Pacific Northwest NATIONAL LABORATORY

PNNL-33792

# Modeling and Market Design Considerations for Conventional and Decarbonized Resources

May 2022

Brent Eldridge
Abhishek Somani

U.S. DEPARTMENT OF ENERGY

Prepared for the U.S. Department of Energy
under Contract DE-AC05-76RL01830

PACIFIC NORTHWEST NATIONAL LABORATORY
*operated by*
BATTELLE
*for the*
UNITED STATES DEPARTMENT OF ENERGY
*under Contract DE-AC05-76RL01830*

**Printed in the United States of America**

**Available to DOE and DOE contractors from the**
**Office of Scientific and Technical Information,**
**P.O. Box 62, Oak Ridge, TN 37831-0062;**
**ph: (865) 576-8401**
**fax: (865) 576-5728**
**email: reports@adonis.osti.gov**

**Available to the public from the National Technical Information Service**
**5301 Shawnee Rd., Alexandria, VA 22312**
**ph: (800) 553-NTIS (6847)**
**email: orders@ntis.gov <https://www.ntis.gov/about>**
**Online ordering: http://www.ntis.gov**

## Abstract

The following paper reviews various market design issues that will need to be reconsidered due to anticipated changes in the resources that supply energy in wholesale electricity markets. As the energy supply continues towards decarbonization, the change in resource technologies will affect the fundamentals of production scheduling and the policies to address supply variability and uncertainty. We show through simple numerical examples that the existing market design approach based on fuel costs will result in $0/MWh prices with intermittent price spikes during reserve shortages, but that incorporating more granular reserve pricing, energy storage participation, and price-responsive demand participation can restore efficient market clearing with reasonable pricing outcomes. The market design approach fundamentally shifts from fuel based to opportunity cost based. We briefly review alternative market design frameworks that can support this shift, including intraday markets, decentralized markets, flexibility options, and swing contracts. Market designs may also be required to accommodate or support various out-of-market policies and agreements; ideally, these external factors can be integrated into the market design to facilitate efficient exchanges across longer time scales, between other markets, and in support of public policy goals. The paper concludes by discussing how decarbonization trends may affect the design and use of production cost models for short term operations and long term planning studies.

# Contents

## Acknowledgements

We thank Ben Hobbs, Diane Baldwin, Jesse Holzer, Travis Douville, Xueqing Sun, Jeremy Twitchell, Xinda Ke, Hayden Reeve, Kostas Oikonomou, Sohom Datta, Brittany Tarufelli, Allison Campbell, Juan Bedoya Ceballos, Nathalie Voisin, Dhruv Bhatnagar, Marcelo Elizondo, Patrick Maloney, Di Wu, Saptarshi Bhattacharya, Michael Kintner-Meyer, Rebecca O'Neill, and all other PNNL staff who participated in the Zero-Marginal Cost Market Design Workshop.

# I. Introduction

The following white paper discusses a broad range of issues surrounding short-term operational modeling and market design of organized wholesale electricity markets in the context of grid decarbonization. It describes typical characteristics of how and why the market is currently operated and what sorts of issues might arise as a result of a changing generation resource mix. The issues discussed in the paper are intended to spur further conversations about what directions might be most beneficial for electricity market design and what roles researchers at PNNL might play in proposing new solutions.

Today's organized wholesale electricity markets, called Independent System Operators (ISOs), were built around a market design that caters to the efficient short-term operation of a fleet of conventional generators that serve generally passive electricity consumers. In pursuit of climate and environmental goals, state and federal policies have targeted decarbonization of the electricity sector in ways that will fundamentally reshape the design of wholesale electricity markets and will result in considerably more low and zero marginal cost resources with variable and uncertain energy output. This paper is intended to frame future market design issues and to foster discussion of potential solutions.

Replacement of conventional generators with renewable wind and solar creates a number of fundamental changes in electricity supply. Conventional generators have relatively low capital investment cost compared to variable fuel costs, a certain and definite operating range, flexibility to ramp up or down to meet demand, and the ability to provide reserves in case of generator outages. In contrast, renewable generation typically has high capital cost with zero variable fuel cost, uncertain power output, and limited ability to provide flexibility or reserves. This change in generator characteristics will almost certainly require complementary market participation from energy storage technologies and demand-side flexibility.

Section II provides motivating examples to describe the fundamental market design issues that arise when electricity supply becomes dominated by low and zero marginal cost resources. Section III discusses how changes in resource technologies may affect the operational decisions made by electricity markets and gives a brief overview of the distinct operational attributes of renewable generators as well as energy storage and demand-side flexibility. Section IV describes a selection of market design alternatives that may be able to address the economic and operational concerns raised by new resource technologies. Section V addresses various issues that are external to ISO market design but nonetheless affect decisions made by the market operator, market settlements, and investment decisions made by market participants. These issues include power purchase agreements in which an energy producer and an energy consumer enter a long-term agreement that hedges them against the real-time spot price of electricity determined by the ISO. This section also discusses how carbon pricing and other federal and state environmental regulations are reflected in ISO market designs, as well as the impact of retail electric tariffs on ISO market participation. Finally, Section VI considers the impact of various market design concepts on short-term operations and long-term planning models.

Readers will notice that discussion of capacity markets is a major omission from this paper. The paper is largely focused first on production cost modeling and second on the relation between production cost modeling and market design issues. Capacity markets are a major topic for market design, especially in the context of a decarbonized grid, but the topic is too complex to adequately cover in this paper. Tarufelli (2022) provides recent discussion of capacity markets and their potential application in transactive energy systems and distribution networks.[1]

# II. Motivating Examples

The issues raised in this paper are motivated by anticipated changes in the status quo market operations that might be summarized as primarily using the market to schedule conventional generation to match demand and setting prices equal to the marginal cost of the most expensive generator. As more conventional generation is replaced by renewables, there is a concern about how to keep supply and demand balanced and how to set market clearing prices.

The following numerical examples illustrate this status quo and then show how the integration of storage and demand-side participation in markets will help alleviate the anticipated stresses on system operations. The examples are not intended as any kind of proof that "technology X will solve these future issues," but rather to illustrate the ways that market designs and participation models might change to better coordinate whatever future resources might best serve the future grid. In each of the following model formulations, model parameters are written in uppercase and model variables are written in lowercase.

## A. Example 1: Status quo, satisfying fixed demand with conventional and renewable generation

First, we model a stylized ISO market that primarily relies on renewable wind generation and includes a small amount of backup conventional generation. We include three consecutive settlement periods so that a range of operating conditions can be considered. Table II.1 shows the two generators participating in the market and their maximum output during the three periods:

Table II.1 Generator Attributes

| Generator | Pmax, t=1 | Pmax, t=2 | Pmax, t=3 | Marginal cost | Start-up cost | No-load cost |
|---|---|---|---|---|---|---|
| **Wind** | 75 MW | 65 MW | 55 MW | $0/MWh | $0 | $0/hour |
| **Conventional** | 20 MW | 20 MW | 20 MW | $75/MWh | $250 | $1/hour |

The renewable wind resource provides different output in each period due to its weather-dependence, but it incurs no production costs for the energy it generates. In contrast, the conventional generator has the same energy capacity in each period, but it incurs a marginal cost for each MWh produced, a start-up cost when it goes from an off-line to an on-line state, and a no-load cost for each hour that it is on-line.

Following status quo assumptions, the above generators will be used to serve fixed demand. In addition, the generators must provide operating reserves that are necessary for system reliability.[2] Any solution that does not serve all demand will be considered infeasible, and a

[1] Tarufelli, B., Eldridge, B., Somani, A., "Capacity Markets for Transactive Energy Systems," PNNL-76328. September 2022.

[2] Operating reserves are typically based on the size of the largest potential generator failure. The example simplifies this aspect of operating reserves significantly: first, the amount of procured operating reserves is much smaller than either generator in the example, and second, any operating reserves might

penalty of $10,000/MW is applied to any shortage of operating reserves. Demand and operating reserves requirements are shown in Table II.2.

Table II.2 Fixed Demand and Operating Reserve Requirements

| | t=1 | t=2 | t=3 | Penalty |
|---|---|---|---|---|
| **Demand** | 60 MW | 70 MW | 67 MW | Infeasible |
| **Operating Reserve Requirement** | 10 MW | 10 MW | 10 MW | $10,0000/MW |

Using the data from tables above, an ISO market determines generator production schedules and market clearing prices by solving the following optimization problem:

| | | |
|---|---|---|
| Cost minimization: | $\min GenCost + PenaltyCost$ | |
| Generator costs: | $GenCost = \sum_{t \in T} \sum_{g \in Gen} \left( M_g p_{gt} + S_g x_{gt} + O_g y_{gt} \right)$ | |
| Reserve violation penalty: | $PenaltyCost = P * \sum_{t \in T} slack_t$ | |
| Power balance: | $D_t - \sum_{g \in Gen} p_{gt} = 0,$ | $\forall t \in T$ |
| Operating reserve req.: | $R_t - \sum_{g \in Gen} r_{gt} - slack_t \leq 0,$ | $\forall t \in T$ |
| Generator minimum: | $p_{gt} \geq P_g^{\min} y_{gt},$ | $\forall g \in Gen, \forall t \in T$ |
| Generator maximum: | $p_{gt} + r_{gt} \leq P_g^{\max} y_{gt},$ | $\forall g \in Gen, \forall t \in T$ |
| Startup logic: | $x_{gt} \geq y_{gt} - y_{g,t-1},$ | $\forall g \in Gen, \forall t \in T \setminus T_0$ |
| Nonnegative variables: | $p_{gt}, r_{gt}, slack_t \geq 0$ | |
| Binary variables: | $x_{gt}, y_{gt} \in \{0,1\}$ | |

In the above formulation, $M_g$ is marginal cost, $S_g$ is start-up cost, $O_g$ is no-load cost, $P$ is the operating reserve violation penalty, $D_t$ is the demand at time $t$, $R_t$ is the operating reserve requirement at time $t$, $P_g^{\min}$ and $P_g^{\max}$ are the generator maximum and minimum output, $p_{gt}$ is generator $g$'s dispatch in $t$, $r_{gt}$ is the reserves provided by $g$ in $t$, $x_{gt}$ is generator $g$'s start-up decision in $t$, $x_{gt}$ is generator $g$'s operating status in $t$, and $slack_t$ is the operating reserve violation in $t$. The objective function is a cost minimization, which is the typical paradigm currently used in production cost models that simulate electricity markets. The market's locational marginal price (LMP) is set by the dual variable (shadow price) to the power balance constraint, and the operating reserves price (ORP) is set by the dual variable to the operating reserve requirement constraint. These prices are determined by the incremental change in the objective function per unit increase in the left hand side value of each respective constraint.

