2018 State of the Market Report for the ERCOT Electricity Markets

POTOMAC ECONOMICS

Independent Market Monitor
for ERCOT

June 2019

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Executive Summary

As the Independent Market Monitor (IMM) for the Public Utility Commission of Texas (Commission), Potomac Economics provides this report which reviews and evaluates the outcomes of the Electric Reliability Council of Texas (ERCOT) wholesale electricity market in 2018. It is submitted to the Commission and ERCOT pursuant to the requirement in §39.1515(h) of the Public Utility Regulatory Act (PURA). It includes assessments of the incentives provided by the current market rules and analyses of the conduct of market participants.

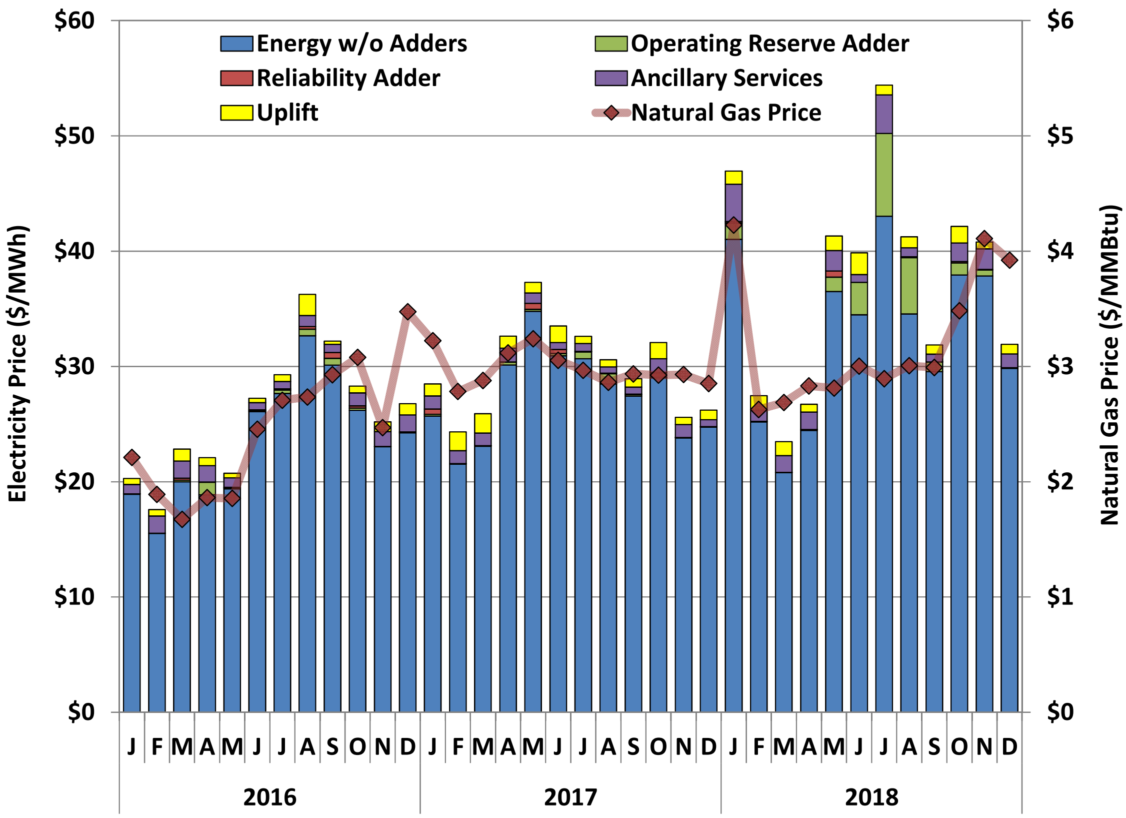
Overall, the ERCOT wholesale market performed competitively in 2018. Key findings and results from 2018 include the following:

- Higher natural gas prices led to higher energy prices in 2018:
  - The ERCOT-wide load-weighted average real-time energy price was $35.63 per MWh in 2018, a 26% increase from 2017.
  - The average price for natural gas was 8% higher in 2018 than in 2017, increasing from $2.98 per MMBtu in 2017 to $3.22 per MMBtu in 2018.
- Contrary to generally held expectations, market conditions were rarely tight, but prices did reach historic peaks. Real-time prices exceeded $3,000 per MWh for approximately 45 minutes, reaching $9,000 per MWh for the first time on January 23 for a duration of about ten minutes.
- A new record peak hour demand of 73,473 MW was set on July 19, 2018. This was a 3.3% increase from previous highest hourly demand which occurred in 2016. Average demand increased by 5.3% from 2017.
- The total congestion costs experienced in the ERCOT real-time market in 2018 were $1.26 billion, an increase of 30% from 2017. A costly, localized constraint in far west Texas was the primary cause of the increase.
- Net revenues provided by the market during 2018 were less than the estimated amount necessary to support new greenfield natural gas generation investment, even in a year with historically low installed reserve margins. Based on the tepid reaction to higher ERCOT market prices, it is unclear whether the mix of new generation additions will be sufficient to meet the growing demands in ERCOT. The Operating Reserve Demand Curve (ORDC), combined with a relatively high offer cap, should increase net revenues when shortages become more frequent.
- Although the market performed competitively, we continue to recommend a number of key improvements to ERCOT's pricing, resource commitment process, and dispatch. These improvements are summarized at the end of this Executive Summary.

Review of Real-Time Market Outcomes

Although only a small share of the power produced in ERCOT is transacted in the real-time market, real-time energy prices set the expectations for prices in the day-ahead market and bilateral forward markets where most transactions occur. Unless there are barriers preventing arbitrage of the prices between the real-time and forward markets, prices in the forward markets should be directly related to the prices in the real-time market. The figure below summarizes changes in energy prices and other market costs by showing the all-in price of electricity, which is a measure of the total cost of serving load in ERCOT. The ERCOT-wide price in this figure is the load-weighted average of the real-time market prices from all Load Zones. Ancillary services costs and uplift costs are divided by real-time load and shown on a per MWh basis.[1]

Average All-in Price for Electricity in ERCOT



ERCOT real-time prices include the effects of two energy price adders that are designed to improve real-time energy pricing when conditions warrant or when ERCOT takes out-of-market actions for reliability. Although published energy prices include the effects of both adders, the

1 For this analysis Uplift includes: Reliability Unit Commitment Settlement, Operating Reserve Demand Curve (ORDC) Settlement, Revenue Neutrality Total, Emergency Energy Charges, Base Point Deviation Payments, Emergency Response Service (ERS) Settlement, Black Start Service Settlement, Block Load Transfer Settlement, and the ERCOT System Administrative Fee.

Operating Reserve Demand Curve Adder (operating reserve adder) and the Reliability Deployment Price Adder (reliability adder) are shown separately here from the energy price.

The largest component of the all-in price is the energy cost. The figure above indicates that natural gas prices continued to be a primary driver of energy prices. This correlation is expected in a well-functioning, competitive market because fuel costs represent the majority of most suppliers' marginal production costs. Because suppliers in a competitive market have an incentive to offer supply at marginal costs and natural gas is the most widely-used fuel in ERCOT, changes in natural gas prices should translate to comparable changes in offer prices. Hence, the 8% increase in natural gas prices, combined with increased shortage conditions, contributed to a 26% increase in ERCOT's average real-time energy prices. The average real-time energy price in 2018 included relatively small contributions from ERCOT's energy price adders: $1.97 per MWh from the operating reserve adder and $0.08 per MWh from the reliability adder.

Other cost categories continue to be a small portion of the all-in electricity price. Ancillary service costs were $1.60 per MWh in 2018, up from $0.87 per MWh in 2017 because of the increase in natural gas prices and larger responsive reserve requirements. Uplift costs, including the ERCOT system administrative fee, accounted for $1.08 per MWh of the all-in electricity price, up from $1.02 per MWh in 2017.

Real-Time Energy Prices

Energy prices vary across the ERCOT market because of congestion costs that are incurred as power is delivered over the network.

Average Annual Real-Time Energy Market Prices by Zone

($/MWh)	2011	2012	2013	2014	2015	2016	2017	2018
ERCOT	**$53.23**	**$28.33**	**$33.71**	**$40.64**	**$26.77**	**$24.62**	**$28.25**	**$35.63**
Houston	$52.40	$27.04	$33.63	$39.60	$26.91	$26.33	$31.81	$34.40
North	$54.24	$27.57	$32.74	$40.05	$26.36	$23.84	$25.67	$34.96
South	$54.32	$27.86	$33.88	$41.52	$27.18	$24.78	$29.38	$36.15
West	$46.87	$38.24	$37.99	$43.58	$26.83	$22.05	$24.52	$39.72
($/MMBtu)								
Natural Gas	$3.94	$2.71	$3.70	$4.32	$2.57	$2.45	$2.98	$3.22

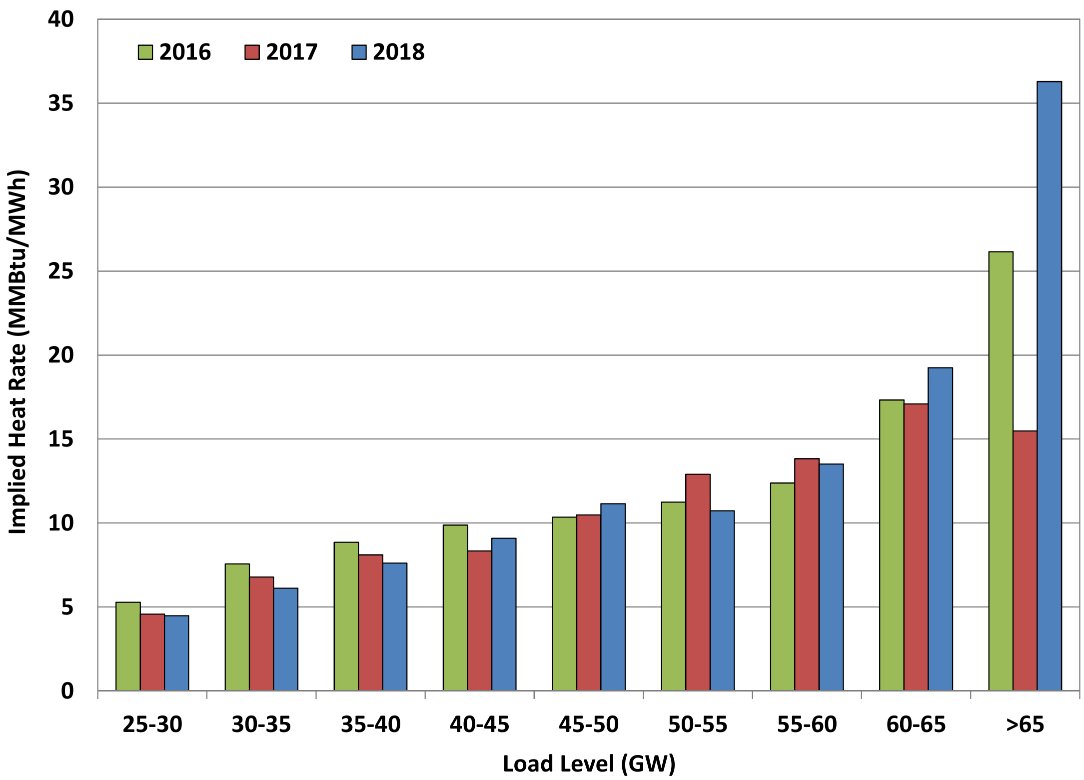
The table above provides the annual load-weighted average price for each zone for the past eight years. The pattern of zonal prices was different in 2018 compared to recent years. Although slightly narrower than in 2017, price spreads were still noticeable across all zones in 2018 because of higher natural gas prices and the increased impacts of transmission congestion. And for the first time since 2014, the West zone had the highest prices, primarily due to multiple localized real-time transmission constraints.

West zone prices relative the ERCOT average have varied through the years. Prior to 2012, West zone prices were lower than the ERCOT average because of wind generation surplus resulting from export limitations. Between 2012 and 2014, load growth caused by higher oil and natural gas production activity resulted in localized import constraints and higher prices. Even as investment in transmission facilities continued, the amount of wind generation additions began to create export limitations and resulting lower prices beginning in 2015. That trend was reversed in 2018 due to multiple real-time transmission constraints.

Non-Fuel Energy Price Changes

To summarize the changes in energy prices related to factors other than fuel cost, an "implied heat rate" is calculated by dividing the real-time energy price by the natural gas price.

Implied Heat Rate and Load Relationship

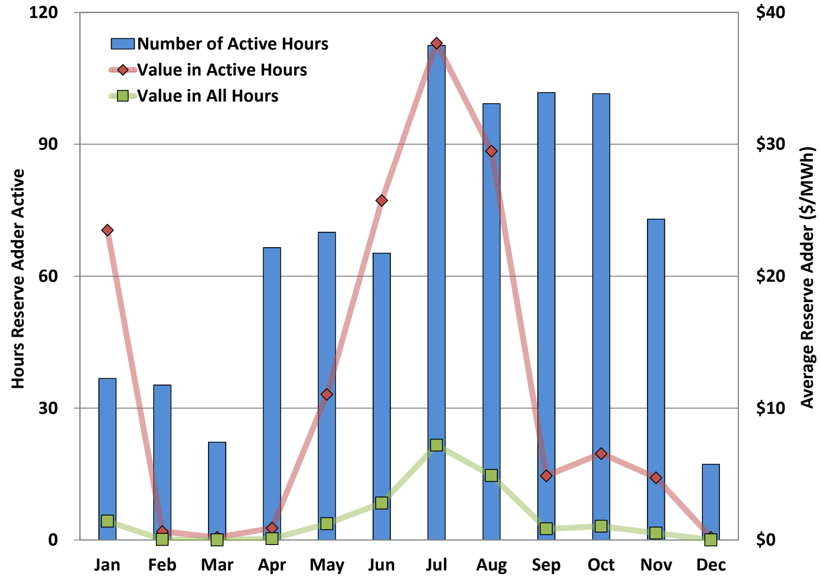


The figure shows the average implied heat rate at various system load levels from 2016 through 2018. In a well-performing market, a clear positive relationship between implied heat rate and load level is expected because resources with higher marginal costs are dispatched to serve higher loads. This relationship continued and the relatively modest effects of increased shortage conditions in 2018 can be seen at the highest load levels.

Energy Price Adders

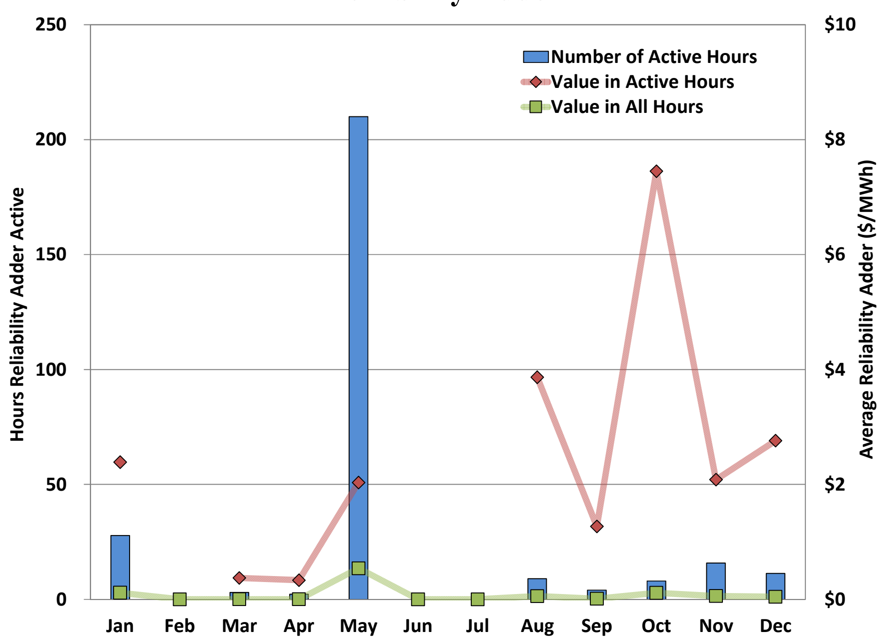
The contributions of the energy price adders were relatively small in 2018 although greater than in recent years due to slightly more frequent occurrences of shortage conditions. Overall, the operating reserve adder contributed $1.97 per MWh or 5.5% to the annual average real-time energy price.

Operating Reserve Adder



The next figure shows the impacts of the reliability adder. The reliability adder reflects the incremental costs of reliability actions taken by ERCOT, including reliability unit commitments (RUC) and deployed load capacity.

Reliability Adder

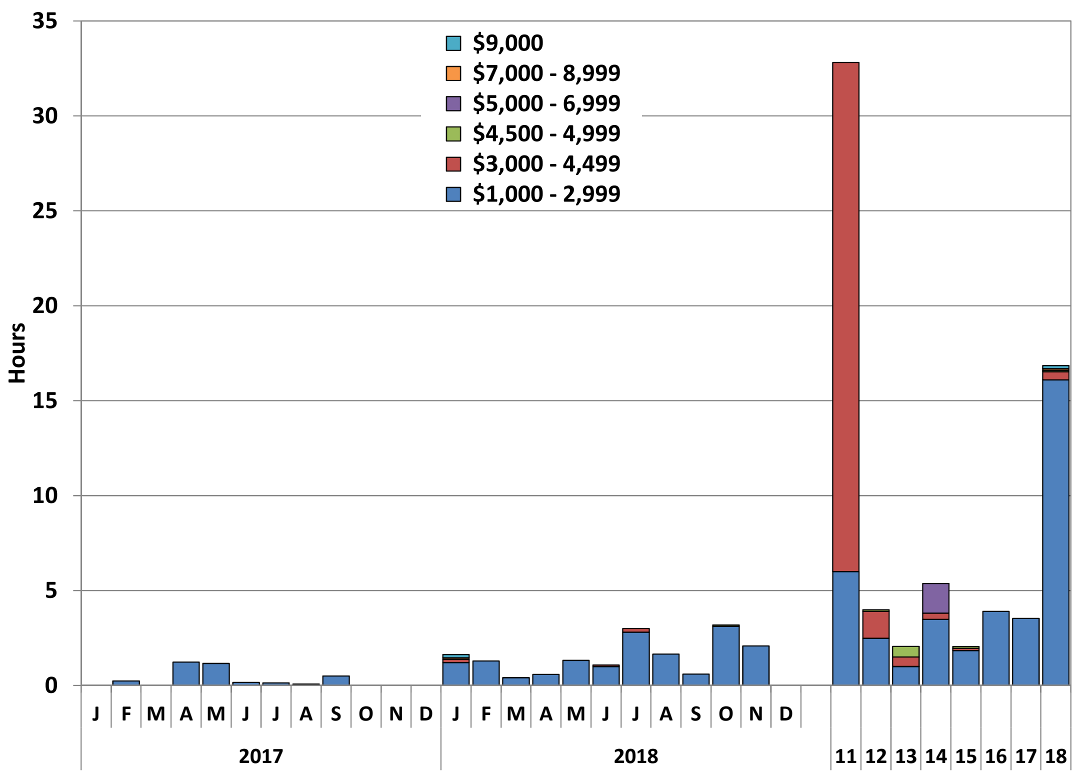


The reliability adder was non-zero for 291 hours, or 3% of the time in 2018, most of which occurred in May. The reliability adder had very little overall effect on market outcomes in 2018 as its contribution to the annual average real-time energy price was $0.08 per MWh.

As an energy-only market, the ERCOT market relies heavily on high real-time prices that occur during shortage conditions. These prices provide key economic signals that provide incentives to build new resources and retain existing resources. However, the frequency and impacts of shortage pricing can vary substantially from year-to-year.

The figure below shows the aggregate amount of time when the real-time system-wide energy price exceeded $1,000 per MWh, by month for 2017 and 2018. Also provided are annual summaries for 2011 through 2018. This figure shows that high prices occurred more frequently in 2018 than in any previous year since 2011. Prices greater than $1,000 per MWh occurred in nearly 17 hours over the entire year. On January 23, 2018, prices reached the $9,000 system-wide offer cap for the first time in ERCOT's history.

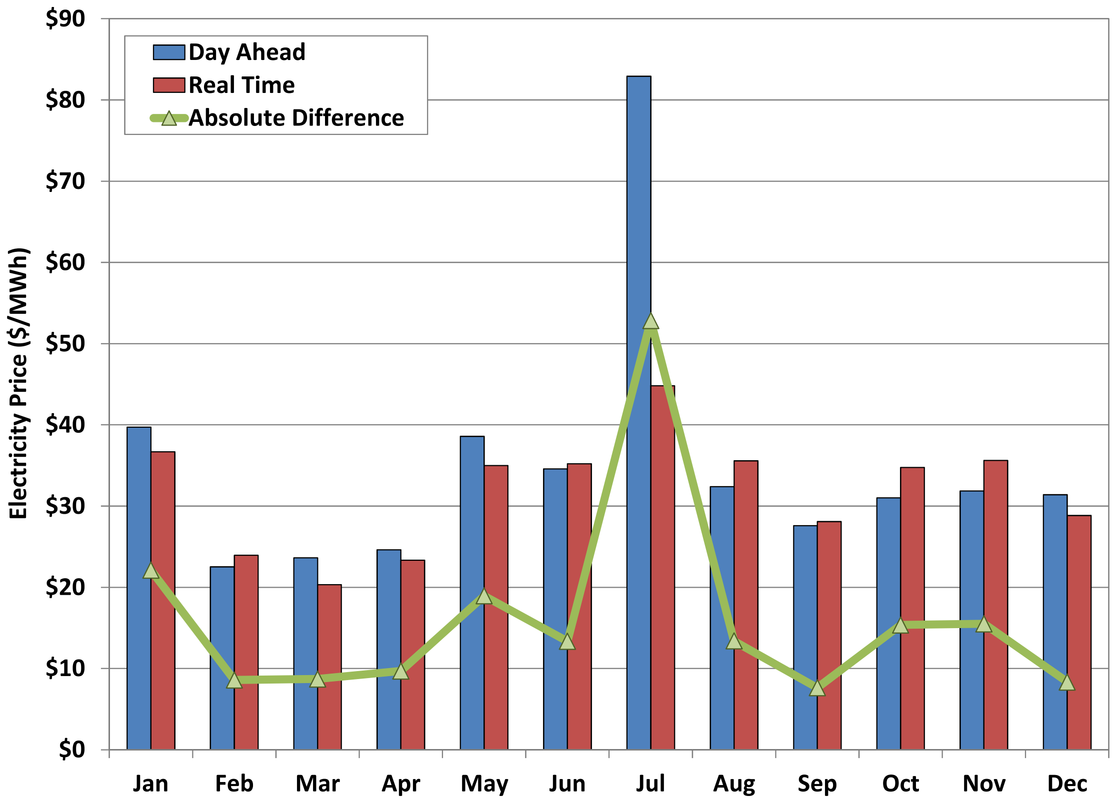
Duration of High Prices



Day-Ahead Market Performance

ERCOT's day-ahead market allows participants to make financially binding forward purchases and sales of power for delivery in real-time. Although all bids and offers are evaluated for the ability to reliably flow on the transmission network, there are no operational obligations resulting from the day-ahead market. These transactions are made for a variety of reasons, including satisfying the participant's own demand, managing risk by hedging the participant's exposure to real-time prices or congestion, or arbitraging the real-time prices. For example, load-serving entities can insure against the higher volatility of real-time market prices by purchasing in the day-ahead market. The day-ahead market helps inform participants' generator commitment decisions. For all of these reasons, the effective performance of the day-ahead market is essential.

Convergence Between Day-Ahead and Real-Time Energy Prices



Day-ahead market performance is primarily evaluated by its convergence with the real-time market because the real-time market reflects actual physical supply and demand for electricity. In a well-functioning market, participants should eliminate sustained price differences on a risk-adjusted basis by making day-ahead purchases or sales to arbitrage the price differences.

Price convergence was evident in all months in 2018 except July, when day-ahead and real-time prices diverged sharply. Day-ahead and real-time prices averaged $35 and $32 per MWh in 2018, respectively.[2] The average absolute difference between day-ahead and real-time prices was $16.21 per MWh in 2018. This represents an increase from $8.60 per MWh in 2017.

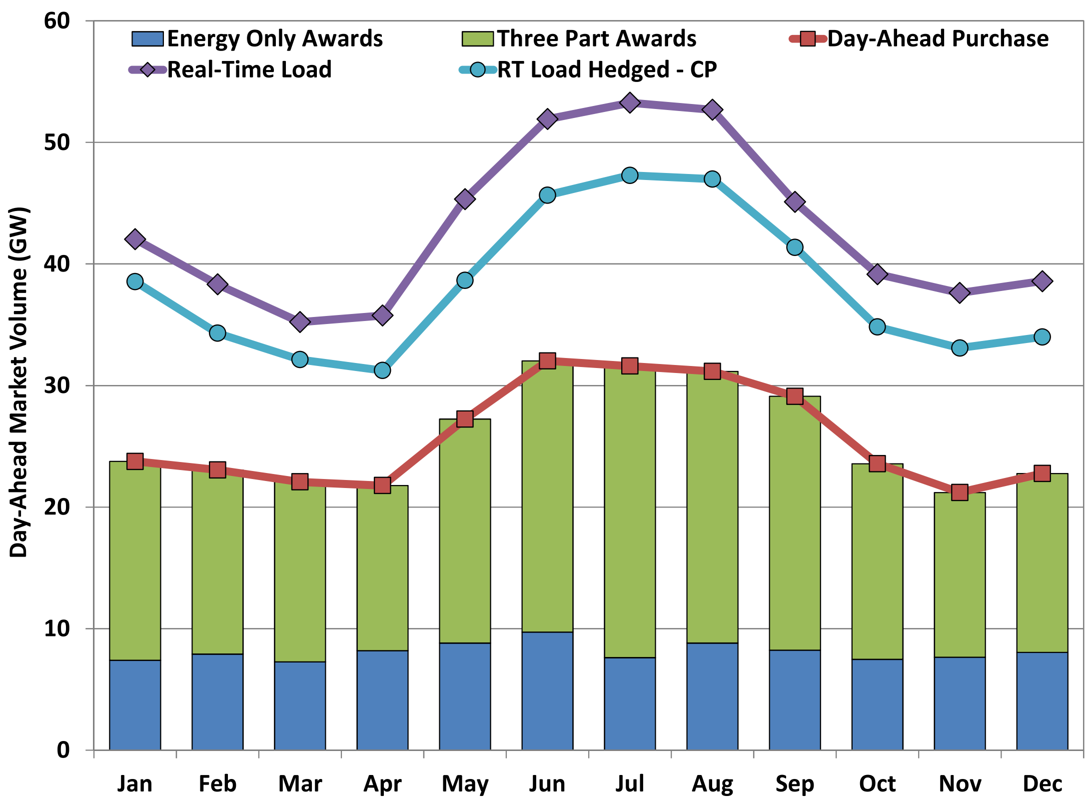
2 These values are simple averages, not load-weighted.

The increased day-ahead premium in 2018 is consistent with expectations because of the much higher volatility of real-time prices. Risk is lower for loads purchasing in the day-ahead market and higher for generators selling day-ahead. The higher risk for generators is associated with the potential of incurring a forced outage and having to buy back energy at real-time prices. This explains why the highest premiums occurred during July 2018 with the highest relative demand and highest prices.

Day-Ahead Market Volumes

The next figure summarizes the volume of day-ahead market activity by month, which includes both the purchases and sales of energy, as well as the volume of Point-to-Point (PTP) obligations that represent the system flows between a Load Zone and other locations.

Volume of Day-Ahead Market Activity by Month



The figure shows that the volume of day-ahead purchases provided through a combination of three-part generator-specific offers (including start-up, no-load, and energy costs) and virtual energy offers was approximately 60% of real-time load in 2018, which was an increase compared to 55% in 2017.

PTP obligations are financial transactions purchased in the day-ahead market. Although PTP obligations do not themselves involve the direct supply of energy, PTP obligations allow a participant to buy the network flow from one location to another.[3] When coupled with a self-scheduled generating resource, the PTP obligation allows a participant to serve its load while avoiding the associated real-time congestion costs between the locations. Other PTP obligations are scheduled by financial participants seeking to arbitrage locational congestion differences between the day-ahead and real-time markets.

To estimate the volume of hedging activity taking place in the day-ahead market, energy purchases are added to the volume of PTPs scheduled by Qualified Scheduling Entities (QSEs) with load that source or sink in Load Zones. This total is shown as the "Real-Time Load Hedged" shown in the figure above. Under this methodology, approximately 64% of QSEs' real-time load was determined to be hedged in the day-ahead market. This is a decrease from 2017 when 82% of QSEs load was determined to be hedged. Although QSEs are the party financially responsible to ERCOT, their financial obligations may be aggregated and held by a Counterparty. When measured at the Counterparty level, the percentage of real-time load hedged increased to 89%, which was similar to the amount seen in 2017.

