Increasing Electricity Demand Flexibility: Approaches and Gaps

Fredrich Kahrl, Ph.D.

January 2023

pacific energy INSTITUTE

Acknowledgments

This paper was authored by Dr. Fredrich (Fritz) Kahrl. The author is grateful for the review and feedback provided by the Pacific Energy Institute Fellows for their collaboration on the ideas that informed this paper.

https://pacificenergyinstitute.org/

Disclaimer

The views and opinions of the authors expressed herein do not necessarily state or reflect those of the Pacific Energy Institute's Fellows and Advisory Board, or their respective organizations.

Contents

Executive Summary

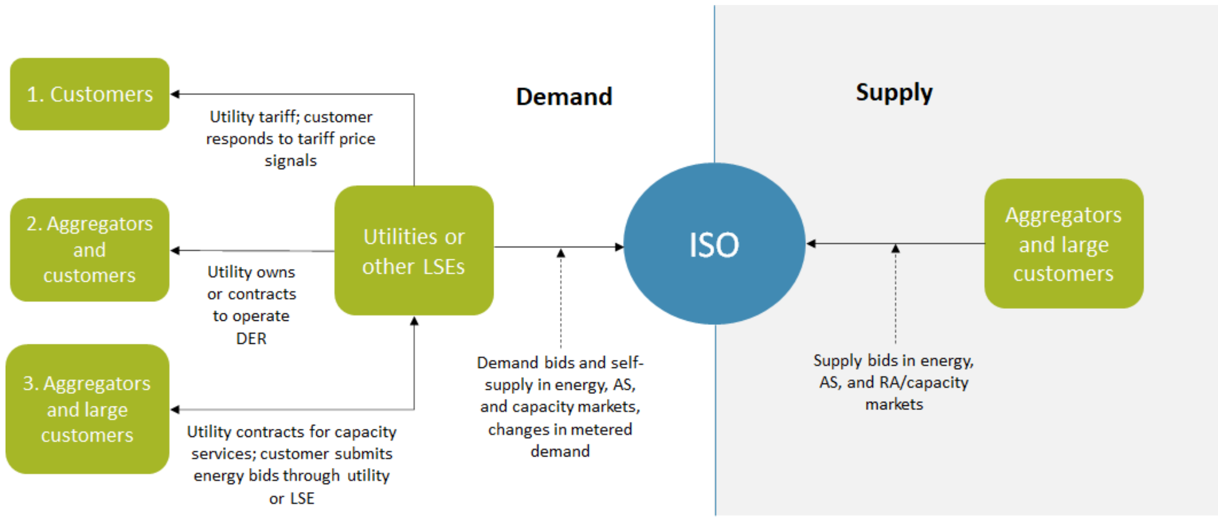
Limited flexibility – in both supply and demand – has been a longstanding feature of electricity systems. As electricity supply becomes more reliant on weather-dependent resources, and as electrification further increases the seasonality of electricity demand, limits in flexibility will lead to higher costs.

Due to recent technological innovation and falling costs, more flexible electricity demand is increasingly feasible. But there are numerous questions about how it might actually work in practice, within the current structure of U.S. wholesale markets. How would distribution-level congestion be managed? How would distributed energy resources (DERs) be credited and compensated for providing resource adequacy services? How would these resources participate in, and be compensated for, services provided in wholesale energy and ancillary services markets?

This paper examines different models for demand flexibility, focusing on DER participation through the demand, rather than the supply, side of wholesale markets. It investigates four models in detail, using numerical examples to explore how they might work in practice. These four models differ based on which entity is optimizing resource operations in the wholesale market. They include:

(1) **Customer tariff optimized model**, in which customers respond to cost-reflective tariffs set by a load serving entity (LSE)
(2) **Utility market optimized model**, in which a utility owns DER or contracts with DER owners and optimizes DER performance vis-à-vis wholesale markets
(3) **Aggregator or customer market optimized model**, in which an aggregator or large customer optimizes DER performance vis-à-vis wholesale markets by bidding through the utility, potentially with capacity services contracts with the utility
(4) **DSO market optimized model**, in which a functionally independent distribution system operator (DSO) optimizes DER and loads across a distribution area vis-à-vis wholesale markets

The first three of these models are illustrated on the lefthand side of the figure below.



Notes: ISO refers to independent system operator (ISO). This paper only focuses on demand-side participation in ISO markets; hence the supply side is shaded out. In a DSO market optimized model, the DSO would replace the "Utilities or other LSEs" box.

The four models imply very different roles and responsibilities, risks and rewards, technology gaps, and regulatory gaps. Many jurisdictions already have some variant of the time-differentiated rates needed for the customer tariff optimized model, though enabling customers with storage to provide a wider array of flexibility services would require more sophisticated rate structures. In particular, it would require providing incentives for avoiding marginal capacity costs, which involves thorny rate design issues around economic bypass – if customers can avoid marginal capacity costs, how much responsibility should they bear for embedded fixed costs?

The second model (utility market optimized) is also within reach for jurisdictions where utilities already have experience operating demand response (DR) programs in response to wholesale market prices. Utilities would likely address any operating issues with contracted DER up front and would need to have the ability to remotely control contracted resources. This model – which has been employed by utilities from Rocky Mountain Power in Utah to Green Mountain Power in Vermont – involves questions around whether utility ownership and operation of DER will crowd out potential competition and innovation (competition policy), whether DER is best procured through utility programs or competitive solicitations, and whether and how avoided distribution infrastructure costs can be accurately valued in procurement.

The third model (aggregator or customer market optimized) would require a big leap for most utilities. Utilities generally do not have the ability to receive demand bids from customers, to ensure that the operation of DER associated with those bids would not lead to real-time reliability issues, to provide scheduling instructions to DER customers, to measure and handle imbalances, and to settle (pay/charge) these resources at wholesale market prices. The technologies to do these tasks exist – independent system operators (ISOs) already perform these functions on the bulk system – though they would need to be adapted for distribution system use. This third model would likely be used in a limited programmatic context, due to concerns over utilities favoring their own resources in dispatch. If successful, the third model thus likely segues into a DSO model over time.

The DSO market optimized model would require significant changes in distribution planning, interconnection, operations, communications, and regulation. The DSO must be at least functionally independent, meaning that utilities would need to firewall distribution planning and operations from resource procurement, marketing, and ratemaking. It should have some form of an open access tariff, to specify rules for interconnection, operations, and markets. It would need to be able to perform all of the functions in the third model, but would do so in an open access context where any DER owner could opt to be dispatched and settled at wholesale market prices upon interconnection. Even in jurisdictions that have open proceedings that could create DSOs, this fourth model is still at least several years away.

Looking across these four models, a few themes emerge:

- ***DER market participation via the demand side may be more efficient than via the supply side***. DER participation through the demand side of ISO markets is likely to be more efficient than participation on the supply side. Aggregating distribution generation, storage, and demand response is more straightforward on the demand than on the supply side, and participation through the demand side obviates the need for ISOs to dispatch resources on the distribution system.

- ***Different models for demand flexibility are not exclusive***. For instance, a distribution utility could employ the first three models at the same time, though regulators would need to be cautious about the second and third being used in tandem. The DSO model can ultimately provide an umbrella for a variety of models of demand flexibility.

- ***Most jurisdictions do not need a DSO in the near term***. In many jurisdictions, well-designed tariffs may be sufficient to enable desired levels of demand flexibility. The need for more sophisticated models will likely be driven by a combination of growth in DER and economics, for instance in transmission-constrained load pockets in which rapid electrification runs up against capacity constraints on the distribution system.

