www.energyinnovation.org
98 Battery Street, Suite 202
San Francisco, CA 94111
power@energyinnovation.org

WHOLESALE ELECTRICITY MARKET DESIGN FOR RAPID DECARBONIZATION: **VISIONS FOR THE FUTURE**

BY SONIA AGGARWAL AND ROBBIE ORVIS ● JUNE 2019

Competitive wholesale electricity markets are at a turning point. Current market rules and practices were established to manage a system built around large central plant stations generating electricity to meet inflexible demand. Prices and market revenues are tied to generators' production costs, which have historically been largely dependent on the prices of fuels burned in those plants.

TABLE OF CONTENTS

Today's resource mix is changing. Carbon free resources with near-zero production costs such as wind, solar, and energy storage are replacing fuel-burning power plants due to falling prices and government policy. This trend is certain to continue.

These resources differ in several important ways from the fuel-burning power plants around which wholesale electricity markets were originally designed. First, they have near-zero production costs as they don't require any fuel. Second, these new resources are smaller, and can therefore be deployed more rapidly and in smaller increments. Third, they have very different production characteristics than thermal resources, with output tied to the availability of their energy resource, i.e. wind or sunshine. These differences have significant implications for how markets run and how prices and revenue can support a least-cost electricity mix.

The evolving mix of energy resources on the grid and decarbonization trend leads to the following question: "What wholesale market design would provide the best framework for reliably integrating the new, clean resources needed to decarbonize the power system at least cost?"

Thinking on this subject has generally fallen along two pathways. The "Robust Spot Market" suggests tightening up and extending today's markets for energy and services, eliminating capacity markets, and extending today's practice of voluntary decentralized bilateral contracting.

The "Long-Term Plus Short-Term Markets" pathway envisions complementing those more robust energy and services markets with an advanced, centralized, forward market to support needed resources and services.

This paper, the first in a series of three, outlines underlying questions emerging about wholesale market reform and introduces the two following papers describing alternative pathways for markets to evolve. Each market concept is discussed more fully in "[A Decentralized Markets Approach]" and "[Long-Term Markets working with Short-Term Energy Markets]."

THE EVOLUTION OF WHOLESALE ELECTRICITY MARKETS

Wholesale electricity markets evolved in the late 20th century in response to changing conditions on the grid. In particular, these markets were formed during a period of rising rates as the costs for new generation grew while demand growth slowed, as a way to incent more efficient investment in and operation of the electricity system.[1] Because existing generation planning practices resulted in high costs for consumers in certain regions, one of the primary goals of wholesale electricity markets was to lower overall system costs for customers by replacing utility-led energy procurement with private investment based on market principles.

Competitive wholesale electricity markets were originally designed to incorporate some demand-side flexibility in addition to their primary function of managing supply-side generation. However, regulatory, economic, and technological barriers largely hamper the ability of demand resources to participate in early wholesale electricity markets. One of the biggest barriers to participation of demand resources arises from the split in decision-making authority between the Federal Energy Regulatory Commission (FERC, which governs wholesale rates) and states (which govern retail rates). This split authority presents a problem because flexible demand typically needs to arise from the retail level to be incorporated into the wholesale market, but state regulators, keeping the interests of individual consumers top of mind, have typically designed retail rates to be as simple as possible while sharing system costs fairly among users.

Meanwhile, FERC has focused on creating non-discriminatory rates that create an even playing field for sophisticated electricity market participants, no matter how complex and specific those rates get. This disconnect, between simple rates that wash over differences between resources at the retail level, and more complex rates designed to elicit differences between resources at the wholesale level, has made it difficult for wholesale price signals to penetrate the retail system and reach the demand resources that might want to participate in the energy market.

With limited participation from demand resources, the wholesale market has functioned primarily with grid operators dispatching large central station plants to meet unalterable demand. In other words, wholesale electricity market operators have not considered demand a dispatchable resource like supply.

