a feature of the energy transition with a unique—and, perhaps, historically unparalleled—significance on energy security is the conversion of the generator fleet to one that is intermittent. Whereas electricity supply has historically been as reliable as the supply and delivery of fossil fuels, renewable supply is limited by the sun, wind, and other largely unpredictable forces of nature.

This lack of flexibility undermines a key dynamic that has arisen in markets. Despite the volatility of electricity prices, which is greater than that of any other commodity price, far less price-sensitive demand has arisen in wholesale power markets than market designers anticipated. A vast majority of commercial and residential customers consume electricity irrespective of price, enabled by fixed rates offered by utilities and retail energy providers, which shield them from price volatility, for a premium. As a consequence, the supply side of power markets is relied upon to be fully flexible to meet demand, meticulously following its ebbs and flows to maintain balance. Turbines cannot follow the wind and system load at the same time, however, which is a conflict that poses tremendous risks for system stability in a fully decarbonized power system.

The drivers of inflexibility

The changeover of the fleet from traditional thermal generators, such as coal- and gas-fired plants, to renewable ones introduces several energy security vulnerabilities. The first is the increasing inability of the fleet to follow and match system load at all times. The physical laws governing electricity distribution are unforgiving. Supply and demand must be balanced at every instant; otherwise the frequency of the grid's alternating current will diverge from its target (60 Hz in the United States), causing devices to malfunction and ultimately the grid's voltage to collapse. Renewables tend to peak when there is low to moderate load on the system—midday for solar, and overnight for wind—and are not always present in force when load peaks, such as on weekday evenings in the summer, or during a deep freeze in the northeast. When renewables cause generation to exceed load, excess renewables that cannot be consumed by load or storage resources must be curtailed in order to achieve balance, an action without risk to system stability, but one that induces waste. California curtailed 318 GWh of wind and solar in April 2020 alone, enough to power roughly 380,000 homes during that period.[186]

The opposite case, conversely, in which renewable supply fails to meet demand, has significant energy security implications. In these times, the system must rely on complementary resources to fill the gap, sometimes in a matter of minutes. Figure 8 illustrates the "duck curve" phenomenon in California, in which solar production falls off just as evening residential load picks up, requiring complementary resources to ramp up dramatically.

Bulk power systems are designed to support worst-case scenarios, however, which include days with minimal sun and wind. Resource adequacy may, therefore, require storage and flexible capacity on the same order as the renewables for which they cover, a costly duplicate investment in capacity. Flexible nuclear is a contender to play this role, as are hydrogen fuel cells. Green hydrogen, an energy carrier that is extracted from water through hydrolysis, can be produced when renewable capacity exceeds load and consumed when renewables fall short.[187]

Accenture estimates that Europe will require 55–90 GW of flexible capacity by 2030 across six markets:

[186] "Managing Oversupply," California Independent System Operator, May 5, 2020, http://www.caiso.com/informed/Pages/ManagingOversupply.aspx; The typical single-family residence consumes roughly 10 MWh of power per year, a quantity termed a Residential Customer Equivalent (RCE).

[187] "Hydrogen Energy Storage," National Renewable Energy Laboratory, accessed April 29, 2021, https://www.nrel.gov/hydrogen/storage.html.

France, Great Britain, the Netherlands, Germany, Spain, and Ireland.[188] At that time, the latter three countries may require flexible capacity exceeding 70 percent of total generating capacity.[189] Supply imbalance fluctuations will be shorter than they are today, but their amplitudes 55–95 percent greater, requiring increased ramping capability as well as generating capacity.[190]

The second energy security vulnerability imposed by renewables also arises from their intermittence, not in regard to following load, but in regard to the grid's sensitivity to sudden *changes* in load. The stability of bulk power systems relies on the synchrony of their generators, many of which—known as synchronous generators—consist of massive magnets rotating in the vicinity of stationary magnets, driven by steam, water, or wind. The collective mass of these rotating magnets, known as rotors, provides inertia for the grid, enabling it to withstand sharp fluctuations in load or supply. Inverter-based generators such as solar PV and battery storage (including batteries in EVs) do not have a rotating mass, and, therefore, do not naturally contribute inertia to the grid.[191] While wind turbines do, the variability of wind makes their inertial response intermittent and less dependable. A key challenge for a power system driven primarily by renewables, therefore, is where to source inertia for the system or how to operate stably without it.[192]