Ancillary Service Prices

The next table compares the average annual price for each ancillary service in 2018 with 2017. Higher prices for the "up" ancillary services, responsive, non-spin and up regulation, in 2018 are explained by the combination of larger requirements, and expectations for high energy prices as evidenced by high day-ahead energy prices. The decrease in the average price of down regulation is explained by lower opportunity costs of providing that service due to more capacity on line to meet the higher load requirements in 2018.

Average Annual Ancillary Service Prices by Service

	2017 ($/MWh)	2018 ($/MWh)
Responsive Reserve	**$9.77**	**$17.64**
Nonspin Reserve	**$3.18**	**$9.20**
Regulation Up	**$8.76**	**$14.03**
Regulation Down	**$7.48**	**$5.19**

After several years of declining quantities, total requirements for ancillary services increased in 2018. The average total requirement in 2018 was greater than 4,900 MW, an increase from the average total requirement of approximately 4,800 MW in 2017 and roughly equal to the 2016 requirements. The principal cause for the overall increase in ancillary service requirements in 2018 was larger responsive reserve requirements due to an emphasis on ensuring adequate online system inertia.

3 PTP Obligations are equivalent to scheduling virtual supply at one location and virtual load at another.

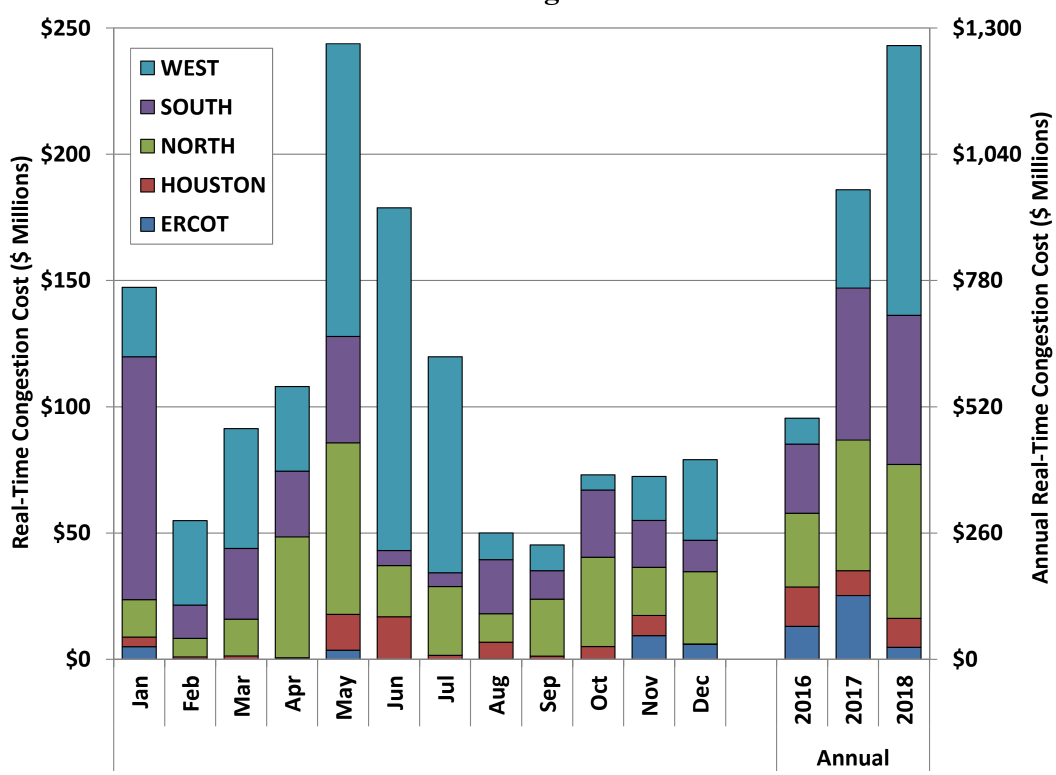
Transmission and Congestion

Congestion arises when the transmission network does not have sufficient capacity to dispatch the least expensive generators to satisfy demand. When congestion occurs, clearing prices vary by location to reflect the cost of meeting load at each location. These nodal prices reflect that higher-cost generation is required at locations where transmission constraints prevent the free flow of power from the lowest-cost resources.

The total congestion costs experienced in the ERCOT real-time market in 2018 were $1.26 billion, a 30% increase from the 2017 value. A costly, localized constraint in far west Texas was the primary cause of the increase. Congestion on the Yucca Switch to Basin 138 kV lines was directly related to high loads associated with increased oil and gas production and generation from solar resources. All zones experienced increased congestion in 2018, though inter-zonal congestion decreased. The decrease in inter-zonal congestion can be attributed to the completion of the Houston Import Project in April 2018.

The next figure displays the amount of real-time congestion costs associated with each geographic zone, with the monthly values of 2018 preceding the annual values for the last three years. Costs associated with constraints that cross zonal boundaries (for example North to Houston) are shown in the "ERCOT" category.

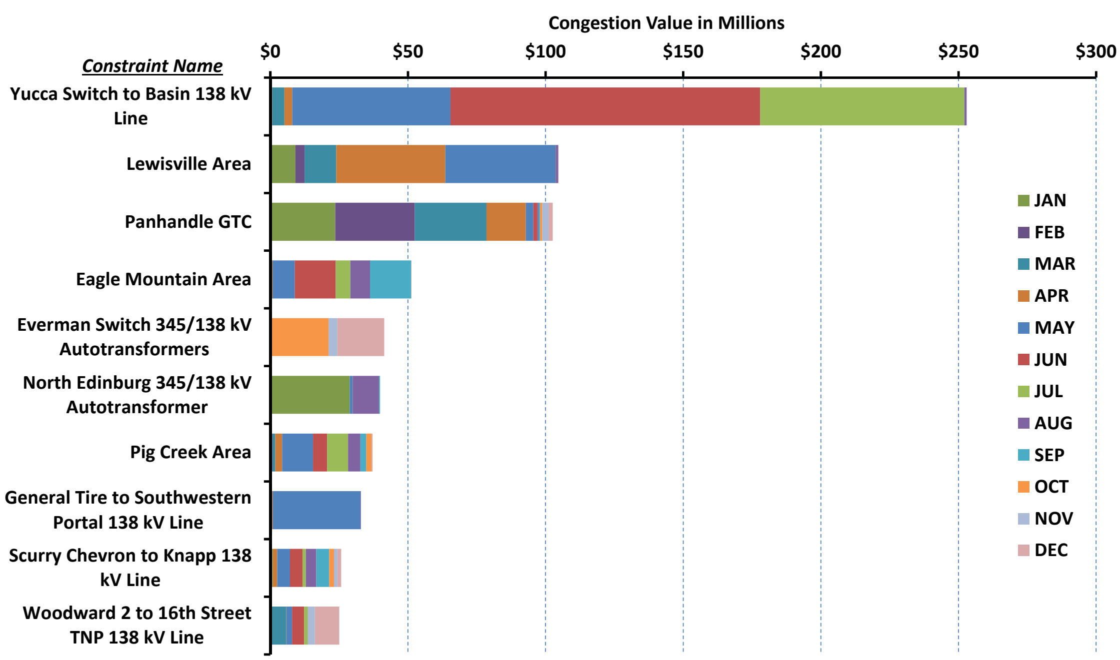
Real-Time Congestion Costs



The months of February, August, and September exhibited the least amount of congestion costs, whereas the remaining months reflected much higher congestion. This 2018 monthly profile is relatively unexpected because shoulder months usually have the highest congestion costs. Shoulder months are when most transmission and generation outages for maintenance and upgrades occur. The increased congestion in January was due to cold weather conditions and higher than expected load conditions in the far west increased congestion costs in May, June and July.

The figure below displays the ten most costly real-time constraints as measured by congestion value. The constraint with the highest congestion cost in 2018 at $253 million was a series of two 138 kV lines connecting Yucca Switch and Gas Pad and on to Basin substation. The majority of the congestion value was generated on the line between Yucca Switch to Gas Pad. The congestion cost associated with Yucca Switch to Gas Pad in 2018 was almost double the most costly constraint in 2017.

Most Costly Real-Time Constraints



The Panhandle constraint dropped from the top valued constraint in 2017 to third in 2018. The Panhandle constraint caused $102 million of congestion in 2018, a 30% decrease from $139 million in 2017. By the end of 2018, there was almost 5 GW of generation capacity in the Panhandle area, 90% of which was wind generation. The Panhandle Generic Transmission Constraint (GTC) limit ranged in value from 2,100 MW to 4,300 MW during 2018. The GTC limit average was 3,500 MW, up by 400 MW from 3,100 MW in 2017, a 12% increase. The Panhandle GTC was active for 14% of the time in 2018, down from 16% in 2017.

Demand and Supply

Load in 2018

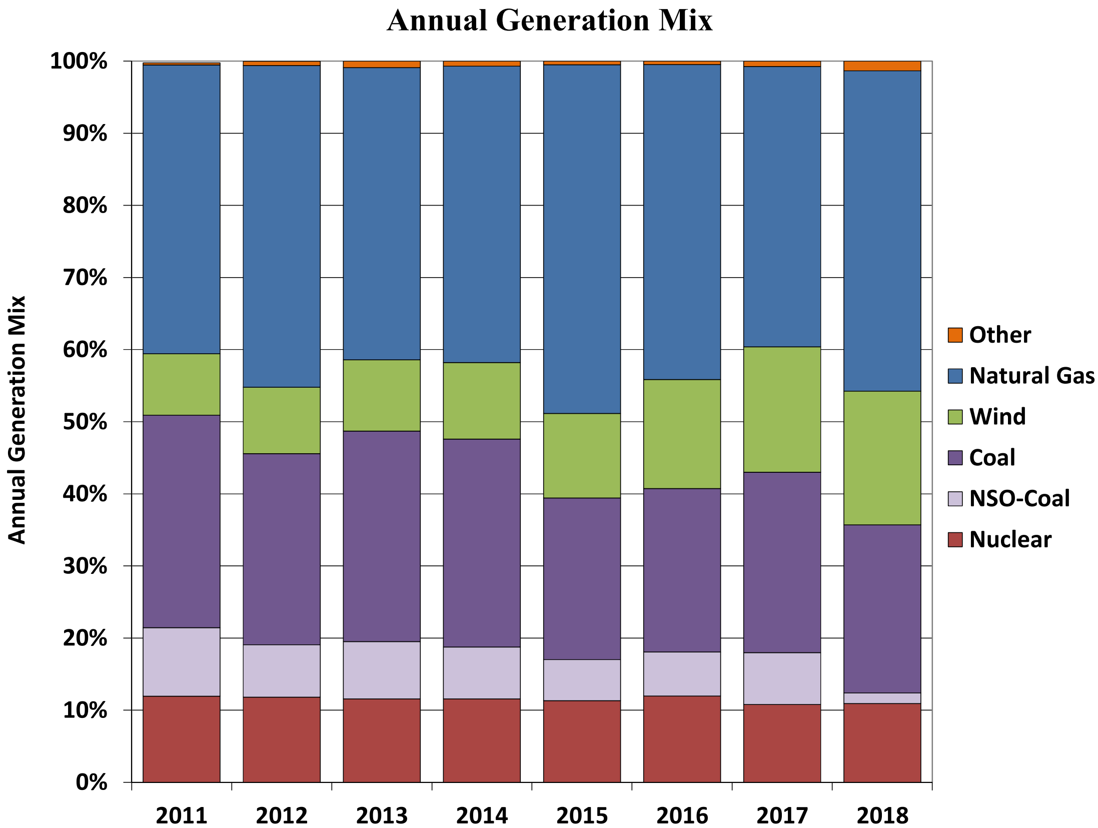
Total ERCOT load in 2018 increased 5.3% from 2017 to total 376.4 TWh. This equates to approximately 2,200 MW per hour on average. All zones showed an increase in average real-time load in 2018 ranging from 2% increase in Houston to 15% increase in the West zone. Continuing robust oil and natural gas production activity in the West zone has been the driver for the high growth experienced recently. Weather impacts on load in 2018 were mixed across the zones. There were fewer annual cooling degree days across all zones. However, for the three summer months of June, July and August, there was a 12% increase from 2017 in the number of cooling degree days in Dallas. For the same time frame, Austin had a 6% increase and Houston was flat.

Summer conditions in 2018 produced a new record peak load of 73,473 MW on July 19, 2018, surpassing the previous ERCOT-wide coincident peak hourly demand record of 71,110 MW set on August 11, 2016. A new winter peak demand record of 65, 915 MW was also set on January 17, 2018. All zones experienced varying increases in peak load ranging from 0.5% increase in Houston to more than 13% increase in the West zone, which continued to experience the highest percentage growth in peak load, due to continuing growth in oil and natural gas production.

Generating Resources

Approximately 3.8 GW of new generation resources came online in 2018, the bulk of which was multiple wind resources with total capacity of 2.2 GW, and an effective peak serving capacity of less than 520 MW. Approximately 150 MW of this wind capacity were re-powered resources. The remaining capacity additions were 670 MW of new combustion turbines and 880 MW of solar resources. A total of nine generation resources, all coal fueled, totaling 5.1 GW, were retired and permanently decommissioned in 2018.

The shifting contribution of coal and wind generation is evident in the figure below showing the percentage of annual generation from each fuel type for the years 2011 through 2018.



The generation share from wind has increased every year, reaching almost 19% of the annual generation requirement in 2018, up from 9% in 2011 and 17% in 2017. The share of generation from coal decreased to 25%, by far the lowest total since well before the advent of a competitive generation market in ERCOT. This figure separately shows the amount of energy provided from seven coal units that were retired at the start of 2018 and two coal units that were retired at the end of the year. The seven coal units provided an average of 7% of the total annual generation requirements from 2011 to 2017, and the two other units provided an additional 2%. In response to the reduction in coal generation, the share of natural gas generation increased again in 2018 to 44%, up from 39% in 2017.

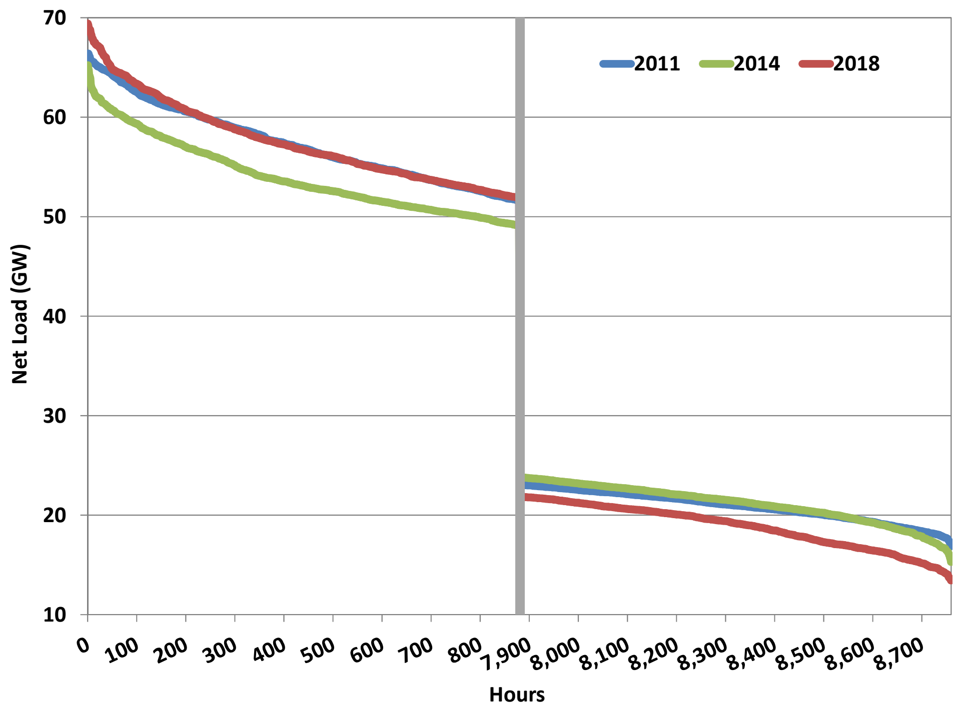
Wind Output

ERCOT continued to set new records for peak wind output in 2018. On December 14, wind output exceeded 19 GW, setting the record for maximum output and on December 27, wind provided nearly 55% of the total load, also a new record.[4] Increasing levels of wind resources in ERCOT have important implications for the net load duration curve faced by the non-wind fleet of resources. Net load is defined as the system load less wind production. The figure below

4 Peak hourly wind generation was 19,168 MW at 12:07 a.m. on December 14, 2018. Instantaneous wind penetration was 54.6% at 4:57 a.m. on December 27, 2018.

shows net load ranked from highest to lowest in GW, with only the highest and lowest deciles displayed.

Top and Bottom Deciles (Hours) of Net Load



The contribution from wind generation results in only modest reductions of the net load relative to the actual load during the highest demand hours, but much larger reductions in the net load in the other hours of the year. Wind generation erodes the total load available to be served by base load coal units, while doing very little to reduce the amount of capacity necessary to reliably serve peak load.

In the hours with the highest net load (left side of the figure below), the difference between peak net load and the 95th percentile of net load has averaged 12.1 GW the past three years. This means that 12.1 GW of non-wind capacity is needed to serve load less than 440 hours per year.

In the hours with the lowest net load (right side of the figure above), the minimum net load has dropped from approximately 20 GW in 2007 to below 13.4 GW in 2018, even with the sizable growth in annual load that has occurred. This trend has put operational pressure on the almost 20 GW of nuclear and coal generation that were in-service in 2018. This operational pressure was certainly one of the contributors to the recent retirement of more than 5 GW of coal in 2018.

Thus, although the peak net load and reserve margin requirements are projected to continue to increase and create an increasing need for non-wind capacity to satisfy ERCOT's reliability requirements, the non-wind fleet can expect to operate for fewer hours as wind penetration increases. This outlook further reinforces the importance of efficient energy pricing during peak demand conditions and other times of system stress, particularly in the context of the ERCOT energy-only market design.

Reliability Commitments

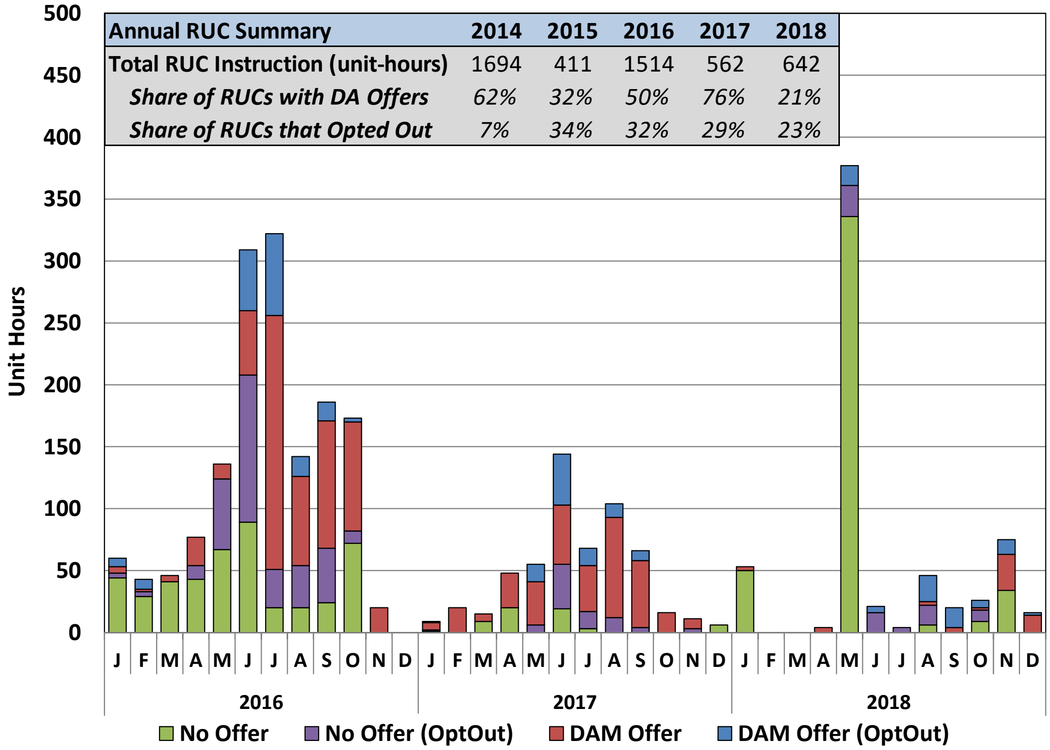
One of the important characteristics of any electricity market is the extent to which it results in the efficient commitment of generating resources. Under-commitment can cause apparent shortages in the real-time market and inefficiently high energy prices; while over-commitment can result in excessive start-up costs, uplift charges, and inefficiently low energy prices.

The ERCOT market does not include a mandatory centralized unit commitment process. The decision to start up or shut down a generator is made by the market participant. ERCOT's day-ahead market informs these decisions, but is only financially binding. That is, when a generator's offer to sell is selected (cleared) in the day-ahead market there is no obligation to actually start that unit. The generator will be financially responsible for providing the amount of capacity and energy cleared in the day-ahead market whether or not the unit operates.

ERCOT continually assesses the adequacy of market participants' resource commitment decisions using a reliability unit commitment (RUC) process, which executes both on a day-ahead and hour-ahead basis. Additional resources may be determined to be needed for two reasons – to satisfy the total forecasted demand, or to make a specific generator available to resolve a transmission constraint.

The next figure below shows RUC activity, by month, indicating the volume of generators receiving a RUC instruction that had offers in the day-ahead market or chose to opt-out of the RUC instruction.

Day-Ahead Market Activity of Generators Receiving a RUC



The number of RUC instructions in 2018 increased somewhat from 2017. The 642 unit-hours of RUC instructions in 2018 represent a 14% increase from the 562 unit-hours in 2017. These 2018 instructions were geographically diverse as well, with 16% to generators in the South zone in a variety of locations: San Antonio, Corpus Christi and the Rio Grande Valley, 2% were to generators in the Houston zone, 24% were to generators in the North zone, and the remaining 58% were to generators in the West zone.

As in 2017, most reliability commitments in 2018 were made primarily to manage transmission constraints (80% of unit-hours). Only 19% of RUC instructions were made to ensure sufficient system-wide capacity and 1% were for voltage support. Most of the RUC instructions in 2018 occurred during the month of May. Congestion in the Permian Basin area of west Texas was caused by higher than expected load and outages, both planned and forced, led to 377 unit-hours of RUC instructions during May.

In 2018, only 21% of the total RUC unit-hours had day-ahead offers. This is lower than expected, considering the incentive to provide day-ahead offers inherent in the RUC claw-back rules. The percentage of RUC unit-hours with day-ahead offers was much higher in 2017 and 2016 at 76% and 50%, respectively. The very low value in 2018 may be explained by the large number of fast starting generators receiving RUC instructions, primarily during May. Sixty percent of the unit-hour instructions in 2018 were for fast starting generators, whereas, since

2014 a more typical share has been 15%. It is not unusual for the decision to commit fast starting units to be made in real-time.

If real-time revenues received by a RUC unit exceed the operating costs incurred by the unit, then excess revenues are “clawed back” and returned to QSEs representing load. A generator receiving a RUC instruction has the choice to “opt out,” meaning it forgoes all RUC make-whole payments in return for not being subject to RUC clawback charges. The percentage of generators receiving RUC instructions in 2018 that chose to opt-out was 23%, similar to the 29% of generators that chose to opt-out in 2017.

During 2018, more than $3.1 million was clawed-back from RUC units while only $0.6 million in make-whole payments were made to RUC units. RUC make-whole payments in 2018 were collected almost exclusively from QSEs that were capacity short.

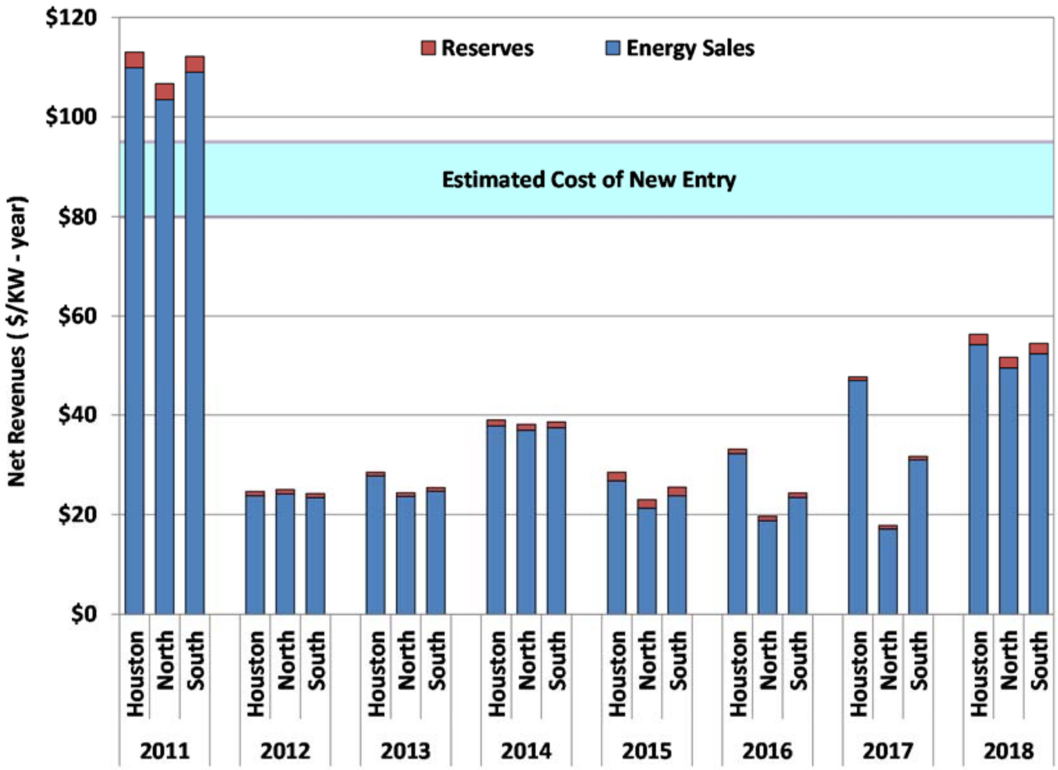
Resource Adequacy

One of the primary functions of the wholesale electricity market is to provide economic signals that will facilitate the investment needed to maintain an adequate set of resources to satisfy the system’s needs. These economic signals are best measured with the net revenue metric, which is calculated by determining the total revenue that could have been earned by a generating unit less its variable production costs. Put another way, it is the revenue in excess of short-run operating costs that is available to recover a unit’s fixed and capital costs, including a return on the investment.

In ERCOT’s energy-only market, the net revenues from the real-time energy and ancillary services markets alone provide the economic signals that inform suppliers’ decisions to invest in new generation or retire existing generation. To the extent that revenues are available through the day-ahead market or other forward bilateral contract markets, these revenues are ultimately derived from the expected real-time energy and ancillary service prices.

The next figure provides an historical perspective of the net revenues available to support investment in a new natural gas combustion turbine, selected to represent the marginal new supply that may enter when new resources are needed. The energy net revenues are computed based on the generation-weighted settlement point prices from the real-time energy market. Weighting the energy values in this way facilitates comparisons between geographic zones, but will mask what could be very high values for a specific generator location. Values for the West zone are excluded because historically lower energy prices make it a less attractive location to site natural gas generation. The figure also shows the estimated “cost of new entry,” which represents the revenues needed to break even on the investment.

Combustion Turbine Net Revenues



Based on estimates of investment costs for new units, the net revenue required to satisfy the annual fixed costs (including capital carrying costs) of a new combustion turbine unit ranges from $80 to $95 per kW-year. Although higher overall in 2018 than any year since 2011, the ERCOT market continued to provide net revenues well below the level needed to support new investment, ranging from below $52 per kW-year in the North Zone to more than $56 per kW-year in Houston.