[1] For more information, see: Aggarwal and Orvis (2015), *Distribution System Optimization: Ready for Takeoff* in Public Utilities Fortnightly, available online at: https://www.fortnightly.com/fortnightly/2015/06/distribution-optimization-ready-takeoff

Within this paradigm, production costs[2] (e.g. fuel, and variable operation and maintenance costs) have determined which power plants are dispatched – those with the cheapest production costs are typically dispatched first, with market operators adding power plants with higher and higher production costs until demand is met. The last, most expensive power plant sets the market clearing price at each specific location, so these production costs have also been the principle factor supporting market prices and power plant revenues.

In this model, called "marginal cost dispatch," power plants can be thought of in three categories, based on their ratio of fixed to production costs. Plants with relatively high fixed costs and low production costs have typically been dispatched first, and therefore most often. Plants in a second category with lower fixed costs and higher production costs have been dispatched only when that "unalterable" demand increases enough to justify paying to run them. Finally, plants with comparatively low fixed costs and very high production costs have been used primarily only during a small number of hours each year – the most extreme demand peaks and system emergencies. Of course, not all plants fall neatly into these categories, but this framework is helpful for thinking about how markets operate.

Today's markets are the product of this marginal cost dispatch paradigm. Marginal cost dispatch has been a good way to introduce competition into a growing electricity system composed of a mix of large baseload power plants with high fixed costs and low production costs (e.g., coal and nuclear) and some more flexible power plants with lower fixed costs and higher production costs (e.g., natural gas). The rules, tradeable products, rates, and software used in wholesale electricity markets were designed around these specific resource profiles and the idea of dispatching large central stations to meet variable, unalterable electricity demand.

WHERE WE ARE TODAY

Today's electricity mix is evolving in ways that depart significantly from the past system in which competitive wholesale power markets were born and built. As prices have fallen and policies have pushed wind, solar, and storage onto the grid, their share of the resource mix has grown. These resources differ in several important ways from the fuel-burning electricity resources of the past.

First, these new resources typically have near-zero production costs. Their costs are almost entirely paid up front, and they are very cheap to run once built. Because dispatch and market clearing prices have typically been tied to production costs, this trait is a significant deviation from the fuel-reliant power plants that have made up most of the electric system in the past.

[2] "Production costs" and "marginal costs" are often used interchangeably to describe real-time generator costs. However, in some instances, "marginal costs" can also refer to the cost of the last "marginal" generator that sets the overall wholesale market clearing price. For clarity, we use the term "production costs" here to describe real-time generator costs. In the rest of this series and other literature, however, note that "marginal costs" may also refer to real-time generator costs.

Second, newer resources tend to have smaller minimum unit sizes—on the order of tens of megawatts (MW) rather than hundreds or thousands. As a result, these resources can be deployed more quickly[3] and in smaller unit sizes. Even if each individual wind power plant is less predictable than each traditional dispatchable coal plant, a fleet of wind power plants might actually be more reliable than the single dispatchable coal plant. This is because probabilistically, ten uncorrelated units that are 100 MW in size are more reliable than a single 1,000 MW unit that could trip off all at once.

Third, these resources have different production characteristics than many existing ones and are already changing how grid operators manage the grid. For example, solar output predictably follows the daily cycle of the sun, requiring grid operators to make other resources available in the evening when solar output drops to zero. Planning and running the grid around resource availability is not a new concept for grid operators – they have always had to plan for nuclear refueling outages, for example – but doing so for a large set of resources on a daily basis is pushing operators to consider new rules and tradeable market products.[4] At the same time, newer resources can provide certain services better or more cheaply than the older ones – consider power electronics inside inverters creating (very) fast "frequency response" (an essential grid service), which can offset the need for some "system inertia" (another essential grid service that has historically been provided automatically, as a product of the spinning mass inside thermal power plants).[5]

In addition to new wind, solar, and storage, the technological barriers that limited demand-side flexibility are rapidly disappearing. Smart thermostats, water heaters, and the "Internet of things" can turn electricity demand into a resource for grid and market operators.