- ***Operational capabilities on the distribution system can grow modularly***. For many jurisdictions, monitoring, communications, and control capabilities for the distribution system can grow in phases over time with expected growth in DER and operational needs. This phased approach to expanding capabilities can reduce the need for large upfront utility investments in the grid. Regulators will need to decide how much to invest in utility operational capabilities and more centralized control of the distribution system, versus letting non-utility DER owners and operators take a more active role in resolving reliability issues.

- ***The key question is where to locate electricity storage***. The most important question for electricity system flexibility more broadly is how much electricity storage to locate in different parts of the system: behind the customer meter, on the distribution system, on the lower or higher voltage transmission systems, or behind the generator meter. For storage to be more cost-effective on the distribution system than the lower and higher voltage transmission systems, the incremental benefits of siting it on the distribution system (reliability and resilience, avoided distribution costs) must be large enough to offset any loss of scale economies.

- ***Well-designed regulation and markets can help to support efficient levels of demand flexibility***. Improvements in regulation and markets can help to guide investments in storage and other flexible resources to locations where they have the highest value for the least cost. For the distribution system, improvements will thus have two components: (1) better valuation for DERs, particularly for reliability and resilience value, real-time locational energy value (real-time LMPs), and avoided distribution, transmission, and RA costs; and (2) a competitive environment for DER. Better valuation and more competition could come through adjustments to DER tariffs, benefit-cost analysis for utility programs, and evaluation frameworks used in utility procurement, or by allowing non-utility DER operators to directly access wholesale markets. Regulators will need to ensure that valuation frameworks clearly differentiate private and public values for reliability and resilience.

Increasing Electricity Demand Flexibility: Approaches and Gaps

Limited flexibility in the magnitude and timing of demand has been a decades-long hallmark of electricity systems. The absence of more flexible demand currently means that electricity systems are often overbuilt — that they have a significant amount of generation and network capacity that is seldom used. The costs of overbuilding have the potential to become much larger with more renewable generation and electrification of transportation and building loads, because of the intermittent nature of wind and solar generation and seasonal and diurnal patterns in transportation and heating energy demand.

Several drivers — changing technologies and changing needs and value — are converging that would allow higher levels of demand side flexibility than have been historically feasible. In terms of technologies, these include the emergence of lower cost battery storage, distributed generation, and real-time monitoring, control, and communications technologies. In terms of needs and value, they include the potential to reduce the scale of electrification-driven distribution infrastructure investments, greater consumer focus on reliability and resilience, and the changing economics of electricity systems with rising levels of solar and wind generation.

Increasing demand flexibility will require changes in utility incentives and regulation, new grid management and communications technologies, and reforms to electricity market designs. This paper describes different approaches to enabling more demand flexibility, provides examples of how they might work in practice, and identifies key gaps for enabling demand flexibility. The focus in this paper is on increasing flexibility through the demand, rather than the supply, side of wholesale markets.

Key Takeaways

- Increasing demand flexibility may not always be more cost-effective than increasing supply-side flexibility, but without creating the operational capabilities, utility incentives, markets, and regulations that enable innovation and growth in distribution-level resources it will be difficult to find a cost-effective balance between demand and supply-side flexibility.
- There are many different, but not exclusive, approaches to increasing demand flexibility, ranging from more sophisticated tariffs to distribution system operator (DSO)-run markets; the key distinguishing factors in these different models are (a) which entity is responsible for optimizing the performance of distribution-level resources, and (b) the level of competition via participation by non-utilities on the distribution system.
- Not all jurisdictions will need a DSO; in some cases, voluntary tariffs for customers with the capability to shift net demand will provide desired levels of demand flexibility. The move toward DSO-like should be driven mainly by a desire for a more open and competitive distribution system, which suggests that it should be paired with reforms to utility procurement and wholesale market operations.
- The future roles of distribution utilities in distribution planning, interconnection, distribution operations, resource procurement, and wholesale market operations are unclear. What utilities are best placed to do will vary by jurisdiction.
- All of the approaches discussed in this paper would need to be enabled by new operational and communications capabilities on the distribution system, changes in utility incentives and programs, procurement, and tariffs; new settlement and billing systems; and innovations in distribution-level and wholesale (ISO) market design.
- For enabling demand flexibility, key near-term priorities areas should include changes in interconnection rules, utility incentives, programs, and procurement, distribution utility monitoring and communications capabilities, and utility settlement and billing systems.
- Energy storage, and battery storage in particular, is the most important new resource for increasing demand flexibility. Whether battery storage is more cost-effectively connected behind the customer meter or to the distribution system vis-à-vis the bulk system (behind the generator meter, to the transmission system) will depend on its local values — avoided distribution investment costs (distribution congestion management), localized transmission congestion management, local resource adequacy, and local or customer reliability and resilience — relative to the economies of scale and other benefits of siting storage on the bulk system.

Models for Demand Flexibility

Demand flexibility is a broad rubric. It means that customers are able to change the level or timing of net demand — demand at the customer meter, net of any self-generation — in response to system needs. Changes in net demand could result from changes in behind-the-meter generation, storage charging and discharging, and load management. Net demand could be positive (assuming demand is a negative value) if the customer is exporting to the network. From the perspective of the physical network, what happens behind the customer meter (or meters) to shape net demand is irrelevant; all that matters is net demand (real and reactive) at the customer meter.

Customers can provide demand flexibility on different timescales: months-ahead, weeks-ahead, day-ahead, hour-ahead, within the hour, and even on a second-to-second basis. Different forms of customer engagement and incentives may be appropriate for different timescales. For the purposes of this paper, a common feature across these different models is that they affect the demand, rather than the supply, side of the ledger. Demand response (DR) that offers into ISO capacity and energy markets, for instance, is thus out of our scope.

There are a large number of potential models for how demand flexibility could be achieved, differing in approaches to interconnection, system operations, and market settlement. The below table describes an illustrative, but not exhaustive, set of these models, organized around which entity is optimizing distribution-level load, generation, and storage; these different models may not be exclusive.

Potential Models for Demand Flexibility

Model	Approach		
	Interconnection	**Operations**	**Settlement**
Customer or ESCO tariff optimized	Customer or ESCO determines interconnection status	Customer or ESCO optimizes DER against a tariff, changes in demand show up in LSE demand forecasts	Customer or ESCO is settled based on an LSE tariff
Utility market optimized	Customer, ESCO, or utility determines interconnection status	Utility directly operates DER within customer or ESCO defined parameters	Utility signs contracts and pays customer or ESCO per contract terms; utility is settled using ISO market prices

Non-utility LSE market optimized	Customer, ESCO, or LSE determines interconnection status	LSE directly operates DER within customer or ESCO defined parameters against market prices and T&D rates; distribution utility may override market dispatch	Utility signs contracts and pays customer or ESCO per contract terms; LSE is settled using ISO market prices
DER aggregator, customer, or ESCO market optimized	Customer, ESCO, or aggregator determines interconnection status	Aggregator bids directly to LSE, LSE incorporates into demand bid, aggregator operates DER, LSE passes through imbalance costs and penalties to aggregator; distribution utility may override market dispatch	LSE settles aggregator with imbalance costs and imbalance penalties or very granular metering and allocation of regulation reserve costs
DSO market optimized	Customer, ESCO, or aggregator determines interconnection status	Customer, ESCO, or aggregator bids directly to DSO, DSO incorporates into demand bid at T-D interface, DSO dispatches DERs subject to distribution security constraints	LSE settles aggregator with imbalance costs and imbalance penalties or very granular metering and allocation of regulation reserve costs

Notes: ESCO refers to energy service companies; DER refers to distributed energy resources; LSE refers to load serving entity, T-D refers to transmission-distribution

Model Examples

Examples help to illuminate the kinds of changes that might be necessary in different possible models. The examples assume that the customer's transmission provider participates in an independent system operator (ISO) run market. They are intended to provide insight on questions around "how would it work" and "where are the gaps"?