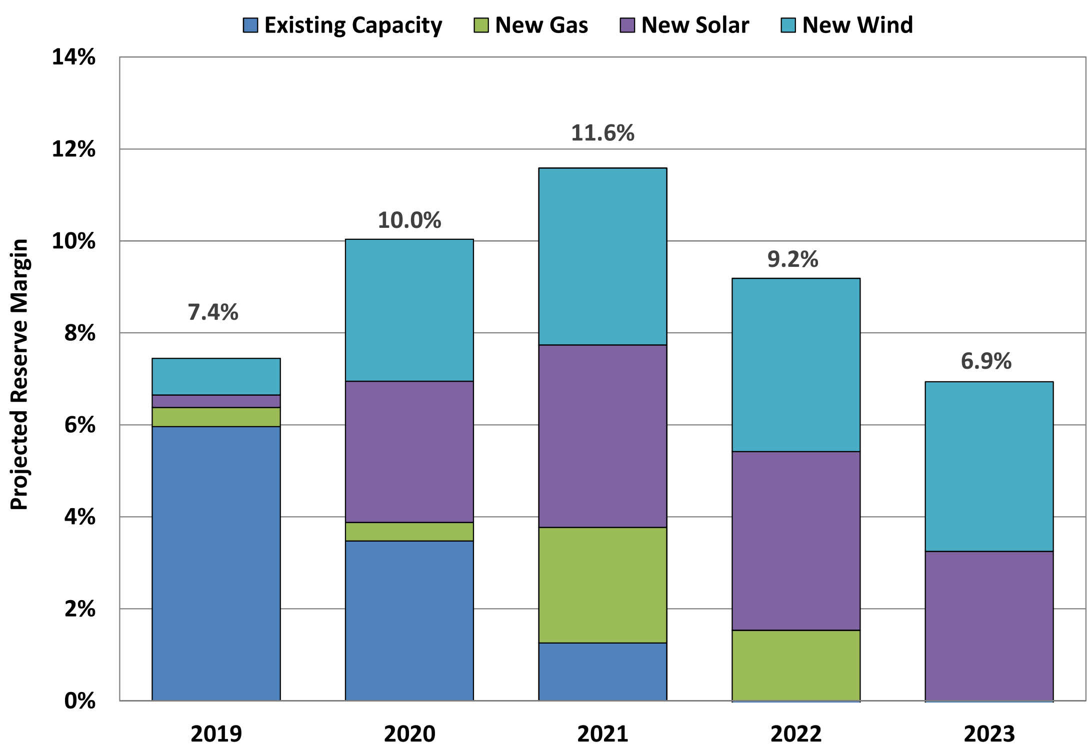
These results are consistent with a shrinking surplus of capacity, which contributed to more frequent shortages in 2018 compared to recent years. In an energy-only market, shortages play a key role in delivering the net revenues an investor needs to recover its investment. Such shortages will tend to be clustered in years with unusually high load or poor generator availability. The results in 2018 do not by themselves raise substantial concern regarding design or operation of ERCOT's Operating Reserve Demand Curve (ORDC) mechanism for pricing shortages. Given the recent generation retirements and continued load growth, 2018 was in fact a year with significantly more occurrences of shortage pricing, with that trend expected to continue in 2019.

The generation-weighted price of all coal and lignite units in ERCOT during 2018 was $33.31 per MWh, an increase from $26.32 per MWh in 2017. Although specific unit costs may vary, index prices for Powder River Basin coal delivered to ERCOT were approximately $2.65 per MMBtu in 2018; remaining at 2017 (and 2015) levels after decreasing to $2.51 per MMBtu in

2016. At these average prices coal units in ERCOT are likely receiving just enough revenue to cover operating costs. It follows then that the decision by Luminant and CPS Energy to retire several coal units is financially justified.

The figure below shows ERCOT's current projection of planning reserve margins and indicates that Texas is heading into the summer months of 2019 with a historically low reserve margin of 7.4%, just over half of ERCOT's previously stated reserve margin goal of 13.75%.[5]

Projected Planning Reserve Margins



Source: ERCOT Capacity, Demand and Reserves Report, December 2018 with Gibbons Creek capacity removed

These reserve margin projections are even lower than those developed in December 2018,[6] due in large part to higher expected seasonal peak loads, additional delays and cancellations of

[5] The target planning reserve margin of 13.75% was approved by the ERCOT Board of Directors in November 2010, based on a one in ten loss of load expectation (LOLE). The Commission directed ERCOT to evaluate planning reserve margins based on an assessment of the Economically Optimal Reserve Margin (EORM) and the Market Equilibrium Reserve Margin (MERM). *See* PUCT Project No. 42303, ERCOT Letter to Commissioners (Oct. 24, 2016). On December 12, 2017, ERCOT published its "Study Process and Methodology Manual: Estimating Economically Optimum and Market Equilibrium Reserve Margins" as part of its ongoing reporting initiative.

[6] See Report on the Capacity, Demand and Reserves in the ERCOT Region (December 11, 2018); http://www.ercot.com/content/wcm/lists/167023/CapacityDemandandReservesReport-Dec2018.pdf; the 2019 summer reserve margin was projected to be 8.1%, a reduction of 2.9 percentage points from the May 2018 CDR report.

planned projects, and the indefinite mothballing of 470 MW at the Gibbons Creek coal unit.[7] The reserve margin is expected to continue to be below the existing target level of 13.75% for the foreseeable future.

Installed reserve margins for summer of 2018 were also historically low. What seem like very low reserves may just be the new normal. Given the overall size of the system and projected growth, a more robust reserve margin may no longer be required to cover load forecast errors and mitigate generator availability risks. Further, with smaller, more distributed generation technologies playing an increasingly important role in ERCOT, the risk associated with generator outages should decrease.

Because the surplus has now disappeared and shortages are likely to be even more frequent in 2019, the economic signals could change rapidly. These short-term market outcomes and price signals, as well as investors' response to these economic signals, will be monitored. This response could cause the planning reserve margins to exceed the forecast shown in the figure.

Analysis of Competitive Performance

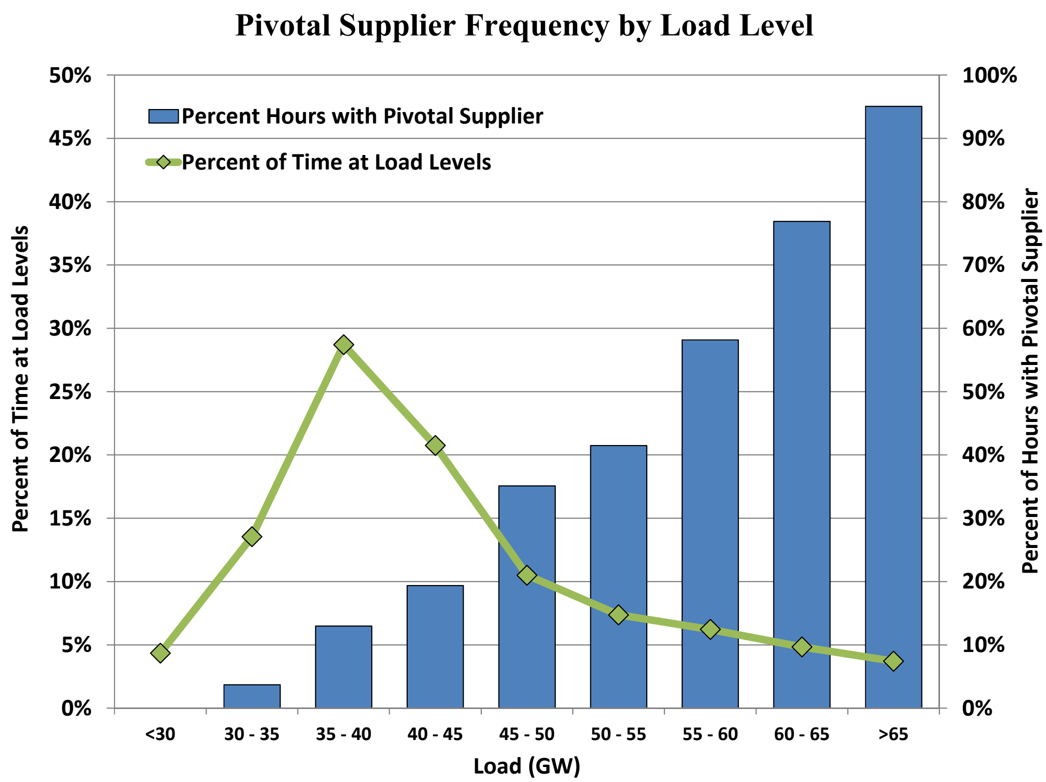
Market power is evaluated from two perspectives, structural (does market power exist) and behavioral (have attempts been made to exercise it).

Structural Market Power

The market structure is analyzed by using the Residual Demand Index (RDI), a statistic that measures the percentage of load that could not be served without the resources of the largest supplier. It assumes the market could call upon all committed and quick-start capacity owned by other suppliers. When the RDI is greater than zero, the largest supplier is pivotal (i.e., its resources are needed to satisfy the market demand). When the RDI is less than zero, no single supplier is pivotal. The RDI is a useful structural indicator of potential market power, although it is important to recognize its limitations. As a structural indicator, it does not indicate whether a supplier may have actually exercised market power, or whether it would have been profitable for a pivotal supplier to exercise market power. Nonetheless, it does identify conditions under which a supplier could raise prices significantly by withholding resources.

The figure below summarizes the RDI analysis by showing the percentage of time at each load level there was a pivotal supplier. The figure also displays the percentage of time each load level occurs.

[7] On December 21, 2018, ERCOT received an NSO from the City of Garland for Gibbons Creek (GIBCRK_GIB_CRG1) indicating that the Resource will be mothballed indefinitely effective June 1, 2019. Gibbons Creek is a 470 MW coal unit located in Grimes County (20 miles southeast of College Station) and owned by the Texas Municipal Power Agency (TMPA), which is an organization jointly owned by four municipalities – the cities of Garland, Denton, Bryan and Greenville.



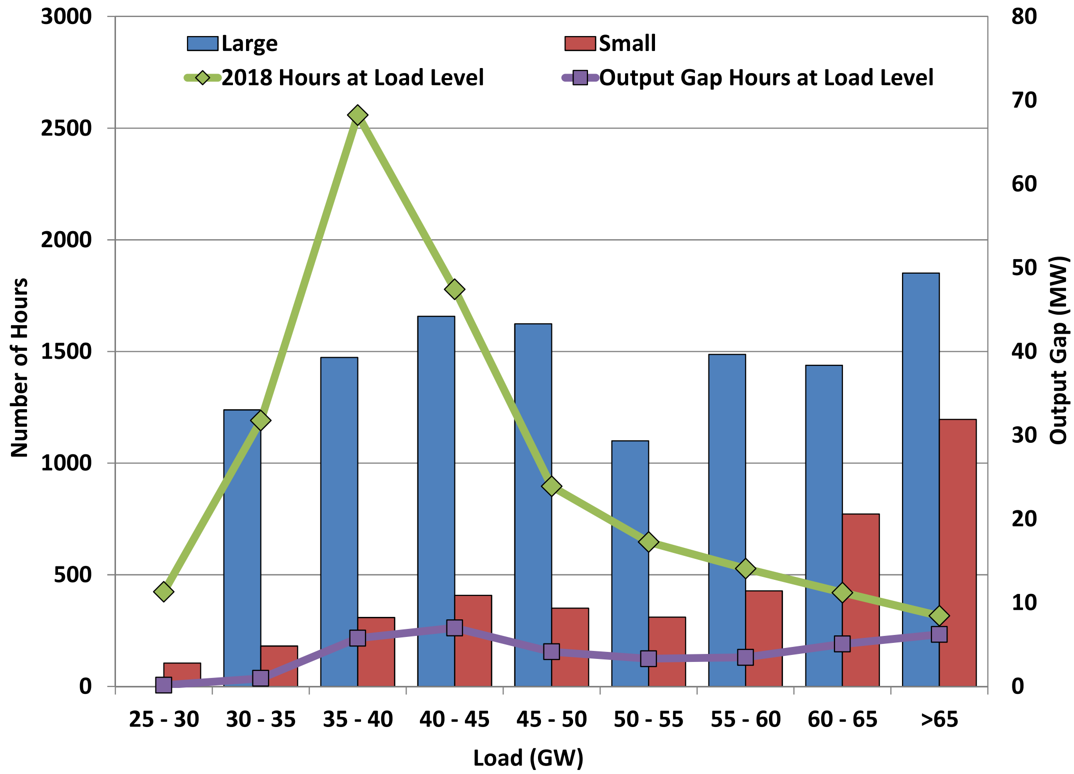
At loads greater than 65 GW there was a pivotal supplier 95% of the time. This is expected because at high load levels, the largest suppliers are more likely to be pivotal as other suppliers' resources are more fully utilized serving the load. Pivotal suppliers existed 30% of all hours in 2018, which was more frequent than in 2017 when pivotal suppliers existed in 25% of all hours. The increase was due to the greater number of high load (>65GW) hours in 2018. Over the past several years, as generation supply ownership in ERCOT has become less concentrated, the fraction of time with a pivotal supplier has decreased from 75-80% of the time down to 25-30% seen in recent years. Even with this reduction, market power continues to be a potential concern in ERCOT, requiring effective mitigation measures to address it.

This analysis evaluates the structure of the entire ERCOT market. In general, local market power in narrower areas that can become isolated by transmission constraints raise more substantial competitive concerns. This local market power is addressed through: (a) structural tests that determine "non-competitive" constraints that can create local market power; and (b) the application of limits on offer prices in these areas.

Evaluation of Conduct

In addition to the structural market power analyses above, actual participant conduct was evaluated to assess whether market participants have attempted to exercise market power through physical or economic withholding. An "output gap" metric is used to measure potential economic withholding, which occurs when a supplier raises its offer prices to reduce its output.

Incremental Output Gap by Load Level and Participant Size – Step 2



The output gap is the quantity of energy that is not being produced by online resources even though the output is economic to produce by a substantial margin given the real-time energy price. A participant can economically withhold resources, as measured by the output gap, by raising its energy offers so as not to be dispatched. A resource is evaluated for inclusion in the output gap when it is committed and producing at less than full output. Energy not produced from a committed resource is included in the output gap if the real-time energy price exceeds that unit's mitigated offer cap by at least $30 per MWh. The mitigated offer cap serves as a proxy for the marginal production cost of energy from that resource.

The figure above shows very small quantities of capacity that would be considered part of this output gap, with only 15% of the hours in 2018 that exhibited an output gap.

These results show that potential economic withholding levels were extremely low in 2018. Output gaps of the largest suppliers are routinely monitored individually and were found to be consistently low across all load levels. These results, together with our evaluation of the market outcomes presented in this report, allow us to conclude that the ERCOT market performed competitively in 2018.

Recommendations

One of the stated objectives for the IMM is to recommend measures to enhance market efficiency. Hence, although we have found the ERCOT markets to have performed well in 2018, we have identified and recommend a number of potential improvements.

One of our longest held recommendations was addressed in early 2019, when the Commission instructed ERCOT to proceed with the implementation of real-time co-optimization. We are very supportive of the Commission's direction and believe real-time co-optimization will bring great benefit to the ERCOT market and the customers it serves. Real-time co-optimization is a major market change which will take much effort and multiple years to accomplish. In the meantime, there are specific aspects of the current market structure which we recommend be addressed.

Reliability Unit Commitment

We offer three recommendations which will improve the reliability commitment process and resulting pricing.

1. Evaluate and improve the Reliability Deployment Price Adder

The current calculation method for the Real-Time On-Line Reliability Deployment Price Adder (RTRDPA) is producing results that are inconsistent with its original intent and should be changed. The existing RTRDPA mechanism is flawed and the IMM believes these flaws should be addressed before expanding the number and types of actions that trigger the RTRDPA.

When the RTRDPA is triggered by reliability unit commitment (RUC) instructions, the intent of the RTRDPA should be to produce prices sufficient to have encouraged a competitive resource to commit and produce energy. We do not support the position held by some that the RTRDPA should produce prices that reflect the total absence of the resource. The current RTRDPA misses achieving this goal in two main ways; first, most RUCs address local reliability concerns and have significant local impact with relatively minor global impacts. The IMM recommends that nodal adders should be reexamined in ERCOT because the local impacts are more important. Second, although the mitigation of resources with RUC instructions is appropriate, the mitigation should in some way reflect the startup and minimum load costs of that Resource.

Further, the IMM believes that reserves should not be paid the RTRDPA. While it makes sense to pay the Operating Reserve Demand Curve (ORDC) adder to reserves due to the added reliability that they afford, the RTRDPA is intended to be a "but for" price adjustment. That is, the RTRDPA is an attempt to pay Resources the price that a competitive market would have produced "but for" the reliability actions taken by the system operator. This "but for" price would have been paid to Resources generating energy in the competitive "but for" scenario, not to reserves.

Finally, the IMM believes that any proposed changes to the RTRDPA calculation should be thoroughly analyzed with past market data and scenarios prior to adoption and deployment. This is a complicated calculation and the risk of unintended consequences is high without significant analysis.

Status: This is a new recommendation.

The flaws in the current calculation method for the RTRDPA are currently being discussed as part of stakeholder deliberations regarding NPRR904, *Revisions to Real-Time On-Line Reliability Deployment Price Adder for ERCOT-Directed Actions Related to DC Ties and to Correct Design Flaws*, introduced in October 2018. This NPRR would revise the categories of ERCOT-directed actions that trigger the RTRDPA to include DC Tie related actions. At a minimum, we recommend improving the methodology for determining the adder prior to expanding the situations in which it would apply.

2. Explore options to consider commitment costs for RUC-committed units

The IMM recommends evaluating RUC mitigation and exploring options to consider commitment costs for RUC-committed units. Mitigation is used to ensure competitive outcomes in noncompetitive situations. The cost to commit a unit should be factored in to the offers and market price that results from ERCOT making a commitment decision. Consideration of commitment costs will make it more likely that prices are at a level that would support the market participant to have committed the unit on their own. Conversely, offers that reflect only marginal energy costs will likely produce prices too low to support self-commitment.

Status: This is a new recommendation.

3. Eliminate the OPTOUT option for RUC-committed Resources

For generators unsure about whether to self-commit, the ability to OPTOUT of RUC instructions provides the incentive for generators to defer the decision to self-commit as long as possible with no risk. The implementation of NPRR744, *RUC Trigger for the Reliability Deployment Price Adder and Alignment with RUC Settlement* served to exacerbate this disincentive to self-commit by allowing generators even more time to OPTOUT once a RUC instruction is received.

ERCOT operators typically have exhibited great restraint, deferring their decisions to issue a RUC instruction to last possible moment. This restraint was recently codified in NPRR864, *RUC Modifications to Consider Market-Based Solutions*, which modified the RUC engine to consider fast-start generators (less than 1 hour start times) as self-committed for future hours. The culmination of these RUC changes has enabled ERCOT to defer supplementary commitment decisions, allowing market participants full opportunity to make their own unit commitment.

The current OPTOUT option provides the incentive for units to not commit by giving those units access to more market information than other Market Participants via the RUC instruction. Because of the disincentive to self-commit, the IMM recommends eliminating the OPTOUT option for RUC-committed Resources.

Status: This is a new recommendation.

Revenue Neutrality Allocation

Revenue neutrality allocation (RENA) causes uplift to the market, which is difficult to hedge against, creating uncertainty for load as to what prices will be. The next two recommendations are intended to reduce RENA.

4. Review PTP obligations linked to options

Point-to-Point (PTP) Obligations with Links to an Option (PTPLO) are cleared as obligations but settled as options. This variance is what often leads to uplift, and its effects are more pronounced now than ever before. The settling of PTPLOs as options dates back to 2011 and adoption of NPRR322, *Real-Time PTP Option Modeling*. This NPRR provided that all PTP Options would be settled in the day-ahead market. If a Non-Opt-In Entity (NOIE) purchases a PTP Obligation in the day-ahead market and owns a PTP Option for the same source and sink pair, and the MWs of the PTP Obligation are not more than the MWs of the PTP Option, then the PTP Obligation would settle in Real-Time as only the positive difference between the sink and source Settlement Point Prices, much like a PTP Option. This NPRR also included a requirement that NOIEs shape the PTP Options going to Real-Time according to their Load forecast and added a new type of Congestion Revenue Right (CRR) to handle the treatment of PTP Obligations that are treated as PTP Options for Settlement purposes.

Since 2017, we have seen significant increases in uplift related to PTPLOs, and this trend continued in 2018. The IMM recommends a review into the effects of NPRR322 and settling PTPLOs as options on RENA

Status: This is a new recommendation.

5. Evaluate and Improve Load Distribution Factors (LDFs) used in the Congestion Revenue Right (CRR) and Day-Ahead Market clearing activities

Load Distribution Factors (LDFs) are developed using historical State Estimator or SCADA data and deal with load levels within a load zone. Per Protocol Section 4.5.1, ERCOT shall generate and maintain the appropriate LDF libraries for use in the day-ahead market and CRR Auctions.

ERCOT updates the LDF libraries by maintaining the existing LDF sets and generating new LDF sets when required, based on significant changes in system-wide load patterns. Updates are regularly required for seasonal load patterns due to weather changes.

Our concern with the current LDF procedure is that it is entirely backwards looking (i.e. using the most recent hot day profile). There are challenges transitioning that process to be more forward looking, such as being able to accurately predict the month that a load is going to come online, but the IMM views those challenges as worth addressing, especially since the CRR balancing account was emptied for the first time in 2018.

This year, due mostly to differences in the West Load Zone day-ahead and CRR load distribution factors, the CRR balancing fund was drawn on heavily in June and depleted entirely in July. While the balancing fund was once again up to its capped value of $10 million by August, another shortfall lowered it in November and at the end of the year it sat at a little less than $8 million.

Status: This is a new recommendation.

Real-time Co-optimization

The benefits expected from real-time co-optimization are substantial. First, jointly optimizing all products in each interval allows ancillary service responsibilities to be continually adjusted in response to changing market conditions. In addition to lowering the cost of satisfying both energy and reserve requirements, the likelihood of having the full capacity of procured reserves is expected to increase, at the same time ensuring that energy is produced at locations where it may be most valuable. Second, real-time co-optimization will improve shortage pricing. The ORDC provides a mechanism for setting real-time energy prices that reflect the expected value of lost load. However, jointly-optimizing the energy and reserve markets would allow this shortage pricing to be more accurate. Other economic benefits will be achieved by allowing all suppliers to participate fully in ERCOT's ancillary service markets. Currently, QSEs without large resource portfolios are effectively precluded from participating in ancillary service markets because of the replacement risk they face having to rely on a supplemental ancillary services market (SASM).

A key part of real-time co-optimization design will be to determine demand curves for each ancillary service. Since all demand curves – energy, ancillary services and transmission – should coordinate in a rational manner, the following recommendation may be considered as part of the efforts underway to design real-time co-optimization for ERCOT. Regardless, we believe the following recommendation is valuable on its own merits.

6. Evaluate Transmission Demand Curves

As the demands curves for each type of reserve service (potentially including locational reserve products in the future) are being evaluated under a co-optimized system, it is a good opportunity to also evaluate transmission penalty curves. Currently there are single values applied to limit the amount spent to resolve a transmission constraint. These Shadow Price Caps vary by voltage level. Much like the concept of demand curves for ancillary services, the value of lowering the flow of electricity on transmission lines increases as the violation amount grows, this should be reflected in energy prices, as it is in some other electricity markets (notably MISO). Given that congestion costs were over $1B in 2018, the IMM recommends that a more nuanced approach to how transmission security affects pricing be evaluated.

Status: This is a new recommendation.

The six remaining recommendations have been described in prior reports

7. Evaluate the need for a local reserve product.

In an energy-only market, all economic signals to support long-term investment and retirement decisions are provided by the energy and ancillary service markets. A substantial component of these economic signals is the prices and revenues generated in shortage conditions. ERCOT's ORDC establishes shortage pricing ERCOT-wide but does not allow for shortage pricing in local areas. Therefore, ERCOT's current market design may support adequate resources in aggregate but may not support adequate resource in some local areas.

In ERCOT's energy-only market, the primary means to ensure that sufficient revenues are provided to satisfy both the market-wide and local resource adequacy needs is to strive for alignment between ERCOT's operating requirements and its planning requirements. In other words, if having sufficient resources to respond to the two largest contingencies is a reasonable planning requirement, it is also likely a reasonable operating requirement. The advantage of defining such an ancillary service product in ERCOT is that it would allow the real-time energy and reserve markets to price local reserve shortages and provide the revenues necessary to satisfy local capacity needs. In doing so, it should eliminate the need to sign out-of-market reliability must-run (RMR) contracts.

Hence, we recommend that ERCOT align its planning requirements and real-time operating requirements and begin evaluating the need for a local reserve product. Changes to the process for determining whether an RMR unit is needed, implemented in NPRR788, were important clarifications. However, if there is a local reliability concern that is best addressed by maintaining additional operating reserves in a specific area, we suggest that ERCOT develop and implement a new local reserve product.

Status: As part of our previous recommendations in Project No. 47199, we offered an approach for implementing a local reserve product that would be constraint-based, incorporating nodal elements, and use non-spinning resources to address the constraint.[8] This proposal requires real-time co-optimization as part of its application, so as real-time co-optimization is implemented, we are prepared to work with ERCOT and market participants to evaluate this proposal or others to address this recommendation.

8. Price future ancillary services based on the shadow price of procuring the service.
In a well-functioning real-time market, the market model will indicate the marginal cost of satisfying any requirement, which is the shadow price of the requirement. This shadow price is the most efficient clearing price for each of ERCOT's ancillary service requirements. Such prices create efficient incentives for participants to offer and provide ancillary services. Hence, we continue to recommend that any new or updated ancillary services be priced on this basis.

Status: Even though the comprehensive redesign of the set of ancillary services was not approved and the development of ancillary services demand curves is currently underway, multiple incremental modifications have been and are being considered. NPRR815, *Revise the Limitation of Load Resources Providing Responsive Reserve (RRS) Service*, which was approved in December 2017 and implemented in mid-2018 1) increased the allowable percentage of responsive reserve service that load resources may provide from 50% to 60%, and 2) specified the minimum amount of primary frequency response (generator provided) as 1150 MW. NPRR863, *Creation of ERCOT Contingency Reserve Service and Revisions to Responsive Reserve*, introduced on January 1, 2018 and ultimately approved on February 12, 2019, modifies the definition Responsive Reserve service (RRS) to include a new frequency response type service (i.e., Fast Frequency Response (FFR)), and added a new Ancillary Service type (i.e., ERCOT Contingency Reserve Service (ECRS)).

These changes may reduce the likelihood that ancillary service provision will be constrained and the IMM will continue to monitor their effects.

9. Price congestion at all locations that affect a transmission constraint.
Since the start of the nodal market, generators greater than 10 MW were considered part of the wholesale market with associated obligations and privileges. Generators less than 10MW and connected to the transmission system are not subject to many of the obligations borne by larger generators. Further, these small facilities are settled at the Load Zone price, not a location-specific nodal price.

This practice may have been adequate for the few number of small generators that existed at the time of nodal market implementation. Currently however, the output of some small generators can significantly affect transmission congestion. When they can relieve a constraint, they would

[8] PUCT Project No. 47199, Comments of Potomac Economics at 2, 8-10 (Sept. 15, 2017).

be paid a much higher price than they are currently. When they aggravate a constraint, they would generally settle at a lower price. Hence, settling with this generator as a zonal prices fails to provide efficient incentive for it to operate in a manner consistent with the reliability needs of the system.

All generators with output that affects a transmission constraint should receive a locational price. Small generators may not have to bear all the obligations of large generation resources, but they should settle in a manner consistent with the effect they have on the system.

Further, ERCOT should price constraints, whether or not a resource currently exists that can effectively solve the constraint. Currently if ERCOT determines that there is no resource that has a significant relieving effect on a constraint that constraint is not considered by the dispatch software and thus does not affect prices.

By pricing all constraints ERCOT would provide a more transparent signal as to where generation should be sited, particularly for smaller, lower cost generation which is less likely to dampen out the congestion pricing. Also, by pricing these constraints the clearing of the day-ahead market and the real-time market would become more similar, as the day-ahead market prices constraints whether there is effective relief or not.

Status: The recommendation to price all constraints in real-time is new.

The recommendation to have all generation receive a nodal price was introduced in 2017. On January 2, 2019, ERCOT introduced NPRR917: *Nodal Pricing for Settlement Only Distribution Generators (SODGs) and Settlement Only Transmission Generators (SOTGs)* that would implement nodal energy pricing for small generators.

10. Modify the real-time market software to better commit load and generation resources that can be online within 30 minutes.

The real-time market efficiently dispatches online resources and sets nodal prices that reflect the marginal value of energy at every location, but ERCOT lacks real-time processes to facilitate efficient commitment and de-commitment of resources that can start quickly (i.e., within 30 minutes). This is a concern because suboptimal dispatch of these resources raises the overall costs of satisfying the system's needs, distorts the real-time energy prices, and affects reliability. For these reasons, other markets have implemented a look-ahead process to optimize short-term commitments of peaking resources. In contrast, ERCOT relies on de-centralized commitment where individual participants bear most of the costs of their own commitment decisions. Because participants lack the information ERCOT has on upcoming conditions and the plans of other participants, this decentralized process will necessarily be less efficient than a fully-optimized real-time process coordinated by ERCOT. Further, as ERCOT attracts more variable wind and solar resources, the value of having access to and optimally utilizing fast-starting controllable resources will grow. Hence, we continue to recommend that ERCOT develop this capability.

<u>Status</u>: We have been recommending this change since the start of ERCOT's nodal market. After taking interim steps to produce non-binding generation dispatch and price projections and then to improve the short term forecasting procedures, ERCOT evaluated the potential benefits of a multi-interval real-time market. This evaluation determined that, because the costs to implement were greater than the projected benefits, moving forward with implementation was not supported at the time.[9] The 2017 finding of insufficient benefits is not surprising given the low-price environment and the level of surplus capacity at the time of the evaluation.

However, with nearly 5 GW of fast-starting generation installed in ERCOT, ever increasing quantities of intermittent renewable resources, and the current lack of surplus capacity, the benefits of improving the short-term commitment process will grow. There is likely a much less costly option available to develop a process to optimize the commitment of fast-starting resources without implementing a full, multi-interval real-time market. Hence, we continue to recommend modifying the real-time market software to better commit load and 30-minute generators.

