Long-Term Resource Adequacy in High Intermittent Renewables Wholesale Electricity Markets: Lessons from California

by
Frank A. Wolak
Department of Economics
Stanford University
Stanford, CA 94305-6072

Transitioning a wholesale electricity market from one based on dispatchable fossil fuel generation resources to one based intermittent renewable generation resources is likely to require a change in a region's long-term resource adequacy mechanism. The traditional capacity-based approach to ensuring sufficient resources to meet demand during all hours of the year is poorly suited to an industry dominated by intermittent wind and solar generation resources. This is because having total installed generation capacity significantly larger than the annual system demand peak no longer guarantees that there are adequate generation resources to meet demand during all hours of the year. With intermittent renewable generation units there can be sustained periods of low levels of intermittent renewable output, as discussed in Tangerås and Wolak (2019) for the case of California. Unless there are enough dispatchable generation units to produce electricity when there are low levels of renewable energy production, it will be necessary to curtail demand involuntarily (i.e., declare rolling blackouts) to maintain system balance.

This paper will use the experience of California during the Summer of 2020 and the rolling blackouts of August 2020 (summarized in the attached slides) to demonstrate the shortcomings of a capacity-based long-term resource adequacy mechanism in an electricity supply industry with a significant share of wind and solar resources. An alternative standardized long-term energy contracting approach to the long-term resource adequacy that addresses these shortcomings will be introduced and its properties discussed. This long-term resource mechanism is outlined in Wolak (2020) which is attached to this document. It is currently under consideration at the California Public Utilities Commission as part of the re-design of the California's long-term resource adequacy mechanism and has been presented to stakeholder groups at a number of markets in the United States and around the world.

Capacity-based long-term resource adequacy mechanisms currently exist in all United States wholesale electricity markets and many countries around the world. The PJM Interconnection, ISO New England, the New York ISO, and Mid-Continent ISO (MISO) all have capacity-based long-term resource adequacy mechanism, as do virtually all the wholesale electricity markets in Europe. As these regions scale up the share of intermittent wind and solar generation resources in their regions many of the shortcomings identified in California are likely to arise, which underscores to adequate this shortcoming in the design of long-term resource adequacy mechanisms.

The Electricity Reliability Council of Texas (ERCOT) is the only wholesale electricity market in the United States that has a comparable share of intermittent renewable energy to that in California. It does not have a capacity-based long-term resource adequacy mechanism. It has a long-term resource adequacy mechanism that shares many features with the standardized long-

term energy contract approach to long-term resource adequacy in Wolak (2020). This paper will also discuss these commonalities with the ERCOT approach to long-term resource adequacy and describe why the ERCOT approach is unlikely to be a viable option for other parts of the United States at the present time.

The remainder of this document describes the rationale for a long-term resource adequacy mechanism in wholesale electricity markets with offer caps. The next section describes the basic features of a capacity market and of its shortcomings for regions with significant intermittent generation resources.

Why A Long-Term Resource Adequacy Mechanism for Electricity?

A first question one might ask is why electricity markets require regulatory intervention to ensure reliability? Consumers want to be able to withdraw electricity from the network when they need it, just like other goods and services. But it is unclear why electricity is so fundamentally different from other products that it requires paying suppliers for production capacity to exist. For example, consumers want cars, but they do not pay for automobile assembly plants. They want point-to-point air travel, but they do not pay for airplanes. They want a loaf of bread, but do not pay for the existence of a bakery. All of these industries are high fixed cost, relatively low marginal cost production processes, similar to electricity supply. Nevertheless, all of these firms earn their return on capital invested by selling the good that consumers want at a price above the variable cost of producing it. Clearly cars, air travel, and bread are in many way essential commodities, yet there is no regulatory invention that ensures that there is sufficient production capacity for these products to meet demand.

So what is different about electricity markets that necessitates the need for a long-term resource adequacy mechanism? The answer lies in how short-term markets for these products operate relative to that for wholesale electricity. This difference is the result of the regulatory history of the electricity supply industry and the technology of electricity metering. The limitation on the level of short-term prices and the way that supply shortfalls are dealt with in wholesale electricity markets creates what Wolak (2013) has been termed a "reliability externality" that requires a regulatory intervention to internalize.

In the market for automobiles, air travel, and even bread, there is no explicit prohibition on the short-term price of the good rising to the level necessary to clear the market. Take the example of air travel. Airlines adjust the prices for seats on a flight over time in an attempt to ensure that the number of customers traveling on that flight equals the number of seats flying. This can result in very different prices for a seat on the same flight, depending on when the customer purchases the seat. A customer that waits too long to purchase a seat faces the risk of an infinite price in the sense that all of the seats on the flight are sold out. This ability to use price to allocate the available seats is also what allows the airline the flexibility to recover its total production costs. Airlines can set low prices to fill flights with low demand and extremely high prices on other flights, or at other times for the same flight, when demand is high.

The ability to use the short-term price to manage the supply and demand balance in the electricity supply industry is limited first by the fact that all North American wholesale electricity

markets have offer caps that limit a supplier's offer price into the wholesale market and/or a price cap that limits the magnitude of the eventual market-clearing price. In addition, historically virtually all electricity supply industries did not have interval meters that can record a customer's 15-minute or hourly consumption throughout the month. Even today, most regions only have mechanical meters that compute the customer's consumption for the entire month as the difference between two consecutive meter readings. With monthly or bi-monthly reading of mechanical meters, it is impossible for the utility to know how much electricity a customer consumed within a given hour of the month.

Although these offer caps and price caps can limit the ability of suppliers to exercise unilateral market power in the short-term energy market, they also reduce the revenues suppliers can receive during scarcity conditions. This is often referred to as the *missing money* problem for generation unit owners (see for example, Joskow (2013)). However, this missing money problem is a symptom of the existence of the *reliability externality*.

This externality exists because offer caps limit the potential downside to electricity retailers and large consumers (able to purchase from the short-term market) delaying their purchases of electricity until real-time operation. Specifically, if the retailer or large consumer knows the price cap on the short-term market is $250/MWh, then it is unlikely to be willing to pay more than that for electricity in any earlier forward market. This creates the possibility that real-time system conditions can occur where the amount of electricity demanded at or below the offer cap is less than the amount suppliers are willing to offer at or below the offer cap. This outcome implies that the system operator must be forced to either abandon the market mechanism or curtail load until the available supply offered at or below the offer cap equals the reduced level of demand, as was the case a number of times during the period January 2001 and March 2001 in California.

Because random curtailments of supply---also known as rolling blackouts---are used to make demand equal to the available supply at or below the bid cap under these system conditions, this mechanism creates a reliability externality because no retailer or large consumer bears the full cost of failing to procure adequate amounts of energy in advance of delivery. A retailer that has purchased sufficient supply in the forward market to meet its real-time energy demand is equally likely to be randomly curtailed as the same size retailer that has not procured adequate amounts of energy in the forward market. For this reason, all retailers and large loads have an incentive to under procure their expected energy needs in the forward market.

Particularly for markets with very low offer caps, retailers have little incentive to engage in sufficient fixed-price forward contracts with generation unit owners to ensure a reliable supply of electricity for all possible realizations of real-time demand. Although this incentive for retailers to rely on a price-capped short-term market is most likely to impact generation units that run infrequently, if the level of demand relative to the amount of available supply is sufficiently large, it can even impact intermediate and baseload units. Because of the expectation of very low prices in the short-term market and the limited prospect of very high prices because of offer caps, retailers may decide not to sign fixed-price forward contracts with these generation unit owners and purchase their energy in the short-term market. By this logic, a short-term energy market with an offer cap always creates an incentive for retailers to delay purchasing some of their energy needs

until real-time, when these caps can be used to obtain this energy at a lower price than the supplier would be willing to sell it in the forward market.

The lower the offer cap, the greater is the likelihood that the retailer will delay their electricity purchases to the short-term market. Delaying more purchases to the short-term market increases the likelihood of the event that insufficient supply will offer into the short-term market at or below the offer cap to meet demand. If a retailer knows that part of the cost of its failure to purchase sufficient fixed-price forward contracts will be borne by other retailers and large consumers because of random curtailment, then it has an incentive to engage in less fixed-price forward contracts than it would in a world where all customers had hourly meters and all customers could be charged hourly prices high enough to cause them to reduce their demand to equal the amount of supply available at that price.

The ERCOT market addresses this problem by having an offer cap on its short-term market that is $9,000/MWh, which is 9 times the value of the offer cap on the California market. There are variety of reasons that ERCOT is able maintain this level of the offer, whereas other markets in the United States are not. The major reason is that ERCOT is not interconnected with the rest of the United States electricity transmission network, so that both its wholesale and retail market are regulated by the Public Utilities Commission of Texas. All other wholesale markets in the United States are subject to federal oversight by the Federal Energy Regulatory Commission and their retail markets are regulated by their respective state Public Utilities Commission.

The above discussion implies that unless the regulator is willing to eliminate or substantially increase the offer cap on the short-term market so that the short-term price can be used to equate available supply to demand under all possible future short-term market conditions, some form of regulatory intervention is necessary to internalize the resulting *reliability externality*. However, if customers do not have interval meters that can record their consumption on an hourly basis, they have a limited ability to benefit from shifting their consumption away from high-priced hours, so raising or having no offer cap on the short-term market would not be advisable. Recall that only a customer's total monthly or bimonthly consumption, not the amount consumed in any hour of the month, can be measured with mechanical meters and monthly or bimonthly meter reading. This means that reducing consumption by one kilowatt-hour at any time in the monthly or bi-monthly billing cycle has the same impact on the customer's bill. Consequently, the customer is likely to reduce demand when it is easiest, rather than when the system is stressed, or the hourly price is highest.

As the above discussion makes clear, relying on a capped short-term energy market price to ensure long-term resource adequacy does not address the reliability externality and leaves both a "missing money" and "missing market" problem. Capacity payment mechanisms are one approach to addressing this reliability externality designed primarily for thermal generation-dominated markets, where the major concern is insufficient generation capacity to meet future demand peaks. In intermittent renewable dominated markets the major reliability concern is more likely to be insufficient renewable energy (i.e. wind and sunshine) to meet a sustained demand for energy, which implies other approaches to addressing the reliability externality may dominate a capacity-based approach for these electricity markets.

As the share of intermittent renewable generation in a wholesale electricity market increases, the magnitude of the reliability externality is also likely to increase. The uncertain availability of wind and solar resources increases magnitude and duration of potential future energy supply shortfalls that must be managed, which implies many more instances when a capped short-term energy market may not yield a sufficient energy supply increase or demand decrease to maintain real-time supply and demand balance.

Capacity Payment Mechanisms and Markets

A major rationale for capacity markets in the U.S. appears to be a holdover from the vertically-integrated regulated monopoly regime when capacity payments compensated generation units for their capital costs, because the regulatory process typically reimbursed unit owners for their variable operating costs through generation unit pooling arrangements.