### 1. Market Clearing

Results from the above market are shown in Table II.3 below.

Table II.3 Example 1 Market Results

| Period: | 1 | 2 | 3 |
|---|---|---|---|
| **Demand (MW)** | 60 | 70 | 67 |
| **Wind (MW)** | 60 | 65 | 55 |
| **Conventional (MW)** | 0 | 5 | 12 |
| **Wind curtailment (MW)** | -15 | 0 | 0 |
| **LMP ($/MWh)** | $0 | $75 | $10,075 |

---

not provide any reliability benefit after a generator failure if they are provided by the failed generator. The example therefore only illustrates some economic properties of ISO markets although it does not accurately portray how reliability is maintained.

| Period: | 1 | 2 | 3 |
|---|---|---|---|
| OpRes (wind, MW) | 15 | 0 | 0 |
| OpRes (conventional, MW) | 0 | 15 | 8 |
| ORP ($/MWh) | $0 | $0 | $10,000 |

The total production cost of this solution is $1,527, and the operating reserve penalty adds an additional $20,000 for a total objective value of $21,527. The available wind capacity is sufficient to serve demand and provide operating reserves in period 1, so both the LMP and the ORP are $0/MWh and $0/MWh. The conventional generator is started in period 2 to serve 5 MW of demand and 5 MW of operating reserves. This causes the LMP to rise to $75/MWh based on the marginal cost of the generator. Once the conventional generator is online, it has no opportunity cost to provide operating reserves, which sets the ORP at $0/MW. In period 3, the reduction in wind output means that the system can no longer procure enough operating reserves. The ORP is set at the penalty cost of $10,000/MWh. The LMP is similarly raised to $10,075/MWh. This price is set by the marginal cost of the conventional generator ($75/MWh) plus its opportunity cost to provide operating reserves instead of energy. Note that, as a result, the wind generator has incentives to convert all of its available capacity to energy, and the conventional generator is indifferent between providing energy or reserves.

### 2. Lesson: Reliance on price spikes

The example shows the three key scenarios that will influence prices and investment incentives in a grid with high renewables. Period 1 shows the system in surplus, with enough available renewable capacity to serve all demand. The LMP at $0/MWh indicates that some amount of the renewable resources must be curtailed. In this case, the curtailed energy can be used as reserves, but the resources receive no payment for either energy or reserves.

Period 3 shows the other extreme, when the system enters a supply shortage that results in high prices. In this case, the shortage only affects operating reserves, but a case with unserved demand (i.e., blackouts) would show similar results. These periods would be relatively rare throughout the year but would be the energy market's primary influence on investment incentives.

Period 2 shows the intermediate case with the backup conventional generator providing enough energy to balance the system and provide operating reserves. The LMP in this scenario is set by the marginal cost of the conventional generator. Because the generator also has fixed costs for start-up and no-load, the standard ISO practice is to provide an out-of-market payment so that the generator recovers its financial loss. These payments (variously called revenue sufficiency guarantees, make-whole payments, or uplift) distort the market's economic signals because they are nontransparent and increase the amount that demand has to pay for energy. Another possibility is also that the backup generation will be on-line at its minimum output in order to provide operating reserves. In this case, the LMP and ORP will both be set at $0/MWh and $0/MWh, even though the conventional resources are dispatched, which may exacerbate the need for out-of-market payments.

## B. Example 2: Using storage to reduce renewable curtailment

Next, we add a storage device to the market. Adding this resource helps to both avoid the operating reserves shortage and reduce curtailment of the wind resource. For simplification, the device is assumed to have 100% round trip charge efficiency. The device parameters are given in Table below.

Table II.4 Storage Attributes

| Resource | Efficiency | Max charge capacity | Max discharge capacity | Max state-of-charge | SoC value, t=1 | SoC value, t=2 | SoC value, t=3 |
|---|---|---|---|---|---|---|---|
| **Battery** | 100% | 10 MW | 10 MW | 20 MWh | $0/MWh | $0/MWh | $40/MWh |

The storage device has a 4-hour charge/discharge cycle due to its respective charge rate, discharge rate, and maximum state-of-charge (SoC). The device also submits a bid for its valuation of SoC during each period, which allows it to tell the ISO that it expects that any stored energy from period 3 could be sold at $40/MWh during some later period. As will become more apparent after Examples 2 and 3, this storage bidding format is simpler than using offer curves for charging and discharging, i.e., analogous to generator offer curves, and this SoC-based offer is better suited for coordinating optimal storage utilization across multiple time horizons.

Storage resources are added to the market optimization problem by updating the constraints below:

| | | |
|---|---|---|
| Net value maximization: | $\max SoCValue - GenCost - PenaltyCost$ | |
| State-of-charge value: | $SoCValue = \sum_{t \in T} \sum_{b \in Bat} V_{bt} e_{bt}$ | |
| Power balance: | $D_t + \sum_{b \in Bat} (p_{bt}^c - p_{bt}^d) - \sum_{g \in Gen} p_{gt} = 0,$ | $\forall t \in T$ |
| Operating reserve req.: | $R_t - \sum_{b \in Bat} r_{bt} - \sum_{g \in Gen} r_{gt} - slack_t \leq 0,$ | $\forall t \in T$ |
| Charging maximum: | $p_{bt}^c \leq P_b^{\mathrm{c,max}} z_{bt},$ | $\forall b \in Bat, \forall t \in T$ |
| Discharging maximum: | $p_{bt}^d \leq P_b^{\mathrm{d,max}} (1 - z_{bt}),$ | $\forall b \in Bat, \forall t \in T$ |
| SoC management: | $e_{bt} = \eta_b p_{bt}^c - p_{bt}^d + e_{b,t-1},$ | $\forall b \in Bat, \forall t \in T$ |
| Reserve SoC: | $r_{bt} \leq e_{b,t-1}$ | $\forall b \in Bat, \forall t \in T$ |
| Reserve capacity: | $p_{bt}^d - p_{bt}^c + r_{bt} \leq P_b^{\mathrm{d,max}}$ | $\forall b \in Bat, \forall t \in T$ |
| Nonnegative variables: | $p_{gt}, p_{bt}^c, p_{bt}^d, e_{bt}, r_{gt}, r_{bt}, slack_t \geq 0$ | |
| Binary variables: | $x_{gt}, y_{gt}, z_{bt} \in \{0,1\}$ | |

Constraints for generator costs, reserve violation penalty, and generator minimum and maximum remain unchanged from the previous formulation. New variables are introduced for battery power charging $p_{bt}^c$, power discharge $p_{bt}^d$, energy state-of-charge $e_{bt}$, operating reserves $r_{bt}$, and charging status $z_{bt}$. Additional model parameters include the value of stored energy $V_{bt}$ at time $t$, maximum charging capacity $P_b^{\mathrm{c,max}}$, maximum discharging capacity $P_b^{\mathrm{d,max}}$, and round trip energy storage efficiency $\eta_b$. Note that the SoC bid represents the value (positive) of energy stored in the device, so the objective function is now formulated as a maximization of net value, i.e., SoC value minus production costs and reserve penalties. That is, the objective is no longer a cost minimization but a value maximization. The LMP and ORP are calculated from the dual variables of the power balance and operating reserve requirement constraints, i.e., the same as before.

### 1. Market Clearing

Results from the Example 2 market are shown in Table II.5 below.

Table II.5 Example 2 Market Results

| **Period:** | **1** | **2** | **3** |
|---|---|---|---|
| **Demand (MW)** | 60 | 70 | 67 |

| Period: | 1 | 2 | 3 |
|---|---|---|---|
| **Wind (MW)** | 70 | 65 | 55 |
| **Conventional (MW)** | 0 | 5 | 2 |
| **Battery (MW)** | -10 | 0 | 10 |
| **Wind curtailment (MW)** | -5 | 0 | 0 |
| **LMP ($/MWh)** | $0 | $75 | $75 |
| **OpRes (wind, MW)** | 5 | 0 | 0 |
| **OpRes (conventional, MW)** | 20 | 15 | 18 |
| **OpRes (battery, MW)** | 0 | 10 | 0 |
| **ORP ($/MWh)** | $0 | $0 | $0 |

The key change provided by the storage device is that it charges 10 MWh of previously curtailed wind energy from period 1 and the storage device delivers the 10 MWh in period 3 to avoid the operating reserves shortfall and reduce the reliance on the expensive conventional generator. LMP remains at $0/MWh in period 1 but falls to $75/MWh in period 3. The conventional generator is also started up earlier, in period 1, to provide operating reserves. However, this generator is able to supply a surplus of operating reserves in each period, so the ORP is now set to $0/MWh in each period.

## 2. Lesson: The need for updated reserve products

This example primarily shows the benefits of storage for smoothing out the intermittency of renewable generation, but it is also worth pointing out a few finer points about operating reserves that the example also demonstrates. First, we discuss the ability of storage devices to provide operating reserves, the potential benefits of an operating reserve demand curve, and an apparent paradox between conventional generator and storage utilization.

The battery entered period 1 with zero SoC, then charged 10 MWh in the period 1, held this SoC in period 2, and discharged it in period 3. Because the battery didn’t discharge its SoC during period 2, the battery was able to provide operating reserves instead. Requirements are currently evolving for batteries that provide operating reserves. In the simple example, it is only required that the battery has enough SoC from the previous period and enough excess discharge capacity in the current period, but it is not clear whether real-world ISOs would use the same requirements. Operating reserves from conventional generators do not need to specify a duration requirement, but storage devices will need to have clear definitions about the required discharge duration in the event that operating reserves are called by the ISO.

In addition, the storage reserve capacity constraint used above allows storage devices to provide more operating reserves than the device’s discharge capacity. This is possible because a storage device that switches from charging mode to discharging mode has a much larger dispatchable range than its raw discharge capacity. Future market designs will need to address how this wider dispatchable range should be compensated. The issue may not always be straightforward, for example, if the storage device is not continuously dispatchable from charging to discharging and/or if switching from charging to discharging requires a switching time during which the device remains idle (e.g., some pumped hydro storage).