The examples are organized around the interconnection, operations, and settlement categories from **Error! Reference source not found.**. In addition, they also describe resource adequacy (RA) mechanics and considerations for making the approach work in practice. All of the examples include some amount of battery storage.

Customer tariff optimized model

A customer on a utility retail tariff installs PV, battery storage, and load management technologies to reduce electricity costs. The utility tariff is based on marginal costs and benefits and is intended to provide incentives for shifting customer net demand. The utility's service territory already has a significant amount of solar PV generation. The utility tariff charges and pays $0.10/kWh during the day and $0.25/kWh during the early evening, late evening, and early morning.

Interconnection. The customer determines interconnection status, meaning that the customer determines the "firmness" of their level of distribution service. Customers that want firm interconnection are responsible for paying for incremental distribution upgrade costs identified through the interconnection study. Customers that want non-firm interconnection can avoid upgrade costs but may be subject to utility curtailment, with terms and conditions specified in distribution interconnection agreements.

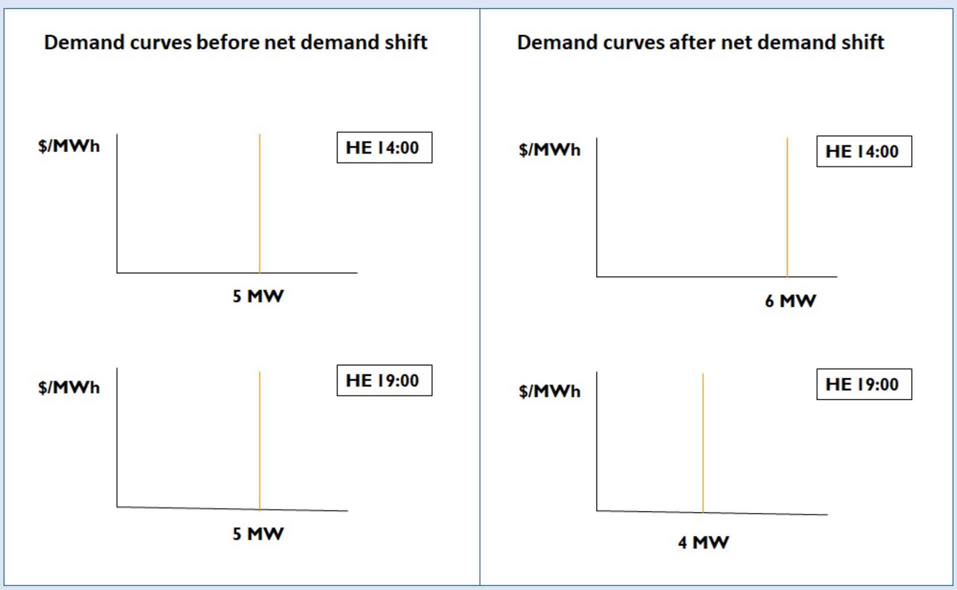
Resource adequacy. Reductions in the utility's peak demand and ISO RA obligations from customer tariff optimization are captured in the ISO's and/or utility's demand forecast over time. To incentivize customers to reduce coincident peak demand, the utility tariff includes consideration of the utility's marginal RA costs.

Market operations. From 1:00 to 2:00 pm (HE1400), the customer has an average of 5 kW of PV, 10 kW of load (already including shiftable load). During HE1900, the customer has no PV and 10 kW of load. The customer has 2 kW of battery storage, with no state of charge constrains in either hour. During HE1400, the customer charges the battery for 2 kW (kWh), giving it an average net demand of 7 kW (= 10 kW load + 2 kW charge – 5 kW PV) at the customer meter. During HE1800, the customer discharges 2 kW, reducing its net demand to 8 kW. The 2 kW shift in net demand due to charging will eventually be reflected in the utility's short-run demand forecasts. Net demand shifting may change the magnitude of the utility's forecast, if large enough, but may not change the shape of the utility's demand curve (see Box).

Changing the magnitude but not the shape of the demand curve

A utility's customers install 1 MW of load shifting capability (battery storage, load controls) under a TOU tariff that has lower prices during the day, during high solar hours, and higher prices in the evening. These customers shift 1 MW consumption to the afternoon (HE 14:00), to take advantage of lower prices, reducing consumption by 1 MW in the evening (HE 19:00). Over time, the utility incorporates this change in its ISO day-ahead demand bid.

Before the load shift, the utility submitted a vertical demand bid curves to the ISO for 5 MW in HE 14:00 and HE 19:00. After the load shift, the utility still submits vertical demand bid curves, but the amount of its bid changes from 5 MW to 6 MW in HE 14:00, reflecting increased consumption, and 5 MW to 4 MW in HE 19:00, reflecting decreased consumption.



Market settlement. The customer will reduce its costs by $0.30/h (= 2 kW × [$0.25/kWh - $0.10/kWh]) through shifting its net demand from, for instance, the early evening to HE1400. The utility will save on ISO energy, AS costs, and possibly RA and transmission costs and potentially distribution investment costs, savings that will be approximately captured in the tariff. For instance, if the utility faces a day-ahead locational marginal price (LMP) of $30/MWh in HE1400 and $50/MWh in HE1800, the 2 kW shift in the utility's demand from HE1800 to HE1400 reduces its cost by $0.40/h (= 2 kW × [$0.50/kWh - $0.30/kWh]). Tariffs will generally not incentivize changes in net demand within the hour, so presumably the utility will not save on its real-time energy charges. If, on average, the utility saves more from demand shifting than it rewards customers through the tariff, other utility customers and potentially shareholders will benefit.

Considerations. Utility tariffs that encourage net demand shifting could have a range of possible designs, from unbundled tariffs like New York's Value of DER (VDER) to more simple bundled tariffs that are based on some measure of the component elements: RA capacity, energy, transmission, distribution, and others. The utility can localize some components of the tariff: zonal RA capacity, LMP energy, and forecast-based distribution value. For regulators, the principal challenge of these kinds of tariffs will be in allocating transmission and distribution charges. Specifically, to what extent do tariff designs allow customers to economically avoid transmission and distribution charges (economic bypass). Economic bypass may reduce long-run costs for all customers, but in the short-run may shift costs from customers that are able to avoid fixed charges to those that are not.

Utility market optimized model

The utility signs annual ($/yr.) contracts with customers to provide demand flexibility, under which the utility has the right to operate customer-owned distributed generation, storage, and load management technologies. The utility operates these resources to reduce its ISO capacity, energy, AS, and transmission costs, and to reduce its distribution costs. For a given day, the utility anticipates that it will have 5 MW / 10 MWh of contracted battery storage capacity at node A and 5 MW / 10 MWh at node B. Within node A, the utility has distribution-level constraints and is operating some of the 5 MW to reduce distribution net loads.

Interconnection. There are multiple possible approaches to interconnection status. If distribution-level resources are contracted programmatically, the utility may waive any distribution upgrade costs, assuming that the utility would optimize resources to avoid any reliability violations. The utility may also give customers the option of paying for any distribution upgrade costs, in exchange for higher contract payments. For instance, utility contracts could match procurement with what is feasible using available distribution capacity.