All of the eastern U.S. wholesale markets with formal capacity markets---PJM, ISO-New England, and the New York ISO---started as power pools where groups of vertically-integrated utilities agreed to dispatch their generation units centrally in order to reduce the variable cost of serving their respective loads. The state-level regulatory process would compensate each utility for the variable cost of serving its load either from its own units or from the units owned by other utilities. The capacity costs of each utility's generation units would be recovered through a dollar per installed kilowatt capacity payment set by each utility's state regulator.

It is important to emphasize that in a wholesale market regime all generation unit owners have the opportunity to earn the market-clearing price which is typically above the generation unit's average variable cost when the unit is operating. In this way, the generation unit earns a return to capital during each hour it produces electricity. This paradigm for earning a return on capital from the difference between the market price and the firm's average variable cost of production provides the appropriate signal for investment in new productive capacity in all workably competitive industries. There is little reason to expect that it could not work in the wholesale electricity industry with an active demand side.

All U.S. regional capacity markets have the following characteristics. All retailers in the region are assigned firm capacity obligations equal to some multiple of their annual peak demand--typically in the range of 110 to 120 percent, depending on the region. The capacity procurement obligations for retailers often also entail sub-regional or local requirements.

All generation units in the region are then assigned a firm capacity quantity equal to the amount of energy the unit can produce under all possible system conditions. For the case of a thermal generation unit, this magnitude is typically equal to the nameplate capacity of the unit multiplied by the availability factor of the unit. For hydroelectric resources, this figure is typically equal to the amount of energy unit was able to produce under the historically worst hydrological conditions.

For solar and wind resources, determining a defensible firm capacity figure is extremely challenging. Specifically, if one takes seriously the view that the firm capacity is the amount of energy a resource can supply under all possible system conditions, then the firm capacity of a solar

or wind resource should be zero, because solar resources cannot produce when the sun does not shine and a wind resource cannot produce energy when the wind is not blowing. Consequently, the process of assigning firm capacity figures to solar and wind resources has a significant ad hoc component. To give an example from California, each 1 MW solar photovoltaic capacity was assigned a firm capacity 0.27 MW for August of 2020, even though these resources produced virtually no electricity during the late evening when rolling blackouts occurred in August 14 and 15.

References Cited

Joskow, Paul (2013). "Symposium on 'capacity markets'". In: Economics of Energy and Environmental Policy2.2, pp. V–VI.

Tangerås, Thomas, and Wolak, Frank A. (2019) "Locational Marginal Network Tariffs for Intermittent Renewable Generation," available at http://www.stanford.edu/~wolak

Wolak, Frank A. (2013). "Economic and Political Constraints on the Demand-Side of Electricity Industry Re-structuring Processes". Review of Economics and Institutions, 4.1, p. 42

Wolak, Frank A. (2020) "Market Design in a Zero Marginal Cost Intermittent Renewable Future," available at http://www.stanford.edu/~wolak.

Evidence from California on Challenges Facing Electricity Supply Industries with a Significant Share of Intermittent Renewables

Frank A. Wolak
Professor of Economics
Director, Program on Energy and Sustainable Development
Stanford University
wolak@zia.stanford.edu

1

Outline of Presentation

- How did California arrive at current situation?
  - 33% Renewables Portfolio Standard (RPS) by 2020 and $2.2 Billion for Rooftop Solar under California Solar Initiative (CSI)
  - Import dependence—25%-33% of annual consumption
  - Capacity-based long-term resource adequacy
  - No dynamic pricing of retail electricity
- What happened on August 14 and 15, 2020 and how prevent it from happening again

2

CA's Renewable Energy Goals

- California has almost 8,000 MW of distributed solar installed
  - Most has been installed since 2007 under California Solar Initiative (CSI)
  - CSI provided $2.167 billion to support distributed solar installations
    - CSI funded by electric ratepayers through higher retail prices
- California has 33% Renewables Portfolio Standard (RPS) by 2020 and 60% by 2030
  - Above-wholesale-market-price costs of qualified renewable energy included in retail price

3

CA's Renewable Energy Goals

- Between 2013, first of year of the 33% RPS compliance period, and 2019 California reduced
  - Natural gas fired-generation capacity by 8,500 MW
  - Nuclear generation capacity by 2,250 MW—San Onofre Nuclear Generation Station (SONGS) was retired
  - Total reduction of 10,750 MW in dispatchable capacity
- Dispatchable generation replaced with
  - 8,200 MW of solar photovoltaic generation capacity
  - 324 MW solar thermal generation capacity
  - 188 MW wind generation capacity
  - Total increase of 8,712 MW in intermittent capacity
- Nuclear capacity produced with ~90 percent capacity factor and natural gas capacity could produce with least at 75 percent capacity factor
- Solar and wind resources produce at ~20 and ~25 capacity factors, respectively

4

CA's Energy Supply

- California is relies on imports for 25 to 33 percent of consumption annually
  - Pacific Northwest supplies hydroelectric energy in early summer
  - Desert Southwest supplies coal and natural gas-fired generation
- California has a capacity-based long-term resource adequacy mechanism
  - Each generation resource is assigned a firm capacity value
    - Amount of energy generation unit can provide under stressed system conditions
  - All load-serving entities much purchase sufficient firm capacity to cover their peak demand plus a reserve requirement
- Firm capacity (FC) of natural gas or nuclear generation unit relatively straightforward to compute
  - FC = Annual Availability Factor x Nameplate Capacity of Unit

5

Firm Capacity of Intermittent Resources

- Firm capacity of hydroelectric resource typically based on historically lowest level of annual energy output
  - Past performance no guarantee of future performance
- Firm capacity of solar or wind resource extremely challenging to compute
  - If stressed system conditions occur when it is dark, firm capacity of solar generation unit should be zero
  - If stressed system conditions occur when wind is not blowing, firm capacity of wind generation unit should be zero
- Assignment of firm capacity to intermittent renewable resources has a significant political component
  - Recent study by three CA investor-owned utilities estimated effective load carrying capability (ELCC) of solar PV at ~5 percent of nameplate capacity
    - *2020 Joint IOU ELCC Study*, prepared by Astrape Consulting
- Conclusion: Firm capacity approach to long-term resource adequacy poorly suited to regions with high shares of intermittent renewable energy

6

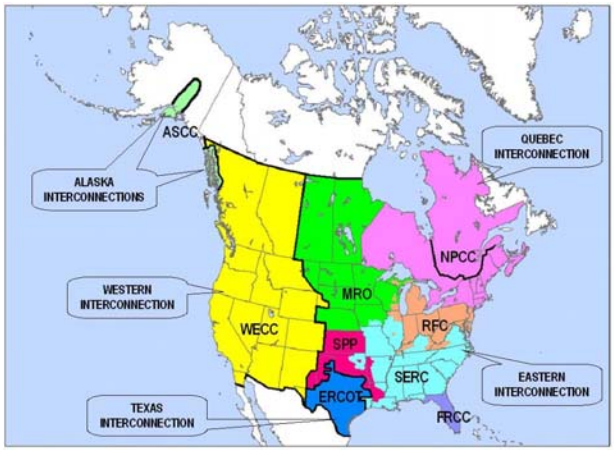
California's Import Dependence

- More than 18,000 MW of transfer capacity between California and neighboring states
  - Significant potential
- Neighboring states have priority access to electricity produced by generation units owned by utilities in their states
  - Implication: When temperatures in the western US are uniformly high, California may not receive sufficient imports without advance purchases of energy
- California is part of Western Electricity Coordinating Council (WECC) that comprises all states and Canadian provinces west of Continental Divide
- California's import dependence poorly suited to firm capacity-based long-term resource adequacy mechanism



7

North America's Interconnections



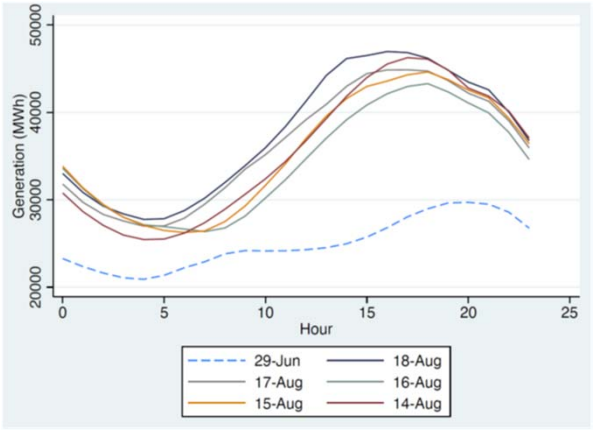
8

California's Retail Market Policies

- All customers of three large investor-owned utilities—Pacific Gas and Electric, Southern California Edison, and San Diego Gas and Electricity—have interval meters
  - Meter records customer's consumption on a 15-minute basis
- No dynamic retail pricing plans offered for residential customers
  - Dynamic prices vary with real-time system conditions in wholesale market
  - Time-of-use prices are NOT dynamic prices because customer is charged same price during *peak* and *off-peak* periods of day, regardless of real-time price of wholesale electricity
- In regions with increasing share of intermittent renewables, demand must shift across of the day maintain real-time supply and demand balance
  - Andersen, Hansen, Jensen, and Wolak (2019) "Can Incentives to Increase Electricity Use Reduce the Cost of Integrating Renewable Resources?" (on web-site)

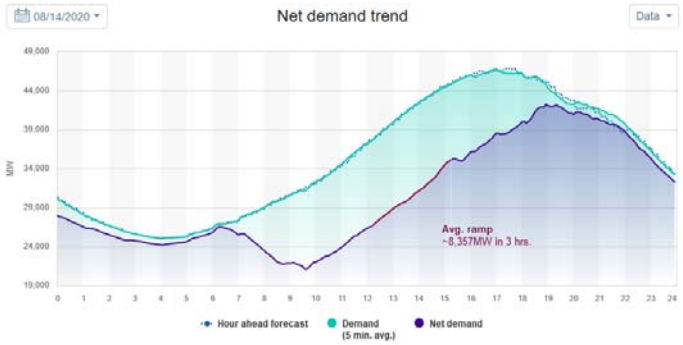
9

The Rolling Blackouts of 8/14/20-8/15/20 (Hourly Demand in MWh)