ORPs are set at $0/MW in each period in the above example, but it could be beneficial for the market to set intermediate prices when operating reserves are close to but not less than the required quantity. For instance, the market may value additional operating reserves because they reduce the probability that the market will be short of reserves, or short of supply, during a future period. In the example, the market is able to procure 25 MW, 25 MW, and 18 MW in each respective period, but the ORP is $0/MWh in each. The $10,000/MWh operating reserves penalty only sets price when the system is short of reserves. Setting an operating reserve demand curve with values between $0 and $10,0000/MWh would help remove this steep increase and allow prices to rise and fall according to the amount of reserves.

The example also provides an apparent paradox with the utilization of conventional generation and storage resources. Generally, it should be expected that increased integration of storage resources will decrease the utilization of conventional resources since the storage will allow previously curtailed renewable energy to substitute for conventional generation during other periods. However, the conventional generation in this example must be committed for three periods instead of two in Example 1. This occurs because period 1 would otherwise be short of operating reserves; the conventional generator is committed in period 1 at its minimum operating level so that it can provide 20 MW of reserves. Storage cannot provide these reserves because it enters period 1 with no SoC, and wind cannot provide enough reserves because it is being used to charge the storage. Although total dispatch from the conventional generator is less in this example than in Example 1, additional constraints like a minimum output above zero could result in more conventional dispatch despite the additional storage resource.

Regarding the benefits and incentives to provide storage, this example shows that the additional storage resource was able to reduce the wind curtailment from 15 MWh to 5 MWh. The new device is also rewarded for its participation since it buys 10MWh in period 1 at $0/MWh and sells it in period 3 for $75/MWh for a profit of $750. Of course, this is much less profit than the battery owner may have expected if it had expected the LMPs and ORPs from Example 1 to prevail, in which case the battery owner would have expected a profit of $100,750. The fact that storage eliminated the operating reserves shortage also eliminated the price spike that may have signaled the need for new resources. The ability to avoid reserve shortages may be economically sensible, so the market design must take care to ensure that suppliers do not have an incentive to withhold capacity (or investment in new capacity) in order to maintain the high prices that occur during supply shortages. This incentive also exists in today's market designs and is unavoidable to some degree. Market monitoring and offer mitigation is required. Implementation of a gradually sloping operating reserves demand curve may help alleviate new market power concerns the costs of fuel-based resources are less likely to influence ORPs.

In a broader sense, the ability of storage (or other reserve resources) to provide operating reserves may require explicitly modeling the device's capabilities during scenarios with generator outages. This framework would essentially replace the deterministic optimization problem used here with a stochastic optimization problem. Current optimization techniques do not scale well enough to solve the size of stochastic problem required. It would also require a fundamentally new market design to determine market settlements in a scenario-based system.

### C. Example 3: Future grid with flexible demand and storage opportunity bidding

In the previous two examples, the conventional generator was necessary to maintain a feasible dispatch with sufficient operating reserves to maintain reliability. Participation from demand was ignored. The following example shows how demand participation can eliminate the need for conventional generation entirely while improving the total economic surplus created by the market.

The previously fixed demand data now participates by submitting a bid value for the energy consumed. Demand can also offer operating reserves. In contrast to the conventional generators, which have no opportunity cost to provide reserves, demand may offer reserves at higher prices to reflect the potential reduction in reliability (i.e., demand that provides reserves will be curtailed first if there is a supply shortage). For simplicity, we assume two segments of demand: a municipal load with relatively high marginal value and a flexible industrial load whose value is determined by the spot price of the goods produced by the factory. Parameters are shown below in Table II.6.

Table II.6 Demand Attributes

| | Total Load (MWh) | | | Marginal Value ($/MWh) | | | Reserve Cost ($/MW) | | |
|---|---|---|---|---|---|---|---|---|---|
| | t=1 | t=2 | t=3 | t=1 | t=2 | t=3 | t=1 | t=2 | t=3 |
| **Muni** | 40 | 50 | 47 | 900 | 900 | 900 | 100 | 100 | 100 |
| **Flex** | 20 | 20 | 20 | 50 | 50 | 50 | 10 | 10 | 10 |

Demand side resources are added to the market optimization problem by updating the constraints below:

| | | |
|---|---|---|
| Market surplus maximization: | $\max BidValue - SoCValue - GenCost - PenaltyCost$ | |
| Demand bid value: | $BidValue = \sum_{t \in T} \sum_{d \in Dem} (B_{dt} p_{dt} - C_{dt} r_{dt})$ | |
| Power balance: | $\sum_{d \in Dem} p_{bt} + \sum_{b \in Bat} (p_{bt}^{c} - p_{bt}^{d}) - \sum_{g \in Gen} p_{gt} = 0,$ | $\forall t \in T$ |
| Operating reserve req.: | $R_t - \sum_{d \in Dem} r_{dt} - \sum_{b \in Bat} r_{bt} - \sum_{g \in Gen} r_{gt} - slack_t \leq 0,$ | $\forall t \in T$ |
| Load maximum: | $p_{dt} \leq P_d^{\max},$ | $\forall d \in Dem, \forall t \in T$ |
| Load reserves: | $p_{dt} - r_{dt} \geq 0,$ | $\forall d \in Dem, \forall t \in T$ |
| Nonnegative variables: | $p_{gt}, p_{dt}, p_{bt}^{c}, p_{bt}^{d}, e_{bt}, r_{gt}, r_{dt}, r_{bt}, slack_t \geq 0$ | |

The above demand side participation model uses a conceptually simple model with no non-convexities or complementarity constraints. Each segment of demand submits a linear price and quantity bid that represents the most they are willing to pay for energy at time $t$, $B_{dt}$. The maximum load quantity is $P_d^{\max}$. In addition, there is a reserve cost offer, $C_{dt}$, which represents the price discount that the customer requires in order to provide operating reserves (i.e., to be curtailed first in case of a generator outage). Reserves provided by demand are capped by the total amount of load served. Whereas Example 2 maximized a "net value" objective that consisted of stored energy value minus generation costs, the above formulation maximizes total market surplus. This change in terminology reflects that all market participants are relying on the market to schedule their energy production and consumption. That is, energy demand was previously fixed, but can now be optimized by the market operator. The LMP and ORP are calculated from the dual variables of the power balance and operating reserve requirement constraints, i.e., the same as before.

### 1. Market Clearing

Results from the above market are shown in  below.

Table II.7 Example 3 Market Results

| **Period:** | **1** | **2** | **3** |
|---|---|---|---|
| **Demand (MW)** | 60 | 70 | 67 |
| **Wind (MW)** | 70 | 65 | 55 |
| **Conventional (MW)** | 0 | 0 | 0 |
| **Battery (MW)** | -10 | 0 | 0 |

| Period: | 1 | 2 | 3 |
|---|---|---|---|
| **Wind curtailment (MW)** | -5 | 0 | 0 |
| **Flex load curtailment (MW)** | 0 | -5 | -12 |
| **LMP ($/MWh)** | $10 | $50 | $50 |
| **OpRes (wind, MW)** | 5 | 0 | 0 |
| **OpRes (conventional, MW)** | 0 | 0 | 0 |
| **OpRes (battery, MW)** | 0 | 10 | 10 |
| **OpRes (flex load, MW)** | 5 | 0 | 0 |
| **ORP ($/MWh)** | $10 | $10 | $10 |

The addition of flexible demand completely eliminates the need for conventional generation. The key difference is that the flexible industrial load is able to provide reserves in period 1 instead of the conventional generator, and it curtails some of its load instead of relying on the conventional generator in periods 2 and 3. As a result, the demand is shaped to match available supply instead of shaping supply to match demand.

Prices in the above example are set by demand and by the opportunity costs associated with energy storage. In period 1, the LMP is $0/MWh due to the oversupply of wind energy, and the ORP is $10/MWh since the reserves are procured from the flexible industrial load. In periods 2 and 3, the LMP is set at $50/MWh because the flexible load curtailment is the marginal system resource. The ORP is again set at $10/MWh; however this price is not based on reserves from demand. The period 2 and 3 ORP is now determined by the opportunity cost of energy storage. Recall that the battery's SoC bid specified that it's SoC value at the end of period 3 was $40/MWh. The storage device has a $10/MWh incentive to discharge since the LMP in periods 2 and 3 is $50/MWh, so the ORP is set at $10/MWh to ensure that the battery's incentives are consistent with providing reserves instead of energy.

### 2. Lesson: Non-zero pricing with zero fuel costs

This example shows Four important concepts for an energy system with zero production costs. Prices in this example are determined by a combination of consumer demand and opportunity costs. First, we will explain the importance of a demand-side participation model that allows segmentation of energy consumers into groups with different energy bid valuations and reliability preferences. Second, the example highlights the need to reconsider reliability criteria used to determine operating reserve requirements as the system transitions away from conventional generation. Then we discuss the relationship between opportunity costs and corresponding LMP and ORP. Last, this section discusses the role of opportunity cost valuations to help coordinate efficient resource utilization across overlapping market clearing horizons.

The demand-side participation model in this example differs from typical demand response because it is based on the measured quantity of energy consumed by demand rather than the (unmeasured) energy not consumed in comparison to a consumer's baseline demand. The latter paradigm is required by FERC Order 745 for certain types of demand response, but ISO markets are also allowed (and do) have participation models that are similar to the former (see Section IV.C for additional discussion of Order 745). In the model presented for Example 3, the demand submits a bid value that reflects the maximum they are willing to pay for energy, and demand submits a reserve offer that reflects the price discount that they require to be used for operating reserves.

Consumers benefit in Example 3 compared to the results in Examples 1 and 2, even though there are load curtailments. Calculations for the total benefits to load are shown in Table II.8, below.

Table II.8 Comparison of Consumer's Surplus

| | Example 1 | Example 2 | Example 3 |
|---|---|---|---|
| **Muni economic surplus (energy)** | 47*(900-0)<br>+ 50*(900-75)<br>+ 47*(900-10075)<br>= -$347,675 | 47*(900-0)<br>+ 50*(900-75)<br>+ 47*(900-75)<br>= $122,325 | 47*(900-10)<br>+ 50*(900-50)<br>+ 47*(900-50)<br>= $124,280 |
| **Flex economic surplus (energy)** | 20*(50-0)<br>+ 20*(50-75)<br>+ 20*(50-10075)<br>= -$200,000 | 20*(50-0)<br>+ 20*(50-75)<br>+ 20*(50-75)<br>= $0 | 20*(50-10)<br>+ 20*(50-50)<br>+ 20*(50-50)<br>= $800 |
| **Operating reserves cost** | 15*0 + 15*0 + 8*10000 = $80,000 | 25*0 + 25*0 + 18*0 = $0 | 10*10 + 10*10 + 10*10 – 5*10 = $250 |
| **Total consumer surplus** | -$280,000 | $122,325 | $124,830 |

Each segment of demand receives its highest economic surplus from the results in Example 3. Notably, they each receive a very large *negative* surplus in Example 1 because the operating reserves penalty is set much higher than their actual valuation of reliability. In real markets, the operating reserves is based on econometric estimates for the value of lost load, but the penalty does not typically reflect the differential reliability valuations across diverse types of consumers. The flexible load received no net surplus from energy in Example 2, even though the price is $0/MWh in period 1, because the prices in periods 2 and 3 are greater than the load's willingness to pay for energy.