11. Evaluate policies and programs that create incentives for loads to reduce consumption for reasons unrelated to real-time energy prices, including: (a) the Emergency Response Service (ERS) program and (b) the allocation of transmission costs.

Any incentives that cause market participants to take actions that are inconsistent with the real-time prices will undermine the performance of the market and its prices. These concerns are heightened when these actions are taken under peak or emergency conditions because the ERCOT market relies on efficient pricing under such conditions to motivate efficient long-term resource decisions by participants. By curtailing load in response to incentives or programs that are not aligned with the real-time energy market, supply is uneconomically reduced and the real-time market is adversely affected. The following two aspects of the ERCOT market raise these concerns.

ERS Program. A load that wishes to actively participate in the ERCOT market can participate in ERS, provide ancillary services, or simply choose to curtail in response to high prices. Participating in ERS greatly limits a load's ability to provide ancillary services or curtail in response to high prices. Given the high budget allotted and the low risk of deployment, ERS is an attractive program for loads. Because the ERS program is so remunerative, we are concerned that it is limiting the motivation for loads to actively participate and contribute to price formation in the real-time energy market.

---

9 *See* PUCT Project No. 41837, *PUCT Review of Real-Time Co-Optimization in the ERCOT Region*, ERCOT Report on the Multi-Interval Real-Time Market Feasibility Study (Apr. 6, 2017).

Transmission Cost Allocation. Transmission costs in ERCOT are allocated on the basis of load contribution in the highest 15-minute system demand during each of the four months from June through September. This allocation mechanism is routinely referred to as four coincident peak, or 4CP. Transmission costs have doubled since 2012, significantly increasing an already substantial incentive to reduce load during probable peak intervals in the summer. ERCOT estimates that as much as 1500 MW of load were actively pursuing reduction during the 4CP intervals from 2016 to 2018.[10]

Load curtailment to avoid transmission charges may be resulting in price distortion during high load periods because the response is targeting peak demand rather than responding to wholesale prices. Even with higher prices in 2018, reductions were observed during June, July, and August at times with wholesale prices less than $40 per MWh.

Status: The Commission made no changes to the ERS program or transmission service rates in 2018.

12. Reserve the inclusion of marginal losses in ERCOT locational marginal prices for post-implementation of real-time co-optimization in ERCOT.

When electricity is produced in one location and consumed at another location, the electricity flows through the transmission system and some of it is lost. The transmission losses vary depending on the distance the electricity is traveling and the voltage of the lines it must flow over. Ideally, the real-time dispatch model should recognize the marginal losses that will result from dispatching units in different locations and set prices accordingly. Recognizing marginal losses will allow the real-time market to produce more from a higher-cost generator located electrically closer to the load, thus resulting in fewer losses. Optimizing this trade-off in the real-time dispatch lowers the overall costs of satisfying the system's needs.

The ERCOT market is unique in its treatment of transmission losses. Marginal losses are not included in ERCOT real-time energy prices and the costs of losses are collected from loads on an average basis. This approach may have been reasonable at the time ERCOT was implementing its initial real-time energy markets because generators were located relatively close to load centers. However, as open access transmission expansion policies and other factors have led to a wider dispersion of the generation fleet across the ERCOT footprint, the failure to recognize marginal losses in the real-time dispatch and pricing has led to larger dispatch inefficiencies and price distortions. Therefore, we are now recommending that the ERCOT real-time market be upgraded to recognize marginal losses in its dispatch and prices.

10 *See* ERCOT, 2018 Annual Report of Demand Response in the ERCOT Region (Mar. 2019) at 7, available at http://www.ercot.com/services/programs/load.

Accompanying this change, a revenue allocation methodology will need to be developed because marginal loss pricing results in the collection of more payments for losses than the aggregate cost of losses. This occurs because the marginal losses are always larger than the average losses (i.e., losses increase as more power flows over the transmission system). Most other RTOs in the U.S. recognize marginal losses and may provide examples of allocation approaches that could be used in ERCOT.

Status: In early 2019, the Commission declined to direct ERCOT to implement marginal losses. Although assigning marginal transmission losses was recognized as common in other markets and an efficient way to account for losses, the Commission does not believe implementation was cost effective at this time. The IMM agrees that assigning marginal transmission losses is an efficient way to account for losses. The market synergies of including marginal losses in system dispatch decisions after real-time co-optimization goes live in ERCOT include both production cost savings and reductions in consumer costs as well as increased reliability and more accurate and efficient price formation mechanisms. The IMM remains an advocate for the implementation of marginal losses as a best practice in energy market design.

I. Review of Real-Time Market Outcomes

Although only a small share of the power produced in ERCOT is transacted in the real-time market, real-time energy prices set the expectations for prices in the day-ahead market and bilateral forward markets where most transactions occur. Unless there are barriers preventing arbitrage of the prices between the real-time and forward markets, prices in the forward markets should be directly related to the prices in the real-time market (i.e., real-time prices and forward prices should converge over the long-run). Hence, low prices in the real-time energy market will translate to low forward prices. Likewise, price spikes in the real-time energy market will increase prices in the forward markets. Because of the interaction between real-time and all forward prices, the importance of real-time prices to overall market performance is much greater than might be inferred from the proportion of energy actually paying real-time prices. This section evaluates and summarizes electricity prices in the real-time market during 2018.

A. Real-Time Market Prices

The first analysis evaluates the total cost of supplying energy to serve load in the ERCOT wholesale market. In addition to the costs of energy, loads incur costs associated with ancillary services and a variety of non-market based expenses referred to as "uplift." An average "all-in" price of electricity has been calculated for ERCOT that is intended to reflect wholesale energy costs, as well as these additional costs.

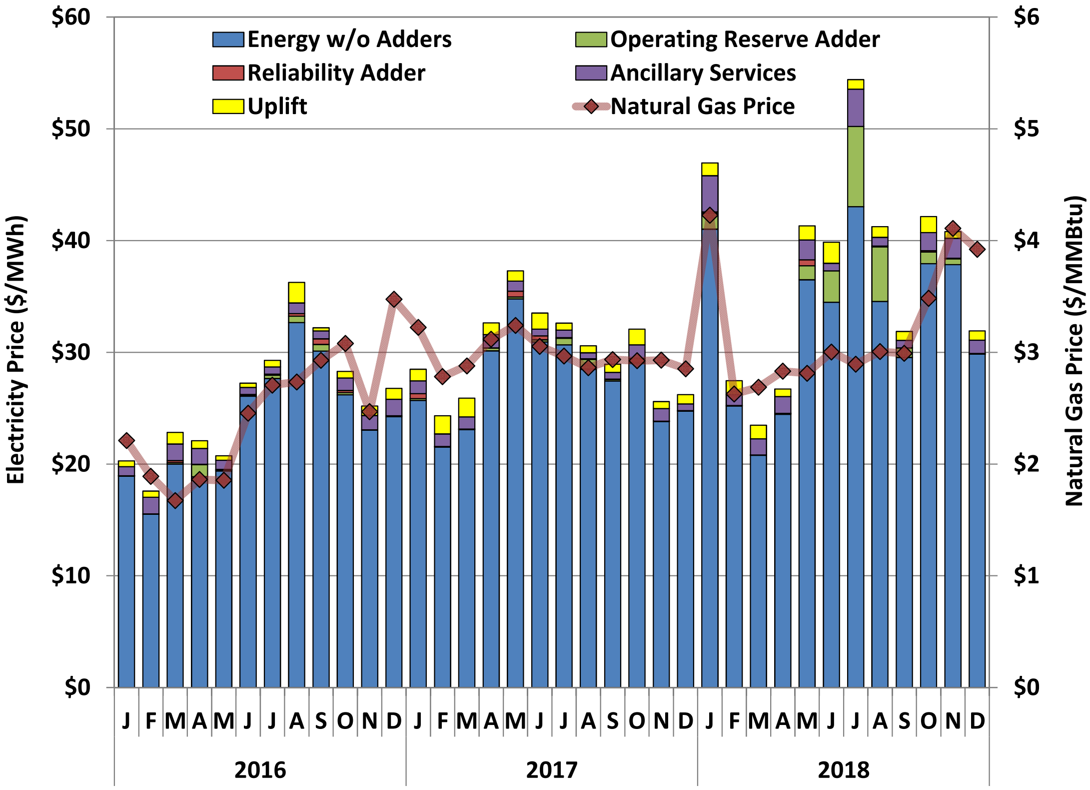
Figure 1 summarizes changes in energy prices and other market costs by showing the all-in price of electricity, which is a measure of the total cost of serving load in ERCOT for 2016 through 2018. The ERCOT-wide price in this figure is the load-weighted average of the real-time market prices from all zones. Ancillary services costs and uplift costs are divided by real-time load and shown on a per MWh basis.[11]

ERCOT real-time prices include the effects of two energy price adders that are designed to improve real-time energy pricing when conditions warrant or when ERCOT takes out-of-market actions for reliability. Although published energy prices include the effects of both adders, the Operating Reserve Demand Curve Adder (operating reserve adder) and the Reliability Deployment Price Adder (reliability adder) are shown separately here from the energy price. The operating reserve adder was implemented in mid-2014 to account for the value of reserves based on the probability of reserves falling below the minimum contingency level and the value of lost load. The reliability adder was implemented in June 2015 as a mechanism to ensure that reliability deployments do not distort the energy prices. The reliability adder is calculated using

11 For this analysis Uplift includes: Reliability Unit Commitment Settlement, Operating Reserve Demand Curve (ORDC) Settlement, Revenue Neutrality Total, Emergency Energy Charges, Base Point Deviation Payments, Emergency Response Service (ERS) Settlement, Black Start Service Settlement, Block Load Transfer Settlement, and the ERCOT System Administrative Fee.

a separate pricing run of the dispatch software, removing any reliability unit commitments (RUC) or deployed load capacity and recalculating prices. When the recalculated system lambda (average load price) is higher than the initial system lambda, the increment is the adder.

Figure 1: Average All-in Price for Electricity in ERCOT



The largest component of the all-in price is the energy cost. The figure above indicates that natural gas prices continued to be a primary driver of energy prices. This correlation is expected in a well-functioning, competitive market because fuel costs represent the majority of most suppliers' marginal production costs. Because suppliers in a competitive market have an incentive to offer supply at marginal costs and natural gas is the most widely-used fuel in ERCOT, changes in natural gas prices should translate to comparable changes in offer prices. The average natural gas price in 2018 was $3.22 per MMBtu, up approximately 8% from the 2017 average price of $2.98 per MMBtu. ERCOT average real-time energy prices increased over 26% from $28.25 per MWh in 2017 to $35.63 per MWh in 2018. As described later, increased shortage conditions in 2018 compounded the effects of higher natural gas prices on higher real-time energy prices.

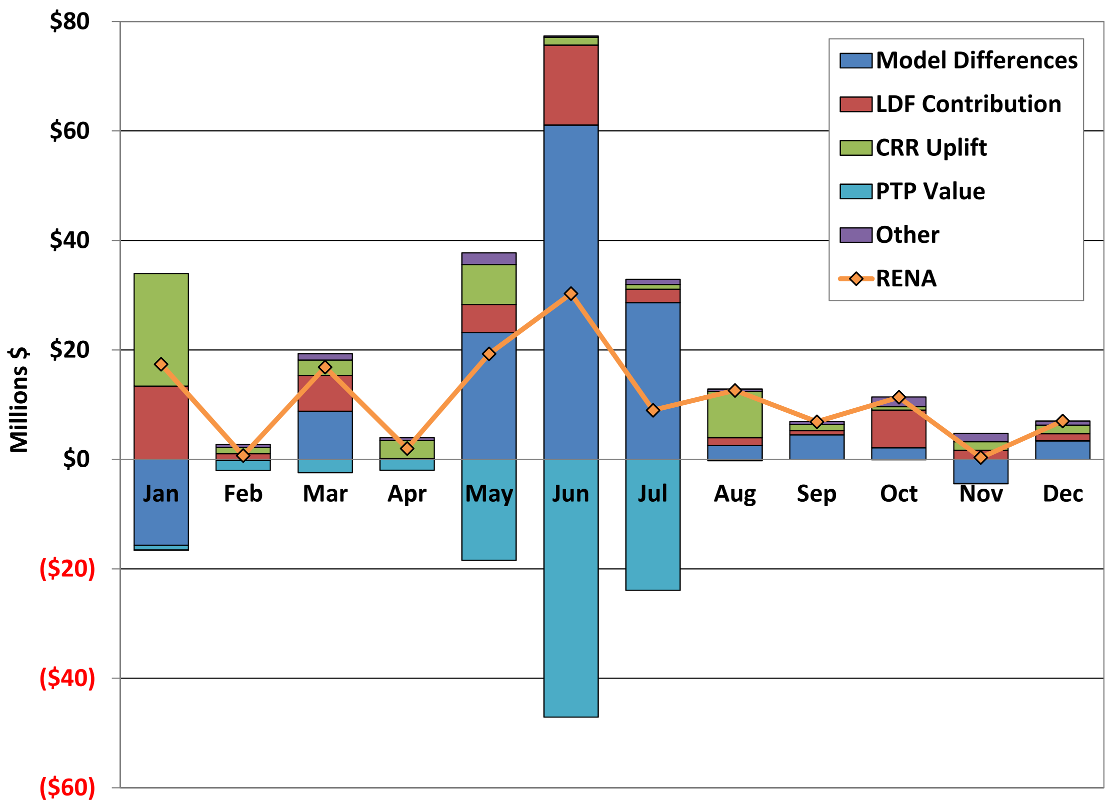
The average real-time energy price in 2018 included relatively small contributions from ERCOT's energy price adders: $1.97 per MWh from the operating reserve adder and $0.08 per MWh from the reliability adder. These values were a departure, more fully explained later, from the levels seen in 2017; $0.21 and $0.16 per MWh, for reserve and reliability adder, respectively.

The highest monthly average operating reserve adder for 2018 was $7.19 per MWh in July, while the highest monthly average reliability adder was $0.54 per MWh in May.

Other cost categories continue to be a relatively small portion of the all-in electricity price. Ancillary services costs were $1.60 per MWh in 2018, up from $0.87 per MWh in 2017 due to several reasons described in Section II: Day-Ahead Market Performance.

Uplift costs accounted for $1.08 per MWh of the all-in electricity price in 2018, up from $1.02 per MWh in 2017. In the context of providing the total cost of serving load in ERCOT, these values include both the ERCOT system administrative fee and the program costs for Emergency Response Service (ERS), which are assessed to all loads. The total amount of uplifted costs in 2018 was approximately $408 million, up from $365 million in 2017. There are many costs included as uplift, but the largest components are the ERCOT system administrative fee ($209 million or $0.56 per MWh), ERS program costs ($50 million or $0.13 per MWh) and the revenue neutrality allocation (RENA), which totaled $134 million or $0.35 per MWh in 2018.

Figure 2: ERCOT RENA Analysis



Most of the increase in uplift costs in 2018 was due to the increase in RENA, which increased almost 40 %, from $96 million ($0.27 per MWh) in 2017 to $134 million ($0.35 per MWh) in 2018. Figure 2 above provides an analysis of uplift and RENA in 2018.

Several factors can contribute to RENA uplift, including 1) transmission network modelling inconsistencies between day-ahead and real-time market during market clearing; 2) differences in load distribution factors used in day-ahead and real-time; 3) settlement of day-ahead Point-to-Point (PTP) obligations linked to options; 4) manual corrections that occur when the clearing price of PTP obligations in the day-ahead market is higher than the submitted bid price; and 5) setting a price floor in the real-time market at -$251 per MWh.

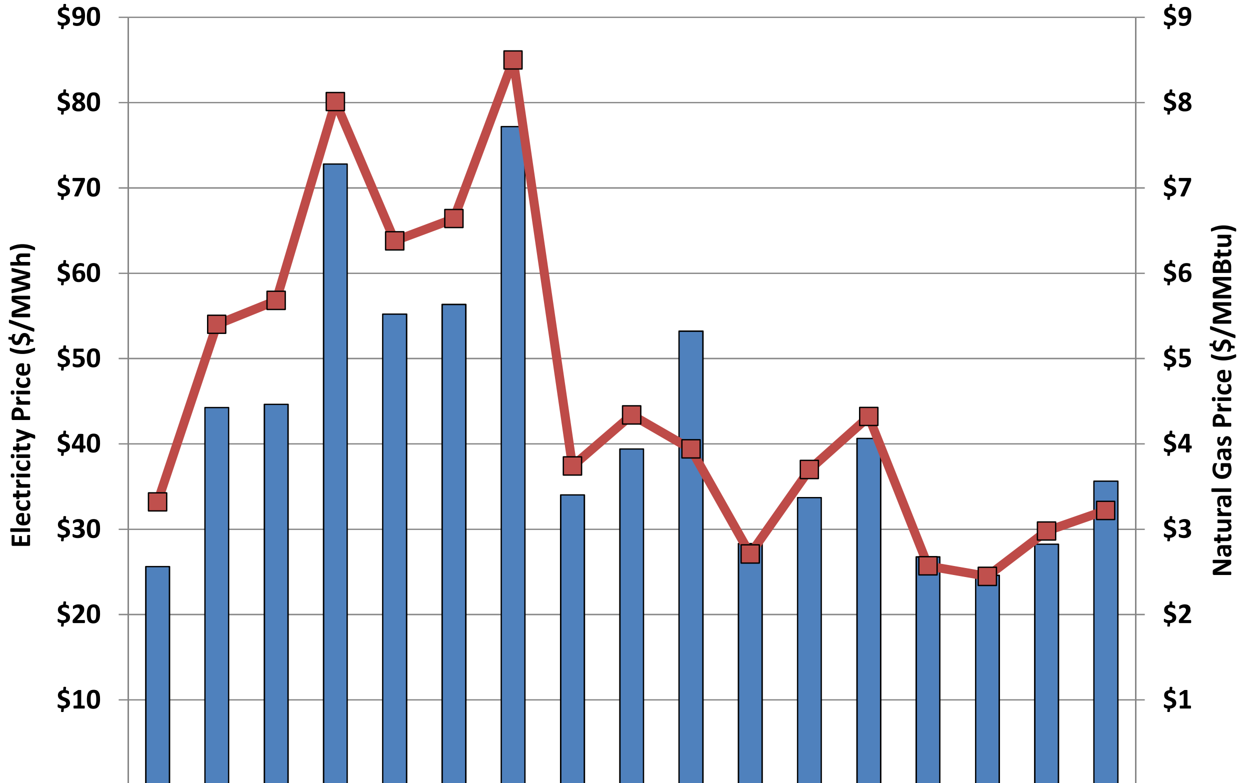
More detailed studies show that almost all the RENA uplift occurred in market hours when there was transmission congestion. The three factors contributing most to RENA uplift in 2018 were the settlement of day-ahead PTP obligations linked to options, network model differences between day-ahead market and real-time market, and load distribution factor inconsistencies between day-ahead and real-time market. The amount of RENA uplift associated with the settlement of day-ahead obligations linked to options was $51 million, the uplift amount resulted from the inconsistencies of load distribution factors was $55 million, and the uplift amount from the transmission model differences in day-ahead and real-time was approximately $114 million. These uplifted values were offset by $97 million in negative uplift related to the manual adjustments made to the clearing price of PTP obligations.

Figure 2 above shows that RENA uplift from the settlement of day-ahead PTP obligations linked to options, described as CRR Uplift, was relatively high in January, May and August, and the uplift from transmission modelling differences was significant in May, June and July. It is worth noting that while the manual corrections of those PTP obligation awards in the day-ahead market usually generated a negative uplift for each month in 2018, the magnitude was significant in May, June and July, which partially offset the larger uplifts resulting from the transmission system modelling differences in those three months. Uplift from the contributions of load distribution factor differences between day-ahead and real-time, described as LDF Contribution, was positive for all the months in 2018, with the most notable contributions in January, June and October.

The task of maintaining accurate and consistent load distribution factors across all markets is a difficult one, made more difficult in areas with large amounts of localized load growth. These are exactly the types of areas that draw higher levels of market interest. To the extent ERCOT is unable to of maintain accurate and consistent load distribution factors across all markets, RENA impacts will persist.

Figure 3 below provides additional historic perspective on the ERCOT average real-time energy prices as compared to the average natural gas prices in each year from 2002 through 2018.

Figure 3: ERCOT Historic Real-Time Energy and Natural Gas Prices



	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
ERCOT	25.64	44.26	44.64	72.79	55.22	56.35	77.19	34.03	39.40	53.23	28.33	33.71	40.64	26.77	24.62	28.25	35.63
Gas	3.32	5.40	5.68	8.01	6.38	6.64	8.50	3.74	4.34	3.94	2.71	3.70	4.32	2.57	2.45	2.98	3.22

Like Figure 1, Figure 3 shows the close correlation between the average real-time energy price in ERCOT and the average natural gas price. Such relationship is consistent with expectations in ERCOT where natural gas generators predominate and tend to set the marginal price. The most noticeable exception occurred in 2011, when energy prices were affected by shortage conditions. Real-time prices in 2018 also showed the impacts of increased scarcity, but they were not as large as in 2011.

Energy prices vary across the ERCOT market because of congestion costs that are incurred as power is delivered over the network. Figure 4 shows the monthly load-weighted average prices in the four geographic ERCOT zones during 2017 and 2018. These prices are calculated by weighting the real-time energy price for each interval and each zone by the total load in that interval. Load-weighted average prices are most representative of what loads are likely to pay, assuming that real-time energy prices are, on average, generally consistent with bilateral or other forward contract prices.

Figure 4: Average Real-Time Energy Market Prices by Zone

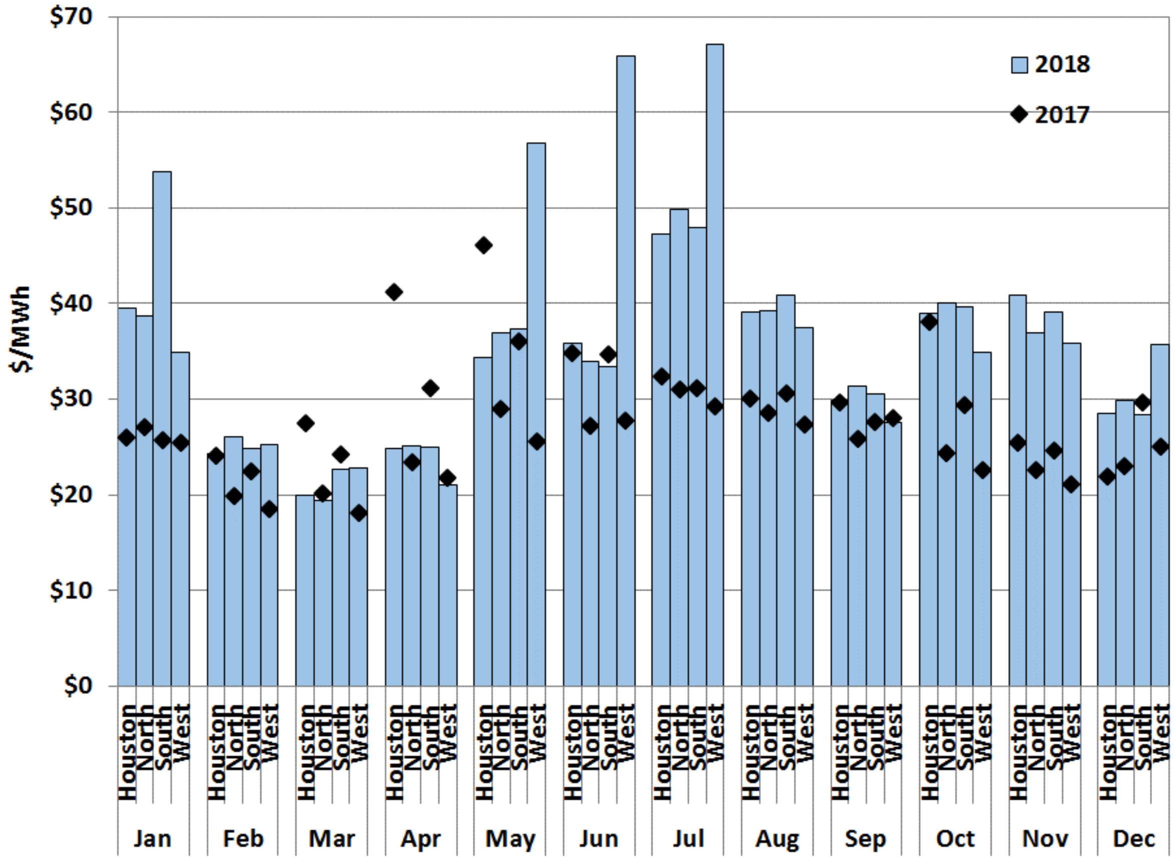


Table 1 provides the annual load-weighted average price for each zone for the past eight years and includes the annual average natural gas price for reference.

Table 1: Average Annual Real-Time Energy Market Prices by Zone

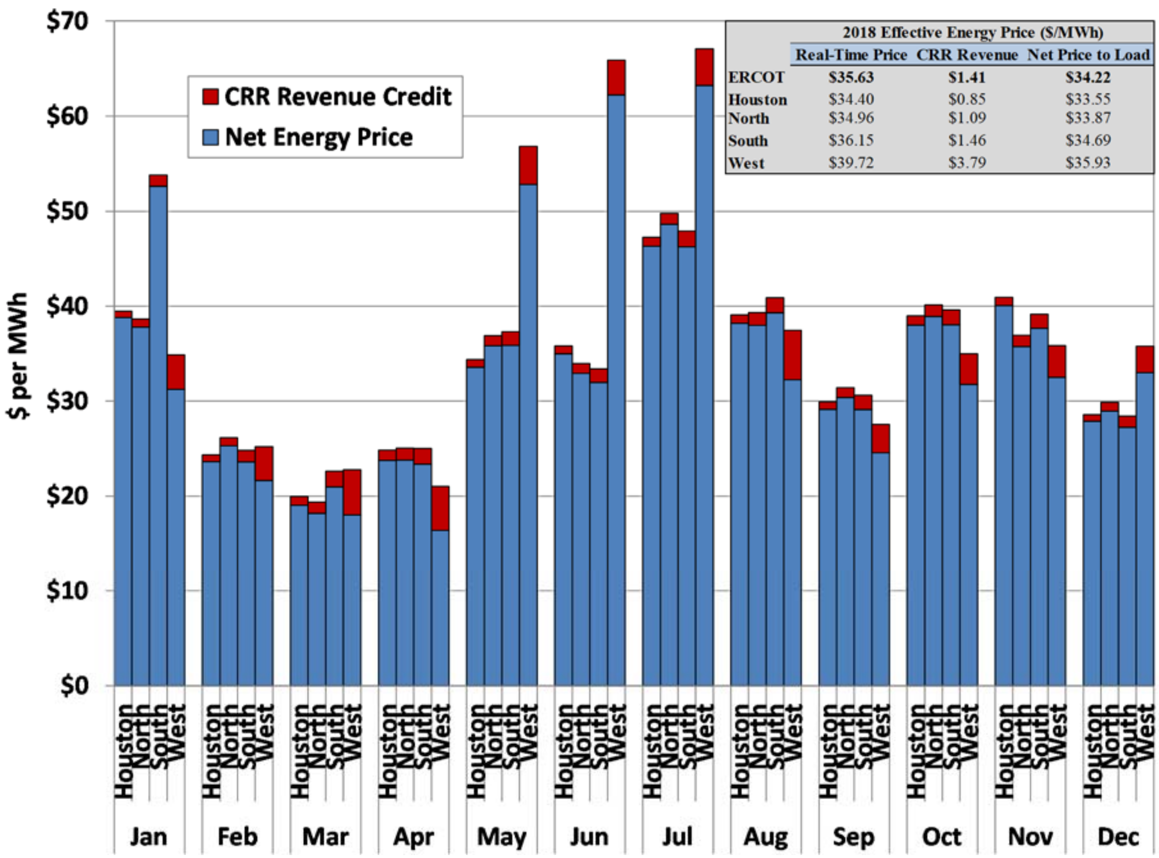
($/MWh)	2011	2012	2013	2014	2015	2016	2017	2018
ERCOT	**$53.23**	**$28.33**	**$33.71**	**$40.64**	**$26.77**	**$24.62**	**$28.25**	**$35.63**
Houston	$52.40	$27.04	$33.63	$39.60	$26.91	$26.33	$31.81	$34.40
North	$54.24	$27.57	$32.74	$40.05	$26.36	$23.84	$25.67	$34.96
South	$54.32	$27.86	$33.88	$41.52	$27.18	$24.78	$29.38	$36.15
West	$46.87	$38.24	$37.99	$43.58	$26.83	$22.05	$24.52	$39.72
($/MMBtu)								
Natural Gas	$3.94	$2.71	$3.70	$4.32	$2.57	$2.45	$2.98	$3.22

The pattern of zonal prices was different in 2018 compared to recent years. Constraints on the ability to import generation led to the Houston zone being the highest priced zone in 2016 and 2017. In 2018, the Houston zone exhibited the lowest prices due to the completion of the Houston import transmission project. Although slightly narrower than in 2017, price spreads were still noticeable across all zones in 2018 because of higher natural gas prices and the increased impacts of transmission congestion. For the first time since 2014, the West zone had the highest prices, primarily due to multiple localized real-time transmission constraints.

