

Retail Rates, Load Growth, and the Rising Value of Demand-Side Solutions

By Brent Nelson, PhD, Senior Managing Director of Market Intelligence
May 18, 2026



*This publication is the **fifth and final** in a series from Ascend Analytics that considers the implications of rapid load growth on US capacity markets, [high capacity prices as the new normal](#), the risks that new realities create for business-as-usual strategies, and the opportunities that this paradigm shift enables for well-planned new entry resources.*

How Rate Design, Energy Efficiency, and BTM Generation Can Support Affordable Load Growth

Behind-the-meter (BTM) solar generation, energy efficiency improvements, and manufacturing offshoring have combined to reduce load over the last two decades, lessening the importance of efforts to reduce demand peaks. However, an impending return to load growth is causing a renewed focus on peak demand and grid utilization across all U.S. power markets. As the costs of serving demand growth rise and energy affordability becomes a growing problem, demand-side resources' historical challenges with customer acquisition costs and first-cost barriers may no longer apply, particularly if retail rate reform is implemented to better align customer bills to grid costs. The value of and opportunity for demand-side solutions has never been higher.

Key Takeaways

- **The economic benefits of decreasing peak demand are higher than ever as a return to load growth is increasing the cost of meeting demand peaks.** These conditions may enable business models and technologies that were previously non-viable, warranting a fresh evaluation of demand-side solutions.
- **Retail rate design reform is needed to incentivize better alignment between load and supply.** Many retail rates still reflect legacy metering solutions and are misaligned with grid costs and needs.
- **BTM resources are well-positioned to capitalize upon retail rates that lag behind wholesale market realities.** Although improved retail rate designs are needed, current legacy rates can offer higher capacity value to BTM resources than equivalent resources participating in wholesale markets.

- **Winter reliability risks enhance the value proposition for energy efficiency measures in the buildings sector.** Energy efficiency measures that reduce heating demand without duration limits are particularly well-suited to long-duration winter peaks, while also providing additional occupant benefits.

High Capacity Prices and Costs Increase the Value of Decreasing Peak Demand

As [capacity prices rise](#) and the value for generators becomes increasingly concentrated in capacity revenues and performance during narrow periods in energy markets, the costs of serving peak demand also become increasingly concentrated.

In regulated markets, demand growth beyond what can be served with existing resources requires utilities to invest in new generation and/or transmission and distribution (T&D) infrastructure that must then be added to the rate base, with the associated costs recovered from customers. While rate design can be used to allocate costs between customer classes, the binary triggering of new investment even for small increases in load can lead to significant cost increases for serving load as illustrated in **Figure 1**. The cost of serving demand growth can be even higher in deregulated markets, where capacity revenues can rise to the level needed to support new entry but would [then be paid across the entire supply stack](#). In contrast, increasing load without increasing peak demand improves utilization of existing infrastructure, reducing costs for everyone.

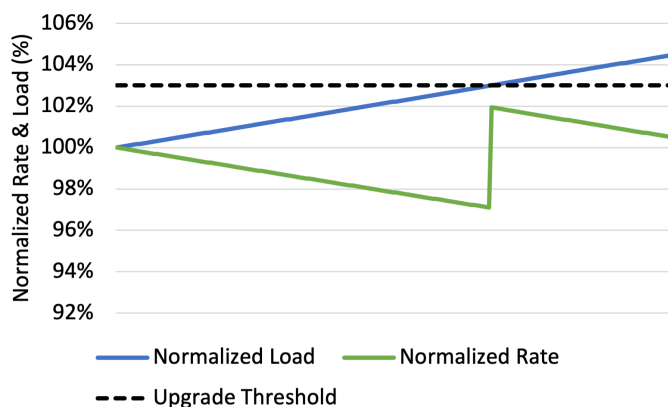


Figure 1. Conceptual Illustration of Step-Change Cost Increase When Load Growth Crosses an Upgrade Threshold

While the costs of serving growing peak demand are rising, the benefits of load reductions are amplified by several factors. Reserve margins and T&D losses mean that load reductions reduce supply obligations by an additional ~20%. Where winter reliability risks result in low accreditation of even gas generation, reduced supply needs are further amplified.