Serious technological changes are hitting the electricity grid, but the concomitant changes in market incentives and rules are lagging behind. As it stands today, electricity demand *can* be increasingly flexible, but precious little has been done to access that flexibility. As new technologies come online at an ever-increasing pace, it's worth taking a closer look to see whether existing wholesale market structures are equipped to handle today's technology.

---

[3] For example, Tesla installed a 100 MW battery, the world's largest, in under 100 days following blackouts in Australia. See: https://www.greentechmedia.com/articles/read/tesla-fulfills-australia-battery-bet-whats-that-mean-industry#gs.1t4hvf

[4] For more information, see: https://energyinnovation.org/wp-content/uploads/2017/10/A-Roadmap-For-Finding-Flexibility-In-Wholesale-Power-Markets.pdf

[5] If the power system had originally been designed around inverter-based resources with advanced power electronics, the grid might not depend on system inertia at all, as conventional frequency response and system inertia derive from the inherent characteristics of conventional generators with spinning mass inside them. Similarly, the fact that inverter-based resources don't currently automatically provide system inertia does not mean they are incapable of it; their growing presence has exposed the value of system inertia given our current grid design, and technology will respond to the need, provided the right price signals or standards are in place. New resources change the landscape of what grid services are required, and expand possibilities for which resources can provide them.

WHERE WE ARE HEADED

Whether because of dramatically lower prices[6] or government policy[7], the electricity system will continue to decarbonize through the addition of new zero carbon resources.

Many of the constraints on today's system arose from the needs of fuel-burning resources that dominated the system when markets were first created: for example, unit commitment and minimum run rates. But these constraints are increasingly at odds with new resources entering the system. For example, evidence is growing that commitment windows and minimum run rates are routinely overstated by fuel-burning generators, limiting grid flexibility – a service becoming ever more crucial as the electricity mix decarbonizes. In other words, wholesale electricity markets will need to modernize to support decarbonized electricity systems.

In addition to the adjustments needed to ensure well-functioning wholesale electricity markets in the context of a changing mix of energy resources, many market observers are asking if more fundamental changes to wholesale market design may be needed.

Before answering the question of whether basic fixes can do the trick or whether more fundamental changes are needed, the first task is to get clear on what we need modern wholesale electricity markets to do.

TEN PRINCIPLES FOR MODERN WHOLESALE ELECTRICITY MARKETS

The ten principles below are intended to ensure technology neutrality and achievement of power system goals at least cost, and are repeated in slightly different form in each of the two following papers in this series. Wholesale electricity markets should:

1) Accommodate rapid decarbonization, including eliminating barriers to participation of zero carbon resources.
2) Support grid reliability, so the incremental costs of reliability do not exceed the amount customers would knowingly be willing to pay for, or do not exceed incremental benefits.
3) Promote short-run efficiency through optimized dispatch of the lowest-cost resource mix, and using existing and emerging technologies to manage reliability and congestion.
4) Facilitate demand-side participation and grid flexibility.
5) Promote long-run efficiency – including efficient, competitive entry to *and* exit from the market – under conditions of significant uncertainty.
6) Minimize the exercise of market power and manipulation.

---

[6] In much of the US, renewable energy prices – which are expected to continue dropping – are so low that it is cheaper to build new renewable plants than continue operating existing thermal plants. See; https://www.forbes.com/sites/energyinnovation/2018/12/03/plunging-prices-mean-building-new-renewable-energy-is-cheaper-than-running-existing-coal/#5c78ed1a31f3

[7] For example, California; Hawaii, New Mexico, and Washington, D.C. have all passed legislation mandating a 100% clean electricity mix by 2050 at the latest. Several other states, including Arizona, Minnesota, Illinois, and New York are considering similar standards as well.

7) Minimize the potential for distortions and interventions that would prevent or limit markets' ability to achieve efficient outcomes, consistent with the public interest (including overarching public interest in a sustainable environment and economy).
8) Enable adequate financing of resources needed to deliver cost-effective reliability, based on an efficient allocation of risk (i.e., those that can best mitigate risk should bear it) that prevents customers from bearing the cost of poor investment decisions made by private investors.
9) Be capable of integrating new technology as electricity needs evolve, and adapting as technology changes.
10) Have designs that are readily and realistically implementable.