The flexible load only receives a positive economic surplus when it is able to submit its bid into the market, which allows it energy consumption to be scheduled only when sufficient renewable generation is available. The load curtailment also causes the LMP to rise to the demand's bid level, which allows the renewable generation to receive compensation from the energy market even though the conventional generator is out of the market. The municipal load, which was valued at $900/MWh, has no change in their schedule and also benefits from lower prices. As markets evolve to allow wider participation from demand-side resources, the participation model used in Example 3 may need to be revised to allow greater detail in how the demand's loads are modeled; the simple example above provide a starting point for how a more detailed demand-side participation model might be formulated.

Demand that is cleared as a reserve is still scheduled to consume energy but would be among the first loads curtailed in event of a system reliability event (e.g., generator outage). In a system dominated by renewable generation, reliability events are more likely to be caused by weather forecast errors than equipment failures, which may require a fundamental reconsideration of the reliability criteria that are currently used to determine operating reserve requirements. In addition, current practices typically allocate the cost of operating reserves to demand on a *pro rata* basis, but that standard practice may no longer make sense if reserves are primarily provided by the demand side. In such a future system, it may make more sense to allocate the cost of operating reserves to the intermittent supply-side resources that cause supply shortages when actual production is less than forecast. If this were the case, then the allocation of operating reserve costs could compel renewable resources to offer their energy at costs about

$0/MWh since resources with the most intermittency would need to offer higher costs to compensate for the additional reserve costs that they would impose.[1]

As alluded to in Example 2, the steep increase in prices when the system becomes “short” of operating reserves will make less and less sense as conventional generators become less dominant supply-side resources. It may make more sense to segment supply resources based on anticipated intermittency (e.g., probability of energy delivery) and to segment demand-side resources based on reliability valuations (e.g., acceptability of an outage probability). This framework would be very similar to the flexibility option idea proposed by Spyrou, *et al.,*[2] and discussed in Section III.C.2, and earlier ideas such as priority service (Chao and Wilson, 1987).[3]

Next, we discuss the importance of opportunity costs in setting market prices in a system with no fuel costs. Opportunity cost calculations often betray otherwise good intuitions or “rules” that people apply to market clearing prices. One such “rule” is that the LMP will be $0/MWh any time that a zero-cost resource is curtailed. Period 1 provides a counterexample and shows that, in fact, the effect of a renewable curtailment is that the LMP and ORP will be equal. This equality occurs because the curtailed renewable energy is being used to provide reserves.[4] In this case, both the LMP and ORP are set by the reserve offer of the flexible industrial demand. The causation of prices flips directions in periods 2 and 3, where the LMP is set by the price of a load curtailment. In that case, the ORP is determined by an opportunity cost calculation; a $40/MWh spread is necessary so that the battery will hold its SoC instead of delivering energy.

As more demand participates in the market, it will become much more common for energy market prices to reflect opportunity costs that are external to the energy market. For example, the flexible industrial customer’s reserve cost might reflect a combination of hard financial accounting (e.g., expected costs associated with being unable to deliver finished goods) and human preferences (e.g., the inconvenience of shutting down plant operations unexpectedly). Neither cost is currently understood to affect prices in electricity markets.

Opportunity costs also allow the market to coordinate between overlapping market clearing horizons. When different planning horizons are not coordinated, the market may run into issues such as not having the planned amount of stored energy. Good coordination across different time frames will become more important as more storage technologies enter the market. Many demand-side loads may behave like storage, as well, since a drop in energy consumption in one period may necessitate additional energy consumption in another period. Opportunity cost bidding allows market participants to express these “now or later” type decisions to the market operator.

[1] Reserves are currently mostly provided by supply-side generation, but note that in this example, the reserves are provided by consumers in the form of interruptible, i.e., less reliable, service. Therefore, “additional reserves” can be understood to be effectively the same as “lower reliability” to the extent that specific consumers are willing to accept the lower reliability in return for lower costs.

[2] Spyrou, E, M. Cai, Y. Liu, Y. Zhang, B. Hobbs, H. Geman, Y. Ma, R. Hytowitz, E. Ela, P. Hines, M. Almassalkhi, J. Kaminsky. “An Integrated Paradigm for the Management of Delivery Risk in Electricity Markets: From Batteries to Insurance and Beyond.” ARPA-E. Accessed May 3, 2022. https://www.arpa-e.energy.gov/sites/default/files/2021-02/NREL_PERFORM%20Kickoff_Final.pdf

[3] Chao, H. P., & Wilson, R. (1987). Priority service: Pricing, investment, and market organization. *The American Economic Review*, 899-916.

[4] Although reserves from renewable energy are uncommon in today’s markets, there is no fundamental reason why renewables can’t provide reserves. The potential for non-delivery is likely the major hurdle, but this can be addressed by appropriately derating the resource capacity that is used for reserves.

In the case of the battery in Example 3, this decision is whether to discharge the battery now (in periods 2 or 3), or to wait until after the current market clearing horizon. To help simplify how the battery operator might calculate its opportunity cost, suppose that it only considers that the system might be in energy surplus (LMP = $0/MWh) or an operating reserve shortfall (LMP = $10,000/MWh), and that it only considers what will happen in period 4 (i.e., there is no opportunity to discharge in period 3 and then recharge in period 4 for a later price spike). Then, the battery's storage valuation of $40/MWh reflects that it considers a $40/$10,000 = 0.4% probability of an impending operating reserve shortage. In a real market, such bidding behavior would reflect many independent forecasts of future conditions and would provide market participants with a high quality "crowd" solution for how to operate the system now to help avoid problems later. Example 3 reflects this idea since the battery is charged in period 1 but does not subsequently discharge the device, which cannot occur in production cost models that do not assume any exogenous opportunity costs.

# III. Short-term operations and modeling

Similar to the examples above, ISOs use various optimization-based software tools to determine resource schedules, dispatch quantities, and market clearing prices. Generator scheduling is the process of determining at what times a particular generator will produce electricity and is determined by the ISO's Security Constrained Unit Commitment (SCUC) software. Specific dispatch quantities are determined by a similar software tool called Security Constrained Economic Dispatch (SCED). Market clearing prices for electricity, called locational marginal prices (LMPs), are calculated by shadow prices (i.e., dual variables) that reflect the cost to serve an additional unit of demand at each location in the electric grid. Prices for other products such as energy reserves, also called ancillary services, are similarly determined by their shadow prices in the optimization software used by ISOs. The rest of this section describes how various resources are modeled in typical design of SCUC and SCED software and how new technologies and new operational challenges might lead to changes to modeling frameworks and market designs.

## A. Modeling of conventional generation

### 1. Deterministic unit commitment

SCUC software is designed around the operational constraints of conventional generators. Namely, it allows the market operator to efficiently schedule conventional generators that typically have considerable start-up times, minimum operating levels, and minimum run times. These features create non-convexity in the optimization problem, which results in computationally difficulties for the market scheduling software. In this section, we discuss how ISO scheduling software models the short-term operations of conventional generators.

Standard economics textbooks often use a continuous offer price and quantity curve to model market supply functions, and similarly, a continuous bid price and quantity curve to model market demand. Supply and demand are represented at a single point in time and at a single location. This standard economic model is too simplified for electricity markets due to the need to schedule generators over a 24-hour horizon and to maintain the reliability of an electric grid that spans a wide geographic area.

Rather than using the simple economic model of intersecting supply and demand curves, SCUC and SCED software instead solves a deterministic unit commitment problem that optimizes generator schedules while respecting the constraints for network reliability and generator start-up times, minimum operating levels, minimum run times, and fixed start-up and operating costs. The main inputs to SCUC and SCED typically consist of generator costs and feasible operating capabilities that are submitted as multi-part offers to the ISO, the demand forecast data for a specific planning horizon, and the transmission network topology. In addition to the simple price and quantity curves used in economics textbooks, ISOs also make extensive use of a multi-part offer format that can convey the minimum operating levels, minimum up and down times, and fixed costs aspects that are common features of conventional generators. In contrast, new types of resources – such as wind, solar, and battery storage – may have distinct characteristics that are not very well characterized in neither simple price and quantity curves nor the multi-part offer formats described above.

Similarly, the demand forecast data used in the day-ahead SCUC typically assumes a 36-48 hour horizon. This time horizon accommodates typical start-up notification times and minimum run times required for many conventional generators. Solving the SCUC problem over this long of a time horizon allows operators to ensure that such generators are not given start-up instructions when they are not really needed or when it would be inefficient to keep them online for their entire minimum run time. The time horizon is also built into the two-settlement (i.e., real-time and day-ahead) market design used by all of the ISOs in the US. The two settlement system allows generators to hedge their production schedules in the day ahead market and then make efficient adjustments in the real time imbalance market. If conventional generation is increasingly replaced by other resources, there may be opportunity to reassess the prevalence of the day-ahead hedging mechanism in favor of other hedging tools over different planning horizons that would be more relevant to the new resource types.

Because the day-ahead SCUC model solves a deterministic scheduling problem, it is important for the ISO to procure reserves that provide system flexibility to respond to real-time uncertainties.[1] Traditional reserve products such as operating reserves are required for system reliability by the North American Electric Reliability Corporation (NERC). These reserves ensure that enough generating capacity is online so that the loss of any generating facility (e.g., the largest generator) can be quickly replaced by other generators. Increasingly, ISOs such as CAISO, MISO, and SPP have implemented procurement mechanisms for additional reserves called ramping products that provide additional flexible generator capacity that is needed to respond to fluctuations in renewable generation. In contrast to the NERC requirements, the most severe generation contingencies will increasingly come from fluctuations in weather, which may have a combined effect larger than the largest conventional generator. Likewise, the reserve product definitions are premised on flexibility provided by conventional generators, yet electricity consumers may also be able to provide the needed flexibility, for example, by throttling the power to a hydrogen electrolysis plant or the amount of computer resources being used in a data center. Additional integration of energy storage technologies may also provide motivation to reassess the definition and procurement of reserve products, discussed in the next subsection.