Figure 2 shows a quantitative illustration of this amplification. Reserve margin and T&D losses mean that a MW of new load requires ~1.2 MW of accredited capacity. Serving this demand with a gas combustion turbine would require ~2 MW of nameplate capacity if the gas unit is accredited at 61% (as in the PJM 2027/2028 capacity auction¹). Therefore, each MW of demand reduction can result in a ~2x reduction in needed nameplate capacity. With the cost of new gas generating capacity rapidly escalating, the value of demand reduction and these amplifications only grows larger.

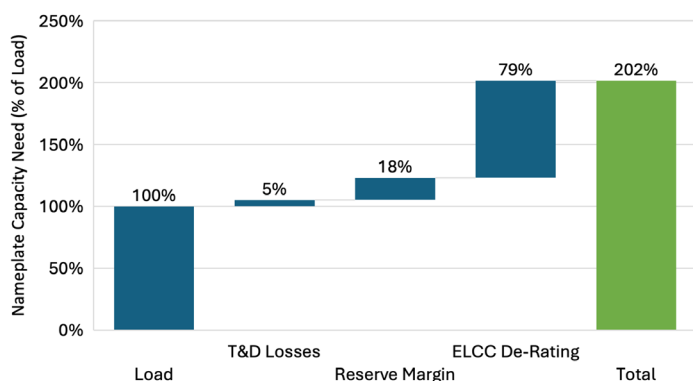


Figure 2. Required Relative Nameplate Capacity to Serve an Increase in Load with an NGCT

Retail Rates Can Drive Change but Need to Better Reflect System Costs

Most U.S. grid costs are fixed and are driven by the need to serve peak demand conditions, as **Figure 3** shows for Pacific Gas & Electric (PG&E) in California. For PG&E, fixed costs (transmission, distribution, and capacity) have historically constituted ~80% of the total retail cost, and Ascend forecasts this to grow closer to 90% by 2050. Despite this dominance of fixed costs, retail rate design in much of the country consists of a fixed volumetric price (¢/kWh) for residential customers along with a combination of a fixed volumetric price (¢/kWh) and a monthly peak demand charge (\$/kW) based on the maximum demand during the given month for commercial and industrial customers.

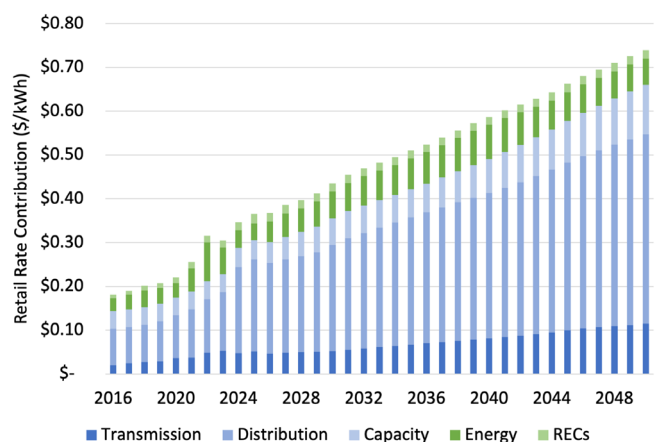


Figure 3. Retail Rate Contribution by Source for Pacific Gas Electric in CAISO (Historical and Ascend Forecast)

This billing structure reflects the legacy of simple meters that could not determine the time at which consumption occurred. However, the broad rollout of advanced metering infrastructure (AMI) enables retail rates that can more accurately reflect grid costs, including different rates by peak period and demand charges based on coincident system peaks or minimum system reserve conditions rather than on-site monthly peaks.