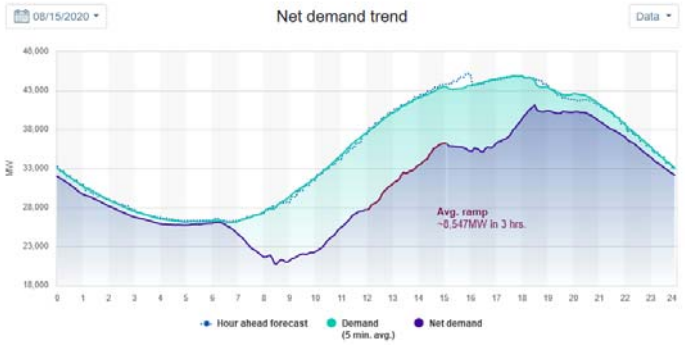
10

The Rolling Blackouts of 8/14/20-8/15/20



11

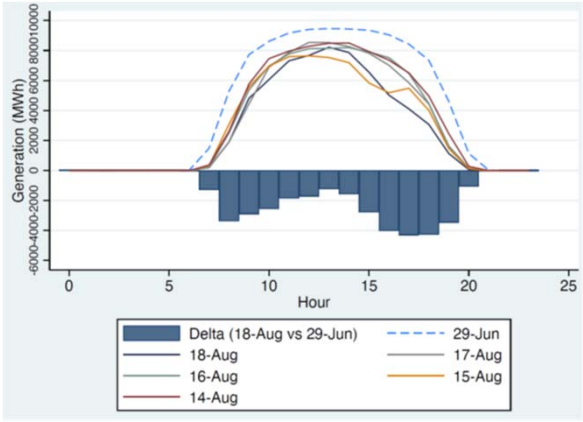
The Rolling Blackouts of 8/14/20-8/15/20



12

The Rolling Blackouts of 8/14/20-8/15/20

(Hourly Production of Grid-Scale Solar Energy)



13

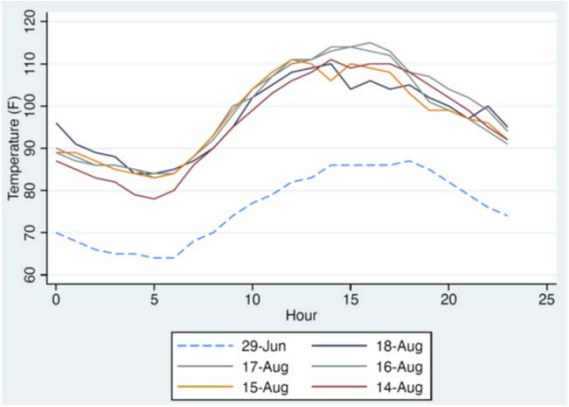
Solar Production in California

- June 29, 2020 is an ideal day for solar production in California
  - Panels have maximum efficiency for converting light into electricity at a temperature of 77º F
- Hot days with significant particulate matter in the air are not ideal for solar production
- What explains almost 20% reduction in solar production relative to ideal conditions on August 14 to 18, 2020?

14

The Rolling Blackouts of 8/14/20-8/15/20

(Hourly Temperature in Barstow, CA)



15

Solar Panels and Temperature

- PV panels are rated at 77º F temperature
  - Convert light into electricity
- Efficiency of panels declines linearly with every degree of temperature above 77º F
- On-site electricity consumption on high temperature days likely to be greater than on lower temperature days
  - Air conditioning load
- Both factors lead to lower net injections to grid from solar PV units
  - Explains less net production from solar units on August 14-18 versus June 29

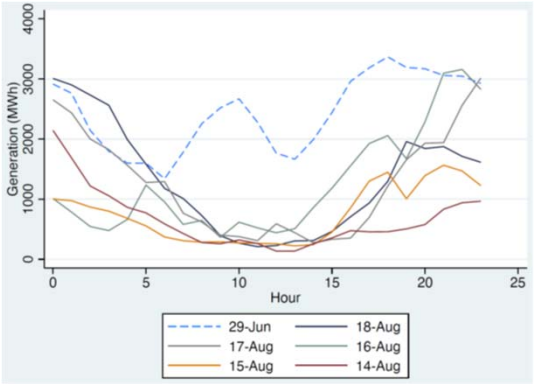
16

Wind Production and Temperature

- Wind production on extremely hot days unlikely to be very high
  - Wind occurs because of temperature differentials between locations
  - If it is hot everywhere, there is likely to be very little wind
  - Higher wind production on lower temperature days
- Wind production likely to be greatest at beginning and end of the daylight hours

17

Wind Production and Temperature (Hourly Wind Output)



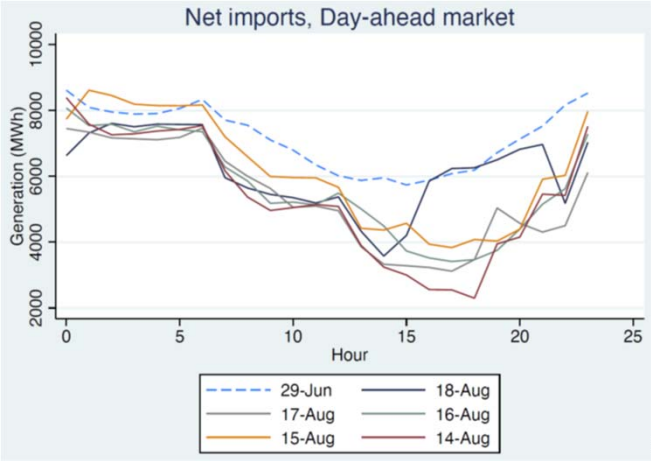
18

Imports and Temperature

- Recall that neighboring control areas have priority for output of generation units in their state
- California load-serving entities can purchase this energy in a fixed-price forward contract to ensure that it is supplied to California
- California can also purchase energy in real-time market
  - Only if price California is willing to pay is higher than price other control areas are willing to pay
  - Some prices outside of California were higher than offer cap on California ISO's real-time market during August 14 and 15
- Important lesson—Offer caps on California market can reduce real-time supply to state during stressed system conditions

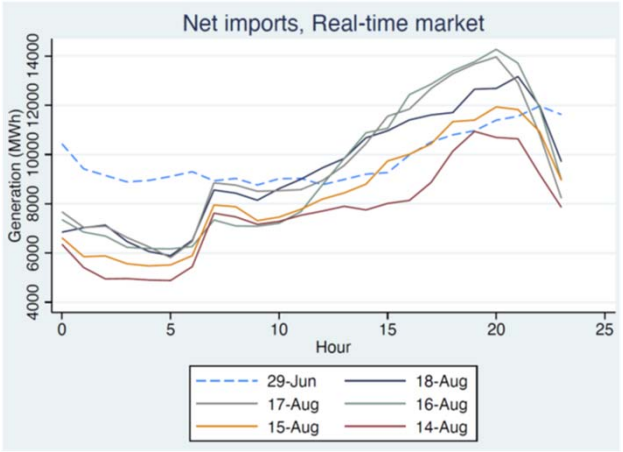
19

Imports and Temperature



20

Imports and Temperature



21

Imports and Temperature

- Difference between August 14 and 15 and August 16 to 18 is that California was able to obtain more imports in real-time market
- A substantial amount of generation capacity exists in the WECC
  - Owners of these units need a financial incentive to turn units on and sell energy to California
  - Events of August 14 and 15 demonstrated California was willing to pay high price for needed energy
- September 5 and 6 heat wave in WECC led to real-time prices during late evening close to $1,000/MWh
  - Annual average wholesale prices in 2019 was slightly less than $40/MWh

22

Conclusions from Analysis

- Replacing dispatchable generation capacity with intermittent generation capacity is very risky
- California has two options to meet real-time demand with less solar and wind energy
  - Increase imports, which can be difficult if entire WECC region is hot
  - Reduce demand, which is difficult because of no customers pay according to dynamic prices
- California needs **energy-based** long-term resource adequacy mechanism
  - Buy necessary energy in forward market to ensure it is committed to California market
    - Buying energy on spot market on hot days is likely flying stand-by on Thanksgiving
  - Wolak (2020) “Market Design in a Zero Marginal Cost Intermittent Renewable Future,” (on web-site)

23

Questions/Comments

For more information

http://www.stanford.edu/~wolak

24

Market Design in a Zero Marginal Cost Intermittent Renewable Future

by

Frank A. Wolak
Director, Program on Energy and Sustainable Development (PESD)
Holbrook Working Professor of Commodity Price Studies
Department of Economics
Stanford University
Stanford, CA 94305-6072
wolak@zia.stanford.edu

Current Draft: October 15, 2020

The basic features of an efficient short-term wholesale market design do not necessarily need to change to accommodate a significantly larger share of zero marginal cost intermittent renewable energy from wind and solar resources. A large share of controllable zero marginal cost generation does not create any additional market design challenge relative to a market with a large share of controllable positive marginal cost generation. Regardless of the technology, generation unit owners must recover their fixed costs from sales of energy, ancillary services, and long-term resource adequacy products.

A larger variance in the hourly amount of energy produced by intermittent resources is the primary market design challenge associated with a zero marginal cost renewable future. The past ten years in California have demonstrated that as the amount of wind and solar generation capacity increases, the variance in hourly energy produced by these resources increases. This increase in supply uncertainty also increases short-term price volatility, which can finance investments in storage and other technologies that allow consumers to shift their withdrawals of grid-supplied energy away from periods when little wind and solar energy is being produced.

An increased risk of large intermittent energy shortfalls and short-term price volatility implies a greater need for risk management activities. Greater short-term intermittent energy supply risk is likely to require accounting for more transmission and generation operating constraints in the day-ahead and real-time energy markets, as well as purchasing more operating reserves and creating additional ancillary service products. Because controllable generation units are likely to have to start and stop more frequently to make up for unexpected renewable energy shortfalls, there will be a greater need to develop short-term pricing approaches that recover the associated start-up and minimum load costs.

The potential for sustained periods of low intermittent energy production creates both a medium and long-term energy supply risk that requires a new long-term resource adequacy mechanism. The traditional capacity-based approach is unlikely to be the least cost mechanism for ensuring that future demand for energy is met. In a zero marginal cost intermittent future, wind and solar resources must hedge their energy supply risk with controllable generation resources in order to maintain long-term resource adequacy. Cross hedging between these technologies accomplishes two goals. First, it can provide the revenue stream necessary for fixed cost recovery by controllable generation units. Second, it ensures that there is sufficient controllable generation capacity to meet demand under all foreseeable future system states, with a high degree of confidence.

The remainder of this article first describes the key features of an efficient short-term wholesale market design: a multi-settlement locational marginal pricing (LMP) market with an automatic local market power mitigation (LMPM) mechanism, which is the standard market design for all short-term markets in the United States (US). This section concludes with a discussion of modifications of this basic design likely to be necessary to accommodate a larger share of intermittent renewables.