The physical characteristics of renewables are not the only impact these resources have on a bulk power system. Utility-scale renewable plants are frequently sited far from population centers, due to the cost of land and the availability of the relevant natural resource. Bringing this power to population centers necessitates increased, and sometimes dedicated, long-distance transmission lines, which implicitly add to the cost of the power. Today's transmission capacity provides insufficient headroom for the anticipated growth in renewables, and transmission planning processes have not been up to the task, most notably in the context of multi-region projects.[193] For example, only fifteen thousand circuit miles of transmission are planned in the next six years, compared to forty thousand built in the last decade, despite the increased need.[194] In SPP and ERCOT, renewables are already reaching transmission capacity.[195] This limitation does not apply to distributed solar, which is sited on the distribution system, conveniently proximal to loads. But for utility-scale solar projects, which have grown faster than residential- and commercial-scale projects from 2015 through 2020, as well as utility-scale wind, transmission-capacity limitations generally imply generation-capacity limitations.[196]

Cybersecurity is a perennial risk to power systems, but it will take on a new character in a world powered, to a large degree, by customer-owned DERs, such as residential and community solar, stationary batteries, EVs, and smart thermostat-enabled air conditioners. Unlike centralized resources, isolated on proprietary utility and energy market communication networks, these resources are exposed to the public Internet, and rely on commercial software platforms and device owners to maintain their security. Bot-net attacks have already demonstrated the vulnerability of Internet-of-Things (IoT) devices to cyber exploits, but the capability of DER bot-nets to deploy electrical power, as well

[188] Sander van Ginkel, et al., "Flex and Balances. Unlocking value from demand-side flexibility in the European power system," Accenture, 2018, https://www.accenture.com/_acnmedia/Accenture/Conversion-Assets/DotCom/Documents/Global/PDF/Dualpub_26/Accenture_Flex_Balances_POV.pdf#zoom=50.

[189] Ibid.

[190] Ibid.

[191] Samuel C. Johnson, et al., "Evaluating Rotational Inertia as a Component of Grid Reliability with High Penetrations of Variable Renewable Energy," *Energy* 180, 2019, 258–271.

[192] Paul Denholm, et al., "Inertia and the Power Grid: A Guide Without the Spin," National Renewable Energy Laboratory, May 2020. https://www.nrel.gov/docs/fy20osti/73856.pdf.

[193] "Wholesale Market Barriers to Advanced Energy—And How to Remove Them," Advanced Energy Economy, May 2019.

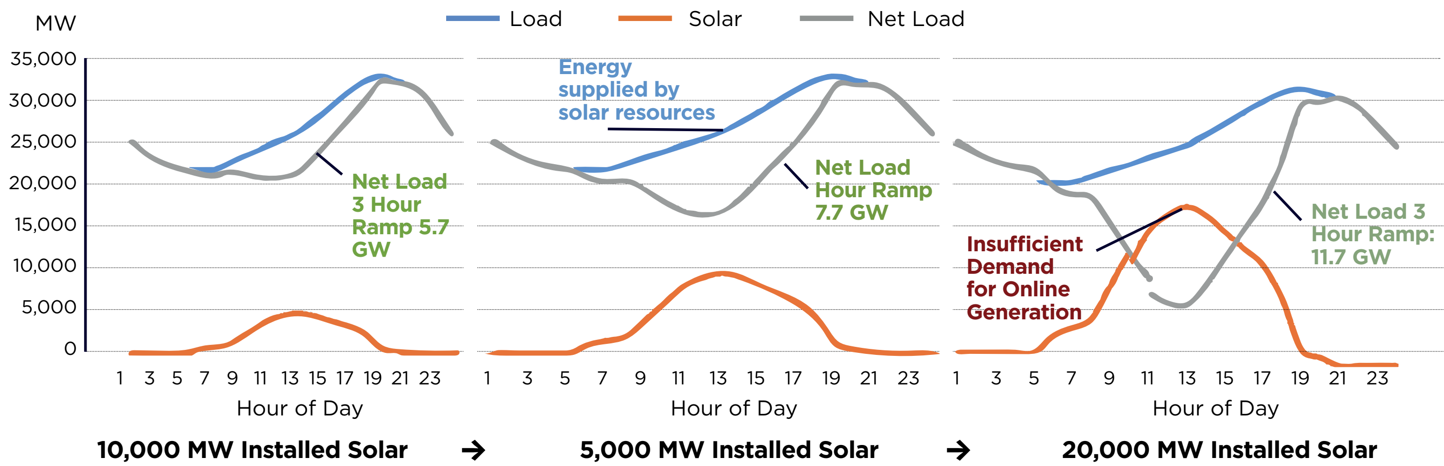
[194] "2019 Long-Term Reliability Assessment," NERC, 2019.

[195] Ibid.

[196] "Electric Power Monthly," US Energy Information Administration, February 2020, https://www.eia.gov/electricity/monthly/epm_table_grapher.php?t=epmt_1_01_a.

*The material in this section was adapated into a shorter issue brief: Ben Hertz-Shargel, "Ensuring Energy Security in a Renewables World," *Atlantic Council*, February 4, 2021, https://www.atlanticcouncil.org/in-depth-research-reports/issue-brief/ensuring-energy-security-in-a-renewables-world/.