While the potential effectiveness of alternative rate designs has been articulated for years² and these AMI-enabled rate structures are becoming more common, implementation remains variable and the rates are often poorly designed. Even where peak period rates have been implemented, the spread between rates has sometimes been too small to drive meaningful changes in consumer behavior, with off-peak rates only marginally lower than the flat rate (as is presently the case in the Xcel Colorado service territory, where Ascend is headquartered). Even coincident demand charges do not necessarily align to the drivers of new capacity needs, which are driven by

minimum system reserve conditions that are increasingly decoupling from gross load peaks due to solar buildout. Applying more thoughtful rate structures could bring pronounced benefits for grid utilization, which has been the subject of growing attention from rate reform advocates and large load customers in recent years.³ Improving grid utilization would be easier if utility bills allocated costs in a way that created incentives for load to better utilize the grid. Smart loads, energy efficiency measures, demand response programs, and consumer behavioral shifts all become easier to support and implement when their true value to the grid can be monetized. **Figure 4** illustrates how much load growth could theoretically be enabled across U.S. markets with more peak demand flexibility, though incentivizing that flexibility remains the challenge.

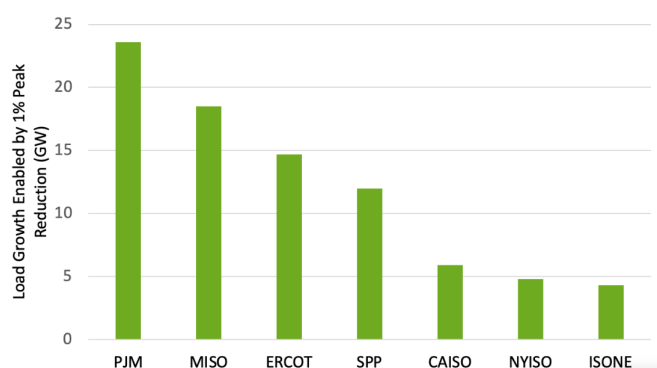


Figure 4. Estimated Load Growth Enabled by Flexible Load from Study by Duke Nicholas Institute.⁴

Some opponents of more complex retail pricing structures express concerns about equity or customers' inability to understand such rate structures.⁵ However, these criticisms often fail to account for the fact that the notion of time-varying pricing has become a fixture throughout the U.S.: for example, consumers are no longer surprised by surge pricing for rideshares, matinee pricing for movies, dynamic roadway tolls, and variable prices for parking, flights, concerts, hotels, and other goods during periods of high demand. While the population has grown accustomed to flat volumetric pricing for electricity, it is highly probable that ratepayers would adapt quickly to the concept of time-dependent pricing, when it is already present in so many other areas of the economy.

BTM Generation Will Hold Extra Benefits for as Long as Retail Rate Designs Lag Behind True System Cost Contributions

While obsolete retail rate designs do not create the right incentives for load, they can create advantages for BTM resources relative to front-of-the-meter (FTM) resources participating in wholesale markets. As noted previously, reductions in metered demand result in an amplified

reduction in supply needs and corresponding capacity cost allocations. Additionally, while capacity markets and accreditation bodies may give limited accreditation value to duration-limited resources like demand response or storage, BTM storage and peak shifting can effectively receive 100% accreditation by reducing on-site peak demand.

For example, the four coincident peak (4CP) program in ERCOT assigns demand charges to load based on consumption during the four monthly system 15-minute peaks from June through September, leading to large savings for sites that can deploy BTM resources during these periods regardless of consumption during the remaining periods. Moreover, critical reliability conditions (and supply addition needs) are increasingly moving to long, flat, overnight peaks during severe winter storms, which are much less well-served by storage and demand response. These potential misalignments led the Texas Senate to call for a re-evaluation of whether the 4CP program appropriately allocates costs as part of SB6,⁶ though the outcome of this re-evaluation remains unclear at the time of this writing.

Similar situations arise in other U.S. power markets. In PJM, demand charges are based on consumption during the five highest system peak hours during the year (typically during the summer). However, wholesale accreditation is largely tied to severe winter conditions when reserves are at their lowest and peaks are longest. In NYISO, the New York VDER program allows distributed storage to receive 100% of capacity revenues if it dispatches during the single highest peak hour while giving utility-scale storage a rapidly declining accreditation value in capacity markets.