Resource adequacy. Reductions in the utility's peak demand and ISO RA obligations from utility optimization of resources are captured in the ISO's and/or utility's demand forecast over time. The utility pays for avoided capacity charges in its contract. In cases where the utility operates resources to avoid distribution reliability violations that are not coincident with system peak, the utility will not count these resources toward its forecasted resource adequacy obligation and will compensate these resources for distribution but not bulk system capacity value. In cases where distribution and bulk system capacity needs are coincident, the utility would compensate resources for both.

Market operations. On a day-ahead basis, the utility anticipates that it will need to charge the batteries at 5 MW (ignoring, for simplicity, losses) during HE1400 and HE1500 to avoid overloading a distribution substation within node A. The utility fully discharges the batteries in node A the evening before (HE1800 and HE1900) during high priced hours and commensurately decreases its expected ISO day-ahead demand bid by 5 MW, from 350 MW to 345 MW during both hours. During HE1400 and HE1500, the utility charges the batteries at node A for 5 MW in each hour and increases its expected day-ahead demand bid by 5 MW, from 300 MW to 305 MW, in HE1400 and HE 1500. The utility's real-time demand (for simplicity) is 345 MW in HE1800 and HE1900 and 255 MW in HE1400 and HE 1500.

For batteries at node B, the utility optimizes against ISO energy market prices, subject to any distribution-level constraints. On a day-ahead basis, the utility expects prices to be lowest in the night in HE0300 and HE0400 and expects prices to be highest in the late evening in HE1800 and HE1900. The utility increases its day-ahead demand bid (charging) by 5 MW in HE300 and HE400 and decreases its demand bid (discharging) by 5 MW in HE1800 and HE1900. During the day, it becomes clear that the highest priced hours will be in HE1900 and HE2000 and the utility discharges the batteries in these hours rather than in HE1800 and HE1900.

Market settlement. For simplicity, day-ahead and real-time loads and prices are the same in all hours except for HE2000, where real-time prices are higher than day-ahead prices. The below table shows ISO settlement for the utility in each hour described above, based on assumed day-ahead and real-time prices in the table. The ISO only settles the utility for its cleared demand; it neither dispatches nor settles the batteries. This illustrates that the benefits from using distribution-level resources on the demand side are in lower ISO charges, rather than income. In this case, the utility is able to come out positive in the energy market in its use of batteries for distribution services (discharges at $50/MWh and charges at $30/MWh), but there may be instances where the utility would be forced to pay a net cost in the energy market to ensure provision of distribution services. Although LMPs might differ between node A and B, the utility only pays an aggregated LMP, which means that it does not have a direct incentive to help manage transmission congestion.

Market Settlement of Utility Optimized Resources

HE	Utility DA demand (MW) Without storage	Utility DA demand (MW) With storage	Utility RT demand (MW) With storage	Utility RT demand (MW) Without storage	Node	ISO market price ($/MWh) DAP	ISO market price ($/MWh) RTP	Utility Settlement ($/hr)	Net storage cost ($/hr)
1800*	350	345	350	345	A	$50	$50	$17,250	-$250
1900*	350	345	350	345	A	$50	$50	$17,250	-$250
0300	250	255	250	255	B	$10	$10	$2,550	$50
0400	250	255	250	255	B	$10	$10	$2,550	$50
1400	300	305	300	305	A	$30	$30	$9,150	$150
1500	300	305	300	305	A	$30	$30	$9,150	$150
1800	350	345	350	350	B	$50	$50	$17,500	$0
1900	350	345	350	345	B	$50	$50	$17,250	-$250
2000	350	350	350	345	B	$50	$100	$17,000	-$500

The utility makes fixed payments of $10/kW-mo. to battery owners in node A, including distribution but not resource adequacy value; it makes $8/kW-mo. payments to battery owners in node B, including resource adequacy value but no distribution value. During the month, the utility will pay $90,000 (= 5 MW × $10/kW-mo. + 5 MW × $8/kW-mo.) for battery services. These monthly payments should be less than the total benefits (cost savings) that the utility realizes in the ISO markets and in providing distribution services over a longer time horizon (e.g., one year, or five years).

Considerations. Utility contracts with distribution-level resources are ultimately at the discretion of a monopsony utility, which means that payments will likely need to be made on the basis of commission-approved avoided costs, or that utility procurement of services from these resources will need to be competitive. Competition might be through competitive solicitations, though these will tend to depress value for resource owners relative to utility customers and shareholders. Competition might also be through providing alternative pathways to market, for instance through DER aggregation and supply side participation in ISO markets, which would provide resource owners with choice — if utility contracts are less lucrative than other options, resource owners could choose other market providers.

Utility optimization implies that utilities need to have the physical infrastructure, software tools, and incentives to maximize resource value. For instance, in this case utilities would need to have the ability to identify distribution constraints on short timescales and the confidence to operate battery storage to relieve those constraints. Utilities would also need to have the ability to operate distribution-level resources against ISO market prices, and to sign contracts that balance value to customers and value to resource owners. Some utilities may have extensive experience operating programmatic DR in ISO markets, whereas other utilities might have no experience with distribution-level resources in these markets.

DER aggregator, customer, or ESCO (non-utility) market optimized model

A DER aggregator develops a 10 MW virtual power plant that will contract for capacity services with the utility and will provide day-ahead energy services through the utility. The DER aggregator submits day-ahead net energy demand bid curves to the utility. The utility ensures that the DER aggregator’s bids are consistent with its capacity needs and, within these constraints, incorporates the energy bid curves into its ISO day-ahead demand curve. The utility sends cleared day-ahead schedules to the DER aggregator. The DER aggregator is responsible for any real-time imbalance charges and will optimize its storage and DR resources to avoid these charges.

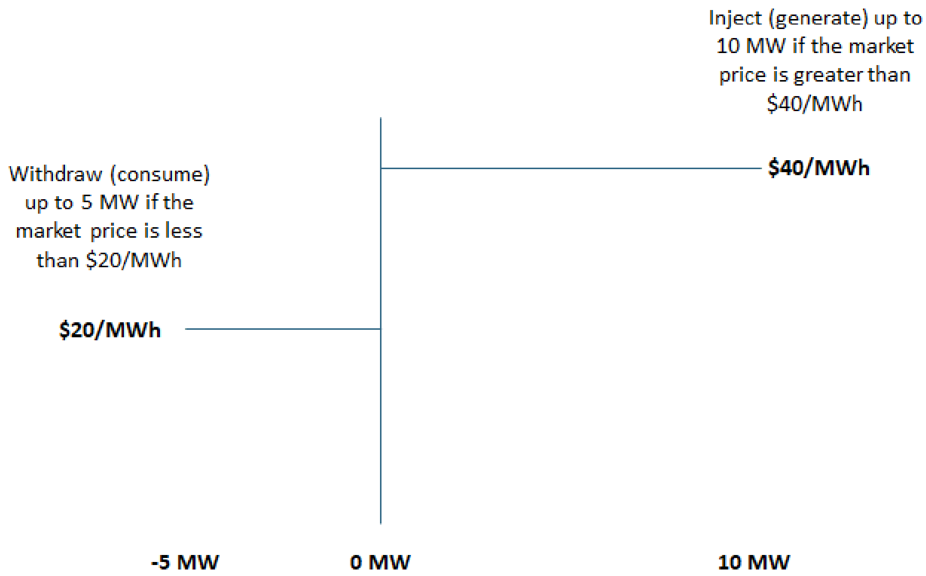
Interconnection. As with the preceding example, there are multiple possible approaches to interconnection status. The DER aggregator’s resources may already have gone through the interconnection process, in which case their interconnection status will already be determined. The utility and DER aggregator may also determine interconnection status as part of the contracting process.