Figure 8: Duck curve illustration



Solar production during the day masks load on the system, resulting in minimal net load that must be met by other resources. As late-day solar production declines, other resources must ramp quickly to replace it, even as demand itself is ramping toward the evening peak.

Reproduction: NERC 2019 Long-Term Reliability Assessment

as computational power, introduces novel energy security risk.[197]

Not all DER exploits will originate from Internet hacks. Global supply chains represent significant vulnerabilities for device manufacturers, whose hardware components and firmware programming are exposed to backdoor infiltration and sabotage. NERC has focused on the risk to bulk power system equipment, such as that used in transmission lines and within substations, leading to reliability standards that address supply chain risk management.[198] But, the risk persists—and is, in fact, wider and more decentralized—for IoT devices, which are produced by countless manufacturers and parts vendors across the globe, and whose role in grid operations may not even be known to ISOs and utilities.[199]

Extreme weather events also represent a perennial risk to the grid, which will persist in a decarbonized system. As discussed previously, conventional thermal generation can fail in both the extreme hot and cold: high temperatures prevent the power plant from evacuating sufficient heat during the steam cooling cycle, causing the plant to trip offline, and cold temperatures can cause mechanical failures and even fuel to freeze.

Renewables are not immune to extreme weather themselves: the efficiency of solar PV systems decay with increased ambient temperatures, and the smoke from forest fires have a particularly acute effect on production. Power forecasting firm Amperon, in coordination with the Australian Energy Market Operator, studied the effect of bush fires on twenty solar plants during the summer of 2019–2020 and found a 4.1 percent mean decrease in production over a two-month period,

[197] Josh Fruhlinger, "The Mirai Botnet Explained: How IoT Devices Almost Brought down the Internet," CSO Online, March 9, 2018, https://www.csoonline.com/article/3258748/the-mirai-botnet-explained-how-teen-scammers-and-cctv-cameras-almost-brought-down-the-internet.html; Lindsey O'Donnell, "Security Glitch in IoT Camera Enabled Remote Monitoring," *Threatpost*, July 27, 2018, https://threatpost.com/security-glitch-in-iot-camera-enabled-remote-monitoring/134504/.

[198] "Cyber Security Supply Chain Risks," North American Electric Reliability Corporation, May 17, 2019.

[199] Muhammed Junaid Farooq and Quanyan Zhu, "IoT Supply Chain Security: Overview, Challenges, and the Road Ahead," Tandon School of Engineering, New York University, July 21, 2019, arXiv:1908.07828v1.

Figure 9: Actual and forecasted upward ramps in CAISO

In 2021, three-hour ramps are expected to exceed 30 percent of total system demand, requiring significant flexible ramping capacity.

Solar supply, 2018	11,800 MW
1-hour upward ramp record set in March, 2018	7,545 MW
3-hour upward ramp record set in March, 2018	14,777 MW
3-hour upward ramps forecasted by March, 2021	17,000 MW
Total demand forecasted in 2021	54,629 MW

Source: NERC 2019 Long-Term Reliability Assessment

a massive loss in energy.[200] CAISO saw a loss of up to a third of solar production at points during the wildfires that plagued California during September 2020.[201] A recent study in *Nature Energy* examined the effects of extreme weather on renewable generation and demand under various climate change scenarios, and found up to a 16 percent drop in power supply reliability.[202] These are reminders that there is no free lunch for grid reliability, whether it is powered by legacy resources or advanced renewables.

Energy security through market-driven transmission investment

Increased transmission capacity is a clear requirement for the continued development of utility-scaled renewables projects. These projects leverage economies of scale to produce power at lower cost than residential- and commercial-scale facilities, and will play an important role in meeting clean energy targets.

In addition to bringing far-flung generation to load centers, such as metropolitan areas, transmission development can address local energy security in these areas as well. Even when a transmission path exists between regional resources and a load center, the capacity of the transmission lines may not be sufficient to carry the needed power during peak times, leading to congestion in the network. Congestion raises energy prices in the constrained area, and can become a bottleneck to such a degree that the system operator is compelled to take out-of-market actions to address it, such as dispatching a polluting and uneconomic resource, or even initiating a long-term RMR contract with such a resource. Part of FERC's basis for approving CAISO's request for unprecedented flexibility in procuring RMR resources was that CAISO would consider transmission investments as a lower-cost alternative, and that RMRs would be "a measure of last resort."[203]