BTM storage also offers the added benefit of supporting speed to power, which is crucial as data centers race to come online. With new supply unable to keep up with the appetite for new demand, data centers can accept interruptible/curtailable grid service and deploy BTM storage when grid service is interrupted. More common but shorter-duration summer peaks can be met with BTM storage even while wholesale storage has its capacity value eroded by less frequent but longer-duration winter-driven accreditation conditions.

Winter Reliability Risks Enhance the Value of Energy Efficiency

As winter reliability risks become more prominent in driving critical grid conditions, demand reduction resources hold additional benefits. While storage and demand response both run into duration limits, and even

BTM gas runs into outage and fuel disruption risks, passive energy efficiency measures structurally reduce demand peaks without a duration limit or customer override risk. These reductions yield permanent benefits that are amplified by the reserve margin, T&D, and accreditation benefits noted above.

Because most winter peaks occur overnight when heating is the primary energy demand driver, the most valuable energy efficiency measures will be those that reduce the electricity required to provide heat. Improved building insulation and efficient windows reduce heating needs while simultaneously increasing comfort for building occupants. Reducing thermal leakage also increases resilience and comfort during the power outages that can often occur during severe weather due to T&D disruption. Smart thermostats enable coordinated heating operation with reduced peaks. Replacing electrical resistance heat with heat pumps, or replacing less efficient heat pumps with more efficient ones, allows the same heating loads to be met with lower power demand. While these measures can often be economically competitive under current retail rate designs, a retail rate design that more accurately reflects the cost of serving peak demand will only increase the economic benefits that peak-reducing energy efficiency measures provide.

The Value and Need for Peak Reduction has Never Been Higher

As load growth returns to U.S. power markets, energy costs are being driven up by the T&D upgrades and new supply needed to meet growing peak demand, even though these resources will only be needed for a relatively narrow range of conditions and hours. In contrast, adding load without increasing (or even while reducing) peaks would increase the utilization of existing infrastructure and reduce electricity prices, with the reduced supply needs substantially larger than the demand reductions.

Thus, the value of reducing demand during critical grid conditions has never been higher. This value needs to be reflected in retail rate design reform so that consumers are incentivized to align consumption to grid needs and so that solutions that provide demand reduction are appropriately valued.

Interested in Learning More?

[AscendMI™ \(Ascend Market Intelligence\)](#) delivers proprietary power market forecasts that have been trusted in hundreds of projects and resource planning activities, supporting over 300 gigawatts and \$25 billion in project assessments. Please [contact us](#) to learn more.

¹ <https://www.pjm.com/-/media/DotCom/planning/res-adeq/elcc/2027-28-bra-elcc-class-ratings.pdf>

² <https://rmi.org/wp-content/uploads/2017/04/A-Review-of-Alternative-Rate-Designs-2016.pdf>

³ <https://www.utilizecoalition.org/>

⁴ Norris, T. H., T. Profeta, D. Patino-Echeverri, and A. Cowie-Haskell. 2025. Rethinking Load Growth: Assessing the Potential for Integration of Large Flexible Loads in US Power Systems. NI R 25-01. Durham, NC: Nicholas Institute for Energy, Environment & Sustainability, Duke University <https://nicholasinstitute.duke.edu/publications/rethinking-load-growth>

⁵ https://rmi.org/wp-content/uploads/2018/07/RMI_Better_Rate_Design_2018.pdf

⁶ <https://legiscan.com/TX/text/SB6/2025>

Ascend Analytics is the leading provider of market intelligence and analytics solutions for the power industry.

The company's offerings enable decision makers in power supply, procurement, and investment markets to plan, operate monetize, and manage risk for renewable energy and storage assets. From real-time to 30-year horizons, their forecasts and insights are at the foundation of over \$50 billion in project financing assessments.

Ascend provides energy market stakeholders with the clarity and confidence to successfully navigate the rapidly shifting energy landscape.

Visit ascendanalytics.com for more information.