[1] Federal Energy Regulatory Commission, "Energy and Ancillary Services Market Reforms to Address Changing System Needs." *Technical Conference Regarding Energy and Ancillary Services Markets.* Staff paper. Docket No. AD21-10-000. September 2021.

## 2. Stochastic unit commitment

Stochastic market clearing models may provide another avenue to improve the current process based on deterministic SCUC.[1] Rather than solving SCUC based on point forecast data in each time interval, stochastic approaches consider a range of possible scenarios that might unfold in future time periods so that the expected cost of operations is minimized. In contrast to solutions from deterministic SCUC, stochastic-based solutions only provide a dispatch schedule in the first immediate period, which is followed by a scheduling policy in future periods that depends on which scenario is realized. Adoption of a stochastic market clearing model has the potential to result in better utilization of resource flexibility to balance the uncertainty of renewable resources. However, the approach also creates many practical challenges and hurdles before it can be implemented as a market clearing framework.

A key issue with straightforward adoption of a stochastic market clearing model is the lack of consistency between the set of scenarios that are explicitly modeled and the actual set of scenarios that are realized in the real-time market. Since any real-time scenario that is actually realized is unlikely to have been in the explicit set of modeled scenarios, there is a fundamental lack of clarity regarding how to cash out financial positions that were based on scenarios that did not occur. The use of virtual bidders in the day-ahead market is currently considered somewhat of a "crowd solution" for likely outcomes of the real-time market.[2] Stochastic market clearing models would need to find alternative means to determine the set of modeled scenarios, which could create a conflict of interest if performed centrally by the ISO.

It is not clear if this task is easily decentralizable as in the case of virtual bidding.[3] Nonetheless, various groups have proposed new approaches for integrating stochastic frameworks or stochastic-like approaches into electricity market clearing models. Zavala, *et al.* shows that the stochastic approach can work, in principle, although they admit that the market implementation and optimization software could be cumbersome in practice.[4] The ARPA-E PERFORM program is currently funding projects that use various methods to integrate risk into electricity market clearing processes.[5] For example, one project proposes implementing a priority service model for demand response in conjunction with a two-stage stochastic optimization that forms the core of the real-time energy market.[6] Short of implementing an explicit stochastic modelling framework, the framework under development by NREL and Johns Hopkins University proposes a flexibility auction that matches the uncertainty and flexibility characteristics of various market participants to clear specially designed options contracts without substantially burdening the

---

[1] Möst, Dominik, and Dogan Keles. "A survey of stochastic modelling approaches for liberalised electricity markets." *European Journal of Operational Research* 207, no. 2 (2010): 543-556.

[2] Mather, Jonathan, Eilyan Bitar, and Kameshwar Poolla. "Virtual bidding: Equilibrium, learning, and the wisdom of crowds." *IFAC-PapersOnLine* 50, no. 1 (2017): 225-232.

[3] Bjørndal, Endre, Mette Bjørndal, Kjetil Midthun, and Asgeir Tomasgard. "Stochastic electricity dispatch: A challenge for market design." *Energy* 150 (2018): 992-1005.

[4] Zavala, Victor M., Kibaek Kim, Mihai Anitescu, and John Birge. "A stochastic electricity market clearing formulation with consistent pricing properties." *Operations Research* 65, no. 3 (2017): 557-576.

[5] ARPA-E. "Performance-based Energy Resource Feedback, Optimization, and Risk Management." US Department of Energy. Accessed May 3, 2022. https://arpa-e.energy.gov/technologies/programs/perform

[6] Chao, Hung-po. "Stochastic Market Auction Redesigned Trading System (SMARTS)." ARPA-E PERFORM Kick-off Meeting, December 2020. Accessed May 3, 2022. https://arpa-e.energy.gov/sites/default/files/2021-02/ETA_PERFORM%20Kickoff_Final.pdf

market clearing optimization software.[1] To be sure, there are many approaches to integrating uncertainty and flexibility characteristics into electricity market clearing frameworks, but to date there is no widespread agreement or best-practices on the topic.

## 3. Electricity market products

The main commodity traded in electricity markets is electric energy, measured in megawatt-hours (MWh). The electric energy commodity a ubiquitous aspect of electricity market design proposals by Fred Schweppe *et al.* beginning in the late 1970s through the 1980s that inspired the energy market restructuring that took place in the US throughout the 1990s and early 2000s.[2] Each MWh produced by a generator can be easily measured and has a cost associated with it that can be calculated by the generator's heat rate and the cost of fuel. Similarly, the amount of energy consumed by load is also easily measured, and these measurements easily correspond to the MW dispatch instructions, integrated over time, that ISOs send to generators.

Although MWh's are convenient for defining a standard electric energy commodity for trading, future system needs could become decoupled from a simple price per MWh. For a concrete example, the charging of an electric vehicle is not only a specific amount of MWh but also a time period within which the vehicle owner would like the vehicle to be charged. If, in addition to the simple MWh quantity, the vehicle owner is also able to specify the charging interval, this additional information can be used by a grid operator to choose the precise times that the vehicle will receive charge and will therefore contribute to grid flexibility. The value of this flexibility is not expressed in a simple price per MWh. This concept has also been explored in a more general sense through the concepts of power paths and swing contracts as proposed by Tesfatsion (2020), in which the market is designed for a specific process or capability instead of a simple measurable commodity.[3] Even in the case of power paths or similar proposals, the price per MWh is not removed from the market design but is supplemented with other payments and mechanisms that compensate resource flexibility.

In addition to the operating and ramping reserve products discussed in the previous section, other ancillary services such as reactive capability/voltage support, primary frequency response/inertia, and black start capability are currently a relatively minor economic concern in wholesale electricity markets, but they may become more important as conventional generators retire. Conventional generators provide AC power by inducing an electric current from the kinetic energy of a spinning rotor. AC power and the spinning mass respectively provide reactive power capability for voltage support and the ability to quickly respond to primary frequency deviations through the rotor's inertial response. Inverter based resources, such as renewables and battery storage, are also able to provide these services but must do so artificially through the design of power electronics. Whereas the inertial response of conventional generators is fairly well understood, having many power electronics devices

---

[1] Spyrou, E, M. Cai, Y. Liu, Y. Zhang, B. Hobbs, H. Geman, Y. Ma, R. Hytowitz, E. Ela, P. Hines, M. Almassalkhi, J. Kaminsky. "An Integrated Paradigm for the Management of Delivery Risk in Electricity Markets: From Batteries to Insurance and Beyond." ARPA-E. Accessed May 3, 2022. https://www.arpa-e.energy.gov/sites/default/files/2021-02/NREL_PERFORM%20Kickoff_Final.pdf

[2] Schweppe, Fred C., Michael C. Caramanis, Richard D. Tabors, and Roger E. Bohn. "Spot Pricing of Electricity." 1988.

[3] Tesfatsion, Leigh. *A New Swing-Contract Design for Wholesale Power Markets*. John Wiley & Sons, 2020.

providing artificial inertia may potentially result in complex interactions that may be more difficult to control.

Black start capability refers to the ability to re-energize the electric grid after a collapse and typically requires the controlled start-up of conventional generators in a specific sequence to prevent unintended line flow violations that could cause further damage to the grid.  Because renewable resources are less controllable and battery storage systems may have a limited energy supply, it may be challenging to procure sufficient black start capability in a future grid with significant conventional resource retirements.  As the traditional suppliers of these services begin to diminish, system operators may need to rethink how to efficiently procure these services from new resources.

In addition to adding to the computational complexity of the optimization problem, non-convexity also typically creates a duality gap in market clearing solutions, which implies a lack of uniform market clearing prices.[1]  This lack of market clearing prices creates a major hurdle for decentralized market coordination because there may be no set of contracts where no subset of market participants can form an independent coalition that is better off than when they are part of the broader market (i.e., the “empty core” problem in economics).  ISO markets help reduce this market coordination issue by providing various side-payments, variously called uplift, make-whole, or revenue sufficiency payments, to generators that are part of the optimal production schedule but would not otherwise have the correct price signals to be dispatched optimally.  That is, the ISO solves a centralized optimization problem that would otherwise face substantial coordination problems in a fully decentralized context.

Other difficulties of conventional generators also point in the direction of using centralized market mechanisms for wholesale electricity markets.  Although some specifics of these issues may change in the future, it is unlikely that they would disappear as the grid begins to rely on new technologies.  According to O’Neill *et al.* (2002),[2] these reasons include:

- Public good aspects of reliability and transmission expansion,
- Efficient pricing of transmission congestion, and
- Market power mitigation.

The public good aspect of reliability is that additional investment in generation capacity might be undertaken to ensure reliability for consumer loads that are most essential, but all consumers benefit from the additional reliability that results from the additional capacity investment.  Without centralized capacity mechanisms, there may be underinvestment in resources needed to maintain the capacity levels that provide a socially optimal level of reliability.  This picture of reliability, however, is premised on the use of conventional generation to provide all necessary capacity.  As conventional generators begin to retire, it may become more appropriate to adopt

[1] O'Neill, Richard P., Anya Castillo, Brent Eldridge, and Robin Broder Hytowitz. "Dual pricing algorithm in ISO markets." *IEEE Transactions on Power Systems* 32, no. 4 (2016): 3308-3310.
[2] O’Neill, Richard P., Udi Helman, Paul M. Sotkiewicz, Michael H. Rothkopf, and William R. Stewart. "Regulatory evolution, market design and unit commitment." *The Next Generation of Electric Power Unit Commitment Models* (2002): 15-37.

consumer-centric approaches, such as the priority service model used for implementing demand response programs.[1,2]

Transmission congestion exists due to the physical limitations of transmission capacity and the lack of controllability of power flow due to Kirchhoff's laws.[3] When a transmission element can no longer support additional power flow, LMPs are adjusted to be higher in areas where additional supply is limited by the presence of transmission congestion and lower in areas where additional supply would risk damage to the transmission constraint. These changes in LMPs to reflect transmission constraints is referred to as congestion pricing. Congestion pricing results in a surplus of money collected from load compared to the payments to generators. Financial transmission right (FTR) auctions reallocate this surplus and provide a geographic or spatial hedge against price differences across different parts of the network. FTRs are typically allocated to firms initially as Auction Revenue Rights (ARRs) based on historical generation and load portfolios and then optionally sold in the FTR auction. Changes in generation portfolios may incline more though into revisions to FTR market designs, especially considering tradeoffs between remote wind and solar generation versus distributed energy resources (DERs) that are placed close to load.