Resource adequacy. If the utility contracts with the DER aggregator for resource adequacy capacity, the utility will stipulate availability requirements in the contract and will ensure that energy bids do not conflict with resource adequacy obligations. The utility will compensate the DER aggregator for resource adequacy services based on its avoided costs.

This implies a different approach, and in particular a more permanent reduction in demand, than has been the case historically with utility interruptible DR programs; it more closely resembles utility load control programs. There could be multiple approaches to resource adequacy contracting for DER aggregators. For instance, for existing customers the utility could baseline historical use and contract for a fixed reduction in demand in certain periods by setting a maximum demand level. For new customers, the utility could determine the capacity amount as the difference between a maximum interconnection level (kW) and maximum demand levels set by the utility during certain periods. Setting maximum demand levels could avoid challenging issues around regular baselining with storage. Utilities could still operate interruptible DR programs in tandem with capacity service contracts with DER aggregators.

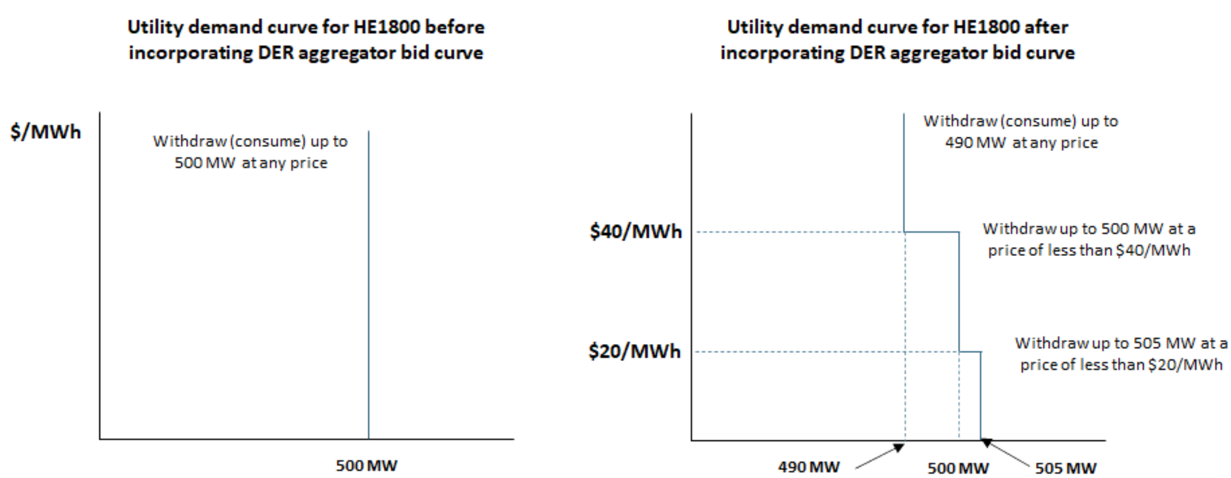
Market operations. On a day-ahead basis, the DER aggregator submits energy bid curves to the utility for each hour the following day. For instance, for HE1800 the aggregator submits the below bid curve to the utility.

Aggregator Bid Curve Submitted to the Utility



The utility screens bid curves for potential distribution violations and resource adequacy obligations and incorporates these bid curves into its aggregate ISO demand bid curve, but in HE1800 there are no constraints. For HE1800, the utility had a vertical demand curve of 500 MW. The figure below shows the utility's demand curve for HE1800 before and after incorporating the bid curve from the DER aggregator. Unlike in the "utility optimizes" case above, the DER aggregator bids may change both the amount and shape of the utility's demand curve.

Utility Demand Curve Before (Lefthand Figure) and After (Righthand Figure) Incorporating DER Aggregator's Bid Curve



The day-ahead market for HE1800 clears at $50/MWh, which means that the utility clears 490 MW and the DER aggregator clears +10 MW (net injection). The utility sends an hourly dispatch instruction to the DER aggregator for a 10 MW injection during HE1800.

Market settlement. The utility pays the ISO $24,500/hr (= 490 × $50/MWh) for its day-ahead market settlement during HE1800, and the utility pays the DER aggregator $500/hr (= 10 MW × $50/MWh). The utility's net cost ($25,000/hr = 500 MW × $50/MWh) is the same regardless of how much the DER aggregator clears in the market.

During HE1800, the DER aggregator's average net injection falls to 5 MW because of an equipment outage. The average real-time price during this interval is $100/MWh. The utility's load (assuming unrealistically that it would have otherwise remained constant across the hour) increases to 495 MW and the utility pays a $500/hr (= [495 MW – 490 MW] × $100/MWh) real-time imbalance cost. The utility passes this cost onto the DER aggregator, which means that the DER aggregator will have a net settlement from the utility of $0/hr (= $500/hr - $500/hr) during this hour.

The utility will also settle any capacity (distribution, resource adequacy) services with the DER aggregator. For instance, if the DER aggregator is providing 5 MW of resource adequacy capacity and the contract payment for resource adequacy capacity is $4/kW-mo., the utility will pay the aggregator $20,000 for that month.

Considerations. This model raises several regulatory issues. If the DER aggregator is aggregating generation, storage, and DR, it will be most efficient for the aggregator to effectively become an LSE, in which case the bids submitted to the utility are net demand bids (negative is net withdrawal, positive is net injection). In jurisdictions without competitive retail, allowing aggregators or ESCOs to become LSEs would require significant changes in regulation. The utility would need to be able to separate the virtual power plant from its own load forecast. Aggregators, customers, or ESCOs participating in this model would need to have at least 15-minute metering, to ensure accuracy of imbalance charges. The utility would also likely charge for dispatch services, for instance as a flat monthly fee. The level of this fee would require regulatory oversight.

DSO market optimized model

An independent DSO aggregates supply offers and demand bids from LSEs, DER aggregators, and ESCOs into single net demand bid curves and submit these curves to the ISO. The DSO's roles in distribution planning, interconnection, bulk resource adequacy, and market operations and settlement may vary significantly across jurisdictions. For instance, the DSO may or may not be responsible for ensuring that ISO resource adequacy requirements for its distribution area are met. The ISO, rather than the DSO, may be responsible for settling loads and resources that clear the ISO market. The description here focuses on one possible implementation of an independent DSO.