West zone prices relative to the ERCOT average have varied through the years. Prior to 2012, West zone prices were lower than the ERCOT average because of a wind generation surplus resulting from export limitations. Between 2012 and 2014, load growth caused by higher oil and natural gas production activity resulted in localized import constraints and higher prices. Even as investment in transmission facilities continued, the amount of wind generation additions began to create export limitations and resulting lower prices beginning in 2015. That trend was reversed in 2018 due to multiple real-time transmission constraints. More details about the transmission constraints influencing zonal energy prices are provided in Section III: Transmission Congestion and Congestion Revenue Rights.

Another factor influencing zonal price differences is Congestion Revenue Right (CRR) auction revenue distributions. These are distributed to Qualified Scheduling Entities (QSEs) representing load, based on a zonal and ERCOT-wide monthly load-ratio share. The CRR auction revenues have the effect of reducing the total cost to serve load borne by a QSE. Figure 5 below shows the effect that this reduction has on a monthly basis, by zone.

Figure 5: Effective Real-Time Energy Market Prices

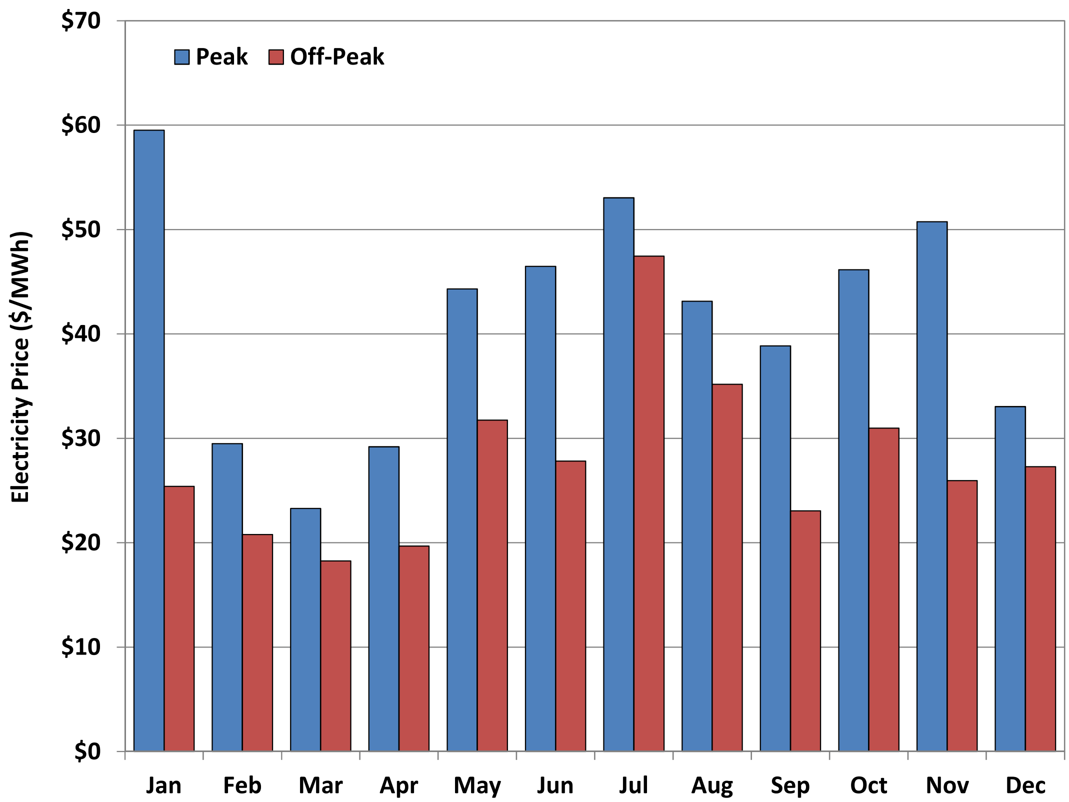


	2018 Effective Energy Price ($/MWh)		
	Real-Time Price	CRR Revenue	Net Price to Load
ERCOT	**$35.63**	**$1.41**	**$34.22**
Houston	$34.40	$0.85	$33.55
North	$34.96	$1.09	$33.87
South	$36.15	$1.46	$34.69
West	$39.72	$3.79	$35.93

With the CRR auction revenue offset included, the ERCOT-wide load-weighted average price rose by $7.03 per MWh to $34.22 per MWh in 2018, compared to $27.19 per MWh in 2017. Zonal differences were far less pronounced in 2018 compared to 2017, even as prices and credit increased across all zones.

Real-time energy prices not only vary by location, but also by time of day. Figure 6 shows the load-weighted average real-time prices in ERCOT for the categories of Peak and Off-Peak for each month in 2018. The Peak block includes hours ending 7-22 on weekdays; the Off-Peak block includes hours ending 1-6 and 23-24 on weekdays and all hours on weekends. These pricing blocks align with the categories traded in forward markets.

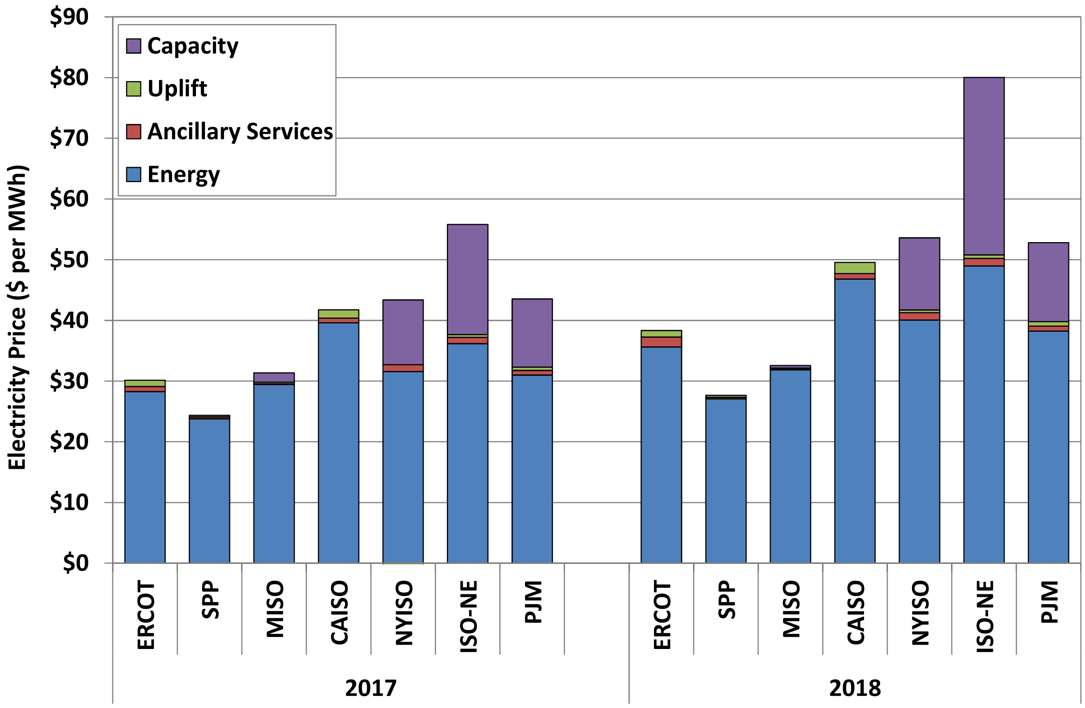
Figure 6: Peak and Off-Peak Pricing



As expected, Peak hours were higher priced than Off-Peak hours for every month in 2018. The monthly difference ranged from a minimum of $5.03 per MWh in March to a maximum of $34.10 per MWh in January. The large difference in January was due to two settlement intervals with very high prices during the morning of January 23. Excluding the effects of those intervals, reduces the difference to $26.55. The average difference between monthly Peak and Off-Peak pricing was $13.63 per MWh.

To provide additional perspective on the outcomes in the ERCOT market, Figure 7 below compares the all-in price in ERCOT with other organized electricity markets in the United States: Southwest Power Pool (SPP), Midcontinent ISO (MISO), California ISO (CAISO), New York ISO (NYISO), ISO New England (ISO-NE), and the PJM Interconnection.

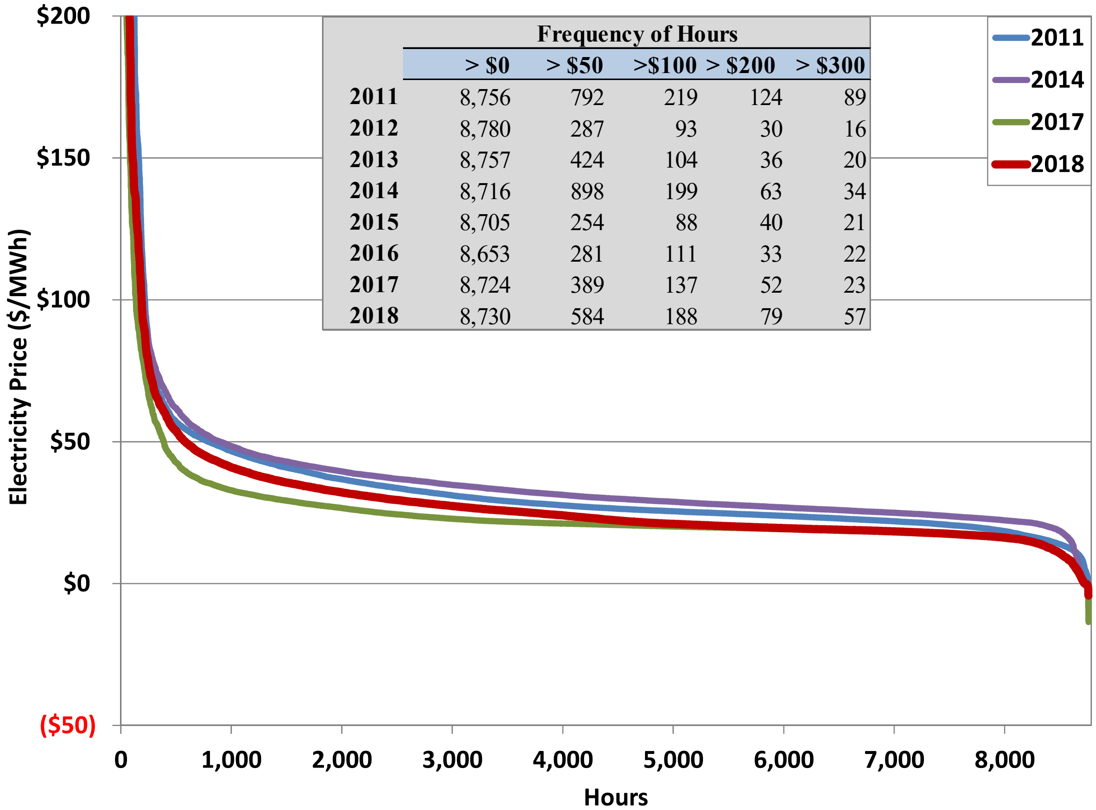
Figure 7: Comparison of All-in Prices Across Markets



The figure shows the average cost (per MWh of load) in each market, separated into the components energy, ancillary services (reserves and regulation), capacity markets (if applicable), and uplift. The ERCOT uplift amounts may appear larger than some other markets due to the inclusion of the administrative fee in the ERCOT values. The costs of funding the ERCOT organization are more than half of the ERCOT uplift amount. It is not clear from the data collected if similar costs are included in the values for other RTOs. Figure 7 also shows that all-in prices were higher across all U.S. markets in 2018. The increase ranged from modest in MISO to a sizable increase to the capacity component in ISO-NE for the second year in a row. Further, prices in ERCOT were lower than most of the other RTOs, the exceptions being MISO and SPP. Lower prices in these two markets may be explained by lower natural gas prices in MISO and proportionally more wind generation in SPP.

A price duration curve indicates the number of hours (shown on the horizontal axis) that the price is at or above a certain level (shown on the vertical axis). Figure 8 below shows price duration curves for the ERCOT energy market for 2018 and 2017, and includes 2014 and 2011 for historical context. The prices in this figure are the hourly ERCOT average prices derived by load weighting the zonal settlement point prices.

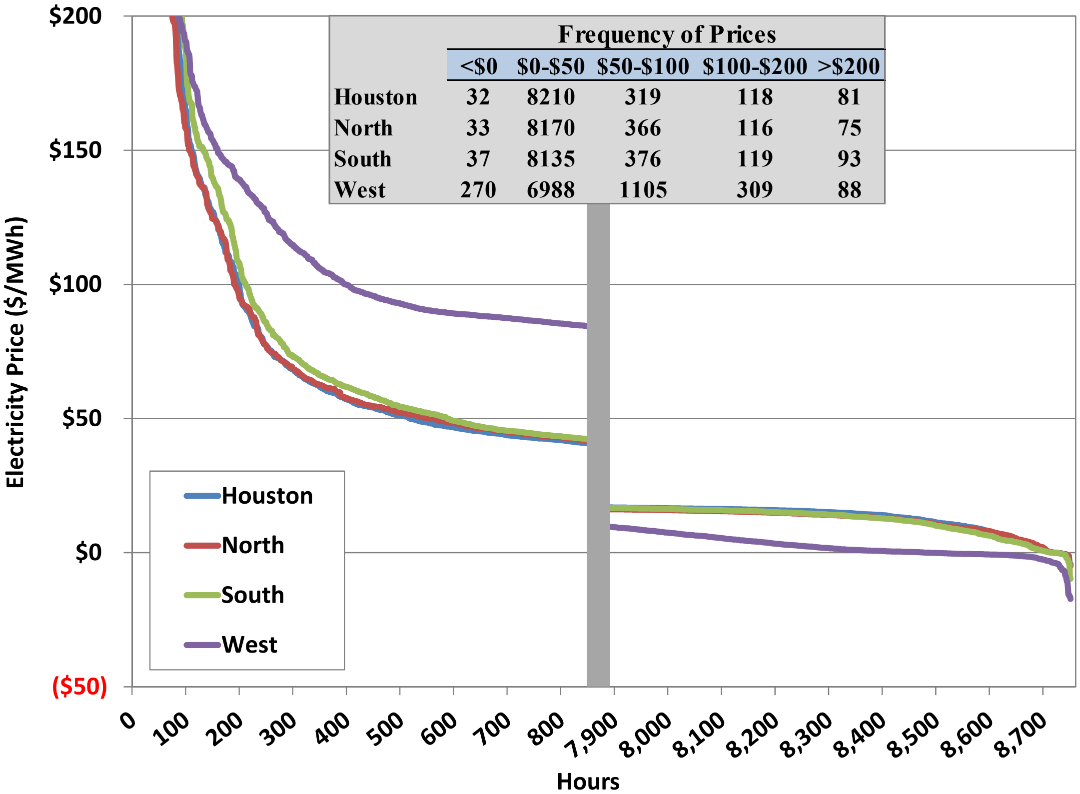
Figure 8: ERCOT Price Duration Curve



Negative ERCOT-wide prices may occur when wind is the marginal generation. More installed wind generation and additional transmission infrastructure has led to increased occurrences of negative prices over the past few years, reaching a high of 131 hours in 2016. That trend reversed in 2017, when there were 36 hours with ERCOT-wide prices at or below zero. In 2018, there were 30 hours with ERCOT-wide prices at or below zero.

To more closely examine the variation in zonal real-time energy prices, Figure 9 below shows the top and bottom 10% of the duration curves of hourly average prices in 2018 for the four zones. Compared to the other zones, both low and high prices in the West zone were noticeably different from the other zones. The lowest prices in the West zone were much lower than the lowest prices in the other zones and also noticeably higher than high prices in the other zones. The differences on both ends of the curves can be explained by the effects of transmission congestion. Constraints limiting the export of low-priced wind and solar generation explain low prices, whereas localized constraints limiting the flow of electricity to the burgeoning oil and gas loads explain the higher prices.

Figure 9: Zonal Price Duration Curves

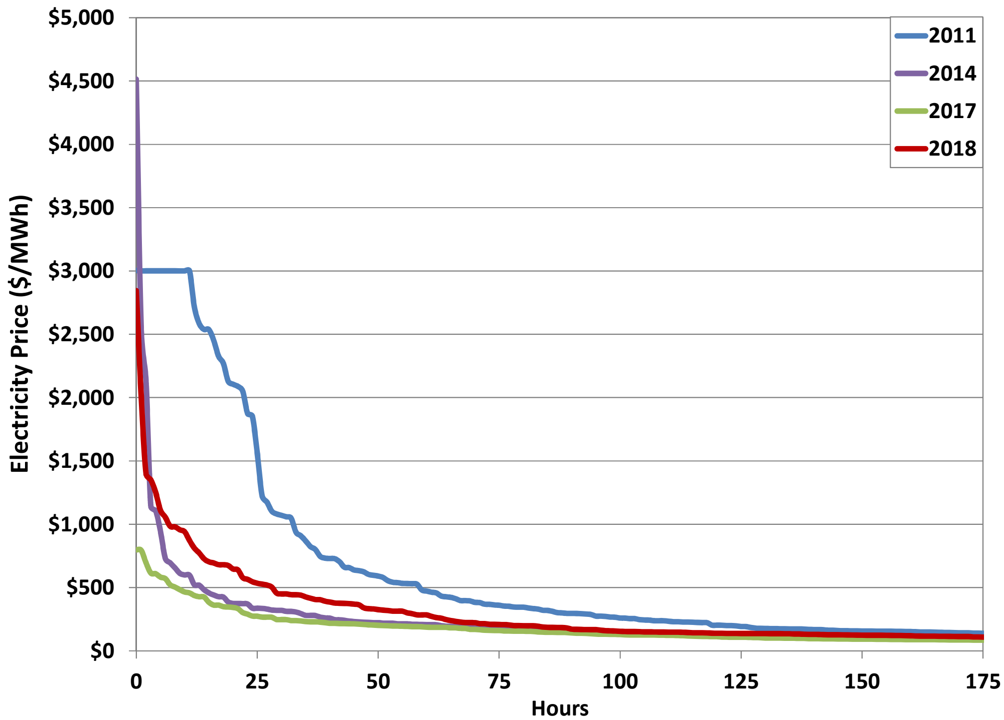


	<$0	$0-$50	$50-$100	$100-$200	>$200
Houston	32	8210	319	118	81
North	33	8170	366	116	75
South	37	8135	376	119	93
West	270	6988	1105	309	88

More details about the transmission constraints influencing zonal energy prices are provided in Section III: Transmission Congestion and Congestion Revenue Rights.

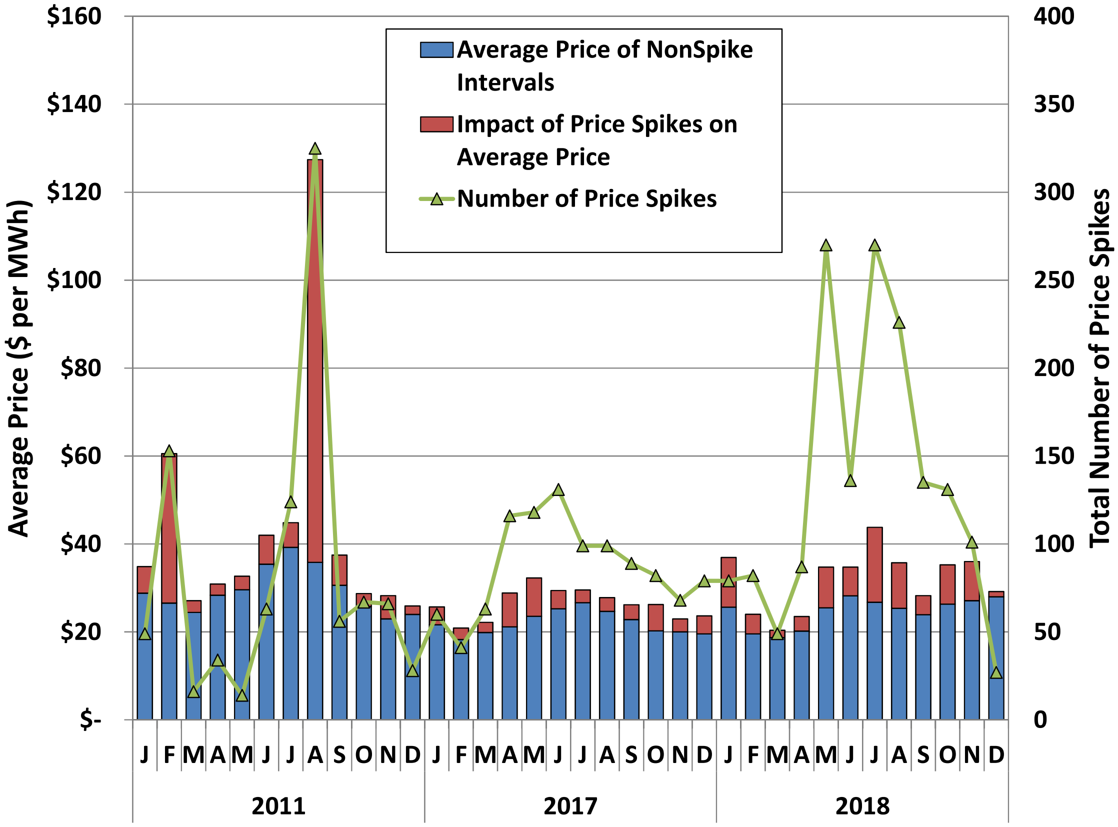
Figure 10 compares prices for the highest-priced 2% of hours in 2018 with 2017. Years 2014 and 2011 are also included for historical context. Energy prices for the top 100 hours of 2011 were significantly higher, while all subsequent years have followed an almost identical pattern through 2017. In 2018, energy prices for the top 100 hours rose slightly from previous years. The higher prices in 2011 were due to high loads leading to more shortage conditions in that year. Although the peak load in 2011 has been exceeded since 2015, generation additions during the intervening years have meant that shortage conditions continue to be rare, even with the historically low installed reserve margin in 2018.

Figure 10: ERCOT Price Duration Curve – Top 2% of Hours



To better observe the effect of the highest-priced hours on the average real-time energy price, the following analysis focuses on the frequency of price spikes in the real-time energy market, as presented in Figure 11. For this analysis, price spikes are defined as intervals when the load-weighted average energy price in ERCOT is greater than 18 MMBtu per MWh multiplied by the prevailing natural gas price. Prices at this level typically exceed the marginal costs of virtually all on-line generators in ERCOT.

Figure 11: Average Real-Time Energy Price Spikes

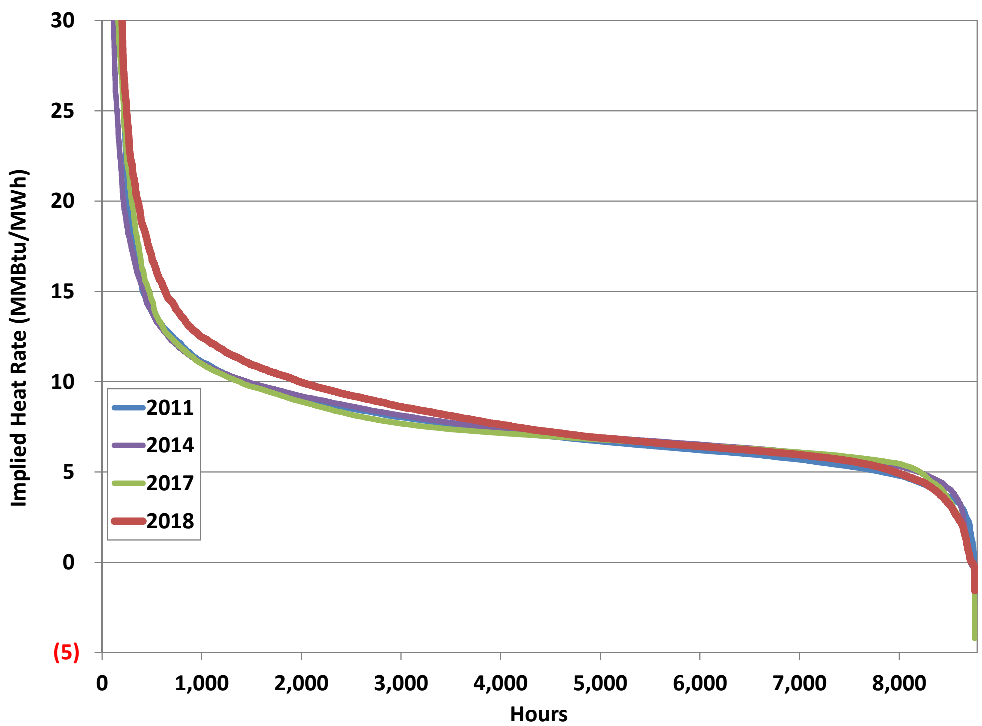


Price spikes were much more frequent and impactful in 2018 compared to 2017 because of larger contributions from the operating reserve adder during reduced reserve availability. The overall impact of price spikes in 2018 was $7.28 per MWh, or 30% of total price. The pattern of price spikes from 2011 is also shown for comparison. Price spikes in 2011 were less frequent and much more impactful, resulting from the periods of extreme weather experienced in both February and August of that year.

B. Real-Time Prices Adjusted for Fuel Price Changes

Although real-time electricity prices are driven largely by changes in fuel prices, natural gas prices in particular, they are also influenced by other factors. To summarize the changes in energy price that were related to these other factors, an "implied heat rate" is calculated by dividing the real-time energy price by the natural gas price. Figure 12 and Figure 13 show the load-weighted, hourly average real-time energy price adjusted to remove the effect of natural gas price fluctuations. The first chart shows a duration curve where the real-time energy price is replaced by the marginal heat rate that implied by assuming natural gas was always on the margin.

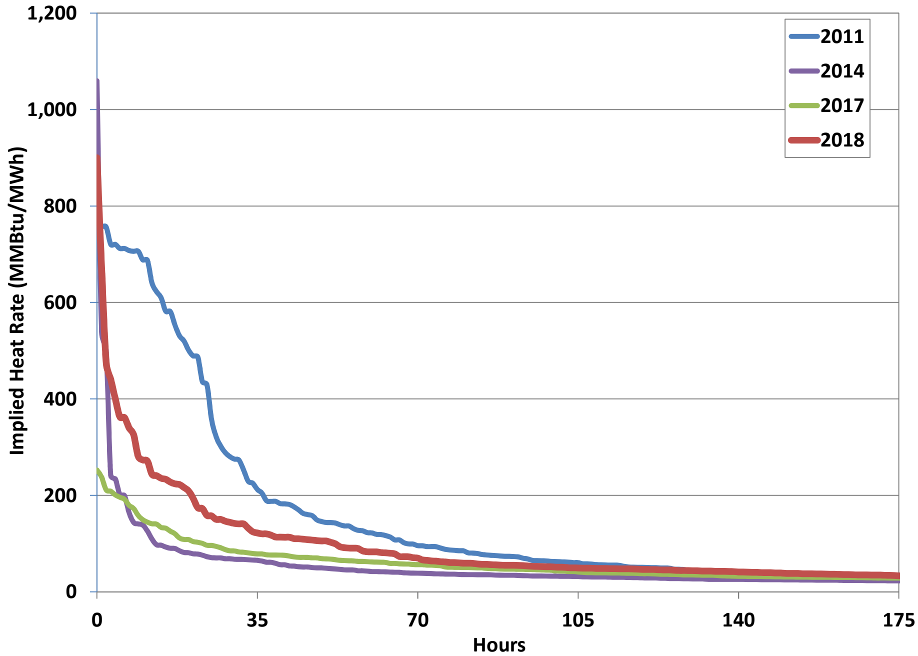
Figure 12: Implied Heat Rate Duration Curve – All Hours



Implied system-wide heat rates for 2018 were somewhat higher on the high end of the curve as compared to other years. This rise in heat rates is explained by lower reserves in 2018, which caused a larger contribution of shortage pricing in high priced hours.

Figure 13 shows the implied marginal heat rates for the top 2% of hours in 2018 and 2017, with years 2014 and 2011 included for historical context. The implied heat rate duration curve for the top 2% of hours in 2018 was higher than 2017 because of the increased contribution of shortage pricing. Even with the increased contribution from shortage pricing in 2018, implied heat rates did not rise to the extreme levels seen in 2011, a year with extreme and record breaking heat and drought.

Figure 13: Implied Heat Rate Duration Curve – Top 2% of Hours



To further illustrate these differences, Figure 14 shows the implied marginal heat rates on a monthly basis in each of the ERCOT zones for 2018 and 2017. This figure is the fuel price-adjusted version of Figure 4 in the prior subsection. Implied heat rates in 2018 were equal to or higher in all zones in 2017, with the largest increase in the West zone.