The second half of the paper describes a new long-term resource adequacy mechanism for the efficient short-term market design for an electricity supply industry with a large share of zero marginal cost intermittent renewables. I first explain why a wholesale electricity market requires a long-term resource adequacy mechanism. I then describe a mandated standardized long-term contract approach to long-term resource adequacy that provides strong incentives for intermittent renewable resource owners to hedge their energy supply risk with controllable generation resource owners. I argue that this mechanism ensures long-term resource adequacy at a reasonable cost for final consumers while also allowing the short-term wholesale price volatility that can finance investments in storage and other load-shifting technologies necessary to manage a large share of intermittent renewable resources.

1. Short-Term Market Design

More than twenty-five years of international experience with wholesale electricity market design has identified four crucial features of an efficient short-term market design. First is the extent to which the market mechanism used to set dispatch levels and locational prices is consistent with how the grid and generation units actually operate. Second is a financially binding day-ahead market that prices all transmission and generation unit operating constraints expected to be relevant in real-time. Third is an automatic LMPM mechanism that limits the ability of a supplier to influence the price it receives when it possesses a substantial ability to exercise market power. Fourth are retail market policies that foster active participation of final demand in the wholesale market.

The early US wholesale market designs in the PJM Interconnection, ISO New England, California, and Texas employed simplified versions of the transmission network configuration and generation unit operating constraints. Similar market designs currently exist throughout Europe and the rest of the world. They set a single market-clearing price for an hour or half-hour for an entire control area or for large geographic regions, despite the fact that in real-time there are often generation units with offer prices below this market-clearing price not producing electricity and

units with offer prices above this market-clearing price producing electricity. This outcome occurs because of the location of demand and available generation units within the region, and the configuration of the transmission network prevents some of these low-offer-price units from producing electricity and requires some of the high-offer-price units to supply electricity.

This approach to short-term market design provides incentives for suppliers to take actions to exploit the fact that " in real-time physics wins," rather than offer their resources into the day-ahead market in a manner that minimizes the cost of meeting demand at all locations in the grid in real-time. Instead, suppliers take actions in the simplified day-ahead market that allow them to profit from knowing they will be needed (or not needed) in real-time because of transmission and generation unit operating constraints.

1.1. Locational Marginal Pricing

Starting with PJM in 1998 and ending with Texas in late 2010, all US wholesale markets adopted a multi-settlement locational marginal pricing (LMP) market design that co-optimizes the procurement of energy and ancillary services and includes an automatic local market power mitigation mechanism built into the market software. This design has a day-ahead financial market which satisfies the locational demands for energy and each ancillary service simultaneously for all 24 hours of the following day. A real-time market then operates using the same network model as the day-ahead market adjusted to real-time system conditions. Deviations from purchases and sales in the day-ahead market are cleared using these real-time prices. Both of these markets price all relevant transmission network and other relevant operating constraints on generation units. As I discuss below, this market design can foster active participation of final demand in the wholesale market.

Only generation unit output levels that are physically feasible will be accepted in both the day-ahead and real-time markets. Prices for the same hour vary depending on whether the location is in a generation-deficient or generation-rich region of the transmission network. The locational marginal price or nodal price at a given location is the increase in the minimized value of the "as-offered costs" of serving the locational demands for energy and all ancillary services as a result of a one-unit increase in the amount of energy withdrawn at that location in the transmission network. The price of each ancillary service is equal to the increase in the optimized value of the objective function as a result of a one-unit increase in the demand for that ancillary service.

The recent experience of many European countries with significant wind and solar resources indicates that the cost of making the final schedules that emerge from their zonal markets physically feasible is likely to get even larger as the amount of intermittent renewable generation capacity increases. According to the European Network of Transmission System Operators for Electricity, in 2017 these costs were over 1 billion euros in Germany, more than 400 million euros in Great Britain, over 80 million euros in Spain, and approximately 50 million euros in Italy.

1.2. Multi-Settlement LMP Market

A multi-settlement LMP market has at least a day-ahead forward market and a real-time market, each of which employs the same market-clearing mechanism. The day-ahead market typically allows generation unit owners to submit three-part offers to supply energy: start-up costs, minimum load costs, and an energy offer curve. These are used to compute hourly generation schedules and ancillary service quantities and LMPs for energy and ancillary services for all 24 hours of the following day. A generation unit will not be accepted to supply energy in the day-ahead market unless the combination of its offered start-up costs, minimum load costs and energy production costs are part of the least as-offered-cost solution to serving the hourly locational demands for all 24 hours of the following day.

The energy schedules that emerge from the day-ahead market do not require a generation unit to produce the energy sold or a load to consume the energy purchased in the day-ahead market at a given location. Any production shortfall relative to a day-ahead generation schedule must be purchased from the real-time market at that location. Any production greater than a generation unit's day-ahead schedule is sold at the real-time price at that location. Any additional consumption beyond a load's day-ahead energy schedule is paid for at the real-time price at that location, and the surplus of a day-ahead schedule relative to actual consumption is sold at the real-time price at that location.

1.3. Mitigating Local Market Power

The configuration of the transmission network, the level and location of demand, as well as the level of output of other generation units can create system conditions in which almost any generation unit or group of generation units has a significant ability to exercise unilateral market power. The constrained-on generation problem is an example of this phenomenon. The unit's owner knows that it must be accepted to supply energy regardless of its offer price. Without an LMPM mechanism, there may be no limit to the offer price the unit's owner could submit and be

accepted to supply energy. During the first summer of the California market, when there was no formal LMPM mechanism, suppliers submitted extremely high offers for energy and ancillary services when these system conditions arose.

This logic is why market power mitigation mechanisms typically used in Europe and other industrialized regions and initially employed in the US, that designate in advance the offers of certain generation units for mitigation for an entire year, miss many instances of the exercise of substantial unilateral market power.

An automated LMPM mechanism built into the market software that relies on actual system conditions to determine whether any supplier has a substantial ability and incentive to exercise unilateral market power is likely to be significantly more effective. This regulator-approved administrative procedure determines: (1) when a supplier has an ability to exercise local market power worthy of mitigation, (2) the value of the supplier's mitigated offer price, and (3) the price the mitigated supplier is paid. It is increasingly clear to regulators around the world, particularly those that operate markets with a finite amount of transmission capacity and significant intermittent renewable generation capacity, that an automatic LMPM mechanism is a necessary feature of any short-term market design.

Because these LMPM mechanisms are built into the market software of all US markets and automatically mitigate the offers of suppliers deemed to have substantial ability to exercise unilateral market power, they are effective at preventing the exercise of significant local market power with little disruption to the operation of the short-term market.

1.4. Benefits of a Multi-Settlement LMP Market

A multi-settlement LMP market design can facilitate the active participation of final consumers in the wholesale market and reduce both the input fuel and total variable cost of producing the same amount of thermal energy relative to a multi-settlement zonal market design. The presence of an automatic LMPM mechanism and make-whole payments that guarantee start-up, minimum load, and energy cost recovery for the day for all generation units committed to operate in the day-ahead market reduces the incentive for suppliers to exercise unilateral market power. An expected profit-maximizing supplier with no ability to exercise unilateral market power will submit an offer price equal to its marginal cost because make-whole payments ensure recovery of their start-up, minimum load, and energy costs.

Because day-ahead purchases are firm financial commitments, a retailer can sell energy

purchased in the day-ahead market at the real-time price by consuming less than its day-ahead energy schedule. This eliminates the need for the regulator to set an administrative baseline relative to which a retailer sells demand reductions. The day-ahead market also allows retailers and large consumers to submit price-sensitive bid curves into the day-ahead market to reduce the market-clearing price and the quantity of energy they purchase in the day-ahead market.

1.5. Modifications for Large Scale Intermittent Renewables Deployment

A multi-settlement LMP market design is capable of managing a generation mix with a significant share of intermittent renewables. However, some modifications are likely to be needed as the share of intermittent renewable resources increases. Additional operating constraints will need to be incorporated into the day-ahead and real-time market models for reliable system operation with an increased quantity of intermittent renewables.

There is also likely to be a need to introduce additional ancillary services to accommodate a larger share of intermittent renewable energy. For example, California introduced a fast-ramping ancillary service product that compensates controllable generation units for not supplying energy during certain hours of the day in order to have sufficient unloaded capacity to meet the rapid increase in net demand (the difference between system demand and renewable generation) in the early evening when the state's solar resources stop producing.

Because controllable resources are likely to have to start and stop more frequently as the share of intermittent resources increases, implementations of convex hull pricing and other market-clearing mechanisms that limit the magnitude of make-whole payments will need to be developed.

2. Resource Adequacy with Significant Intermittent Renewables

Why do wholesale electricity markets require a regulatory mandate to ensure long-term resource adequacy? Electricity is essential to modern life, but so are many other goods and services. Consumers want cars, but there is no regulatory mandate that ensures enough automobile assembly plants to produce these cars. They want point-to-point air travel, but there is no regulatory mandate to ensure enough airplanes to accomplish this. Many goods are produced using high fixed cost, low marginal cost technologies, similar to electricity supply. Nevertheless, these firms recover their cost of production, including a return on the capital invested, by selling their output at a market-determined price.

So, what is different about electricity that requires a long-term resource adequacy mechanism? The regulatory history of the electricity supply industry and the legacy technology for metering electricity consumption results in what I call a *reliability externality*.

2.1. The Reliability Externality

Different from the case of wholesale electricity, in the market for automobiles and air travel there is no regulatory prohibition on the short-term price rising to the level necessary to clear the market. Airlines adjust the prices for seats on a flight over time in an attempt to ensure that the number of customers traveling on that flight equals the number of seats flying. This ability to use price to allocate the available seats is also what allows the airline to recover its total production costs.

Using the short-term price to manage the real-time supply and demand balance in a wholesale electricity market is limited by a finite upper bound on a supplier's offer price and/or a price cap that limits the maximum market-clearing price. Although offer caps and price caps can limit the ability of suppliers to exercise unilateral market power in the short-term energy market, they also reduce the revenues suppliers can receive during scarcity conditions. This is often referred to as the *missing money* problem for generation unit owners. However, this missing money problem is only a symptom of the existence of the "reliability externality."

This externality exists because offer caps limit the cost to electricity retailers of failing to hedge their purchases from the short-term market. Specifically, if the retailer or large consumer knows the price cap on the short-term market is $250/MWh, then it is unlikely to be willing to pay more than that for electricity in any earlier forward market. This creates the possibility that real-time system conditions can occur where the amount of electricity demanded at or below the offer cap is less than the amount suppliers are willing to offer at or below the offer cap. This outcome implies that the system operator must be forced to either abandon the market mechanism or curtail load until the available supply offered at or below the offer cap equals the reduced level of demand, as occurred a number of times in California between January 2001 and April 2001, and most recently on August 14 and 15, 2020.