There are important outstanding questions regarding the development of transmission capacity. Transmission facilities are typically procured through administrative planning processes, such as PJM's Regional Transmission Expansion Planning, or Europe's Ten Year Network Development Plan, which look at system needs over various time horizons to identify needed investments. Once a need is identified and approved by the system operator's board, competitive project bids are solicited and paid on a cost basis. Long-term planning over horizons of five years or more can increase energy security, but often fails to take into account alternative, comparatively short-term investments in local generation capacity.[204] It may be most cost-effective to develop a solar plant or a grid-scale battery storage facility in a congested area, for instance, or even solicit demand response, rather than invest in transmission to import additional power. It is,

200 Geert Scholma and Ydo Wexler, "Attenuation of Large-Scale Solar PV Production by Bushfire Smoke in South-East Australia," Amperon Holdings, Inc, 2020, https://amperon.co/case-studies/Attenuation-of-Large-Scale-Solar-PV-Production-by-Bushfire-Smoke-in-South-East-Australia.pdf?.

201 Peter Behr, "Solar Power Plunges as Smoke Shrouds Calif.," *E&E News*, September 11, 2020, https://www.eenews.net/stories/1063713459.

202 A. T. D. Perera, et al., "Quantifying the Impacts of Climate Change and Extreme Climate Events on Energy Systems," *Nature Energy* 5, 2, 2020, 150–159.

203 Order Accepting Tariff Revisions, 168 FERC, paragraph 61,199, September 27, 2019.

204 Hogan and Pope, "Priorities for the Evolution of an Energy-Only Market Design in Texas."

therefore, important that market transmission planning processes evolve to work around, rather than preempt, market-driven generation investments, and let energy market price signals do their work.

A limitation of transmission planning processes is that they centralize transmission investment. While ISOs and RTOs provide a vital public good by assessing and addressing reliability-based transmission needs, they do not facilitate economic-based transmission project development: for instance, a transmission line introduced between load zones in order to arbitrage congestion-based price differences between them, which provides market value. In a 2013 policy statement, FERC made clear its support for decentralized, market-based investment, enabling transmission developers to contract directly with loads or generators that stand to benefit from their investment.[205] In situations where the beneficiaries of a transmission investment may be too broad to contract with directly, developers might instead contract with the ISO/RTO directly, through standard transmission operating agreements. In either case, the impetus for the transmission investment is a market opportunity, rather than a reliability exigency. Facilitating such merchant transmission investments—which, unlike administrative planning processes, are closely in tune with energy market prices and alternative investment options—should be an element of power market reform. Market-based transmission investment is not a substitute for reliability-driven centralized procurement, but it is a valuable complement.

Regional markets: Security through market size

Transmission networks enable long-distance power flow, but their effectiveness is limited by the reach of the markets they serve. The US power grid is legally segmented into so-called balancing authority areas, each administered by a balancing authority—such as an ISO, RTO, or monopoly utility—tasked with ensuring stability through the balance of supply and demand. Imports and exports of power across balancing authority areas are permitted, but are not optimized like power flows within a wholesale market territory. This limits the effectiveness of excess generation in one area to serve excess demand in another, an opportunity enabled by long-distance transmission.

Regional power markets, such as CAISO's EIM, address this deficiency by co-optimizing load and generation across balancing areas. As an imbalance market, the EIM was designed as a real-time market only, optimizing power flows that are not committed by day-ahead schedules or long-term bilateral agreements. This includes facilitating the purchase of excess wind generation in the northwest by customers on the California coast, and the export of excess California solar to loads in Arizona. Since its inception in 2014, the EIM has avoided more than 1.2 million MWh of renewables curtailment, reduced close to 550,000 tons of CO_2, and generated more than $1.1 billion in gross benefits for its members.[206]

Regional markets offer energy security benefits in addition to cost savings. Their wide footprint increases the likelihood that a lull in wind or solar in one locale will be offset by a surplus in another, given natural variations in weather. By the same token, they enable a higher penetration of renewables than would otherwise be possible in fragmented and less coordinated networks.[207] Additionally, the impact of a generator tripping offline or a sudden spike in demand is reduced when there is a vast network of regional generators from which to import. For these reasons, the regional market model has become popular outside of the United States as well, including the European Union's Internal Electricity Market, the Central American Electricity Market, the Australian National Electricity Market, and the West African Power Pool.[208] Islands such as Great Britain, Ireland, and Hawaii cannot enjoy the benefit of regionalization, however, owing to their predominant (or complete) electrical isolation.

Despite early successes, there is opportunity for this model to evolve and expand. CAISO is planning to extend the EIM with a day-ahead market, in order to better coordinate regional unit commitment and power scheduling based on day-ahead load forecasts.[209] Ironically, day-ahead markets have a much greater impact on real-time power flows than real-time markets do, given how much power flow is scheduled a day

[205] Allocation of Capacity on New Merchant Transmission Projects and New Cost-based, Participant-Funded Transmission Projects, 142 FERC, paragraph 61,038, January 17, 2013.