When power is traded bilaterally, transactions across pre-determined transmission zones are limited by an Available Transmission Capacity (ATC) calculation, which attempts to approximate the effect of transmission limits and congestion pricing in an ISO market. No ATC calculation can simultaneously ensure that all physically feasible inter-zonal transactions will be under the ATC limit and that all transactions that are under the ATC limit will not violate physical transmission constraints. As a result, reliance on ATC for decentralized bilateral transactions cannot generally support efficient market outcomes.

Market mitigation, as well, will be impacted by changes in resource mix. Traditional methods of market power monitoring include various tests and indices like the pivotal supplier test, conduct and impact tests, and Herfindahl-Hirschman Index (HHI). These tests and indices are premised on the supply-side of the market. Likewise, generators that fail various market mitigation procedures will typically have their market offers modified to a resource-specific cost-based offer. Demand will need to play a larger role in balancing the uncertainties of renewable generation, and therefore will have a larger role in setting market prices. An abundance of various consumer types from diverse sectors (e.g., steel production, hydrogen electrolysis, data centers, etc.) could potentially usher in a high degree of market competitiveness that reduces the need for market monitoring. Conversely, the demand-side part of the market could also become dominated by a small number of electricity supply firms and result in similar market monitoring needs as exists today.

## B. Modeling of new technologies

As discussed above, existing electricity markets are designed around the physical characteristics and operational needs of conventional generators. New resources, such as

---

[1] Chao, Hung-po, Shmuel S. Oren, Stephen A. Smith, and Robert B. Wilson. "Priority service: Market structure and competition." *The Energy Journal* 9, no. Special Issue 2 (1988).

[2] Oren, Shmuel S. "A historical perspective and business model for load response aggregation based on priority service." In *46th Hawaii International Conference on System Sciences*, pp. 2206-2214. IEEE, 2013.

[3] Hogan, William W. "Contract networks for electric power transmission." *Journal of regulatory economics* 4, no. 3 (1992): 211-242.

renewables, energy storage, and flexible demand participation, may have distinct characteristics that will motivate fundamental changes in the modeling and decision making made by electricity market operators. In particular, the 24-48 hour unit commitment problem may be less applicable to these resources, and other sorts of scheduling problems could be more appropriate.

In the case of solar and wind generation, the key issue is weather uncertainty. In short-term operations, solar energy may be very easy to accurately forecast on a sunny day but could be very difficult to forecast over a short horizon, e.g., <1 hour, during cloudy days. Wind resources have similar forecasting difficulties over short horizons. Scheduling these resources in the day-ahead market based on the best available forecast data opens the resource owners to financial risk if the weather forecast turns out to be inaccurate. As a result, many renewable resources do not meaningfully participate in the day ahead market and simply show up in the real-time market. Over longer periods, the total energy output of renewable resources is easier to forecast than the minute-to-minute availability needed in real-time markets.

Energy storage is inherently more controllable than renewables and can provide useful flexibility to balance uncertain renewable output. Like conventional generators, efficient storage utilization requires advance planning that might be best accomplished with a day ahead market. However, most energy storage technologies, with the exception of pumped hydro, operate very differently than conventional generators and will naturally have different operational constraints and key decision variables. Rather than being constrained by minimum run times, operating levels, and fixed operating costs, the operational constraints of energy storage are characterized by state-of-charge management, round-trip efficiency, and distinct charge and discharge statuses. Enhanced modeling to allow more accurate energy storage offers is an ongoing issue, for example, new proposals are undergoing stakeholder processes in CAISO.[1]

Demand response also possesses flexibility characteristics that can be used to balance the uncertainty from renewables. To date, most demand response programs have focused on controls for building HVAC, water heaters, clothes dryers, or other energy-intensive residential and commercial energy uses. These demand-side resources are typically located in low-voltage distribution systems and distributed among many individual end-users. Most ISOs also have a few industrial demand response participants such as energy-intensive aluminum smelting plants and paper mills. Industrial energy consumers have a broad potential for providing many grid services, as recent case studies have shown for various chemical plants[2,3] and data centers.[4] Industrial processes may be ideal candidates for exploring further participation in electricity markets because their operations are already driven primarily by economic interests, they are often already connected to the high-voltage transmission system, and they may stand to receive significant financial benefits by scheduling production when energy prices are lowest. However, ISO demand response models may not be well-suited for the diverse set of characteristics of different industrial processes, and potential demand

---

[1] CAISO. “Initiative: Energy storage enhancements.” Accessed May 4, 2022. https://stakeholdercenter.caiso.com/StakeholderInitiatives/Energy-storage-enhancements

[2] Otashu, Joannah I., and Michael Baldea. "Scheduling chemical processes for frequency regulation." *Applied Energy* 260 (2020): 114125.

[3] Tsay, Calvin, Ankur Kumar, Jesus Flores-Cerrillo, and Michael Baldea. "Optimal demand response scheduling of an industrial air separation unit using data-driven dynamic models." *Computers & Chemical Engineering* 126 (2019): 22-34.

[4] Radovanovic, Ana, Ross Koningstein, Ian Schneider, Bokan Chen, Alexandre Duarte, Binz Roy, Diyue Xiao et al. "Carbon-aware computing for datacenters." *arXiv preprint arXiv:2106.11750* (2021).

response providers may see significant risks associated with participation in spot energy markets.

It is not immediately clear which operational models or market designs would most efficiently coordinate and schedule the diverse set of resources that might contribute to the future electric grid. New technologies offer a wide range of operational characteristics and have various relevant planning horizons that will often differ from conventional generators and may not be captured by existing market designs. The best ways to integrate these resources into grid operational models and market design is a far-reaching topic that will require careful consideration. The necessary improvements to current practices can certainly be approached as piecemeal modifications to current market designs. In the longer term, comprehensive and fundamental redesigns could be needed but would need to be supported by extensive and rigorous analysis.

# IV. Market design alternatives and modeling frameworks

The two-settlement, day-ahead and real-time market design sits at the foundation of today's electricity markets and is based around the operational characteristics of conventional generators. It is also well-suited to electricity markets due to the regular, daily fluctuations in energy usage and the need to procure enough energy capacity to supply peak daily needs. As renewable resources become more dominant in the energy supply, the day-ahead market could be modified with new products and possibly even new or alternative clearing mechanisms that create more efficient utilization of energy resources. A few of the possible alternatives are described below.

## A. Intraday markets

Because forecasting data becomes more detailed and accurate closer to real-time, one market design solution would be to increase the number of settlement periods between the day-ahead and the real-time markets.[1] Intraday markets would introduce a multi-settlement system into the ISO market design, as opposed to the current two-settlement system. More frequent settlement intervals would allow ISOs to schedule resources as more accurate forecast data becomes available. ISOs currently perform similar procedures through various residual, look-ahead, and real-time unit commitment software tools, but the output of these tools is typically only advisory information for the operator, i.e., not financially binding for market participants. In an intraday market, market participants would be able to send updated offer data to the market operator. As a result, there would be more opportunity to make adjustments to the day-ahead schedules and less need to over-commit conventional generators that might not be needed.

A move towards intraday markets could facilitate alternative commitment and dispatch mechanisms as well. The current two-settlement system relies on the ISO to make most commitment and dispatch decisions, and self-commitments and self-scheduling are mostly considered unimportant towards the economics of the overall system. This system could continue in an intraday market, with the ISO market clearing software being relied upon to determine a larger number of production schedules and market settlements to support the larger number of settlement periods.

[1] Herrero, Ignacio, Pablo Rodilla, and Carlos Batlle. "Enhancing intraday price signals in US ISO markets for a better integration of variable energy resources." *The Energy Journal* 39, no. 3 (2018).

Alternatively, this computational challenge could be avoided by significantly reducing the complexity of existing commitment and dispatch models. ISO-NE's chief economist Matt White has discussed some of the advantages of this approach, relying on forward price discovery to resolve non-convex scheduling decisions instead of centralized optimization.[1] With enough forward trading opportunity, market participants could self-commit and de-commit themselves to maximize profits and internally resolve any non-convex operating constraints. The result may be more efficient than existing methods based on deterministic SCUC, and possibly better than centralized stochastic unit commitment, since uncertainties could be addressed by an ensemble of crowd-based solutions. Decentralized approaches are discussed again in Section IV.B below.

## B. Decentralized markets

Market designs can be decentralized to varying degrees, but electricity markets will almost certainly require some central authority to help manage transmission congestion and maintain grid reliability. The main distinction from today's ISO markets is that the central market authority would not attempt to resolve the internal constraints of individual resources, but rather the resource owners would be responsible for the feasibility of their production schedule. That is, bid and offer formats are greatly simplified in a decentralized market. Decentralized markets can include bilateral markets – in which all traded quantities and prices are determined only by individual buyer and seller – as well as markets with uniform market clearing prices that are determined through a centralized auction mechanism. As previously discussed in Section IV.A above, decentralized frameworks have potential to work very well in markets with intraday trading.

Double auction and clearinghouse markets typically only allow simple price and quantity bids and offers. In the double auction, also called a continuous double auction, participants may post their bids or offers to the market interface or choose to "accept" any of the posted bids or offers at any time.[2] The clearinghouse design, also called a discrete double auction, is similar except that bids and offers are collected and cleared by the clearinghouse at some regular interval. Because the trades are not bilateral, the clearinghouse design typically uses uniform prices and tiebreaker rules. These methods are considered decentralized because the complexities of the market schedule – e.g., weather uncertainty, non-convexities of conventional generators, charging and discharging patterns for batteries – need to be considered by individual market participants submitting bids and offers rather than by the ISO.

Decentralized market designs do not perform many of the existing functions of ISOs, so they may have difficulty in efficiently scheduling the output of block-loaded conventional generators, satisfying the power flow limits of the transmission system, and procuring sufficient reliable capacity. Decisions to adopt a decentralized market design should therefore only be made after sufficient studies and experience show that the market is able to satisfy physical constraints of the grid and of individual resources. This is not a hopeless endeavor; there is historical precedent for decentralized power exchanges in California.[3] The recently-proposed Southeast

---

[1] White, Matt. "Reducing Complexity with Forward Price Discovery: I Power Really 'Unique'?: An Economic Perspective on the Complexity/Simplicity Issue in Electricity Market Design." IEEE PES Conference. Washington, DC. July 2014.
[2] Friedman, Daniel. "The double auction market institution: A survey." *The double auction market: Institutions, theories, and evidence* 14 (1993): 3-25.
[3] Wilson, Robert. "Activity rules for the power exchange." *Report to the California trust for power industry restructuring* 3 (1997).