Because random curtailments of supply—also known as rolling blackouts—are used to make demand equal to the available supply at or below the offer cap under these system conditions, this mechanism creates a "reliability externality" because no retailer bears the full cost of failing to procure adequate amounts of energy in advance of delivery. A retailer that has purchased

sufficient supply in the forward market to meet its actual demand is equally likely to be randomly curtailed as another retailer of the same size that has not procured adequate energy in the forward market. For this reason, all retailers have an incentive to under-procure their expected energy needs in the forward market.

The lower the offer cap, the greater is the likelihood that the retailer will delay their electricity purchases to the short-term market. Delaying more purchases to the short-term market increases the likelihood of insufficient supply in the short-term market at or below the offer cap. Because retailers do not bear the full cost of failing to procure sufficient energy in the forward market to meet their future demand, there is a missing market for long-term contracts for long enough delivery horizons into the future to allow new generation units to be financed and constructed to serve demand under all future conditions in the short-term market. Therefore, a regulator-mandated long-term resource adequacy mechanism is necessary to replace this missing market.

Some form of regulatory intervention is necessary to internalize the resulting reliability externality, unless the regulator is willing to eliminate or substantially increase the offer cap so that the short-term price can be used to equate available supply to demand under all possible future system conditions. This approach is taken by the Electricity Reliability Council of Texas (ERCOT), which has a $9,000/MWh offer cap, and National Electricity Market in Australia, which has a 15,000 Australia Dollars per MWh offer cap. If customers do not have interval meters that can record their consumption on an hourly basis, then they have a very limited ability to benefit from shifting their consumption away from high-priced hours. All that can recorded for these customers is their total consumption between two successive meter readings so they can only be billed based on an average wholesale price during the billing cycle. Therefore, raising or having no offer cap on the short-term market would not be advisable in a region where few customers have interval meters. Even in regions with interval meters, there would be substantial political backlash from charging hourly wholesale prices that cause real-time demand to equal available supply under all possible future system conditions.

Currently, the most popular approach to addressing this reliability externality is a capacity payment mechanism that assigns a firm capacity value to each generation unit based on the amount of energy it can provide under stressed system conditions. Sufficient firm capacity procurement obligations are then assigned to retailers to ensure that annual system demand peaks can be met.

Capacity-based approaches to long-term resource adequacy rely on the credibility of the firm capacity measures assigned to generation units. This is a relatively straightforward process for thermal units. The nameplate capacity of the generation unit times its annual availability factor is a reasonable estimate of the amount of energy the unit can provide under stressed system conditions. For the case of hydroelectric facilities, this process is less straightforward. The typical approach uses percentiles of the distribution of past hydrological conditions for that generation unit to determine its firm capacity value.

Assigning a firm capacity value to a wind or solar generation unit is extremely challenging for several reasons. First, these units only produce when the underlying resource is available. If stressed system conditions occur when the sun is not shining or the wind is not blowing, these units should be assigned little, if any, firm capacity value. Second, because there is a high degree of contemporaneous correlation between the energy produced by solar and wind facilities within the same region, the usual approach to determining the firm capacity of a wind or solar unit assigns a smaller value to that unit as the total MWs of wind or solar capacity in the region increases. These facts imply that a capacity-based long-term resource adequacy mechanism is poorly suited to a zero marginal cost intermittent renewable feature.

2.2. Supplier Incentives with Fixed-Price Forward Contract Obligations for Energy

The standardized fixed-price forward contract (SFPFC) approach to long-term resource adequacy recognizes that a supplier with the ability to serve demand at a reasonable price may not do so if it has the ability to exercise unilateral market power. A supplier with the ability to exercise unilateral market power with a fixed-price forward contract obligation finds it expected profit maximizing to minimize the cost of supplying this forward contract quantity of energy. The SFPFC long-term resource adequacy mechanism takes advantage of this incentive by requiring retailers to hold hourly fixed-price forward contract obligations for energy that sum to the hourly value of system demand. This implies that all suppliers find it expected profit maximizing to minimize the cost of meeting their hourly fixed-price forward contract obligations, the sum of which equals the hourly system demand for all hours of the year.

To understand the logic behind the SFPFC mechanism, consider the example of a supplier that owns 150 MWs generation capacity that has sold 100 MWh in a fixed-forward contract at a price of $25/MWh for a certain hour of the day. This supplier has two options for fulfilling this forward contract: (1) produce the 100 MWh energy from its own units at their marginal cost of

$20/MWh or (2) buy this energy from the short-term market at the prevailing market-clearing price. The supplier will receive $2,500 from the buyer of the contract for the 100 MWh sold, regardless of how it is supplied. This means that the supplier maximizes the profits it earns from this fixed-price forward contract sale by minimizing the cost of supplying the 100 MWh of energy.

To ensure that the least-cost "make versus buy" decision for this 100 MWh is made, the supplier should offer 100 MWh in the short-term market at its marginal cost of $20/MWh. This offer price for 100 MWh ensures that if it is cheaper to produce the energy from its generation units—the market price is at or above $20/MWh—the supplier's offer to produce the energy will be accepted in the short-term market. If it is cheaper to purchase the energy from the short-term market—the market price is below $20/MW—the supplier's offer will not be accepted and the supplier will purchase the 100 MWh from the short-term market at a price below $20/MWh.

This example demonstrates that the SFPFC approach to long-term resource adequacy makes it expected profit-maximizing for each seller to minimize the cost supplying the quantity of energy sold in this forward contract each hour of the delivery period. By the logic of the above example, each supplier will find it in its unilateral interest to submit an offer price into the short-term market equal to its marginal cost for its hourly SFPFC quantity of energy, in order to make the efficient "make versus buy" decision for fulfilling this obligation.

In addition, because all suppliers know that the sum of the values of the hourly SFPFC obligations for all suppliers is equal the system demand, each firm knows that its competitors have substantial fixed-price forward contract obligations for that hour. This implies that all suppliers know that they have limited opportunities to raise the price they receive for short-term market sales beyond their hourly SFPFC quantity. For the above example, the supplier that owns 150 MWs of generation capacity has a strong incentive to submit an offer price close to its marginal cost to supply any energy beyond the 100 MWh of SFPFC energy it is capable of producing. Therefore, attempts by any supplier to raise prices in the short-term market by withholding output beyond their SFPFC quantity are likely to be unsuccessful because of the aggressiveness of the offers into the short-term market by its competitors with hourly SFPFC obligations.

2.3. SFPFC Approach to Resource Adequacy

This long-term resource adequacy mechanism requires all electricity retailers to hold SFPFCs for energy for fractions of realized system demand at various horizons to delivery. For example, retailers in total must hold SFPFCs that cover 100 percent of realized system demand,

95 percent of realized system demand one year in advance of delivery, 90 percent two-years in advance of delivery, 87 percent three years in advance of delivery, and 85 percent four years in advance of delivery. The fractions of system demand and number of years in advance that the SFPFCs must be purchased are parameters set by the regulator to ensure long-term resource adequacy. In the case of a multi-settlement LMP market, the SFPFCs would clear against the quantity-weighted average of the hourly locational prices at all load withdrawal nodes.

SFPFCs are shaped to the hourly system demand within the delivery period of the contract. Figure 1 contains a sample pattern of system demand for a four-hour delivery horizon. The total demand for the four hours is 1000 MWh, and the four hourly demands are 100 MWh, 200 MWh, 400 MWh and 300 MWh. Therefore, a supplier that sells 300 MWh of SFPFC energy has the hourly system demand-shaped forward contract obligations of 30 MWh in hour 1, 60 MWh in hour 2, 120 MWh in hour 3 and 90 MWh in hour 4 for Firm 1 in Figure 2. The hourly forward contract obligations for Firm 2 that sold 200 MWh SFPFC energy and Firm 3 that sold 500 MWh of SFPFC energy are also shown in Figure 2. These SFPFC obligations are also allocated across the four hours according to the same four hourly shares of total system demand. This ensures that the sum of the hourly values of the forward contract obligations for the three suppliers is equal to the hourly value of system demand. Taking the example of hour 3, Firm 1's obligation is 120 MWh, Firm 2's is 80 MWh and Firm 3's is 200 MWh. These three values sum to 400 MWh, which is equal to the value of system demand in hour 3 shown in Figure 1.

These standardized fixed-price forward contracts are allocated to retailers based on their share of system demand during the month. Suppose that the four retailers in Figure 3 consume 1/10, 2/10, 3/10, and 4/10, respectively, of the total energy consumed during the month. This means that Retailer 1 is allocated 100 MWh of the 1000 MWh SFPFC obligations for the four hours, Retailer 2 is allocated 200 MWh, Retailer 3 is allocated 300 MWh, and Retailer 4 is allocated 400 MWh. The obligations of each retailer are then allocated to the individual hours using the same hourly system demand shares used to allocate the SFPFC energy sales of suppliers to the four hours. This allocation process implies Retailer 1 holds 10 MWh in hour 1, 20 MWh in hours 2, 40 MWh in hour 3 and 30 MWh in hour 4. Repeating this same allocation process for the other three retailers yields the remaining three hourly allocations shown in Figure 3. Similar to the case of the suppliers, the sum of allocations across the four retailers for each hour equals the total hourly system demand. For period 3, Retailer 1's holding is 40 MWh, Retailer 2's is 80 MWh, Retailer 3's is 120 MWh, and Retailer 4's is 160 MWh. The sum of these four magnitudes

is equal to 400 MWh, which is the system demand in hour 3.

2.4. Mechanics of Standardized Forward Contract Procurement Process

The SFPFCs would be purchased through auctions several years in advance of delivery in order to allow new entrants to compete to supply this energy. Because the aggregate hourly values of these SFPFC obligations are allocated to retailers based on their actual share of system demand during the month, this mechanism can easily accommodate retail competition. If one retailer loses load and another gains it during the month, the share of the aggregate hourly value of SFPFCs allocated to the first retailer falls and the share allocated to the second retailer rises.

The wholesale market operator would run the auctions with oversight by the regulator. One advantage of the design of the SFPFC products is that a simple auction mechanism can be used to purchase each annual product. A multi-round auction could be run where suppliers submit the total amount of annual SFPFC energy they would like to sell for a given delivery period at the price for the current round. Each round of the auction the price would decrease until the amount suppliers are willing to sell at that price is less than or equal to the aggregate amount of SFPFC energy demanded.

The wholesale market operator would also run a clearinghouse to manage the counterparty risk associated with these contracts. All US wholesale market operators currently do this for all participants in their energy and ancillary services markets. In several US markets, the market operator also provides counterparty risk management services for long-term financial transmission rights, which is not significantly different from performing this function for SFPFCs.

SFPFCs auctions would be run on an annual basis for deliveries starting two, three, and four years in the future. In steady state, auctions for incremental amounts of each annual contract would also be needed so that the aggregate share of demand covered by each annual SFPFC could increase over time. The eventual 100 percent coverage of demand occurs through a final true-up auction that takes place after the realized values for hourly demand for the delivery period are known.