[206] "Western EIM Benefits Report. Third Quarter 2020," California Independent System Operator, October 29, 2020, https://www.westerneim.com/Documents/ISO-EIM-Benefits-Report-Q3-2020.pdf.

[207] David Newbery, et al., "Market Design for a High-Renewables European Electricity System," *Cambridge Working Paper in Economics* 1726, June 2017, https://www.eprg.group.cam.ac.uk/wp-content/uploads/2017/06/1711-Text.pdf.

[208] Arina Anisie, Elena Ocenic, and Francisco Boshell, "Regional Markets. Innovation Landscape Brief," International Renewable Energy Agency, 2019, https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2019/Feb/IRENA_Regional_markets_Innovation_2019.pdf?la=en&hash=CEC23437E195C1400A2ABB896F814C807B03BD05.

[209] "2020 Draft Three-Year Policy Initiatives Roadmap and Annual Plan," Market and Infrastructure Policy, California Independent System Operator, September 30, 2019.

ahead, and so this extension may significantly amplify the EIM's cost savings and energy security benefits. The market continues to grow, with numerous utilities scheduled to join from 2021 through 2022, including major providers Xcel Energy in Colorado and Avista in Washington State and Idaho. It is being challenged, however, by a new Western Energy Imbalance Service Market, led by SPP.[210] Further regionalization in the west must continue, however, as lack of coordination between wholesale markets will introduce severe flexibility costs by the 2030s.[211] Greater transmission capacity will likely also be required to meet long-term decarbonization goals.[212]

As regionalization in the west increases, the absence of basic deregulation and wholesale competition in the southeast becomes even more glaring. This could be addressed by an expansion of MISO and/or PJM, both of which border the region, either via full market integration or a more limited real-time imbalance market, similar to the EIM. Another possibility is a new RTO. A recent study found that, over two decades, such an RTO could generate $384 billion in cumulative economic savings and 285,000 additional jobs compared to business as usual, with the jobs driven by the construction of new battery storage and renewables assets.[213] Either option will require the support of state lawmakers and regulators in the southeast, who would be best served by opening public utility commission dockets to study the potential benefits of regionalization to their ratepayers.

The hunt for flexibility

The averaging effects of regional markets can mitigate the intermittency of renewables, but they cannot make them self-sufficient. Complementary flexible resources will be required to cover the gap between baseload and renewable supply and demand, consuming power when renewables overproduce and injecting it when they underproduce. These swings from consumption to production will occur over the span of seconds, minutes, and hours, with little warning, an exacting demand for energy resources, few of which can ramp up and down so quickly.

Much of this flexible capacity will be procured in ancillary service markets, whose products require much greater flexibility on the part of assets than real-time energy products. Examples include frequency regulation, which requires assets to follow production setpoints that change every few seconds, and CAISO and MISO's flexible-ramping products, which compensate assets for the number of MW they are able to ramp up or down from their current economic dispatch setpoint within a five- or fifteen-minute timeframe, respectively. Flexible-ramping products are viewed as a key tool in managing renewable variability: NYISO and NE-ISO have carefully studied the current implementations, and both SPP and its independent market monitor have concluded that it should develop its own product.[214]

In its failed ESI filing, ISO-NE settled on a product, known as Energy Imbalance Reserve (EIR), that is similar to a flexible-ramping product but fits within its proposed framework for ancillary services.[215] As discussed in section 3, all day-ahead ancillary services in the new framework would be procured as call options on real-time energy. In the case of EIR, that energy corresponds to the upward ramping capacity of the resource, and the total reserves procured are equal to the difference (if positive) between the day-ahead forecasted load and the day-ahead cleared load. While FERC has rejected ISO-NE's proposal, that rejection was not on the merits of the novel call-option-based framework, or EIR. While an empirical comparison may not be possible, other ISOs and RTOs would do well to consider EIR as an alternative to flexible-ramping products. Like other ESI products, it has the potential to offer resources greater revenue than traditional products if they are prepared to meet their obligations.

It is worth noting that flexible demand could be a tremendous source of flexible-ramping capacity. Aggregations of heating and cooling loads, EVs, and commercial building loads can be surgically dialed up at the individual kW level, offering more than the precision needed for demand response ramping products. As discussed in section 4, however, price-sensitive demand participation can offer similar value without the market integration cost or complexity.

[210] Robert Walton, "Xcel, 3 Other Colorado Utilities Choose California's Imbalance Market over Southwest Power Pool," *UtilityDive*, December 18, 2019, https://www.utilitydive.com/news/xcel-3-fellow-colorado-utilities-choose-californias-imbalance-market-over/569361/.

[211] Keegan Moyer, "Western Flexibility Assessment," Energy Strategies, NW Energy Coalition Clean & Affordable Energy Conference, December 2, 2019.