Energy Exchange Market (SEEM)[1] is currently seeking FERC approval and is based on an earlier "split savings" pricing rule used in by a power exchange that operated in Florida during the 1980s.[2] The main benefit of these designs is the potential for a very low administrative overhead for running the auction mechanism, but there are also many potential issues related to reliability, real time power balance, and market power mitigation that would need to addressed.

### C. Flexibility options

Stemming from the natural complementarity between uncertain renewable generation and flexible resources like storage and DERs, another idea is to implement a new hedging product called a flexibility option in the day ahead market.[3] Flexibility options would match the uncertainty characteristics of renewable resources with flexibility provided by other resources. The flexibility option product is designed as a hedge in which a renewable generator expresses the amount of energy it will be able to produce at various probability-based tiers. Those uncertain quantities are balanced with flexible capacity from other resources that agree to inject additional energy or curtail demand if needed. Adding this product to the day ahead market would allow wind and solar resources to hedge their risks by receiving payment for expected production in the day-ahead market instead of relying on actual production and volatile spot prices. Flexible resources would benefit as well by receiving an option payment to compensate their flexible characteristics.

### D. Swing contracts

Swing contracts are motivated by similar flexibility-based concerns. Whereas the current two-settlement system is based on energy forward products, swing contracts use the day-ahead market to procure "power paths" from market participants. These power paths are defined mathematically as the set of possible production schedules that a generator would be able to feasibly take over a specific time period. Each resource that is awarded a swing contract receives a two-part payment: one performance component based on how the device is dispatched in real time, and one reservation component that covers any fixed costs that may be needed to ensure that the unit is online and available. The swing contract design is extensively described in recent work by Tesfatsion.[4]

## V. Out-of-market policies and agreements

Wholesale electricity markets do not operate in a vacuum and are often affected by policies and agreements made outside of the market. This section briefly describes the relationship of ISO

---

[1] Utility Dive. 2021. "FERC deems Duke, Southern SEEM proposal 'deficient', sends utilities back to the drawing board." Accessed May 4, 2022. https://www.utilitydive.com/news/ferc-deems-duke-southern-seem-proposal-deficient-sends-utilities-back-t/600015/

[2] Hahn, Robert W., and Mark V. Van Boening. "An experimental examination of spot markets for electricity." *The Economic Journal* 100, no. 403 (1990): 1073-1094.

[3] Spyrou, E, M. Cai, Y. Liu, Y. Zhang, B. Hobbs, H. Geman, Y. Ma, R. Hytowitz, E. Ela, P. Hines, M. Almassalkhi, J. Kaminsky. "An Integrated Paradigm for the Management of Delivery Risk in Electricity Markets: From Batteries to Insurance and Beyond." ARPA-E. Accessed May 3, 2022. https://www.arpa-e.energy.gov/sites/default/files/2021-02/NREL_PERFORM%20Kickoff_Final.pdf

[4] Tesfatsion, Leigh. *A New Swing-Contract Design for Wholesale Power Markets*. John Wiley & Sons, 2020.

market clearing outcomes with power purchase agreements (PPAs), carbon pricing and other environmental regulations, and retail electricity pricing policies.

## A. Power purchase agreements

PPAs allow longer term contracting terms than what is offered in the energy and capacity markets operated by ISOs, often 10 to 20 years in length. There are multiple types of PPAs, and an in-depth review is not attempted here. In simplified terms, PPAs can be structured in energy blocks, in which the energy supplier provides a fixed quantity during specific time periods, or it can be unit contingent, in which the buyer purchases whatever quantity the supplier is able to produce. An energy block PPA, for example, might be useful for a vertically integrated utility that needs additional energy supply to ensure that it can meet demand during daytime peak periods. Unit contingent PPAs might be purchased by utilities seeking to meet state renewable portfolio standards, or, more recently, companies that want to decrease their carbon footprint.[1]

The terms of a PPA are typically based on clearing prices in an ISO market. For example, suppose a buyer and wind farm (the seller) agree to a unit contingent PPA with a $25/MWh price, and the seller produces 100 MWh over a period when the LMP is $30/MWh. The wind farm would receive $3,000 in energy revenue from the ISO. The terms of the PPA would then be for the wind farm to pay $5/MWh, or $500, back to the buyer of the PPA. Conversely, if the LMP were $20/MWh, then the wind farm would receive $500 from the buyer. The outcome is that the wind farm will always receive $25 per MWh that it produces, and the buyer has a partial hedge at $25/MWh.

Both types of PPAs are used for hedging. An energy producer, such as the wind farm, receives a guaranteed cash flow for its energy production that does not depend on volatile spot energy prices, and the buyer of the PPA also benefits by being insulated from price spikes.

Now, consider how a PPA might affect an energy producer's incentives for offering its resource to the ISO. In an energy block PPA, suppose the supplier has a capacity of 20 MW and has agreed to supply 10 MW for one hour for at a price of $25/MWh. If the producer generates 10 MWh to fulfill the contract, it makes a profit of $25 minus its marginal production cost. Alternatively, the producer could also fulfill its contract by purchasing 10 MWh from the ISO. For example, this would be more profitable for the producer if the LMP is less than the power plant's marginal production cost. The breakeven point for the producer is therefore to offer its PPA contract quantity at its marginal cost so that it can either take advantage of cheaper energy from the market when prices are below its marginal cost or increase its output and earn additional profits when prices are above its marginal cost. The PPA in this case may help improve market efficiency since it encourages the producer to offer the PPA quantity the true marginal cost.

A unit contingent PPA, however, is based on the physical output of the plant and does not have this option to fulfill its obligation by purchasing power on the spot market. In contrast to an energy block PPA, a wind or solar farm with a unit contingent PPA will lose revenue any time

[1] Kobus, James, Ali Nasrallah, and Jim Guidera. "The Role of Corporate Renewable Power Purchase Agreements in Supporting Wind and Solar Deployment." *Center on Global Energy Policy, Columbia University* (2021).

the plant's output is curtailed. Effectively, these resources become "must take" resources.[1] Their incentive is to offer their resource at the lowest possible price, the offer floor, to avoid curtailment from the ISO. This could create a scenario where energy spot prices become negative for long periods of time when there is an excess of energy available from renewable resources and energy from these PPA contracts needs to be curtailed. Negative prices imply that consumers should be paid to increase their energy consumption in order to reduce the surplus of renewable energy. If these conditions occurred regularly, it would begin to undermine the financial viability of purchasing a PPA since consumers would be better off relying on low or negative spot prices.

## B. Carbon pricing and other federal and state environmental regulations

Some form of carbon pricing has been implemented in California and in the northeastern states[2] that are part of the Regional Greenhouse Gas Initiative (RGGI). Carbon pricing is conceptually simple to include in the offer-based auction framework used by ISOs. For example, suppose a generator has to purchase $CO_2$ allowances at $10/ton, and it produces 500 kg $CO_2$ per MWh. Then its marginal cost offer would be adjusted as follows:

$$\frac{\$10}{\text{ton CO}_2} * \frac{1 \text{ ton CO}_2}{907 \text{ kg CO}_2} * \frac{500 \text{ kg CO}_2}{\text{MWh}} = \$5.51/\text{MWh}$$

Resources that produce more $CO_2$ per MWh would have a proportionally larger increase to their marginal cost, and carbon-free resources would not have any change to their marginal cost.

However, one issue in organized markets is that carbon pricing is not used in all regions of the US. Carbon leakage occurs when the resources in one area are subject to carbon pricing but resources in a neighboring area are not. Because carbon prices are not uniform, the market software may choose to dispatch more polluting resources in areas that are not subject to carbon prices, resulting in less environmental benefit than desired. CAISO applies a border adjustment to out-of-state generation that is deemed to be used to serve California load in the EIM. In the ISO markets that include RGGI states (i.e., PJM, NYISO, and ISO-NE), carbon pricing adders are only applied to the generators that are required to purchase emissions allowances.[3]

## C. Retail electricity pricing

FERC Order 745 required ISOs to compensate demand response programs for the reduction in power consumption at the same price that is paid to power producers. Specifically, demand response is paid the LMP for each MWh of curtailment from their baseline consumption level.[4] Because the demand response provider is not obligated to purchase the energy before selling

[1] Didsayabutra, Ponpranod, Wei-Jen Lee, and Bundhit Eua-Arporn. "Defining the must-run and must-take units in a deregulated market." *IEEE Transactions on Industry Applications* 38.2 (2002): 596-601.

[2] Connecticut, Delaware, Maine, Maryland, Massachusetts, New Hampshire, New Jersey, New York, Rhode Island, Vermont, and Virginia are members of RGGI.

[3] Butner, M., B. D. Noll, J. Gundlach, B. Unel, and A. Zevin. "Carbon Pricing in Wholesale Electricity Markets: An Economic and Legal Guide." *Institute for Policy Integrity, New York University School of Law.* https://policyintegrity.org/publications/detail/carbon-pricing-inwholesale-electricity-markets (2020).

[4] FERC, "Demand Response Compensation in Organized Wholesale Energy Markets," Docket No. RM10-17-000; Order No. 745. March 15, 2011. Link: https://www.ferc.gov/sites/default/files/2020-06/Order-745.pdf

the response into the wholesale market, their incentives for providing demand response are not simply a factor of the LMP but also involve the retail tariff rate and the consumer's marginal value of energy consumption.  Table 1 below shows how the incentives to dispatch demand response jointly depend on the LMP, the retail rate, G (assumed greater than zero), and the marginal value of energy consumption, V.  A demand response provider has incentive to inefficiently curtail load if the marginal value of energy consumption is greater than the LMP but less than LMP plus the retail rate.

Table IV.1: Combined wholesale and retail market incentives for demand response under FERC Order 745

| Wholesale Price Signal | Wholesale and Retail Price Signal | DR Decision | Market Efficiency |
|---|---|---|---|
| LMP > V | LMP + G > V | Dispatch DR | Efficient |
| LMP > V | V > LMP + G | N/A | N/A |
| V > LMP | LMP + G > V | Dispatch DR | Inefficient |
| V > LMP | V > LMP + G | No DR | Efficient |

Incentives for efficient dispatch of demand response can be restored by compensating demand response at the price "LMP – G", which would align the wholesale and retail incentives shown in Table 1. This would provide the same incentives as if demand was required to purchase an energy forward purchase before demand response can be provided.  There is a broad consensus among economists that the economically efficient price for demand response takes some form of this "LMP – G" scheme.[1]  Although the price G was described above simply as the retail rate, its precise definition is a factor of the expected wholesale price and fixed network costs that make up the retail rate and has been subject to considerable debate in the context of efficient demand response compensation.[2]  Paying demand response the LMP as required by FERC Order 745 therefore overcompensates these resources under current policies.