Consider the following two examples of how the true-up auction would work. Assume for simplicity, the monthly load shares of the four retailers remain unchanged. Suppose that the initial 1000 MWh SFPFC in the above example sold at $50/MWh. However, suppose that actual demand turned out to be 10 percent higher in every period as shown Figure 4 and the additional 100 MWh purchased in the true-up auction sold at $80/MWh. If each firm sold 10 percent more SFPFC

energy in the true-up auction this would yield the hourly obligations for each supplier shown in Figure 5. The hourly obligations for the four retailers are shown in Figure 6. These would clear against the average cost of purchases from the original auction and true-up auction of $52.73. If the realized hourly demands are ten percent lower as shown in Figure 7, the true-up auction would buy back 100 MWh of SFPFC energy. If all suppliers bought back 10 percent of their initial sales at $20/MWh, the resulting hourly obligations would be those shown in Figure 8. The 10 percent smaller hourly obligations of the four retailers are shown in Figure 9 and these would clear against the average cost of the initial auction purchase less the revenues from the true-up auction sales for the required 900 MWh of obligations of $53.33.

As shown in Figures 6 and 9, each purchase or sale of the same annual SFPFC product is allocated to retailers according to their load shares during the delivery month. If three different size purchases are made for the same annual SFPFC product at different prices, then each retailer is allocated its load share for the month of these three purchases. This ensures a level playing field for retailers with respect to their long-term resource adequacy obligation. All retailers face the same average price for the long-term resource adequacy obligation associated with their realized demand for the month.

The advance purchase fractions of the final demand are the regulator's security blanket to ensure that system demands can be met for all hours of the year for all possible future system conditions. If the regulator is worried that not enough resources will be available in time to satisfy this requirement, it can increase the share of final demand that it purchases in each annual SFPFC auction. As shown above, if too much SFPFC energy is purchased in an annual auction, it can be sold back to generation unit owners in a later auction or the final true-up auction.

Cross hedging between controllable generation units and intermittent renewable resources under this mechanism is enforced by tying the amount of SFPFC energy a generation unit owner can sell on an annual basis to the value of their firm energy. The system operator would assign firm energy values for each generation unit using a mechanism similar to what is currently used to compute firm capacity values. Multiplying a unit's MWs of firm capacity by the number of hours in the year would yield the unit's firm energy value, which is the upper bound on the amount of SFPFC energy the unit owner could sell in all auctions for an annual compliance period. Because the firm capacity of a generation unit is defined as the amount of energy it can produce under stressed system conditions, this limitation on annual sales of firm energy implies that intermittent wind and solar resources would sell much less SFPFC energy than the total MWhs they expect to

produce in a typical year, and controllable generation unit owners would sell significantly more SFPFC energy than the total MWhs they expect to produce in an typical year.

In most years, a controllable resource owner would be producing energy in a small number of hours of the year, but earning the difference between the price at which they sold the energy in the SFPFC auction and the hourly short-term market price times the hourly value of its SFPFC energy obligation for all the hours that it does not produce energy. Intermittent renewables owners would typically produce more than their SFPFC obligation in energy and sell the additional energy at the short-term price. In years with low renewable output near their SFPRC obligations, controllable resource owners would produce close to the hourly value of their SFPFC energy obligation, thus making average short-term prices significantly higher. However, aggregate retail demand would be shielded from these high short-term prices because of their SFPC holdings.

2.4. Advantages of SFPFC Approach to Long-Term Resource Adequacy

This mechanism has a number of advantages relative to a capacity-based approach. There is no regulator-mandated aggregate capacity requirement. Generation unit owners are allowed to decide both the total MWs and the mix of technologies to meet their SFPFC energy obligations. There is also no prohibition on generation unit owners or retailers engaging in other hedging arrangements outside of this mechanism. Specifically, a retailer could enter into a bilateral contract for energy with a generation unit owner or other retailer to manage the short-term price and quantity risk associated with the difference between their actual hourly load shape and the hourly values of their retail load obligation. This mechanism provides a nudge to market participants to develop a liquid market for these bilateral contract arrangements at horizons to delivery similar to the SFPFC products. Instead of starting from the baseline of no fixed-price forward contract coverage of system demand by retailers, this mechanism starts with 100 percent coverage of system demand, which retailers can unwind at their own risk.

For the regulated retail customers, the purchase prices of SFPFCs can be used to set the wholesale price implicit in the regulated retail price over the time horizon that the forward contract clears. This would provide retailers with a strong incentive to reduce their average wholesale energy procurement costs below this price through bilateral hedging arrangements, storage investments, or demand response efforts.

There are several reasons why this mechanism should be a more cost-effective approach to long-term resource adequacy than a capacity-based mechanism in a zero marginal cost intermittent

future. First, the sale of SFPFC energy starting delivery two or more years in the future provides a revenue stream that will significantly increase investor confidence in recovering the cost of any investment in new generation capacity.

Second, because retailers are protected from high short-term prices by total hourly SFPFC holdings equal to system demand, the offer cap on the short-term market can be raised in order to increase the incentive for all suppliers to produce as much energy as possible during stressed system conditions. Third, the possibility of higher short-term price spikes can finance investments in storage and load-shifting technologies and encourage active participation of final demand in the wholesale market, further enhancing system reliability in a market with significant intermittent renewable resources.

If SFPFC energy is sold for delivery in four years based on a proposed generation unit, the regulator should require construction of the new unit to begin within a pre-specified number of months after the signing date of the contract or require posting of a substantially larger amount of collateral in the clearinghouse with the market operator. Otherwise, the amount of SFPFC energy that this proposed unit sold would be automatically liquidated in a subsequent SFPFC auction and a financial penalty would be imposed on the developer. Other completion milestones would have to be met at future dates to ensure the unit is able to provide the amount of firm energy that it committed to provide in the SFPFC contract sold. If any of these milestones were not met, the contract would be liquidated.

3. Final Comments

There is no perfect wholesale market design. There are only better wholesale market designs, and what constitutes a better design depends on many factors specific to the region. The long-term resource adequacy mechanism should be coordinated with the short-term market design. Although there is general agreement on the key features of a best-practice short-term market design, many details must be adjusted to reflect local conditions. For this reason, wholesale market design is a process of continuous learning, adaption, and hopefully, improvement. The standardized energy contracting approach to long-term resource adequacy described in this paper is an example of this process. While it has a number of features that are likely make it significantly better suited to a zero-marginal-cost intermittent-renewables-electricity-supply industry, there are many details of this basic mechanism that should be adapted to reflect local conditions.

Further Reading

Graf, C., Quaglia, F. and Wolak, F.A. (2020) "Simplified Electricity Market Models with Significant Intermittent Renewable Capacity: Evidence from Italy (January 2020), available at http://www.stanford.edu/~wolak.

Gribik, P. R., Hogan, W. W., & Pope, S. L. (2007). Market-clearing electricity prices and energy uplift. *Cambridge, MA*.

Hogan, W. W. (2002). Electricity market restructuring: reforms of reforms. *Journal of Regulatory Economics*, *21*(1), 103-132.

Schweppe, F. C., Caramanis, M. C., Tabors, R. D., & Bohn, R. E. (2013). *Spot pricing of electricity*. Springer Science & Business Media.

Wolak, F. A. (2020) Wholesale Electricity Market Design, in *Handbook on the Economics of Electricity Markets*, J.M. Glanchant, P.L. Joskow, and M. Pollitt (editors), Cambridge University Press.