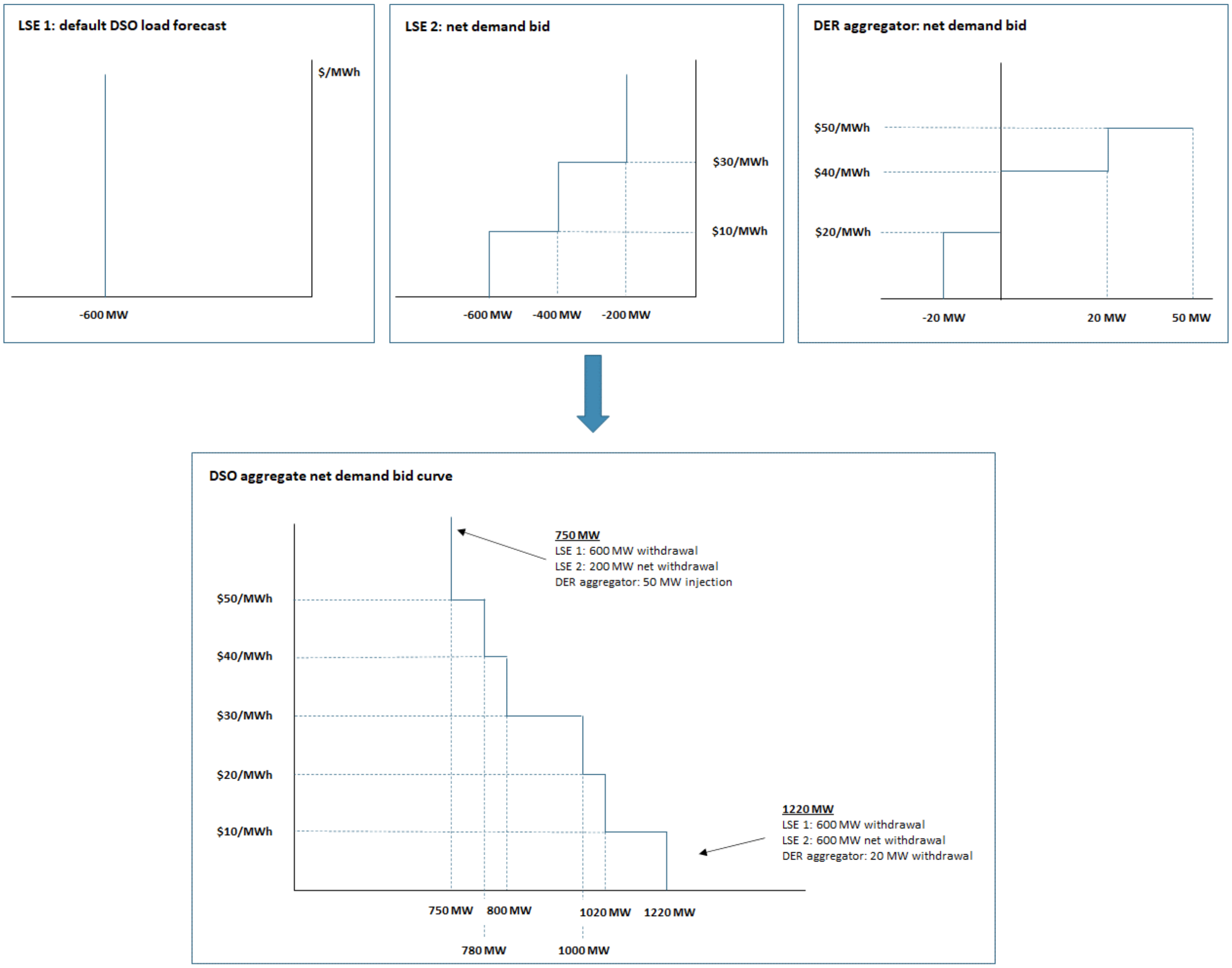
Interconnection. The DSO runs a non-discriminatory distribution interconnection process, through which it offers different levels of distribution service to new resources and loads above a minimum size threshold. In the interconnection process, resource owners determine the "firmness" of their service level or specify a maximum limit on injections. The DSO relieves distribution system congestion according to service levels. The DSO undertakes a semi-annual distribution planning process to identify reliability, economic, and policy driven investment needs for the distribution system that are made pre-emptively rather than triggered through interconnection.

Resource adequacy. ISOs manage bulk system resource adequacy in the same way that they currently do. LSEs, potentially including distribution utilities, are responsible for resource adequacy obligations. The ISO procures resource adequacy through auctions, through which LSEs are able to manage their resource adequacy charges through either transmission-level or distribution-level resources. Distribution-level resources may not be eligible to provide resource adequacy, for instance because of a conflict with provision of distribution services.

Market operations. The DSO allows distribution market participants to either use the DSO's default forecasting service or submit their own net demand bid curves. The DSO aggregates bids from three entities: LSE 1, which uses the DSO's default service; LSE 2, which submits a net demand bid curve; and a DER aggregator, which also submits a net demand bid curve. The ISO allows DSOs to submit aggregate hourly demand bids in day-ahead and 5-minute demand bids in hour-ahead markets.

The figure below shows day-ahead net demand bid curves from the three entities and the DSO's aggregate net demand bid curve for HE1800. For consistency, the three entities demand (withdrawal) bids lie on the negative part of the x-axis, which means that demand curves are upward, rather than downward, sloping. The DSO aggregates these bid curves into a single, monotonic, downward sloping net demand bid curve for the ISO, incorporating distribution-level security checks.

LSE and DER Aggregator Net Demand Bid Curves and DSO Aggregate Net Demand Bid Curve



For HE1800, the market clears at $45/MWh and the ISO clears 780 MW for the DSO. The DSO sends day-ahead schedules to LSE 2 (200 MW withdrawal) and the DER aggregator (20 MW injection). The DSO submits the same net demand bid curve to the ISO for the hour-ahead market for HE1800 (for simplicity). The hour-ahead market clears at $55/MWh in each 5-minute interval and the ISO clears 750 MW for the DSO in each interval. The real-time market clears at $60/MWh. The DSO sends hour-ahead dispatch instructions to LSE 2 (200 MW withdrawal) and the DER aggregator (50 MW injection). During HE1800, LSE 1 withdraws 605 MW (hourly average metered), LSE 2 withdraws 205 MW (average), and the DER aggregator injects 45 MW (average). The DSO assesses imbalance penalties to all three entities based on real-time LMPs.

As part of the hour-ahead market, the ISO runs an hour-ahead commitment process to address potential commitment needs before the operating hour, using its own forecasts. The ISO clears the real-time (5-minute) market using the ISO's own net demand forecasts. The ISO deals with real-time market imbalances using regulation reserves, allocating the costs of these reserves on a load ratio share basis.[1]

Market settlement. For each market interval (day-ahead, hour-ahead, real-time), the ISO settles the DSO and the DSO settles entities within its distribution area. The table below shows energy market settlement for each entity and for the DSO as a whole (ISO-DSO settlement), where the negative sign indicates withdrawal or a payment to the DSO (or ISO) and a positive sign indicates injection or payment from the DSO (or ISO). The ISO allocates regulation charges to the DSO and the DSO passes these charges on to each entity. The DSO is revenue neutral.

Energy Market Settlement for Each Entity and the DSO

Entity	Schedule or dispatch (MW)			LMP ($/MWh)			Market settlement ($/hr)			
	DA	HA	RT	DA	HA	RT	DA	HA	RT	Total
LSE1	-600	-600	-605	$45	$55	$60	-$27,000	$0	-$300	-$27,300
LSE2	-200	-200	-205	$45	$55	$60	-$9,000	$0	-$300	-$9,300
DAgg	20	50	45	$45	$55	$60	$900	$1,650	-$300	$2,250
DSO	-780	-750	-765	$45	$55	$60	-$35,100	$1,650	-$900	-$34,350

Considerations. Much of the institutional infrastructure for the kind of DSO described in this case does not exist today. The ISOs do not have demand bidding in intraday markets or LMP settlement, both of which only make sense if there is sufficient demand flexibility. The DSO in this case would need to be at least functionally, and more likely organizationally, independent, governed by open access rules and subject to concurrent state and federal regulation. There are many possible permutations for the DSO's roles and responsibilities vis-à-vis the ISO, in terms of resource adequacy, real-time markets, and market settlement; similarly, there are many options for ISO market designs that take advantage of the increased operational flexibility that DSOs can provide. Again, the above case just highlights one possible implementation.

[1] An alternative approach would allow intra-hour net demand bidding by loads through the DSO and allocate regulation costs to DSOs on the basis of deviations from 5-minute dispatch.