To avoid market distortions caused by overcompensation, demand response resources are subject to the Net Benefits Test, which sets a price floor below which the RTO/ISO does not dispatch demand response.[3]  The price threshold for the net benefits test is updated monthly and is calculated by determining the price at which the benefits to consumers become greater than the payments to demand response providers.[4]  By design, the test assesses consumer surplus rather than the total market surplus (the sum of consumer and producer surplus), so it is not capable of determining whether dispatching demand response will improve market efficiency.[5]  Further, the net benefits test uses broadly aggregated metrics that may mis-identify

---

[1] R. L. Borlick, J. Bowring, J. Bushnell, P. A. Centolella, H.-P. Chao, A. Faruqui, M. Giberson, D. Gonatas, S. Harvey, B. F. Hobbs, W. W. Hogan, J. P. Kalt, R. J. Michaels, S. S. Oren, D. B. Patton, C. Pirrong, S. L. Pope, L. E. Ruff, R. Schmalensee, R. J. Shanker, V. L. Smith, and R. D. Tabors. BRIEF OF ROBERT L. BORLICK, JOSEPH BOWRING, JAMES BUSHNELL, AND 18 OTHER LEADING ECONOMISTS AS AMICI CURIAE IN SUPPORT OF PETITIONERS, 2012. Link: [http://www.scotusblog.com/wp-content/uploads/2015/09/2015-09-09-SCOTUS_EconomistsBriefDR.pdf](http://www.scotusblog.com/wp-content/uploads/2015/09/2015-09-09-SCOTUS_EconomistsBriefDR.pdf)

[2] Chao, Hung-po, and Mario DePillis. "Incentive effects of paying demand response in wholesale electricity markets." Journal of Regulatory Economics 43.3 (2013): 265-283.

[3] FERC, Order No. 745. At P. 4.

[4] Das, Chhandita, "Net Benefit Test Methodology – FERC Order 745," Market Information Working Group, New York ISO. December 2019. Link: [https://www.nyiso.com/documents/20142/3832196/745_Methodology_MIWG_NYISO.pdf/](https://www.nyiso.com/documents/20142/3832196/745_Methodology_MIWG_NYISO.pdf/)

[5] Hogan. (2009).

beneficiaries due to, for example, lack of granularity and lack of information about forward contracts that may limit consumer exposure to LMPs.[1] The need for the Net Benefits Test arises because the LMP creates inefficiently high incentives for demand response, and reliance on the net benefits test could likely be avoided by reforming demand response compensation methods.

However, FERC justified Order 745's demand response requirements as a balance of policy judgements that need not strictly follow from textbook economic analysis.[2] Indeed, the market distortion may be limited given the relatively small amount of demand response that currently participates in the market, and the overcompensation may help spur development of new demand-side resources that may be needed in the longer term. On the other hand, the current demand response baseline methodology and need for a Net Benefits Test may overcomplicate market participation and end up discouraging new entry from demand-side resources that are used to simple volumetric retail tariffs.

# VI. Discussion: The role of production cost models

The topic of this discussion paper was precipitated by concerns that today's production cost modeling tools, such as SCUC and SCED, will become increasingly inadequate or unable to model the operations and economics of a fully decarbonized electric grid. Whereas SCUC and SCED are used for market clearing and operations on short time scales, many other types of production cost modeling tools are used to simulate the effects of SCUC and SCED at longer time scales – such as in long term planning models for capacity and transmission expansion. Simple examples in Section II illustrated several market design issues that will arise as the grid begins to depend on new technologies. The key insight is that every resource type has unique operational characteristics, opportunity costs, and intermittency/uncertainties that may require different approaches to both production scheduling and price formation. In addition to the internal complexities of resources in the market, out-of-market policies and agreements also shape the decisions made in electricity markets. Efficient market design requires addressing all of the above aspects in a simple and transparent coordination mechanism. We conclude this paper by discussing how the anticipated increase in low and zero marginal cost resources – and market designs to support them – will affect production cost models for short operations and long term planning.

## A. Short-term operations models

The current, typical, modeling paradigm for short term model of the electricity market might be characterized as a deterministic SCUC model with one-hour time intervals, i.e., the model used for scheduling in the day-ahead market. Real time electricity markets use similar modeling tools, albeit with shorter time intervals and shorter planning horizons. Although these models solve deterministic optimization problems, they typically include various reserve products that procure flexible resource capacity that allows the market to respond to unforeseen generator outages and variability in renewable energy production.

---

[1] Ott, Andrew, "Statement of Andrew L. Ott," Panel discussion at the Commission's Technical Conference, Demand Compensation in Organized Wholesale Energy Markets. Federal Energy Regulatory Commission. Docket No. RM10-17-000. September 13, 2010.

[2] FERC, "Demand Response Compensation in Organized Wholesale Energy Markets," Docket No. RM10-17-000; Order No. 745. At P. 46. March 15, 2011. Link: https://www.ferc.gov/sites/default/files/2020-06/Order-745.pdf

As the electric grid becomes more decarbonized, much of the energy and flexibility currently provided by conventional generators will instead need to be provided by some combination of renewable resources, energy storage technologies, and flexible demand-side participation. Renewable energy has zero marginal cost and therefore is unlikely to be the marginal resource in most hours. When renewable energy does need to be curtailed, LMPs will fall to zero (or possibly negative), which may have economic consequences for what types of resources decide to enter or exit the market. As it pertains to the adequacy of short term operations models, however, renewable resources do not pose any obvious difficulty to conventional SCUC formulations because they do not have the same non-convex physical constraints and cost structure as conventional generators. Periods with surplus renewable generation may see prices fall to or below zero, but such a market outcome would send the correct economic signals to balance supply and demand.

Energy storage devices are more complex to model and may require enhancements to the current SCUC formulations. As previously described in Section III.B, modeling these technologies requires new constraints for state-of-charge, round trip efficiency, and charge and discharge logic that are not considered for conventional generators. Short-term models may face difficulty in efficiently scheduling storage resources if the planning horizon is too short or if the specific storage technology has highly nonlinear charging and discharging behavior. Likewise, because of the intertemporal nature of storage devices, conventional rolling-horizon models used in real-time markets may fail to provide adequate compensation to storage devices. The necessary enhancements to short-term operations models and market design considerations for storage is likely to be an important area for future research.

The ideal role of the demand side in operations models is not very clear. Demand-side resources are comprised of a very wide array of physical characteristics, including many different residential, commercial, and industrial uses for electric energy. Whether demand-side resources should bid using application-specific participation models or more generic bid formats is an open question. It is similarly unclear whether efficient demand-side participation is better supported through participation from individual end-users or through aggregators that act as a go-between for end-users and the market operator. A compounding issue is how or whether existing demand response compensation schemes, especially those required by FERC Order 745, may affect efficient demand-side participation. Almost certainly, demand-side flexibility will be an important aspect of large-scale renewable integration and will be a significant departure from the current modelling paradigm that assumed fixed demand.

As the above types of resources become more dominant in grid operations, there will be a shift in what types of optimization tools will be most appropriate for ensuring reliable and economically efficient wholesale power markets. This may result in new market participation models that need to be implemented in market clearing software and new configurations for the planning horizons, dispatch intervals, and treatment of uncertainty in the market clearing model formulations.

SCUC and SCED models currently use an objective that minimizes production costs. However, this shorthand description eludes broader components of the objective, which also includes benefit from energy consumption, benefits from the procurement of reserves, and penalties for violating the physical limits of transmission lines, transformers, or other elements of the electric grid. When this broader conception of the short-term modeling objective is considered, the objective is more appropriately a maximization of total market surplus. Therefore, as more conventional resources are retired, there will be more opportunity for market clearing prices to reflect values from of a broader range of resources and uses of electric power.

## B. Long-term planning models

Long-term planning models use simplifications of the short-term operations models to approximate and project the economic behavior of the grid through future changes in resource mix or other developments. Details of the non-convex production characteristics of conventional generators are typically simplified with convex approximations, and detailed models of AC power flow and various other operator decisions are also simplified or ignored. Nonetheless, long-term planning models are widely used to provide economic projections of price signals and what types of resources will be utilized based on our assumptions about the future. The following section describes some of the challenges that long-term models may face as conventional resources are replaced with new technologies.

Long-term planning models will face some of the same difficulties as short-term models in characterizing how renewables, storage, and demand-side participation might affect system dispatch and pricing. In the case of long-term models, the use of deterministic optimization algorithms may paint an overly optimistic picture of system dispatch. Deterministic tools may treat hourly renewable output as given rather than uncertain, so these models can utilize resources to always balance supply and demand in the most efficient way. As a result, model outputs might show that expensive backup resources are not needed, reserves that are less valuable, and prices that are less volatile than would occur if the model was able to better reflect the effects of uncertainty. Similarly, storage resources in a deterministic model may appear less valuable than they actually are because the model will tend to not place any value on storage capacity that is not utilized. In reality, storage capacity that is unused in a deterministic model may still have option value due to the presence of forecasting errors.

As described for short-term modeling, the proper role and correct modeling of demand-side participation are significant unknowns. Those issues are compounded in long-term models. A long term model not only needs to account for the flexibility that demand-side might offer in spot markets but also how energy prices might affect long-term shifts in energy usage. Shifts in demand participation might occur in two directions: first, increases in energy usage when wholesale prices are low, and second, decreases in energy usage in response to price spikes. Neither response can be adequately modeled when demand is assumed to be fixed.

Increased demand-side participation in wholesale electricity markets depends on the value of the energy consumed, technological barriers to entry, and regulatory policies. For industrial end-users, the value of electricity might be based on various commodity prices and the cost of running their production processes. The value to commercial end-users may depend on their expected daily profit, i.e., the value of keeping their doors open for business. Residential users may have more subjective values that might be based on the comfort of their AC settings or convenience of when their EV is charged. How should this multitude of information be included in a long-term planning model? More broadly, how should a long-term model reflect the market participation from entities that currently do not meaningfully participate? Possibly the correct answer is “not at all,” and the model outputs should be used to guide the development of new technologies and removal of regulatory barriers that currently prevent wider demand-side participation.

**Pacific Northwest National Laboratory**

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99354
1-888-375-PNNL (7665)

***www.pnnl.gov***