Figure 14: Monthly Average Implied Heat Rates

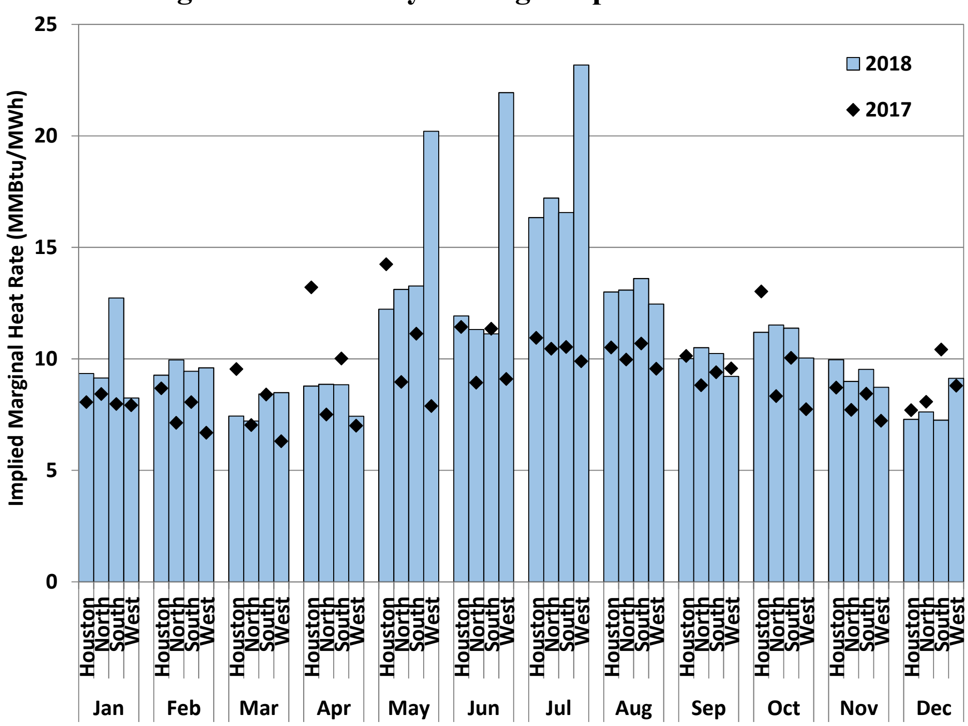


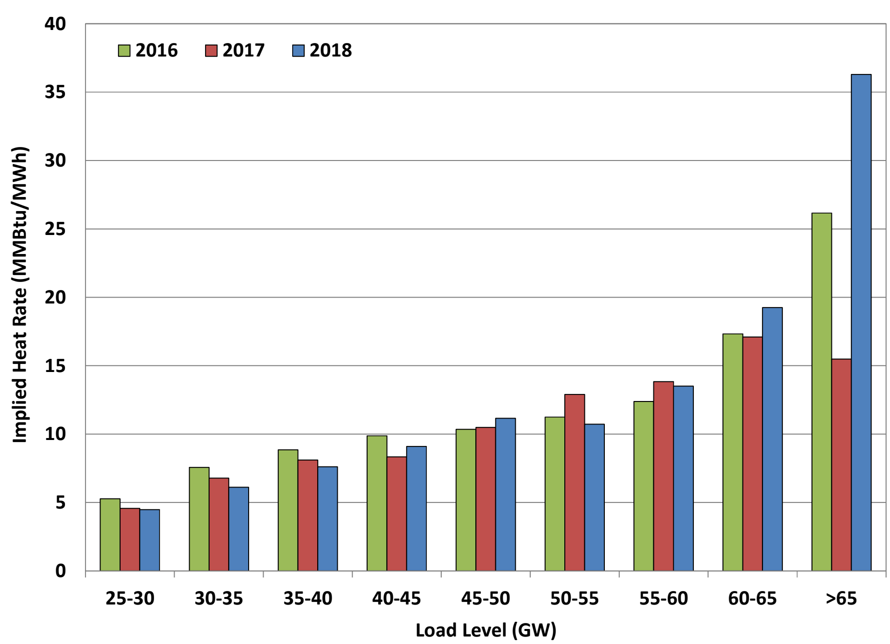
Table 2 displays the annual average implied heat rates by zone for 2011 through 2018. Adjusting for natural gas price influence, Table 2 shows that the annual, system-wide average implied heat rate increased in 2018 compared to 2017. Further, zonal implied heat rates in 2018 were the highest experienced in the nodal market since 2011.[12] Zonal variations in the implied heat rate were greater in 2017 because of the increased influence of transmission congestion. The zonal variations in 2018 were not as pronounced.

Table 2: Average Implied Heat Rates by Zone

(MMBtu/MWh)	2011	2012	2013	2014	2015	2016	2017	2018
ERCOT	**13.5**	**10.5**	**9.1**	**9.4**	**10.4**	**10.1**	**9.5**	**11.1**
Houston	13.3	10.0	9.1	9.2	10.5	10.8	10.7	10.7
North	13.7	10.2	8.9	9.3	10.2	9.7	8.6	10.9
South	13.8	10.2	9.2	9.6	10.6	10.1	9.9	11.2
West	11.9	14.1	10.3	10.1	10.4	9.0	8.2	12.3
($/MMBtu)								
Natural Gas	$3.94	$2.71	$3.70	$4.32	$2.57	$2.45	$2.98	$3.22

The examination of implied heat rates from the real-time energy market concludes by evaluating them at various load levels. Figure 15 below provides the average implied heat rate at various system load levels for years 2016 through 2018.

Figure 15: Implied Heat Rate and Load Relationship



[12] The implied heat rate for the West zone was highest in 2012 due to extreme transmission congestion.

In a well-performing market, a clear positive relationship between implied heat rate and load level is expected because resources with higher marginal costs are dispatched to serve higher loads. This relationship continued and the relatively modest effects of increased shortage conditions in 2018 can be seen at the highest load levels.

C. Aggregated Offer Curves

The next analysis compares the quantity and price of generation offered in 2018 to that offered in 2017. By averaging the amount of capacity offered at selected price levels, an aggregated offer stack can be assembled. Figure 16 provides the aggregated generator offer stacks for the entire year. The largest amount of capacity is actually not offered at any price. All capacity below generators' Low Sustainable Limit (LSL) is not able to respond to dispatch instructions and is a price-taking portion of the offer stack. Dispatchable capacity priced below zero is shown separately for wind and other types of generators. In 2018 the amount of capacity offered at prices less than zero increased by approximately 1,600 MW, roughly half coming from wind generators.

The next largest share of capacity is priced at levels between zero and a value equal to 10 times the daily natural gas price (known as the Fuel Index Price, or FIP). This price range represents the incremental fuel price, the largest component of the marginal cost to generate, for the vast majority of the ERCOT generation fleet. Approximately 500 MW of additional capacity was offered in this price range in 2018. Overall, the average amount of generation capacity offered into ERCOT's real-time market increased by nearly 2,500 MW in 2018. The amount of capacity offered at prices between ten multiplied by FIP and $75 per MWh increased by 330 MW from 2017 to 2018. There were smaller incremental amounts offered at higher prices, such that the percentage of total capacity in each category remained the same in 2018 as in 2017.

Figure 16: Aggregated Generation Offer Stack – Annual

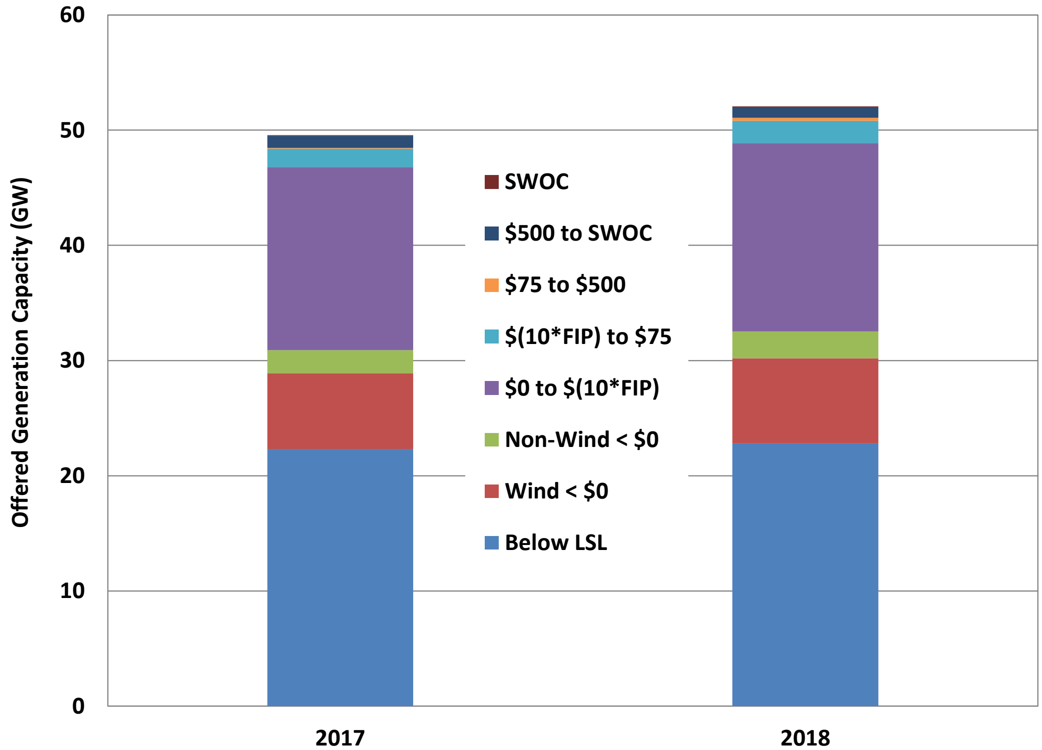
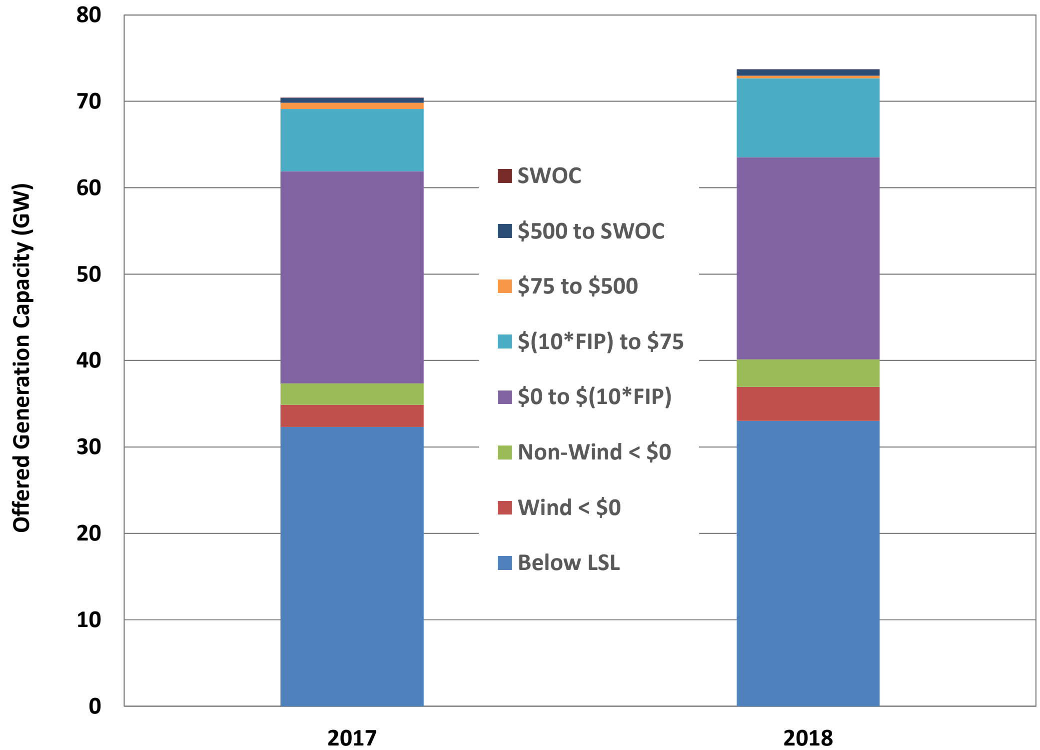


Figure 17 provides a similar comparison focused on the summer months. As shown below, the changes in the aggregated offer stacks between the summer of 2018 and 2017 were somewhat different than those in the annual aggregated offer stacks. The average offer stack for the summer of 2018 was approximately 3,200 MW higher than in the previous summer. The amount of additional capacity offered at negative prices was approximately 2,800 MW, half of which coming from wind units. The largest increase was 1,900 MW of additional capacity offered at prices between $(10*FIP) and $75per MWh. This increase was offset by a reduction of 1,200 MW of capacity offered at prices between zero and $(10*FIP).

Figure 17: Aggregated Generation Offer Stack – Peak Hour



Both the annual and peak hour offer stacks displayed a significant amount of capacity below units' low dispatchable limits in 2018. Because unit output is not dispatchable in this range, it is considered to be "price-taking" and is considered by the dispatch software to have a price of negative $250 per MWh. There has been a steady decrease in the amount of non-dispatchable, price-taking capacity since 2014. Prior to 2014, maximum generation capacity dispatchable based on offer curves was 23%. Since that time, the amount of dispatchable capacity has been steadily increasing. In 2018, the maximum dispatchable capacity was nearly 42%, with almost 24% dispatchable capacity in more than half the intervals. More dispatchable capacity is indicative of more generators competing based on offers, rather than being price-taking. This increase in dispatchable capacity was primarily from wind generation.

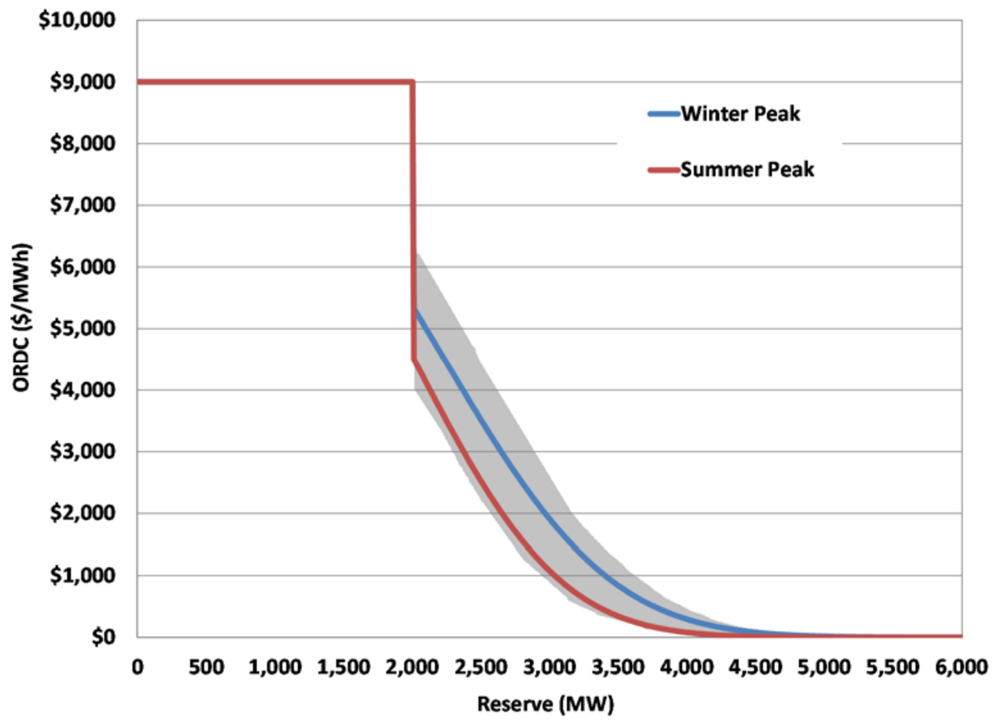
D. ORDC Impacts and Prices During Shortage Conditions

The Operating Reserve Demand Curve (ORDC) is a scarcity pricing mechanism that reflects the loss of load probability (LOLP) at varying levels of operating reserves multiplied by the deemed value of lost load (VOLL).[13] Selected at the time as an easier to implement alternative to real-time co-optimization of energy and ancillary services, the ORDC places an economic value

13 At the open meeting on September 12, 2013, the Commission directed ERCOT to move forward with implementing ORDC, including setting the Value of Lost Load at $9,000

on the reserves being provided, with separate pricing for online and offline reserves. Figure 18 depicts the ORDC curves applicable during winter and summer peak hours in 2018 superimposed over the range of ORDC curves for the year. The curves within the range are determined in advance for four-hour blocks that vary across seasons. This shaded depiction shows the breadth of distribution of the ORDC values across the year. Once available reserve capacity drops to 2,000 MW, prices will rise to $9,000 per MWh for all the ORDC curves.

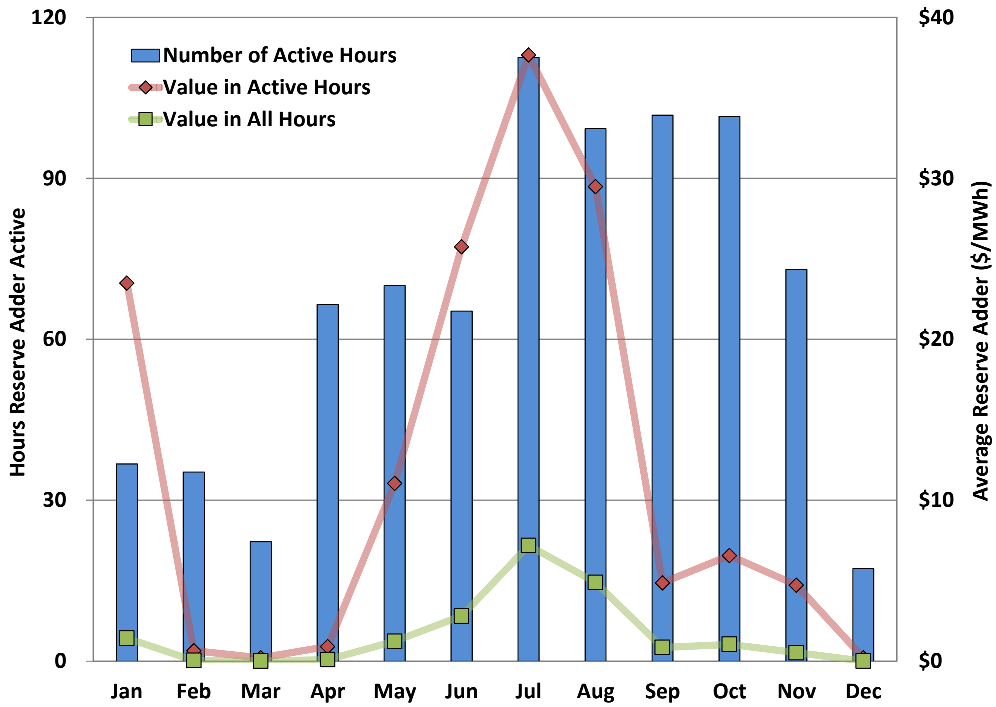
Figure 18: Winter and Summer Peak Operating Reserve Demand Curves



The following two analyses illustrate the contributions of the operating reserve adder and the reliability adder to real-time prices. As described above in Figure 1, the contributions of the energy price adders increased in 2018. The first of the two adders, the operating reserve adder, is a shortage value intended to reflect the expected value of lost load given online and offline reserve levels.

Figure 19 shows the number of hours in which the adder affected prices, and the average price effect in these hours and all hours. This figure shows that in 2018, the operating reserve adder had the largest impacts to price during July. The contribution from the operating reserve adder in 2018 was greater than in recent years due to slightly more frequent occurrences of shortage conditions. Overall, the operating reserve adder contributed $1.97 per MWh, or 5.5% to the annual average real-time energy price of $35.63 per MWh. The contribution from the operating reserve adder to real-time prices in 2018 was larger than in previous years due to the overall lower levels of operating reserves resulting from the lower level of installed reserve capacity. The effects of the operating reserve adder are expected to vary substantially from year to year, and to have the largest effects when low supply conditions and unusually high load conditions occur together and result in sustained shortages.

Figure 19: Average Operating Reserve Adder

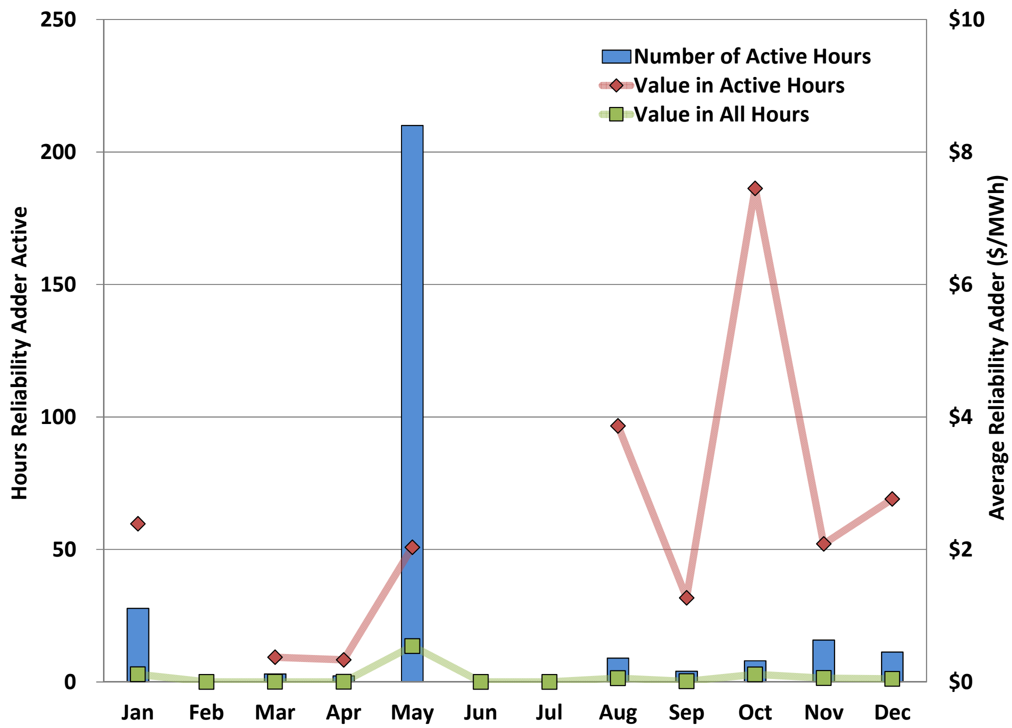


The reliability adder is intended to allow prices to reflect the costs of reliability actions taken by ERCOT, including RUCs and deployed load capacity. Absent this adder, prices will generally fall when these actions are taken.

Figure 20 below shows the impacts of the reliability adder in 2018. When averaged across only the hours when the reliability adder was non-zero, the largest price impacts of the reliability adder occurred during a few hours on October 4, 2018 when multiple resources were issued RUC instructions for system capacity insufficiency. The largest impact to price from the reliability adder occurred in May. Frequent RUC instructions, primarily in the first half of the month, resulted in a positive reliability adder in more than 200 hours, with a contribution to real-time price of $0.54 per MWh. The reliability adder was non-zero for 291 hours, or 3% of the time in 2018, most of which occurred in May. The contribution from the reliability adder to the annual average real-time energy price was $0.08 per MWh. The reliability adder had very little overall effect on market outcomes in 2018.

Weaknesses in the implementation of the reliability deployment adder have been identified and were under review by stakeholders at the end of 2018. The primary flaw identified in the calculation method used to determine the reliability adder was that relaxing the low dispatch limit of all resources made the price adders higher, even when the RUC-instructed resource was dispatched above its low dispatch limit the pricing run. The price adders fluctuated based on interval-to-interval changes in the system, including changes for resources that were not RUC-instructed. Relaxing high and low dispatch limits of resources that were not RUC-instructed was intended to avoid ramp limitations that could exaggerate the pricing impacts of the out-of-market action. Proposed changes to the calculation method of the reliability adder were discussed in 2018 in NPRR904, *Revisions to Real-Time On-Line Reliability Deployment Price Adder for ERCOT-Directed Actions Related to DC Ties*.

Figure 20: Average Reliability Adder

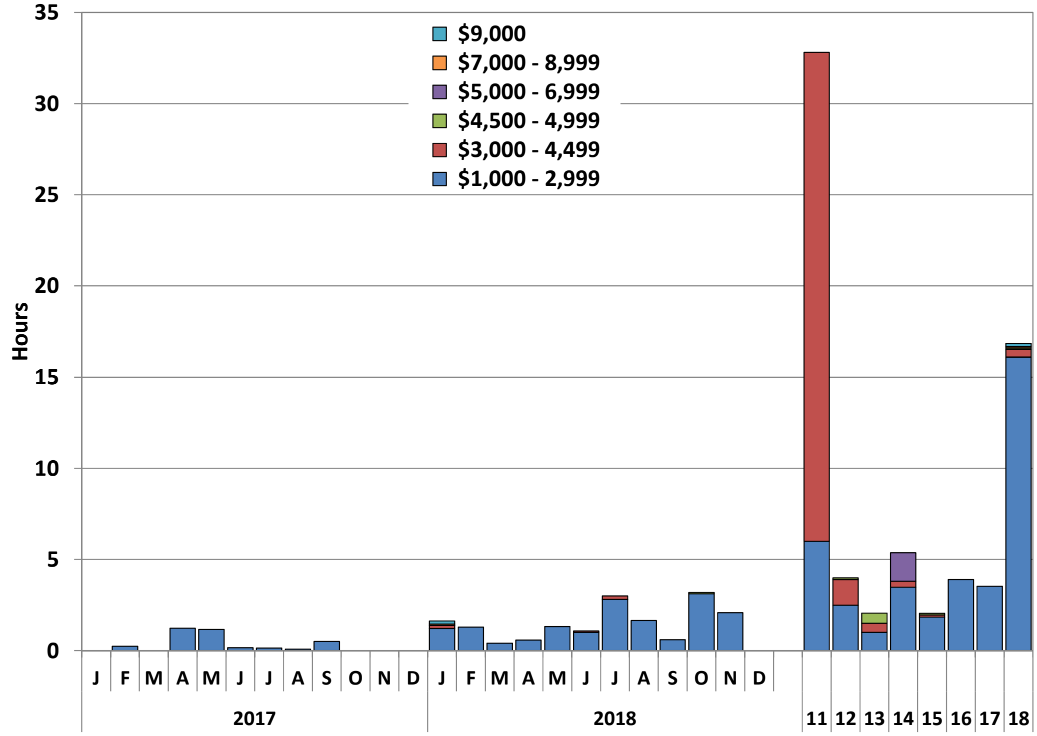


As an energy-only market, the ERCOT market relies heavily on high real-time prices that occur during shortage conditions. These prices provide key economic signals that provide incentives to build new resources and retain existing resources. However, the frequency and impacts of shortage pricing can vary substantially from year-to-year.

To summarize the shortage pricing that has occurred since 2011, Figure 21 below shows the aggregate amount of time when the real-time system-wide energy price exceeded $1,000 per MWh, by month for 2017 and 2018, as well as annual summaries for 2011 through 2018. This figure shows that high prices occurred more frequently in 2018 than in any previous year since 2011. Prices greater than $1,000 per MWh occurred in nearly 17 hours over the entire year. Prices were between $3,000 and $4,499 for roughly half an hour in 2018, while there were also brief occurrences of prices between $5,000 and $6,999 and $7,000 and $8,999.

On January 23, 2018, prices reached the $9,000 system-wide offer cap for the first time in ERCOT's history. The cap was reached during two dispatch intervals for a total of about 10 minutes.

Figure 21: Duration of High Prices

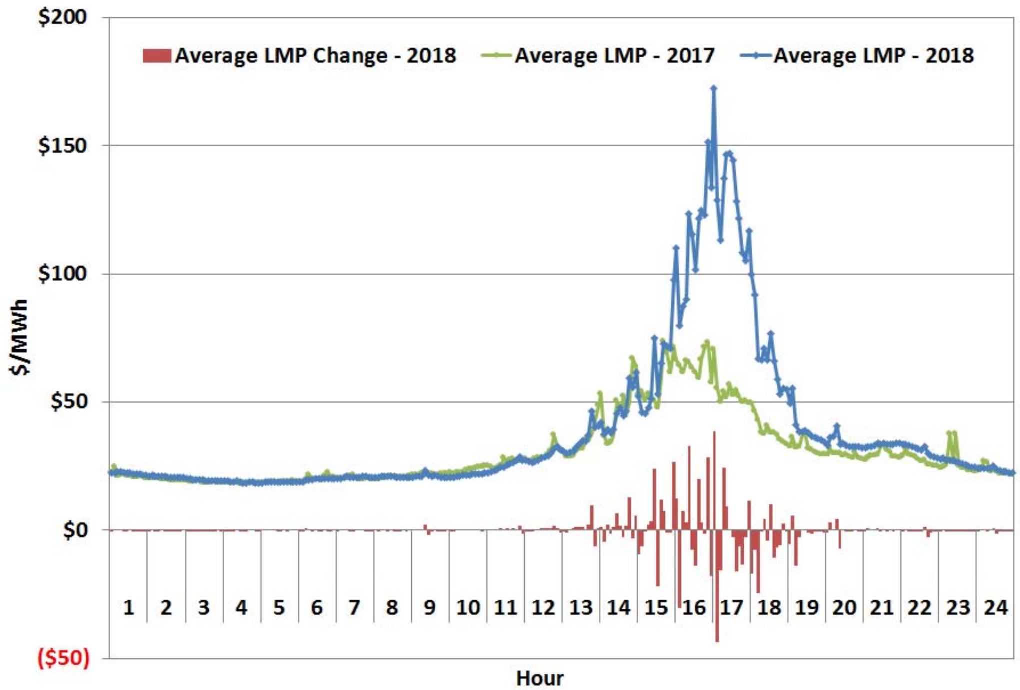


As a comparison, market prices cleared at the then effective cap of $3,000 per MWh for 28.44 hours in 2011. Extreme cold in February 2011 and unusually hot and sustained summer temperatures led to much more frequent shortages in that year.