Key Gaps and Issues

All of the models described in the previous section would require utilities, regulators, and ISOs to address gaps in utility infrastructure and technical capabilities, utility and industry regulation, and tariff and market design. The nature and extent of these gaps will vary by jurisdiction.

Customer tariff optimized model

For most jurisdictions, the customer tariff optimized model would require the least amount of change relative to existing practice. Some jurisdictions may not have the regulatory framework to support interconnection export limits. Utilities in many jurisdictions do not have incentives to propose tariffs that would reduce their resource adequacy, transmission, and other capacity charges, and most jurisdictions do not have tariffs that can incentivize significant net demand shifting from customers. These tariffs could be voluntary and, as an initial step, do not need to be complex. However, they should include some incentives to reduce distribution-level peak demands and system coincident peak demand. This implies providing customers with incentives to reduce capacity (fixed) costs, which in turn may imply some amount of cost shifting onto customers that do not shift their loads.

To address gaps in the customer tariff optimized model, key areas of focus include:

- Revising interconnection rules and processes to facilitate interconnection export limits. California's Rule 21 provides one example of the implementation of export limits in interconnection rules.[2]
- Developing incentive frameworks and expectations for utilities that encourage more capacity and energy cost savings. Incentives could be explicit, such as through performance incentive mechanisms; they could be broader, through revenue or price caps and other forms of performance-based regulation; or they could be through expectations of areas that utilities should focus on during a given rate case cycle.
- Designing new opt-in tariff options. Tariffs may range in complexity from TOU tariffs with simple on-peak and off-peak rates, assuming coincident distribution and system peaks, or they could be more sophisticated, with several TOU periods and a combination of variable and fixed charges. New York's VDER tariff is a useful reference point for tariff design for customers with distributed resources and the issues around transitioning from net energy metering to tariffs.[3]

[2] See https://www.cpuc.ca.gov/rule21/.

[3] See https://www.nyserda.ny.gov/All-Programs/ny-sun/contractors/value-of-distributed-energy-resources.

Fixed cost allocation in tariffs that encourage net load shifting

Consider a utility that has \$0.08/kWh average variable costs and \$500/kW-yr. (kW peak) average fixed costs. The utility has an average system load factor of 0.6, so in volumetric terms its average fixed costs are \$0.095/kWh (= \$500/kW-yr. / [8,760 × 0.6])[A] and its total average cost is thus \$0.175/kWh. The utility determines that its on-peak variable costs are \$0.13/kWh and its off-peak variable costs are \$0.06/kWh, but it surmises that the difference between on-peak and off-peak rates is unlikely to encourage a significant amount of net load shifting.

The utility decides to develop a marginal cost-based rate specific to distribution planning zones. For a transmission-constrained urban zone, the utility has on-peak energy (LMP-based) costs of \$0.20/kWh and off-peak costs of \$0.06/kWh. For this zone, the utility calculates its long-run marginal fixed costs at \$700/kW-yr. (kW system coincident peak), higher than average due to local generation capacity and distribution capacity limits. The utility allocates its marginal fixed costs across all on-peak hours, giving it an on-peak fixed cost rate of \$0.20/kWh.[B] Including variable costs, the utility's TOU rate in this zone will be \$0.40/kWh on-peak and \$0.06/kWh off-peak.

If this tariff is effective and encourages customers to shift their net demands to avoid the much higher on-peak rate, customers that are not on the tariff would also benefit from shifting load: the utility's average variable energy costs would decline, as off-peak energy costs (\$0.06/kWh) are lower than average energy costs (\$0.08/kWh), and the utility would avoid average fixed cost increases, because marginal fixed costs in the zone (\$700/kW-yr.) are higher than average fixed costs (\$500/kW-yr.).

However, if most customers that are on the tariff shift net demand to off-peak periods, the utility would need to recover its \$500/kW-yr. in embedded costs from customers that are not on the tariff, potentially leading to cost shifting. Designing tariffs that balance incentives, marginal benefits, and fair allocation requires in-depth analysis and regulatory judgement.

Notes: [A] This calculation assumes constant load during on-peak hours. [B] This calculation assumes that 40% of hours are on-peak, consistent with the on-peak, off-peak, and average rates in the preceding paragraph. With a smaller number of on-peak hours, the marginal fixed cost rate could be significantly higher.

Utility market optimized model

Some utilities have long used the utility market optimized model, particularly for DR. However, even in cases where utilities have optimized distribution-level resources vis-à-vis wholesale market prices utility programs have often focused on avoiding bulk system charges and not distribution-level charges. In many jurisdictions, utilities lack the monitoring and control capabilities, technical expertise, and staffing to optimize distribution-level resources for distribution and bulk system value.

Additionally, utilities often do not have the planning expertise, regulatory frameworks, and incentives to contract with customer-owned, or invest in utility-owned, distribution-level resources in a way that minimizes utility long-run marginal costs. Most utilities procure services from distribution-level resources through utility programs and, to a lesser extent, competitive solicitations for new resources. Utility programs may have limited customer outreach, customer payments that are based on a limited number of benefits (e.g., avoided average energy costs only), and limited program budgets. Utility bulk system procurement typically only considers a limited number of benefits (energy and resource adequacy capacity), and distribution-level resources often face an aggregation challenge in competitive solicitations. For utility non-wires procurement, there are often a limited number of distribution needs where non-wires alternatives will be viable.

Part of the utility incentive gap may be in the incentives utilities face via their participation in ISO energy and capacity markets. Utility loads are generally settled at aggregated LMPs rather than nodal LMPs, which, combined with the fact that many utilities are transmission owners that receive allocations of financial transmission rights, means that utilities may not have incentives to use distribution-level resources to manage local congestion. ISO markets also do not provide LSEs with opportunities to affect real-time price formation, and many utilities do not use distribution-level resources to actively manage real-time market risk. In ISOs with three-year forward capacity markets (PJM, ISO-NE), the gap between capacity auctions and delivery year may create a disincentive for utilities to contract for capacity services.

A final gap with the utility market optimized model is in the regulatory framework for utility ownership and operation of distribution-level resources. Utility ownership of these resources could displace non-utility ownership, as utilities may have monopoly information on resource quality, siting and interconnection costs, and system value that would disadvantage non-utility projects. Utility operation of these resources could also displace non-utility operation, for instance through DER aggregation in ISO markets or the non-utility market optimized model discussed in the next section. If state regulators are concerned about this crowding out of potential competition, some form of functional unbundling and open access framework would be needed to ensure that utilities provide the information and unbundled services that facilitate competition.

To address gaps in the utility market optimized model, key areas of focus include:

- Encouraging utilities to develop the technical capability to optimize distribution-level resources for distribution and bulk system value, including the ability to identify distribution constraints on short timescales, and the ability to operate distribution-level resources against ISO market prices.
- Revisiting utility programs and procurement to ensure that utilities invest in or procure distribution-level resources when cost-effective, and that utilities are contractually able to operate these resources to maximize their value.
- Examining ISO energy and capacity market designs and the incentives they provide for utility optimization of distribution-level resources.
- Developing regulatory frameworks to support competition in cases where utilities may own and operate distribution-level resources.