[212] Ibid.

[213] Eric Gimon, et al., "Economic and Clean Energy Benefits of Establishing a Southeast U.S. Competitive Wholesale Electricity Market," *Energy Innovation*, August 2020, https://energyinnovation.org/wp-content/uploads/2020/08/Economic-And-Clean-Energy-Benefits-Of-Establishing-A-Southeast-U.S.-Competitive-Wholesale-Electricity-Market_FINAL.pdf.

[214] "State of the Market 2018," Southwest Power Pool Market Monitoring Unit, May 15, 2019, https://spp.org/documents/59861/2018%20annual%20state%20of%20the%20market%20report.pdf.

[215] ISO New England Inc.: Compliance Filing of Energy Security Improvements, Docket No. ER20-1567, April 15, 2020.

Wind turbines in Palm Springs, California. Unsplash/Kai Gradert (@kai)

Despite their value in high-variability environments, markets have not adequately priced ancillary services. In PJM, ancillary services account for less than 7 percent of the revenue brought in through capacity markets, despite the absence of any flexibility requirement in capacity products.[216] In MISO that percentage is higher, but ancillary services still account for only 0.3 percent of the all-in price of electricity.[217] Markets' under-allocation of revenue to ancillary services due to overreliance on forward capacity and out-of-market dispatch is problematic: it may fail to incent sufficient investment in flexible capacity, locking in today's uneconomic (and frequently high-emitting) backup resources for years to come. While adequate today, these resources may not be adequate in a high-variability future driven by near-100 percent renewables, compromising energy security.

Beyond management of price signals, there are actions markets can take to encourage greater flexibility from resources by adjusting market rules. Many were discussed in section 4, and involve creating new participation models and market products to enable advanced technologies such as dispatchable renewables, DER aggregations, flexible load, and advanced nuclear, in order to leverage their full physical capabilities. Another example is improved coordination with natural gas markets. By posting day-ahead unit commitments further in advance of the natural gas day-ahead window, power markets would enable generators to bid their full flexible capacity, in the knowledge that they will have time to estimate and procure exactly as much fuel is necessary to support however much of their bid clears the market.[218]

Another rule change would be to limit, or even eliminate, the practice of resources self-scheduling their generation, rather than participating in economic dispatch.[219] Resources typically self-schedule if they have already contracted with a buyer, such as through a PPA, or if they have a long lead times to start up, and cannot wait for the day-ahead market to close. These resources are inflexible in several respects: they are price taking, rather than price responsive; they increase the risk of network congestion, as their dispatch cannot be optimized; and they frequently have preferred treatment with respect to curtailment. As discussed previously, reducing self-commitment was one of the reasons the SPP's Market Monitoring Unit recommended an extension of its day-ahead market to two

216 "State of the Market Report for PJM, Q1 2020," Monitoring Analytics, LLC., May 14, 2020, https://www.monitoringanalytics.com/reports/PJM_State_of_the_Market/2020/2020q1-som-pjm.pdf.

217 "2018 State of the Market Report for the MISO Electricity Markets," Potomac Economics, June 2019, https://cdn.misoenergy.org/2018 State of the Market Report364567.pdf.

218 Robbie Orvis and Sonia Aggarwal, "A Roadmap for Finding Flexibility in Wholesale Markets," *Energy Innovation*, October 2017.

219 Ibid.

days ahead.[220] Requiring contracted units to participate in economic dispatch would coerce greater flexibility from them, but potentially at the expense of the otherwise-beneficial practice of long-term contracting, and must be carefully studied, from both market and participant perspectives. Carveouts would be required for truly inflexible resources such as conventional nuclear, which provides clean baseload generation, as well as resources such as hydroelectric plants whose actions must prioritize environmental considerations above market ones.

Incremental improvements in resource flexibility and ramping will support decarbonization in the near term. But, as renewable capacity eclipses that of conventional resources, the ultimate challenge will be a lack of grid inertia. Inverters today that enable batteries and renewables to deliver power to the grid are grid following, in that they know only how to follow the alternating current of synchronous generators. In an environment with few synchronous generators, inverters will need to operate in a more challenging grid-forming mode, as they do in an islanded microgrid, in which they act as leaders rather than followers in establishing alternating-current synchrony.[221] Beyond that, their power electronics will need to react near instantaneously to power transients, mimicking physical inertial response.[222] This feat is analogous to what a Segway scooter accomplishes in staying upright, despite the movements of its rider. Inverter designs capable of such “synthetic” or “virtual” inertial response are still in the developmental stage, however, and will need to prove themselves capable of stabilizing the grid at scale, and to the exacting requirements of NERC and other regulators.[223]