Biography

Frank A. Wolak is with Stanford University

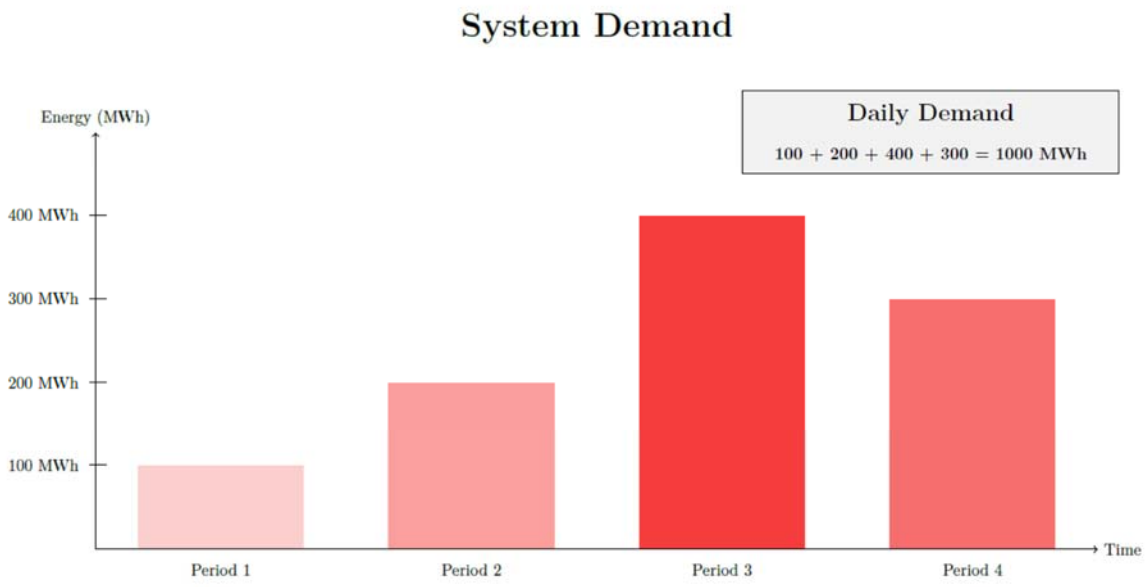


Figure 1: Hourly System Demands

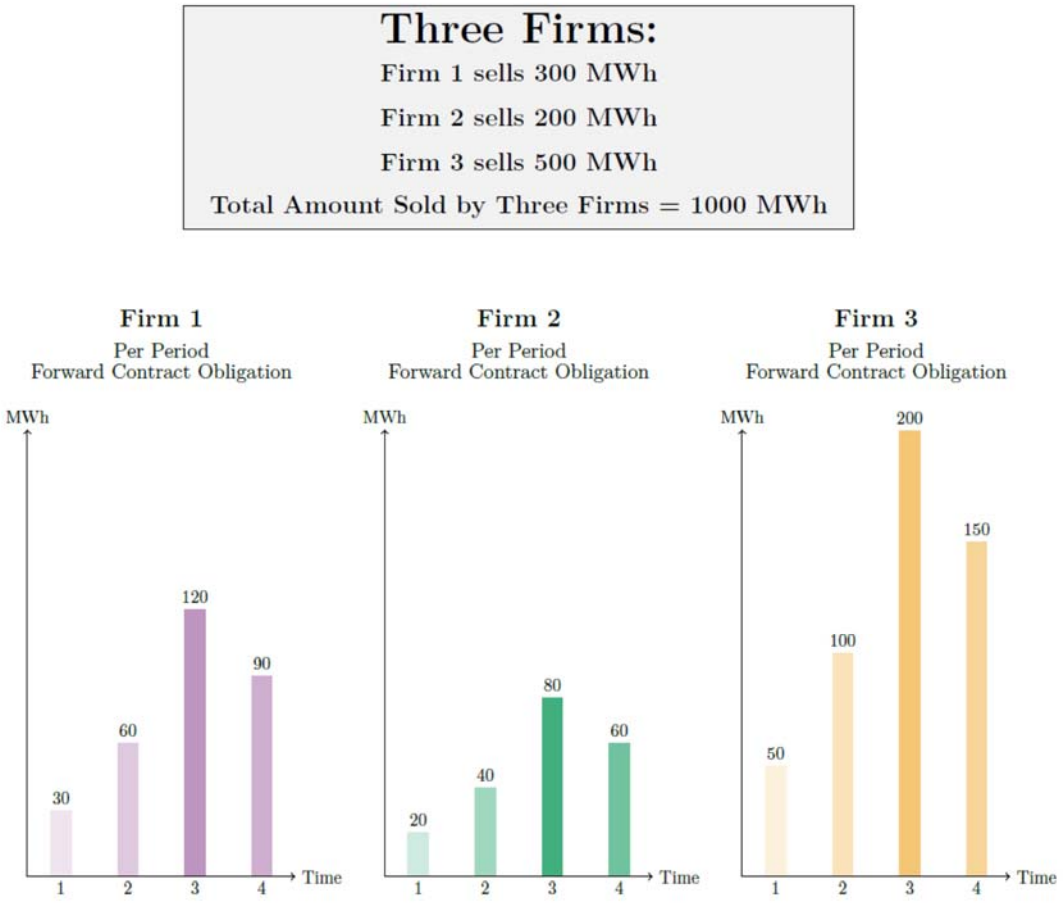


Figure 2: Hourly Forward Contract Quantities for Three Suppliers

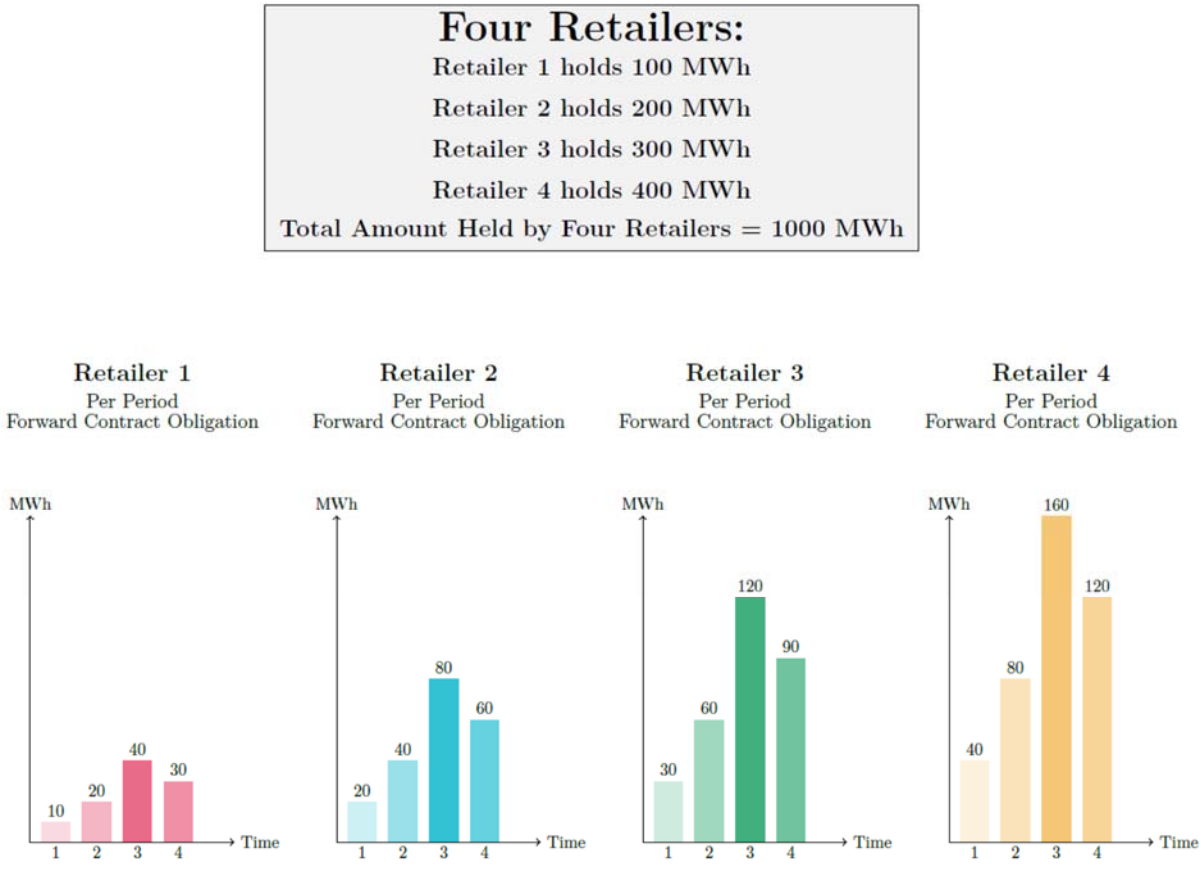


Figure 3: Hourly Forward Contract Quantities for Four Retailers

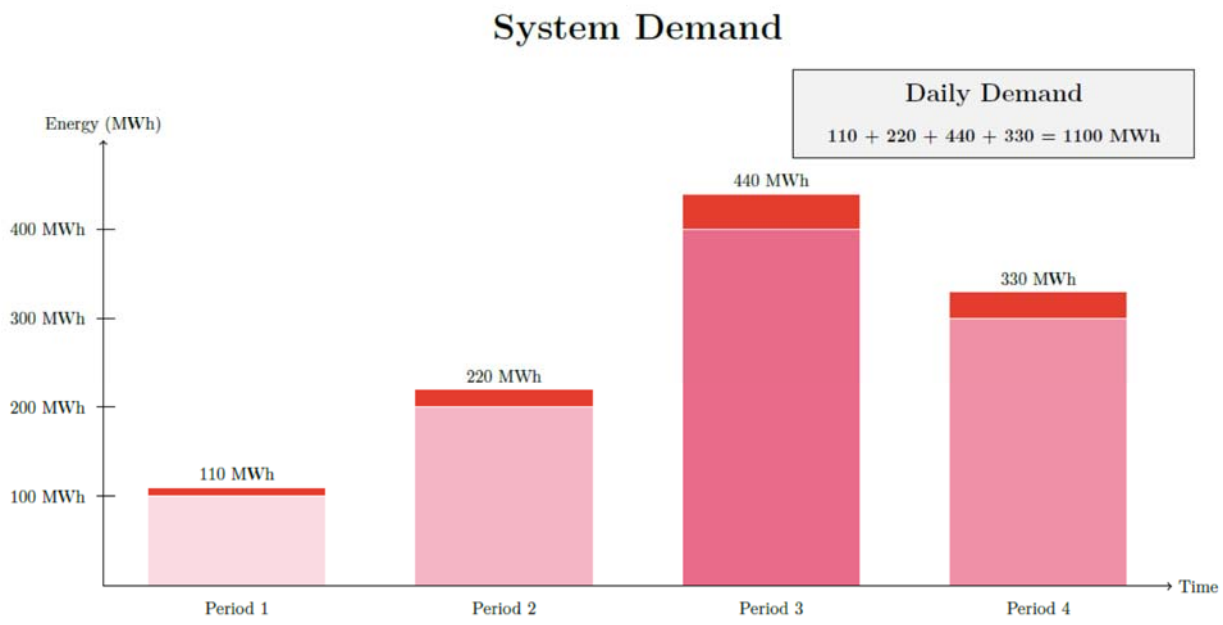


Figure 4: Hourly System Demands (10 Percent Higher)

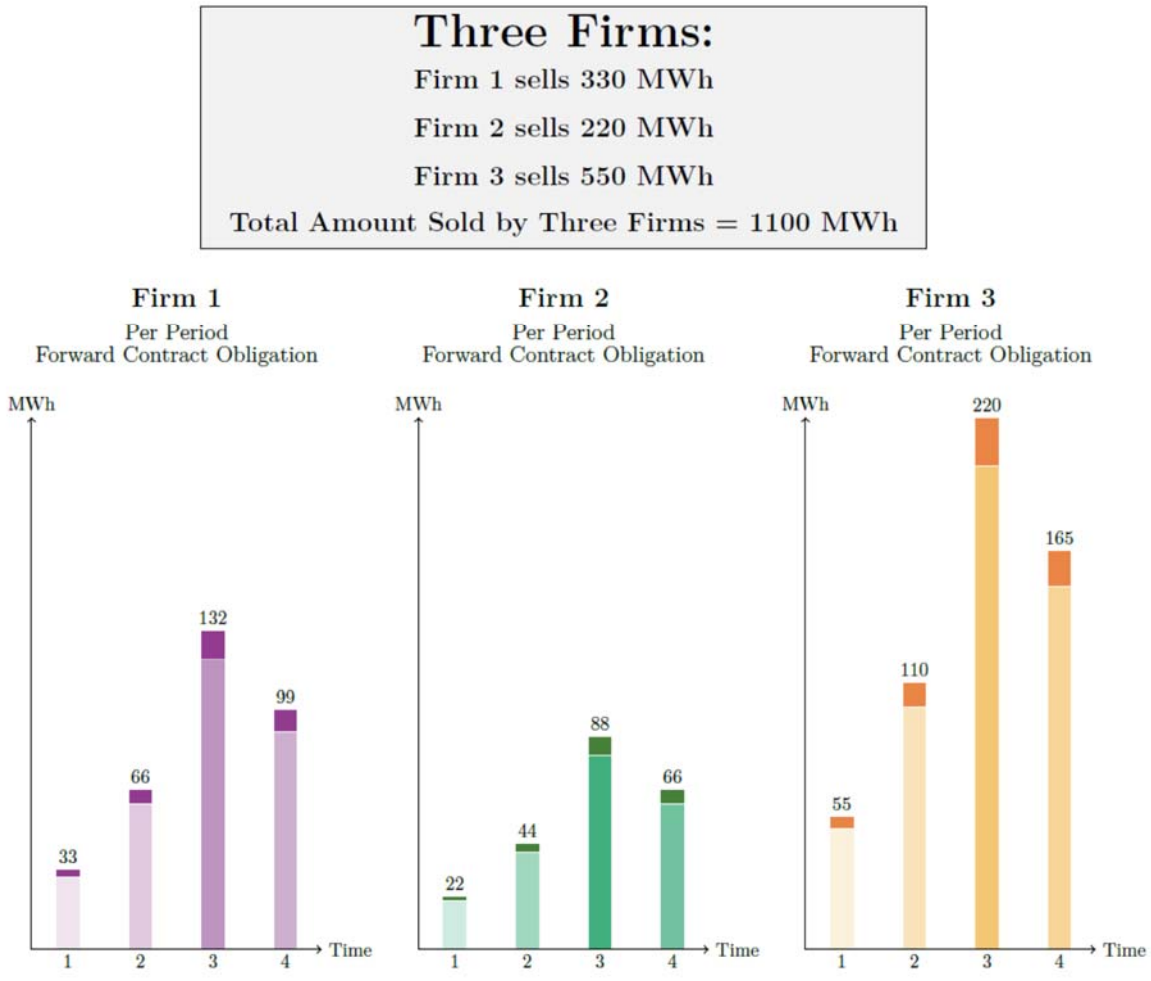


Figure 5: Hourly Forward Contract Quantities for Three Suppliers (10 Percent Higher)