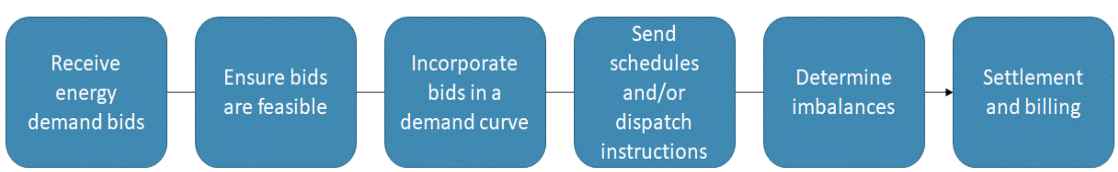
DER aggregator, customer, or ESCO ("non-utility") market optimized model

Most utilities do not have the hardware and software infrastructure in place to enable non-utility market participation of distribution-level resources through the demand side. Allowing more flexible interconnection, beyond static export limits, would help to facilitate this model but is likely not strictly necessary. Few utilities currently have the capability to do so.

A key element of this non-utility market optimized model is likely to be utility procurement of generation, transmission, and distribution capacity services. In jurisdictions with competitive retail, procurement of capacity services could be through a combination of competitive LSEs and transmission and distribution utilities. As in the utility market optimized model, utility procurement of capacity services involves utility incentives. In cases where contracting with non-utilities would displace utility rate base investments, utilities have little inherent incentive to sign contracts with non-utility providers, even when these would lower costs. Even in cases where non-utility providers can offer capacity services at below the utility's going forward fixed costs, utilities may be hesitant to sign contracts due to concerns over the "used and usefulness" of existing assets.

An additional important gap with the non-utility market optimized model is in software and, to a lesser extent, hardware — the equivalent of ISO state estimation and energy and market management software. Utilities would need to have the software capabilities to perform the range of services in the below figure (recall that non-utility providers are participating through the utility's demand curve).

Illustrative Process for Utilities to Incorporate and Settle Bids from Non-Utilities in Demand Curves



There are multiple approaches to implementing the functions in Figure @@, ranging from incremental fixes to more comprehensive solutions. For instance, for small number of projects that are concentrated in specific parts of the distribution system, utilities may not need more comprehensive monitoring and software capabilities, whereas for a larger number of projects distributed throughout the distribution system they likely would.

As with the utility market optimized model, the non-utility model involves considerations around non-discriminatory access to the distribution system. Addressing concerns over access would require some form of functional unbundling for utilities, to ensure that utilities use a "comparable service" framework to screen bids for feasibility, incorporate them into demand bids, send schedules or dispatch, and determine imbalances. For a larger number of projects, utilities would also need some form of credit management system, to guard against the risk of non-utility default on payments, as well as a fixed charge to cover utility costs related to providing market services.

Non-utility actors face similar ISO market challenges around LMP settlement and real-time price formation that utilities do. If utilities are not settled at nodal LMPs they are unlikely to pass on these prices to non-utilities, limiting the ability of non-utilities to reduce local transmission congestion. Similarly, non-utilities are unable to contribute to real-time price formation through the demand side.

To address gaps in the non-utility market optimized model, key areas of focus include:

- Incrementally developing utilities' technical and organizational capabilities to support non-utility market optimization through utility demand curves, including the ability to screen bids for distribution security and contractual requirements, assess imbalance charges, undertake settlement and billing, and levy some form of grid management charge.
- Revisiting utility programs and procurement to ensure that utilities invest in or procure capacity services from non-utility distribution-level resources when cost-effective, and that non-utilities are able to operate these resources to maximize their total value.
- Revisiting retail regulations to assess whether non-utility providers could aggregate loads into virtual power plants.
- Examining ISO energy and capacity market designs and the incentives they provide for non-utility optimization of distribution-level resources from the demand side.
- Developing distribution open access frameworks and functional unbundling requirements for distribution utilities.

DSO market optimized model

DSOs do not yet exist in the U.S. Importantly, a DSO could be housed within a utility. At a minimum, however, DSOs would need to be functionally independent from the utility, to ensure that the DSO provide non-discriminatory services for entities wishing to use the distribution system. In this and other respects, there are important parallels between the organizational and regulatory requirements for a DSO and those for transmission system operators under FERC's transmission open access order (Order 888).

Much of the regulatory framework for a DSO could be developed incrementally from regulatory changes required for other models. DSO interconnection processes could build on increasing amounts of customer flexibility in the tariff, utility, and non-utility models. The utility and non-utility market optimized models would require improvements in utility monitoring, control, and communications capabilities that are directionally consistent with what might be needed for a DSO. These models may require changes in regulation to preserve an open, competitive environment that could evolve into an open access tariff for the distribution system.

That being said, DSO implementation would require a new level of investments and sophistication in distribution operations, communications, interconnection, network planning, market settlement, and billing capabilities, relative to those in the other three models reviewed here.

Conclusions

Demand flexibility is one approach to increase electricity system flexibility, and it may ultimately turn out to be more cost-effective for some applications (e.g., distribution congestion management, local resilience) and not others (e.g., bulk system energy arbitrage, system resource adequacy). But without creating a framework that supports innovation and business models, it will be difficult to find a cost-effective balance between demand and supply-side flexibility.

There are many different approaches for encouraging demand flexibility. These different approaches are not necessarily exclusive (e.g., the customer tariff optimized and utility market models), though some are more compatible with others.

The choice among models will depend on desired levels of change, optimization, and competition. Tariffs tend to require less change but are often less precise in the response that they generate. Utility optimization can be more precise, requires more change, but tends to have limited benefits in competition and innovation. Encouraging participation by more non-utility actors in the ownership and operation of distribution-level resources can provide more precise response and enhance competition but requires more change. In choosing among these models, at some level state regulators will need to decide how much to open up the distribution system to non-utility providers.

Some of the changes needed to address gaps are common across models. These include:

- **Interconnection flexibility** — providing interconnecting customers with more choice in their distribution service levels (e.g., export limits, firm and non-firm rights). Creating the physical and software infrastructure to enable different distribution service levels is an important first step to enabling more market-based approaches congestion management on the distribution system, which will help to address the costs of transportation and building electrification.

- **Utility programs, procurement, tariffs, and incentives** — re-examining utility programs, procurement, tariffs, and rates to ensure these are aligned with least-cost procurement of distribution and bulk system services. This re-examination should also include a review of utility incentives, both specific (e.g., through performance and other incentive mechanisms) and broader (e.g., through rate cases).
- **Utility operational and communications capabilities** — building utilities' capabilities for monitoring the distribution system, anticipating reliability violations, operating distribution-level resources, and communicating with customers through system and operations software, all on short timescales.
- **Wholesale market design** — assessing the potential costs and benefits of changes in ISO market designs that would enable more effective and efficient demand-side participation in energy and capacity markets.
- **Regulatory frameworks** — developing rules for ensuring transparent, non-discriminatory access and operation of the distribution system. This framework could borrow from open access tariffs on the transmission system.

The potential to build on incremental changes to make the more sweeping changes needed for something more like a DSO highlights the value of roadmapping. Roadmaps can help utilities, regulators, and other stakeholders navigate pathways from the present to nearer-term change to potential longer-term changes.

The most important new source of demand flexibility is energy storage, whether through PV + battery systems, standalone storage systems, or EV batteries. Battery storage in particular can be sited throughout the electricity system — behind the customer meter, directly connected to the primary distribution system, directly connected to the sub-transmission or transmission systems, or behind the generator meter. Where batteries are most societally cost-effective will depend on their customer-specific or wholesale value. On the demand side, the most important value of energy storage will be its local value: avoided distribution investment costs (distribution congestion management), localized transmission congestion management, local resource adequacy, and local or customer reliability and resilience. As the examples in this paper illustrate, storage can be a useful lens to think through operational and business models for demand flexibility.