To incentivize the development of this technology while it remains inessential, markets should consider valuing inertial response as an ancillary service. This action would have the added benefit of initiating the long and complex discussion between markets, regulators, and stakeholders of how to measure and compensate for grid inertia. The Australian Energy Market Commission has begun such an investigation, and found that while requiring minimum levels of inertial response is adequate for now, a market for inertial response might be needed in the future in higher-VER environments.[224] In parallel with such market testing, the federal government can accelerate the development of synthetic inertia through Department of Energy research grants and national laboratory research partnerships, such as the one between the National Renewable Energy Laboratory and Pacific Gas & Electric.[225]

---

[220] “Self-Committing in SPP Markets: Overview, Impacts, and Recommendations,” Southwest Power Pool Market Monitoring Unit, December 2019, https://spp.org/documents/61118/spp%20mmu%20self-commit%20whitepaper.pdf.

[221] Benjamin Kroposki et al., “Achieving a 100% Renewable Grid,” *IEEE Power & Energy Magazine*, March 1, 2017, https://ipu.msu.edu/wp-content/uploads/2018/01/IEEE-Achieving-a-100-Renewable-Grid-2017.pdf.

[222] “When the Gears Stop Turning. NREL and PG&E Collaboration Demonstrates Synthetic Inertia,” National Renewable Energy Laboratory, May 30, 2018, https://www.nrel.gov/news/program/2018/when-the-gears-stop-turning.html.

[223] Roy Kuga, et al., “EPIC 2.05: Inertia Response Emulation for DG Impact Improvement,” EPIC Final Report, Pacific Gas and Electric Company, February 20, 2019.

[224] “Frequency Control Frameworks Review,” Final Report, Australian Energy Market Commission, July 26, 2018, https://www.aemc.gov.au/sites/default/files/2018-07/Final report.pdf.

[225] “When the Gears Stop Turning: NREL and PG&E Collaboration Demonstrates Synthetic Inertia,” National Renewable Energy Laboratory, May 30, 2018, https://www.nrel.gov/news/program/2018/when-the-gears-stop-turning.html.

Conclusion and policy recommendations

Decarbonizing the power sector is critical for cities, states, and corporations to achieve their climate goals. While ISOs and RTOs do not have a statutory obligation to support these goals, they must not stand in the way, and must instead operate markets as best they can to support them, consistent with their open-access transmission tariff. They have little choice, ultimately, given the relentlessly falling costs of solar, wind, and battery storage, and the growing demand for clean power generation.

Supporting the transition of the generator fleet from primarily fossil resources to renewables and energy storage will require wide-ranging reforms to energy, ancillary service, and capacity markets. Some of these reforms must be economic, focused on price formation, to ensure that real-time market prices incent optimal long-term investment, and are not distorted by the zero marginal cost of renewables or revenue transfer to capacity markets. Other reforms must be focused on physical energy security, ensuring that the fleet has sufficient flexibility to compensate for the intermittence of renewables and their lack of stabilizing inertia. Key to these latter reforms is dismantling barriers to entry of DERs, flexible loads, and other advanced technologies that can provide flexibility at low cost, and facilitating their participation in markets.

A key challenge for markets, regulations, and resource owners is how distributed resources should participate in markets, and how they can be leveraged to reduce uneconomic, out-of-market dispatch by operators. A paradigm shift might be called for, whereby flexible loads such as HVAC and electric vehicle charging bid as price-sensitive demand, rather than demand response, answering intermittent supply with flexible demand. This participation model precludes high-value ancillary services, such as frequency regulation and frequency response, however, a tradeoff that may not be justified for the most high-performing resources.

Some power market reform questions, such as the relative societal value of supply- versus demand-side participation, remain unanswered and require investigation by markets and regulators. Others, such as the most effective role for markets in addressing generator carbon emissions, have a preponderance of evidence that suggests a clear answer, in this case, integration of a carbon price, rather than ignoring or counteracting it.

In light of this, as well as the foregoing challenges of power market reform, ISOs and RTOs should do the following.

- Consider decentralizing resource adequacy, returning responsibility to load-serving entities and greater authority to states, and changing its focus from installed capacity to flexible capacity. Mandatory centralized capacity markets have introduced bloat into resource adequacy procurement, relying on administrative constructs that are poor substitutes for market mechanisms. They have also interfered unnecessarily with state decarbonization goals. More generally, resource adequacy must become defined in terms of flexible capacity and procured through an array of forward products that satisfy real-time essential reliability needs and are eligible to be provided by advanced technologies.

- Integrate carbon prices into real-time markets imposed by states, including through regional cap-and-trade programs such as RGGI. This is necessary to prevent emissions leakage and indiscriminate cost increases across consumers, which arise when markets either passively accommodate or counteract carbon prices. Both the Western Energy Imbalance Market and NYISO's proposed carbon price integration should be considered as models, in particular, their differing approaches to imports and exports.