E. Real-Time Price Volatility

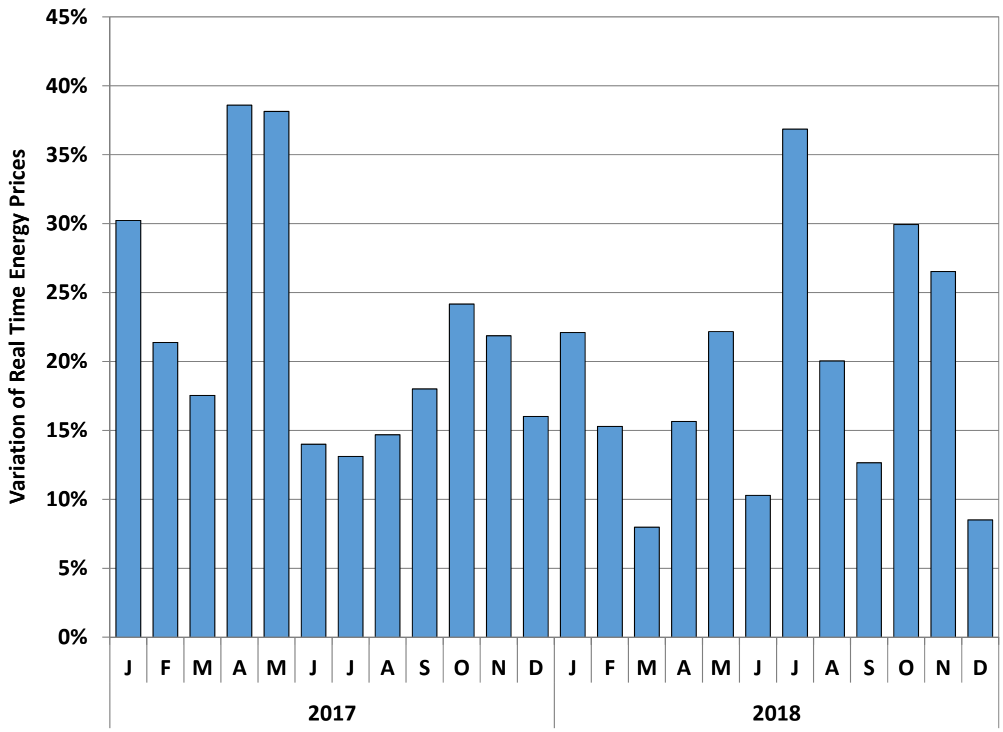
Volatility in real-time wholesale electricity markets is expected because system load can change rapidly and the ability of supply to adjust can be restricted by physical limitations of the resources and the transmission network. Figure 22 below presents a view of the price volatility experienced in ERCOT's real-time energy market during the summer months of May through August. Average five-minute real-time energy prices for 2018 are presented along with the magnitude of change in price during each five-minute interval. Average real-time energy prices from the same period in 2017 are also presented. Comparing average real-time energy prices for 2018 with those from 2017 shows very similar outcomes in off-peak hours but with higher average prices and greater volatility during peak hours of 2018.

Figure 22: Real-Time Energy Price Volatility (May – August)



The average absolute value of changes in five-minute real-time energy prices during the months of May through August, expressed as a percentage of average price, was 6.4% in 2018, compared to 5.5% in 2017. Expanding the view of price volatility, Figure 23 below shows monthly average changes in five-minute real-time prices by month for 2017 and 2018. Even with prices reaching the system-wide offer cap for the first time in January 2018, the highest price variability occurred during July when occurrences of high real-time prices were most frequent.

Figure 23: Monthly Price Variation



For another view of price volatility, Table 3 below shows the variation in 15-minute settlement point prices, expressed as a percentage of annual average price, for the four geographic zones for years 2013-2018.

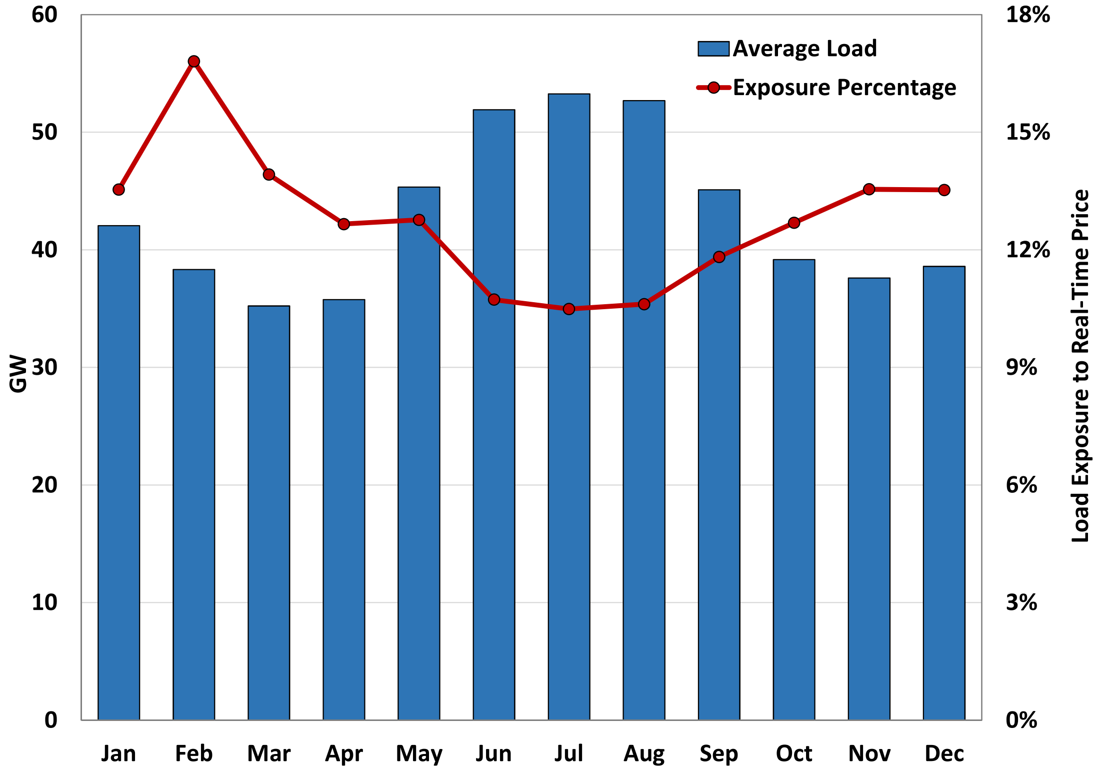
Table 3: Zonal Price Variation as a Percentage of Annual Average Prices

Load Zone	2013	2014	2015	2016	2017	2018
Houston	14.8	14.7	13.4	20.8	24.9	21.5
South	15.4	15.2	14.6	19.9	26.2	23.5
North	13.7	14.1	11.9	15.5	14.8	20.7
West	17.2	15.4	12.9	16.8	17.5	21.8

These results show that price volatility decreased in 2018 for the Houston and South zones and increased in the North and West. Although reduced from 2017, volatility remained highest in the South zone.

To conclude review of the real-time market, Figure 24 below shows the percentage of load exposed to real-time energy prices.

Figure 24: Monthly Load Exposure



This determination of exposure is based solely on ERCOT-administered markets, and does not include any bilateral or OTC index purchases. It is not surprising to find that during the high-load summer months of June, July and August, the smallest portion of load is potentially exposed to real-time prices. Although the overwhelmingly majority of load is not exposed to real-time prices, these prices do form the foundation for all pricing expectations, which inform both supplier and consumer contracting decisions.

II. Day-Ahead Market Performance

ERCOT's day-ahead market allows participants to make financially binding forward purchases and sales of power for delivery in real-time. Offers to sell can take the form of either a three-part supply offer, which allows sellers to reflect the unique financial and operational characteristics of a specific generation resource, or an energy-only offer, which is location specific but not associated with a generation resource. Bids to buy are also location specific. In addition to the purchase and sale of power, the day-ahead market also includes ancillary services and Point-to-Point (PTP) obligations. PTP obligations allow parties to hedge the incremental cost of congestion between day-ahead and real-time operations.

With the exception of the acquisition of ancillary service capacity, the day-ahead market is a financial market. Although all bids and offers are evaluated for the ability to reliably flow on the transmission network, there are no operational obligations resulting from the day-ahead market. Day-ahead transactions may be made for a variety of reasons, including satisfying the participant's own demand, managing risk by hedging the participant's exposure to real-time prices or congestion, or arbitraging the real-time prices. For example, load-serving entities can insure against the higher volatility of real-time market prices by purchasing in the day-ahead market. Finally, the day-ahead market helps inform participants' generator commitment decisions. For all of these reasons, the effective performance of the day-ahead market is essential.

In this section, energy pricing outcomes from the day-ahead market in 2018 are reviewed and convergence with real-time energy prices is examined. The volume of activity in the day-ahead market, including a discussion of PTP obligations, is also reviewed. This section concludes with a review of the ancillary service markets.

A. Day-Ahead Market Prices

One indicator of market performance is the extent to which forward and real-time spot prices converge over time. Forward prices will converge with real-time prices when: (1) there are low barriers to shifting purchases and sales between the forward and real-time markets; and (2) sufficient information is available to market participants to allow them to develop accurate expectations of future real-time prices. When these conditions are met, market participants can be expected to arbitrage predictable differences between forward prices and real-time spot prices by increasing net purchases in the lower priced market and increasing net sales in the higher priced market. This improves the convergence of forward and real-time prices, which should lead to improved commitment of resources needed to satisfy the system's real-time needs.

In this subsection, price convergence between the day-ahead and real-time markets is evaluated. This average price difference reveals whether persistent and predictable differences exist between day-ahead and real-time prices, which participants should arbitrage over the long term.

To measure the short-term deviations between real-time and day-ahead prices, the average of the absolute value of the difference between the day-ahead and real-time price are calculated on a daily basis. This measure captures the volatility of the daily price differences, which may be large even if the day-ahead and real-time energy prices are the same on average.[14]

Figure 25: Convergence Between Day-Ahead and Real-Time Energy Prices

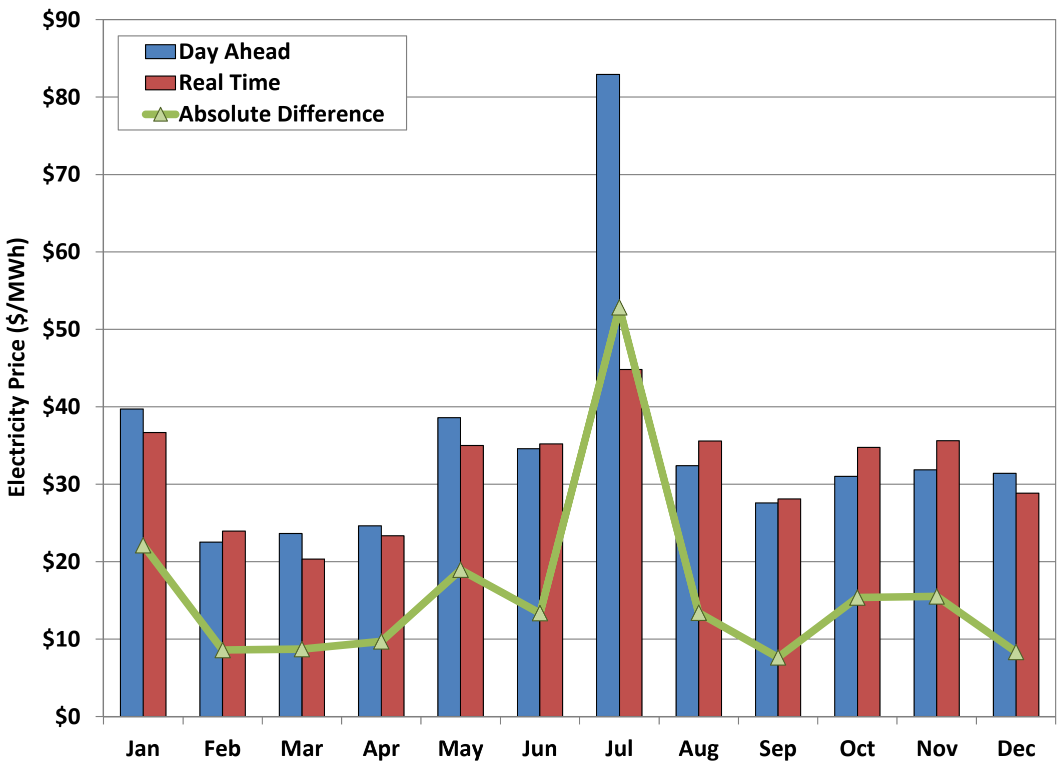


Figure 25 summarizes the price convergence between the day-ahead and real-time markets, by month, in 2018. Day-ahead and real-time prices averaged $35 and $32 per MWh in 2018, respectively.[15] Price convergence was evident in all months in 2018 except July, when day-ahead and real-time prices diverged sharply. Smaller quantities of installed reserves for the summer of 2018 led to expectations of more frequent shortage conditions and produced much higher day-ahead prices compared to recent experience. These expectations for high prices were not manifested in real-time due to higher than usual resource availability.

Real-time energy prices in ERCOT are allowed to rise to levels that are much higher than in other organized electricity markets, which increases risk and would explain a higher day-ahead

14 For example, if day-ahead prices are $30 per MWh on two consecutive days while real-time prices are $20 per MWh and $40 per MWh respectively, the absolute price difference between the day-ahead market and the real-time market would be $10 per MWh on both days, while the difference in average prices would be $0 per MWh.

15 These values are simple averages, rather than load-weighted averages as presented in Figures 1 and 2.

premium in ERCOT. Although six months experienced a day-ahead premium in 2018, it should not be expected that every month will produce a day-ahead premium. The real-time risks that lead to the premiums will materialize unexpectedly on occasion, resulting in real-time prices that exceed day-ahead prices (*e.g.*, in February, June, and August through November).

The average absolute difference between day-ahead and real-time prices was $16.21 per MWh in 2018. This represents an increase from $8.60 per MWh and $7.44 per MWh in 2017 and 2016, respectively. The large difference in July contributed significantly to the overall increase, but even with the effects of July removed, the absolute difference between day-ahead and real-time prices was still higher in 2018 than in 2016 and 2017.

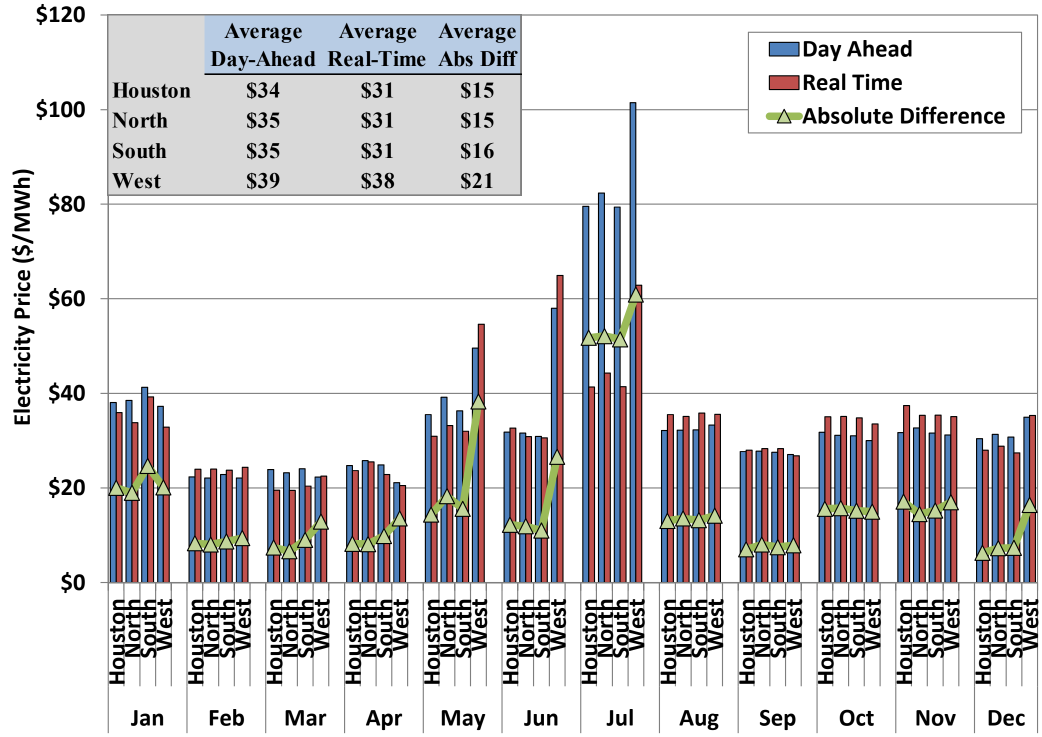
Table 4 displays the average day-ahead and real-time prices, showing the convergence for years 2011 through 2018. The difference in 2018 average real-time and day-ahead prices was larger than in recent years, and at the same level as in 2011.

Table 4: Historic Average Day-Ahead and Real-Time Prices

	Average Day-Ahead Price	Average Real-Time Price
2018	$35	$32
2017	$26	$26
2016	$23	$22
2015	$26	$25
2014	$40	$38
2013	$33	$32
2012	$29	$27
2011	$46	$43

In Figure 26 below, monthly day-ahead and real-time prices are shown for each of the geographic zones. Even with the transmission congestion related to Houston import constraints largely resolved in 2018, overall volatility still increased in 2018, particularly in the West zone. July 2018 witnessed the most pronounced price differences, with an average difference between day-ahead and real-time prices of $52.84 per MWh. Finally, although the average day-ahead and real-time prices were the most similar in the West zone, the average absolute difference in the West zone was the largest. This trend is explained by wide swings in West zone prices, the result of different kinds of transmission congestion constraints in the area.

Figure 26: Day-Ahead and Real-Time Prices by Zone



B. Day-Ahead Market Volumes

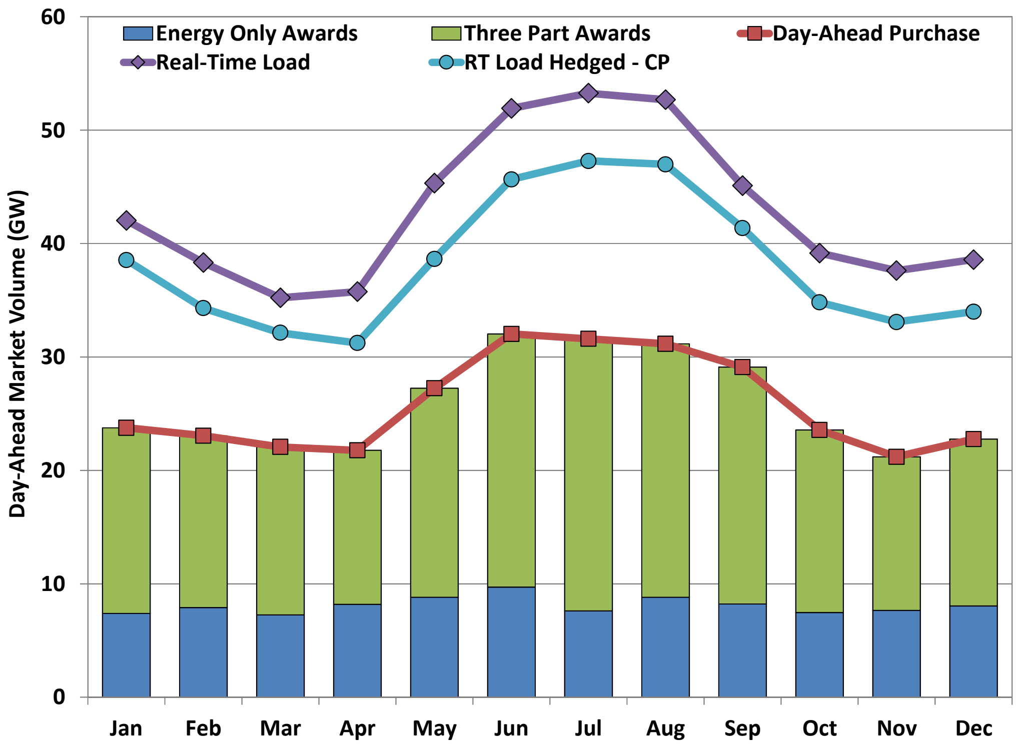
The next figure summarizes the volume of day-ahead market activity by month, which includes both the purchases and sales of energy, as well as the volume of PTP obligations that represent the system flows between a Load Zone and other locations. Figure 27 below shows that the volume of day-ahead purchases provided through a combination of three-part generator-specific offers (including start-up, no-load, and energy costs) and virtual energy offers was approximately 60% of real-time load in 2018, which was an increase compared to 55% in 2017. Although it may appear that many loads are still subjecting themselves to greater risk by not locking in a day-ahead price, other transactions that utilize PTPs are used to hedge real-time prices and congestion.

PTP obligations are financial transactions purchased in the day-ahead market. Although PTP obligations do not themselves involve the direct supply of energy, PTP obligations allow a participant to buy the network flow from one location to another.[16] When coupled with a self-scheduled generating resource, the PTP obligation allows a participant to serve its load while avoiding the associated real-time congestion costs between the locations. PTP obligations are

[16] PTP obligations are equivalent to scheduling virtual supply at one location and virtual load at another.

also scheduled by financial participants seeking to arbitrage locational congestion differences between the day-ahead and real-time markets.

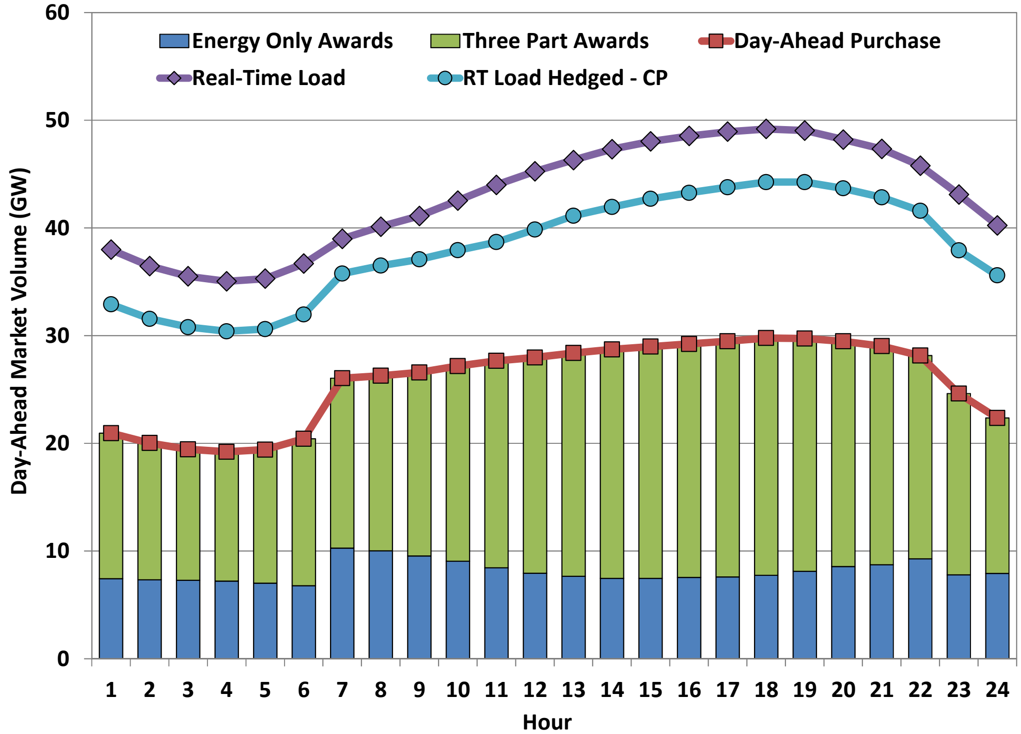
Figure 27: Volume of Day-Ahead Market Activity by Month



Real-time load in ERCOT may be hedged through the day-ahead market, either by purchasing energy in the market or by self-scheduling generation coupled with PTP "transfers" to the load. To estimate the volume of hedging activity, energy purchases are added to the volume of PTPs scheduled by Qualified Scheduling Entities (QSEs) with load that source or sink in Load Zones. This total is shown as the "Real-Time Load Hedged" shown in Figure 27 above. Under this methodology, approximately 64% of QSEs' real-time load was determined to be hedged in the day-ahead market, a decrease from 2017 when 82% of QSEs load was determined to be hedged. Although QSEs are the party financially responsible to ERCOT, their financial obligations may be aggregated and held by a counterparty. When measured at the counterparty level, the percentage of real-time load hedged increased to 89%, similar to the amount seen in 2017.

Figure 28 below presents the same day-ahead market activity data summarized by hour of the day. In this figure, the volume of day-ahead market transactions is disproportionate with load levels between the hours of 7 and 22 (hour ending). Because these times align with common bilateral and financial market transaction terms, the results in this figure are consistent with market participants using the day-ahead market to trade around those positions.

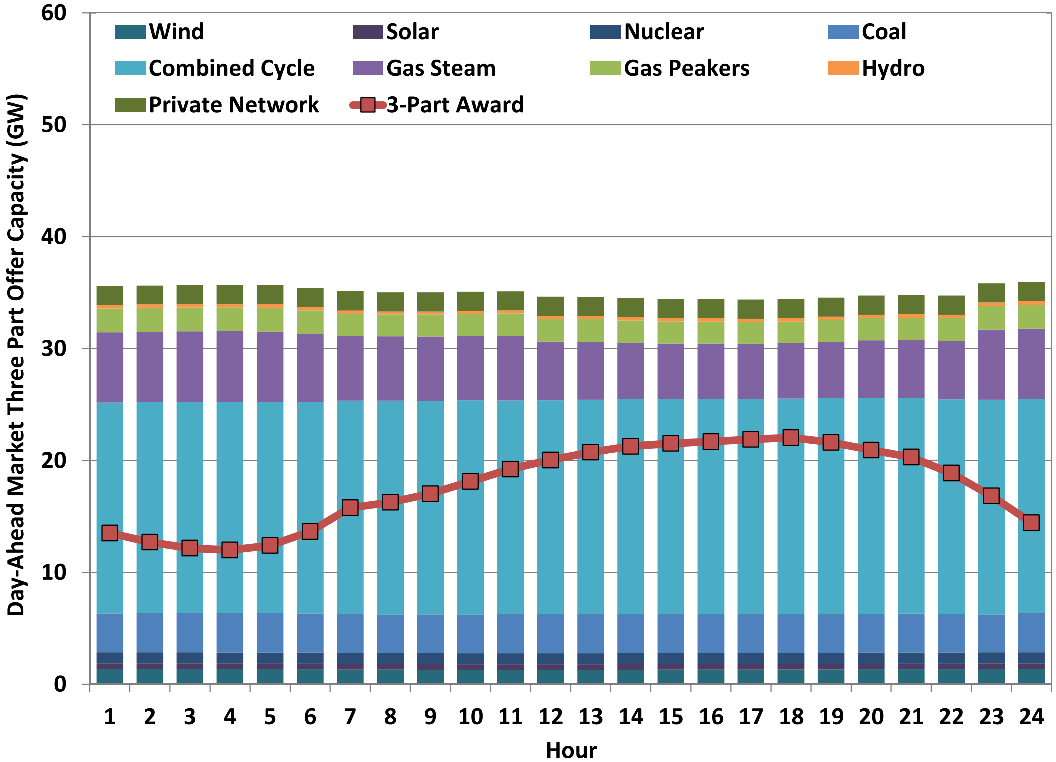
Figure 28: Volume of Day-Ahead Market Activity by Hour



The previous two figures showed that the volume of three-part offers comprised a small part of day-ahead market clearing. To determine whether this was due to small volumes of three-part offers being submitted, the following analysis was performed.

Figure 29 below shows the total capacity from three-part offers submitted in the day-ahead market for 2018. The submitted capacity has been averaged for each month and is shown to be significantly more than the amount of capacity cleared. With the largest share of installed capacity, it follows that combined cycle units are the predominant type of generation submitting offers in the day-ahead market. More importantly, because combined cycle units are most typically marginal units, offering that capacity into the day-ahead market allows a market participant to determine whether its unit is economic.