EMERGING PATHWAYS FOR FUTURE WHOLESALE ELECTRICITY MARKETS

In addition to fulfilling the principles laid out above, future wholesale electricity market solutions must address the following questions, which are central to whether wholesale electricity market solutions can support wide-scale deployment of low cost, low carbon resources:

- Today's market prices are derived from generators' production costs to generate electricity, which for fuel-burning power plants are clearly tied to fuel prices. Renewables and storage typically have no fuel cost. When zero production cost resources form the majority of resources on the system, grid operators will still need to know in real-time which set of resources to dispatch.
  - **How would markets efficiently form real-time prices in a system with large quantities of energy resources with near-zero production cost?**
- A changing set of resources may precipitate changes to the value of products traded in the wholesale markets, altering resources' revenue streams.
  - **How can sufficient investment signals be maintained, and how are new resources efficiently financed as the resource mix evolves?**
- Today's markets are largely oversupplied[8], which mutes grid flexibility price signals. Grid flexibility will become increasingly important as fuel-burning power plants exit the system and are replaced by renewables:
  - **How will markets expose the value of important system characteristics, such as flexibility, through this transition in the energy resource mix?**
- Today's markets struggle to develop and finance transmission, storage, and other resources that support the efficient functioning of the grid.
  - **How are transmission lines; energy storage; and local, non-transmission alternatives efficiently financed and deployed in a future market?**
- Finally, a future market design must be capable of integrating or accommodating policies focused on reducing carbon from the electricity system.

---

[8] See also "Power Markets in Transition: Consequences of Oversupply and Options for Market Operators" in *Current Sustainable/Renewable Energy Reports*, April 2019.

- How is carbon policy addressed in a future market?

These are tough but important questions that don't have a single "right" answer. Reasonable people can—and do—arrive at different answers.

The market proposals most commonly discussed in intellectual circles tend to fall along two pathways, each of which seeks to satisfy the principles outlined above and answer the questions posed here. The first pathway emphasizes improving today's markets for energy and services, eschewing capacity markets, and extending today's practice of voluntary de-centralized bilateral contracting – as described in *A Decentralized Markets Approach*. The second pathway envisions complementing those more robust energy and services markets with an advanced, centralized, forward market to procure needed resources and services as described in *Long-Term Markets working with Short-Term Energy Markets*.

Both pathways agree on important features for modern markets:

- Competitive wholesale electricity markets are a good thing: Trading over a diverse portfolio of resources augments reliability and decreases overall costs, and the larger the market, the greater the benefits.
- Wholesale electricity markets need to work with external (state or federal) policies governing the electricity system, not work against (i.e., mitigate) them.
- Shorter dispatch intervals and multi-period optimization can make markets more efficient.
- The capacity markets in use around the U.S. today, which largely trade capacity without much regard to the operational characteristics of the energy resources being traded, should be fundamentally transformed or eliminated.

At the same time, important differences exist between the two pathways, driven in part by the authors' views on the following questions:

- How big of a risk is political interference in markets?
- How much do we expect the "real world" to behave as theory suggests?
- How strong are the counterparties in markets, and how strong do we expect them to be in the future; i.e., can we expect that utilities or other load-serving entities will be able to buy smart energy resource portfolios, flexible and well-hedged, to serve customers over the long-term?
- What extent can factors other than strict production costs set locational marginal prices; i.e., congestion in the transmission system, ancillary service needs, other opportunity costs? If those other factors do play a substantial role setting locational marginal prices, what is the risk that real-world prices (which may be in-part driven by lumpy retirements) are too low to attract needed flexibility resources or too high to expose their value?
- Is keeping voluntary bilateral markets (which already underlie centralized wholesale electricity markets) decentralized the best approach, or would centralizing and organizing those bilateral contracts be more beneficial?

There is no “right” answer to these questions. Wholesale electricity markets will evolve differently in various regions, but the issues raised in this paper series are extremely important for grid managers to study and deliberately consider as the electricity system decarbonizes.