- Take a multi-pronged approach to shore up scarcity pricing in real-time energy markets. Appropriate scarcity pricing is necessary to address the "missing money" problem for generators, and to reduce operator dependence on out-of-market actions. One priority is to base operating reserve demand curves on customers' value of lost load, rather than assumed resource costs. Another is to ensure that real-time prices are effective for a significantly zero marginal cost fleet, which may require harmonizing long- and short-term markets such that the financing costs that displace fuel costs are reflected in these prices.

- Recognize renewables as dispatchable resources and leverage the vast flexible capacity they offer. While the maximum output of renewables is variable, they have near-surgical control below that level. Markets should develop technology-neutral participation models for variable energy resources that enable both ramping and fast-response ancillary services, while accounting for the opportunity cost of real-time energy.

- Do not impose unreasonable limitations on DER aggregation formation or ability to provide market services. Aggregations should be permitted to span load service entities as well as transmission nodes, provided that distribution factors can be produced. They should not be required to provide metering and telemetry below the transmission node level. ISO/RTO requirements to the contrary should be closely scrutinized by FERC on the basis of technical necessity.

- Embrace statistical estimation-based telemetry methodologies, such as NYISO's alternative telemetry option, in order to facilitate residential DER participation in real-time markets. These methodologies permit smaller resources to share status information with the system operator based on occasional measurements and statistical estimation, rather than costly high-frequency measurements. Such methodologies should be permitted to incorporate DER and smart device data and be approved solely on the basis of demonstrated accuracy, such as by submetering a sample population.

- Permit dual participation in retail and wholesale markets whenever the unbundled services for which DERs earn revenue in each market are non-overlapping. This is the case, in particular, when the purpose of the retail program is distribution system value. New York's Value Stack methodology serves as a reference.

- Study whether flexible load resources would provide greater value in markets as price-sensitive demand, compared to supply-side demand response, as they largely participate today. If this hypothesis is borne out, initiatives to attract greater demand-side participation should be pursued. This includes exploring variants of PJM's Price Responsive Demand participation model, in which resources would be compensated for price sensitivity across all price levels, not simply receive capacity cost offsets for reductions during emergencies.

- Increase regionalization through direct ISO/RTO expansion or regional organized markets, such as CAISO's Western Energy Imbalance Market. Regionalization smooths the intermittency of renewables, reducing wasteful curtailment, and mitigates the risk of generator outages. Vertical utilities across the west should consider joining CAISO or SPP's regional markets, and the southeastern states should consider deregulating, and either joining an RTO or forming a new one. The latter could save hundreds of billions of dollars and generate hundreds of thousands of jobs over two decades.

- Evaluate productizing inertial response as an ancillary service. Investigations would have the added benefit of initiating the long and complex discussion between markets, regulators, and stakeholders of how to ensure grid inertia with a near-100 percent renewable fleet. They should build off the investigation concluded by the Australian Energy Market Commission in 2018 and compare the efficacy of a market product for inertial response to requirements established in interconnection agreements.

About the Author

Ben Hertz-Shargel is a Non Resident Senior Fellow at the Atlantic Council Global Energy Center, with a portfolio including decarbonization, emerging energy technologies, and power market reform. He was most recently head of data science and demand management at Rhythm, a retail clean energy provider, where he was responsible for the company's data, business intelligence, and analytics platform. Prior to that Ben was vice president, advanced grid services and analytics at EnergyHub, where he was responsible for analytics within the company's distributed energy resource management system (DERMS) and energy market operations. At EnergyHub, Ben managed the forecasting and optimization of some of the largest virtual power plants in the world, comprising electric vehicle (EV) chargers, stationary batteries, grid-interactive water heaters, and connected thermostats. Ben came to EnergyHub from ThinkEco, where he was vice president, technology, and prior to that was a member of the quantitative strategies group at Credit Suisse.

Ben is an expert on the intersection between advanced energy technology and energy markets, contributing to reports on the digitalization and decarbonization of the power system published by the Atlantic Council and the Council on Foreign Relations. He serves on the external Advisory Committee of the Alfred P. Sloan Foundation's Energy and Environment Program and has contributed to *Axios* and Utility Dive.

Ben holds a bachelor of arts in computer science from Northwestern, a master of science in mathematics from the Courant Institute at New York University, and a PhD in mathematics from University of California, Los Angeles.

The Atlantic Council is a nonpartisan organization that promotes constructive US leadership and engagement in international affairs based on the central role of the Atlantic community in meeting today's global challenges.

Atlantic Council
1030 15th Street, NW, 12th Floor,
Washington, DC 20005
(202) 463-7226, www.AtlanticCouncil.org