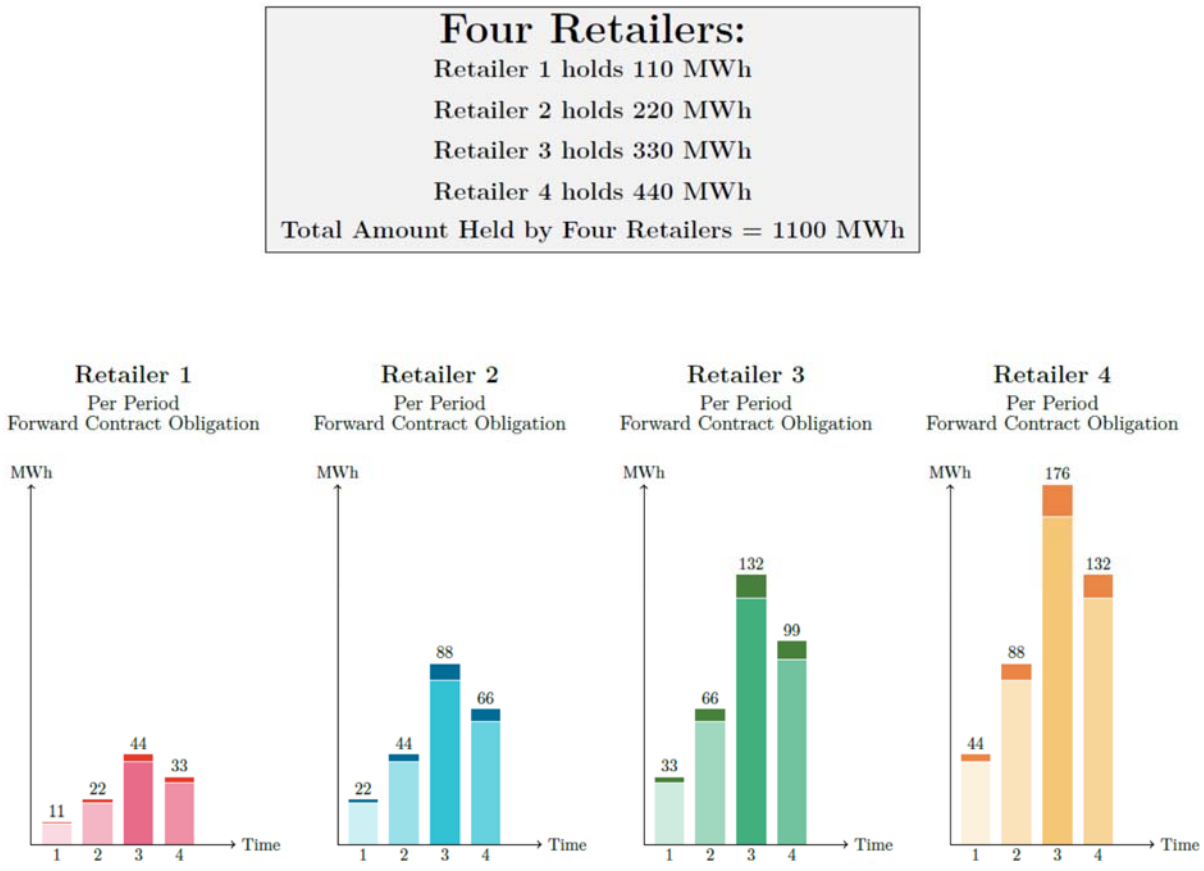


Figure 6: Hourly Forward Contract Quantities for Four Retailers (10 Percent Higher)

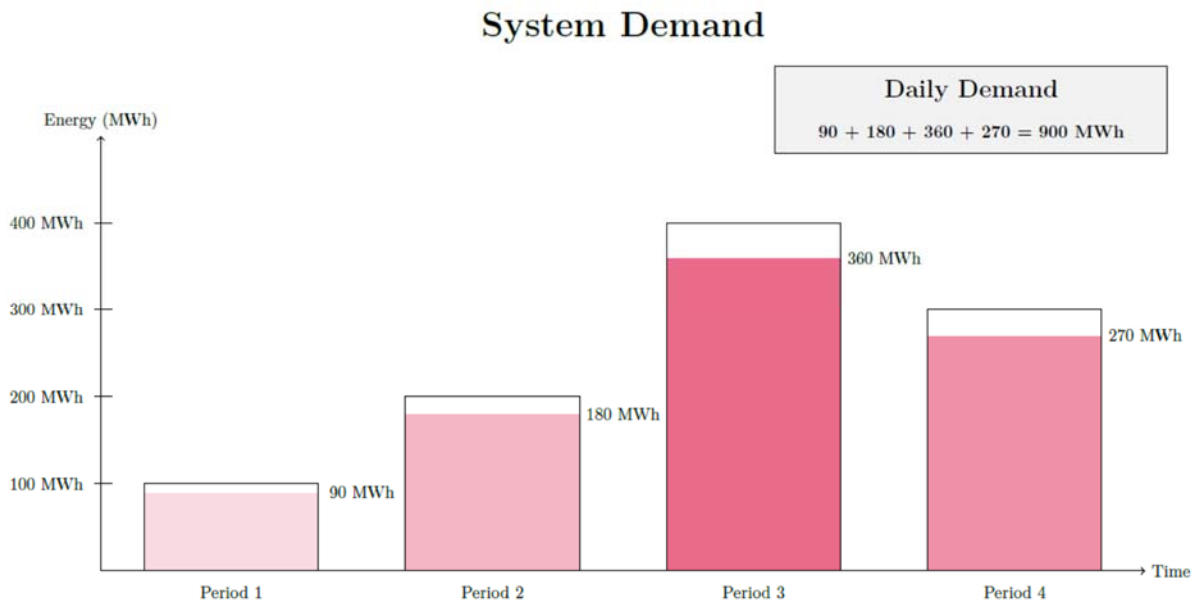


Figure 7: Hourly System Demands (10 Percent Lower)

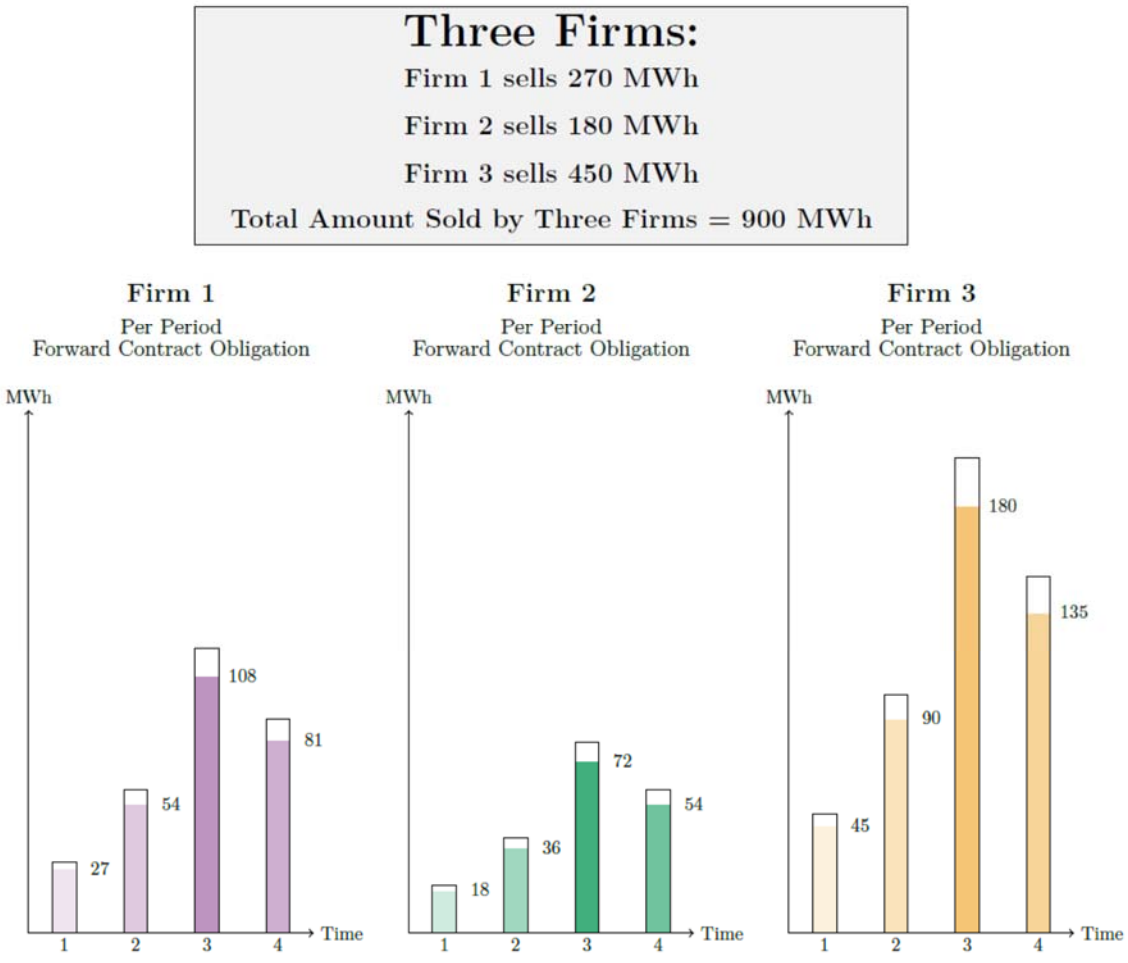


Figure 8: Hourly Forward Contract Quantities for Three Suppliers (10 Percent Lower)



Figure 9: Hourly Forward Contract Quantities for Four Retailers (10 Percent Lower)