Figure 29: Day-Ahead Market Three-Part Offer Capacity

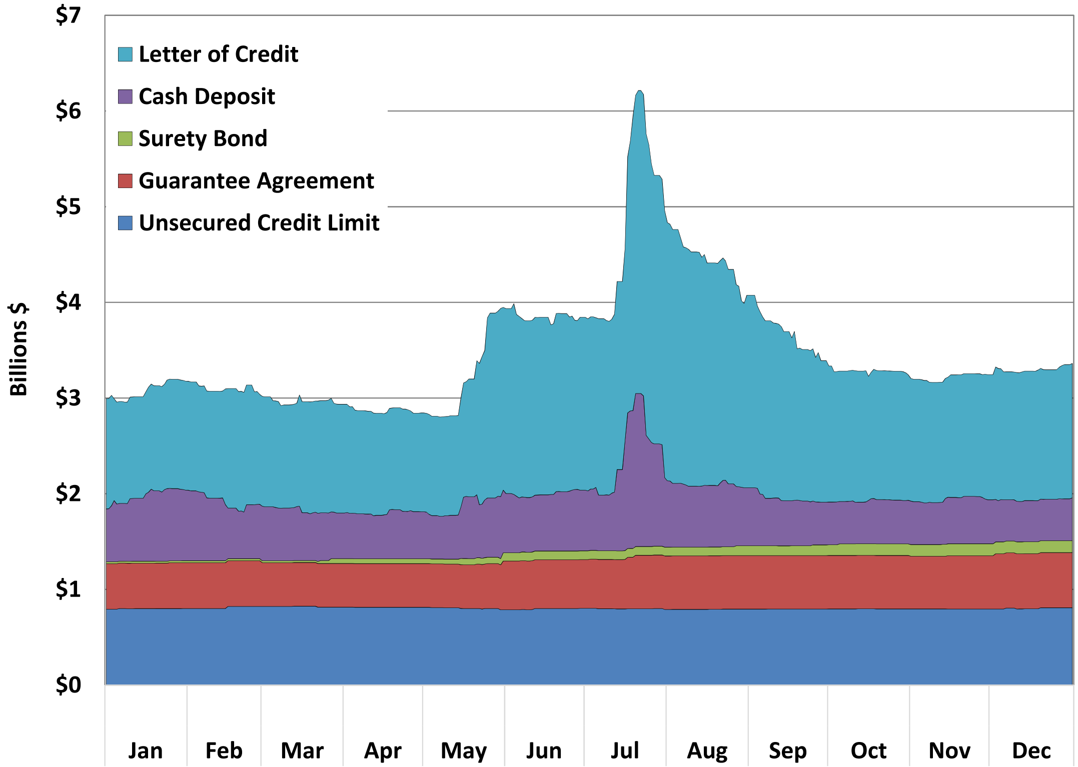


Parties that wish to participate in ERCOT's day-ahead market are required to have sufficient collateral with ERCOT. ERCOT's determination of collateral requirements underwent significant revision in 2018. Previously, ERCOT determined a party's credit exposure based entirely on historical prices. In 2018, ERCOT introduced forward prices as a determinant in calculating collateral requirements.[17] With smaller installed reserves, forward prices were especially high for the summer months of 2018. The effect that forward prices had on the total collateral held by ERCOT throughout the year was quite significant, as shown below in Figure 30.

---

17 NPRR800: *Revisions to Credit Exposure Calculations to Use Electricity Futures Market Prices*

Figure 30: Daily Collateral



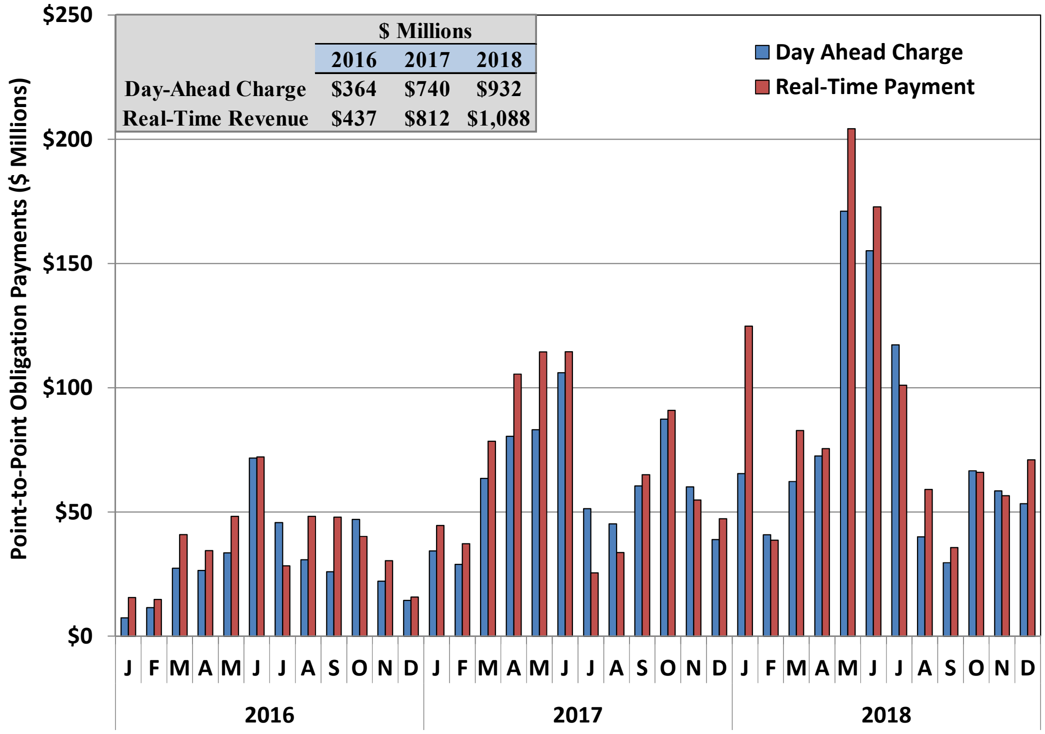
C. Point-to-Point Obligations

Purchases of PTP obligations comprise a significant portion of day-ahead market activity. They are similar to, and can be used to complement Congestion Revenue Rights (CRRs). CRRs, as more fully described in Section III: Transmission Congestion and Congestion Revenue Rights, are acquired via monthly and annual auctions and allocations. CRRs accrue value to their owner based on locational price differences as determined by the day-ahead market.

Participants buy PTP obligations by paying the difference in prices between two locations in the day-ahead market. The holder of the PTP obligation then receives the difference in prices between the same two locations in the real-time market. Hence, a participant that owns a CRR can use its CRR proceeds from the day-ahead market to buy a PTP obligation between the same two points in order to transfer its hedge to real-time. Because PTP obligations represent such a substantial portion of the transactions in the day-ahead market, additional details about the volume and profitability of these PTP obligations are provided in this subsection.

The first analysis of this subsection, shown in Figure 31, compares the total day-ahead payments made to acquire these products, with the total amount of revenue received by the owners of PTP obligations in the real-time market. As prices and total congestion costs have significantly increased in recent years, so have the costs and revenues associated with PTP obligations.

Figure 31: Point-to-Point Obligation Charges and Revenues



As in prior years, the aggregated total revenues received by PTP obligation owners in 2018 was greater than the amount charged to the owners to acquire them. This indicates that, in aggregate, buyers of PTP obligation profited from the transactions. This occurs when real-time congestion is greater than day-ahead market congestion. In eight of twelve months, as well as for the year in total, the acquisition charges were less than the revenues received, implying that expectations of congestion as evidenced by day-ahead purchases were less than the actual congestion that occurred in real-time. During February, July, October, and November these expectations were reversed, as congestion anticipated in the day-ahead market did not materialize in real-time.

The payments made to PTP obligation owners come from real-time congestion rent. The sufficiency of real-time congestion rent to cover both PTP obligations and payments to owners of CRRs who elect to receive payment based on real-time prices are assessed in Section III: Transmission Congestion and Congestion Revenue Rights.

Figure 32: Point-to-Point Obligation Volume

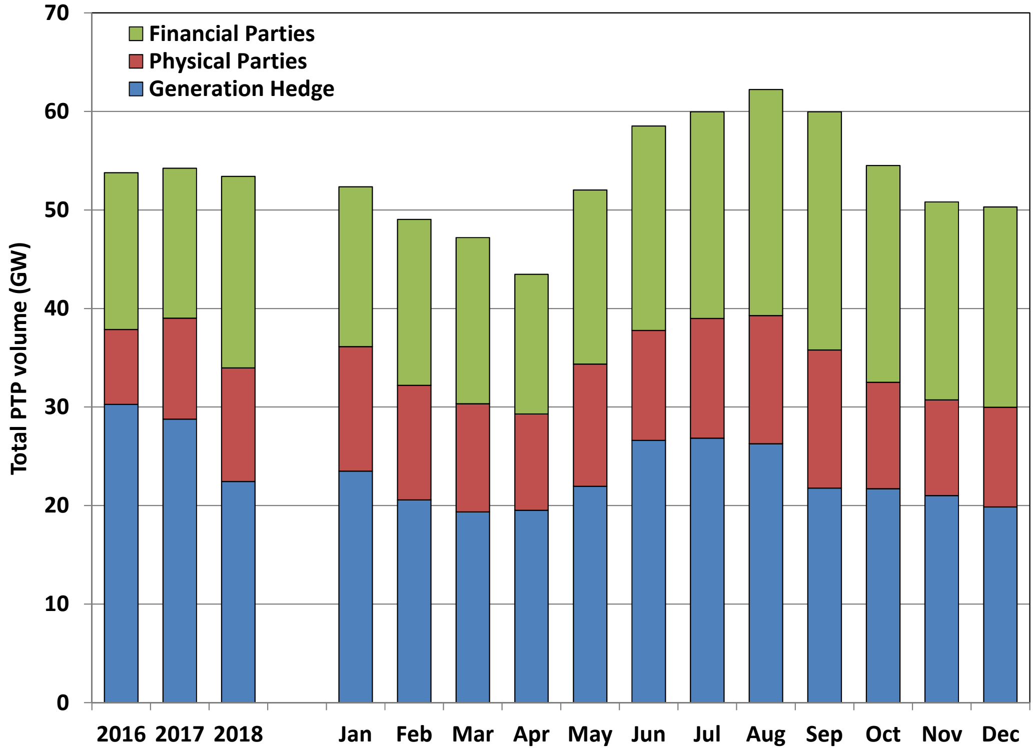
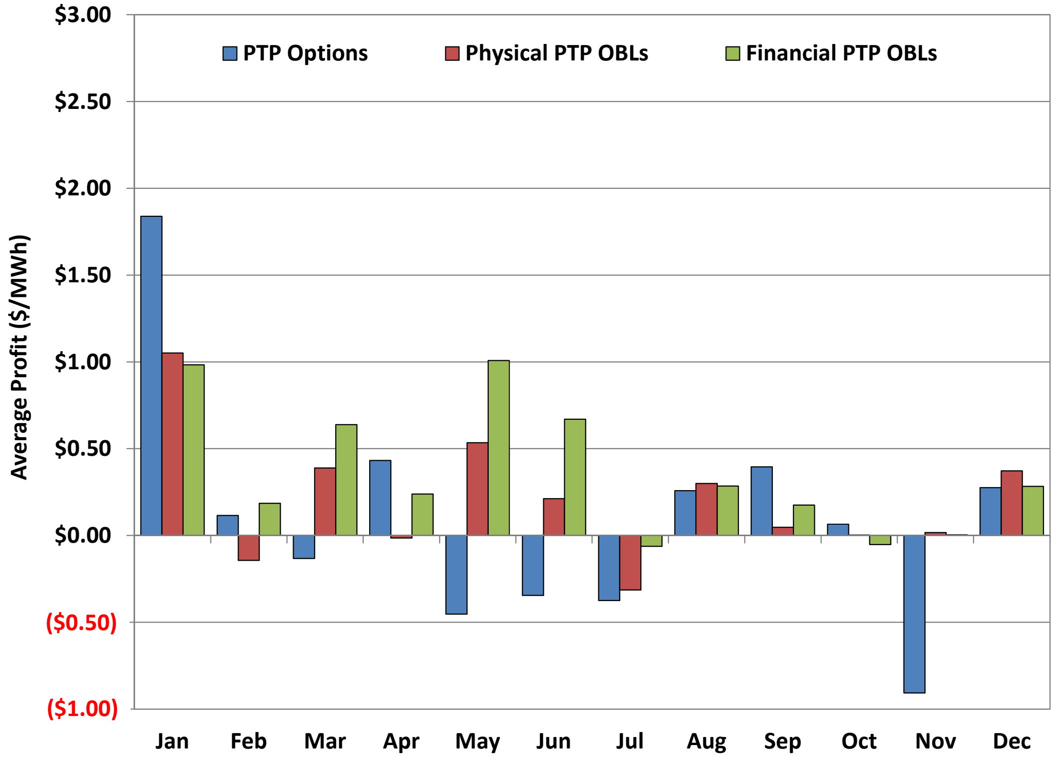


Figure 32 above presents the total volume of PTP obligation purchases divided into three categories. Different from Figure 27 and Figure 28 above, the volumes in this figure do not net out the injections and withdrawals occurring at the same location. Average purchase volumes are presented on both a monthly and annual basis. In contrast to the charges and revenues presented above in Figure 31, the total volume of PTP obligation purchases has been fairly stable for the past three years.

For all PTP obligations that source at a generator location, the capacity up to the actual generator output is considered to be hedging the real-time congestion associated with generating at that location. The figure above shows that generation hedging comprised most of the volume of PTP obligations purchased. The remaining volumes of PTP obligations are not directly linked to a physical position and are assumed to be purchased primarily to arbitrage anticipated price differences between two locations. This arbitrage activity is further separated by type of market participant. Physical parties are those that have actual real-time load or generation, whereas financial parties have neither. Financial parties purchased 36% of the total volume of PTP obligations in 2018, an increase from 28% in 2017 and 30% in both 2016 and 2015.

To the extent the price difference between the source and sink of a PTP obligation is greater in real-time than it was in the day-ahead market, the owner will profit. Conversely, if the price difference does not materialize in real-time, the PTP obligation may be unprofitable. The profitability of PTP obligation holdings for all physical parties and financial parties are compared in Figure 33. Also shown is the profitability of "PTP obligations settled as options," which are instruments available only to Non-Opt-In Entities, shown below as "PTP Options".

Figure 33: Average Profitability of Point-to-Point Obligations



This analysis shows that in aggregate, PTP obligation transactions in 2018 were profitable for the year, yielding an average profit of $0.24 per MWh, double the average profit of $0.12 per MWh from 2017. PTP obligations were profitable during 2018 for all types of parties, with average profits of $0.21 per MWh for physical parties, $0.34 per MWh for financial parties, and $0.09 per MWh for PTP obligations settled as options.

D. Ancillary Services Market

The primary ancillary services are regulation up, regulation down, responsive reserves, and non-spinning reserves. Market participants may self-schedule ancillary services or have them purchased on their behalf by ERCOT. In general, the purpose of responsive and non-spinning reserves is to protect the system against unforeseen contingencies (*e.g.*, unplanned generator outages, load forecast error, wind forecast error), rather than for meeting normal load fluctuations. ERCOT procures responsive reserves to ensure that the system frequency can quickly be restored to appropriate levels after a sudden, unplanned outage of generation capacity. Non-spinning reserves are provided from slower responding generation capacity, and can be deployed alone, or to restore responsive reserve capacity. Regulation reserves are capacity that responds every four seconds, either increasing or decreasing as necessary to fill the gap between energy deployments and actual system load.

Since June 1, 2015, ERCOT has calculated the requirement for responsive reserves based on a variable hourly need. This requirement is determined and posted in advance for the year. ERCOT procures non-spinning reserves such that the combination of non-spinning reserves and regulation up will cover 95% of the calculated Net Load forecast error. ERCOT will always procure a minimum quantity of non-spinning reserves greater than or equal to the largest generation unit during on-peak hours.

Regulation is used to fill the gap between generation dispatched in each interval and the actual moment-to-moment requirements for generation. To the extent the generation dispatched each interval can better reflect expected moment-to-moment changes, the quantity of regulation deployed can be minimized. The determination of the quantity of generation to be dispatched is affected by the 5-minute short-term load forecast, the deployment of regulation in the previous interval and estimates of actual unit ramping capabilities, factors which are routinely evaluated and adjusted. At the very end of 2018, ERCOT began including a new factor in the determination of generation to be dispatched, the short-term (5-minute) wind forecast. Including this factor should serve to improve ERCOT's dispatch of other generators and improve the efficiency of regulation deployments.

After several years of declining quantities, total requirements for ancillary services increased in 2018. The average total requirement in 2018 was greater than 4,900 MW, an increase from the average total requirement of approximately 4,800 MW in 2017 and roughly equal to the 2016 requirements. The principal cause for the overall increase in ancillary services requirements in 2018 was a larger responsive reserve requirement aimed at ensuring adequate online system inertia.

Figure 34 displays the hourly average quantities of ancillary services procured for each month in 2018.

Figure 34: Hourly Average Ancillary Service Capacity by Month

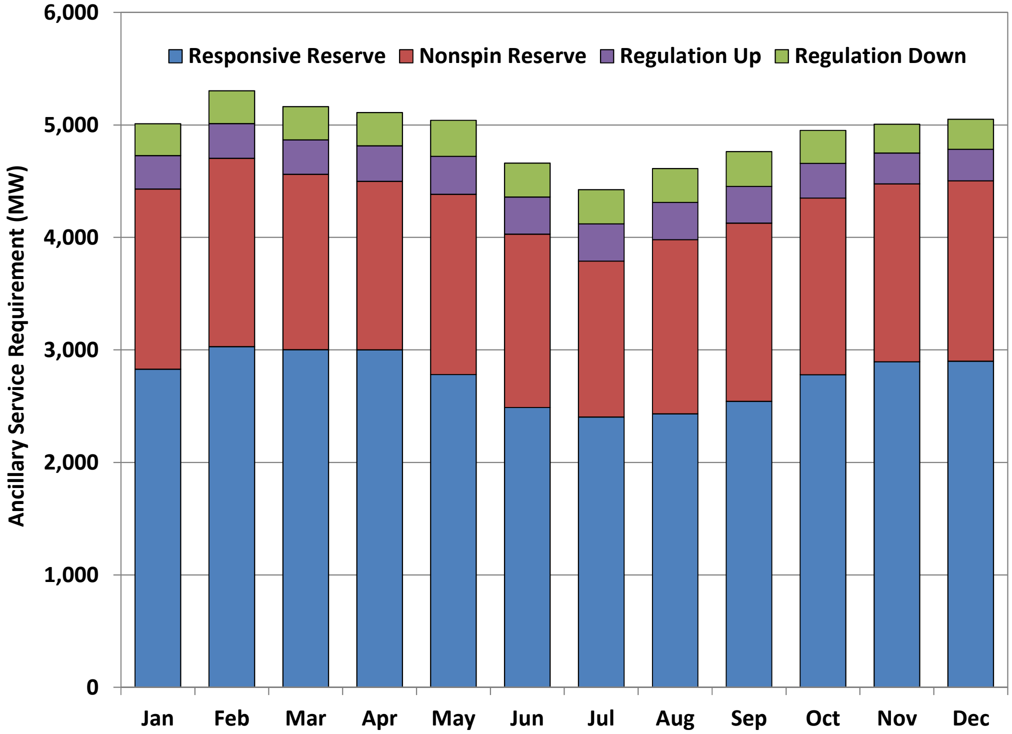


Figure 35 presents an alternate view of ancillary service requirements, displaying them by hour, averaged over the year. In this view the large variation in quantities between some adjacent hours were readily apparent. For example, capacity requirements increased more than 500 MW in hour 7, decreased 244 MW in hour 8 and gradually increased for the next two hours. Hour 22 provided another example of an increase in requirements in the hour prior to a decrease. This pattern was a result of the methodology which sets responsive and non-spinning reserve quantities in four hour blocks, while regulation reserve quantities are set hourly. Although the current ancillary service procurement methodology minimizes the quantities required, smoothing out these discontinuities may reduce or eliminate the occasional ancillary service price spikes.

Figure 35: Yearly Average Ancillary Service Capacity by Hour

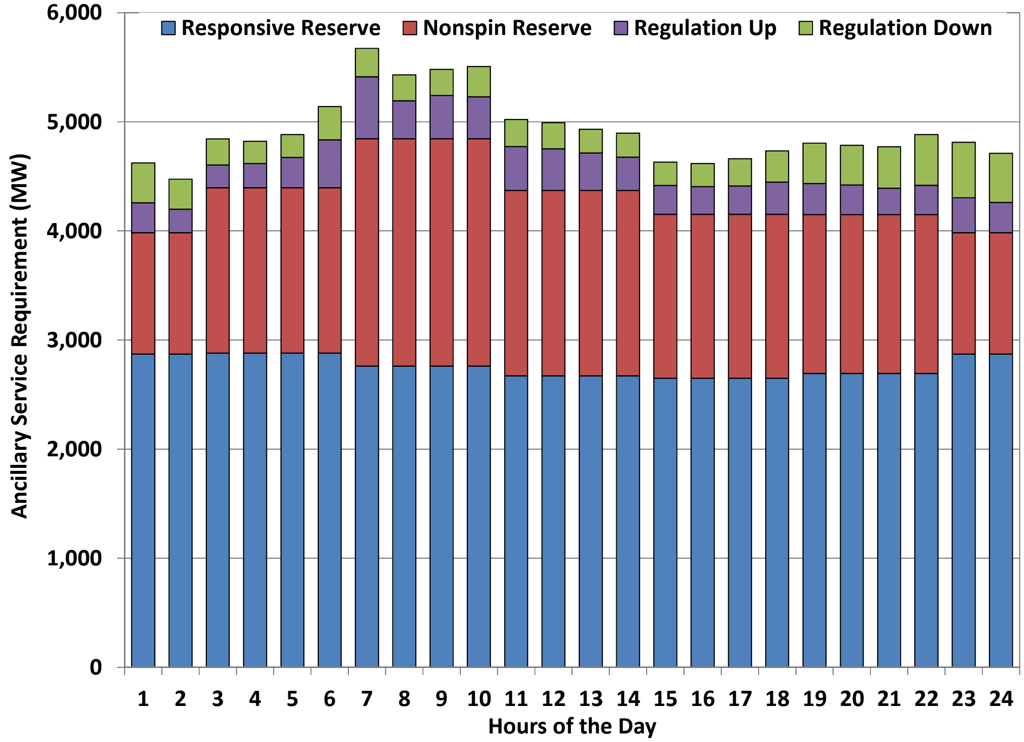
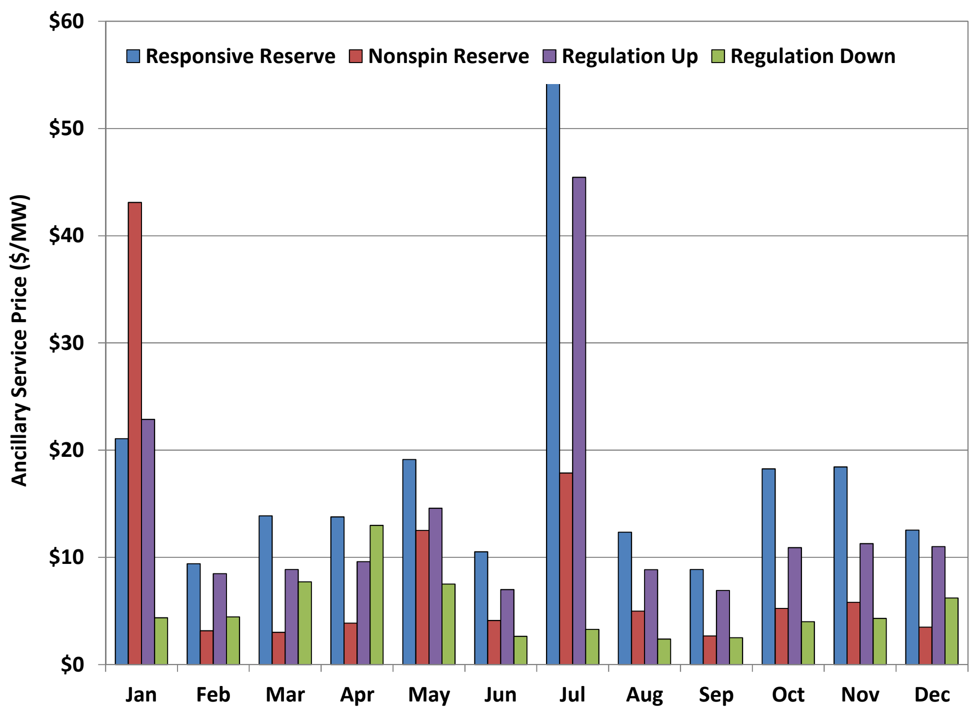


Figure 36 below presents the average clearing prices of capacity for the four ancillary services.

Figure 36: Ancillary Service Prices



The prices for ancillary service were noticeable higher in the months of January and July. These outcomes are consistent with the higher clearing prices for energy in the day-ahead market for those two months because ancillary services and energy are co-optimized in the day-ahead market. This means that market participants need not include expectations of forgone energy sales in their ancillary service capacity offers. Because ancillary service clearing prices explicitly account for the opportunity costs of selling energy in the day-ahead market, ancillary service prices should generally be correlated with day-ahead energy prices.

The high average clearing price for non-spin in January was due to three hours of very high clearing prices that occurred on January 17, during a multi-day period of extreme cold and precipitation. If higher quality ancillary services were allowed to be substituted for lower quality services, non-spin clearing prices for the three hours, and therefore the monthly average, would have been significantly lower.

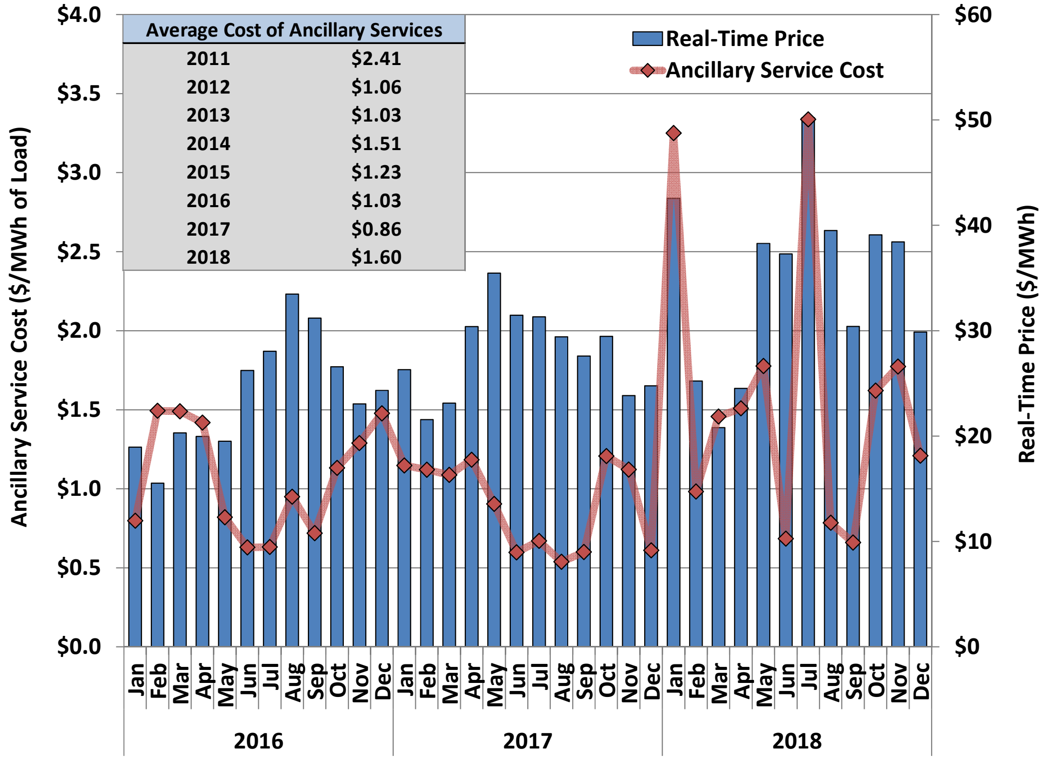
Table 5 compares the average annual price for each ancillary service in 2018 with 2017. In 2018, higher prices for the "up" ancillary services, responsive, non-spin and up regulation, are explained by the combination of larger requirements, and expectations for high energy prices as evidenced by high day-ahead energy prices. The decrease in the average price of down regulation is explained by lower opportunity costs of providing that service due to more capacity on line to meet the higher load requirements in 2018.

Table 5: Average Annual Ancillary Service Prices by Service

	2017 ($/MWh)	2018 ($/MWh)
Responsive Reserve	**$9.77**	**$17.64**
Nonspin Reserve	**$3.18**	**$9.20**
Regulation Up	**$8.76**	**$14.03**
Regulation Down	**$7.48**	**$5.19**

In contrast to the individual ancillary service prices, Figure 37 shows the monthly total ancillary service costs per MWh of ERCOT load and the average real-time energy price for 2016 through 2018.

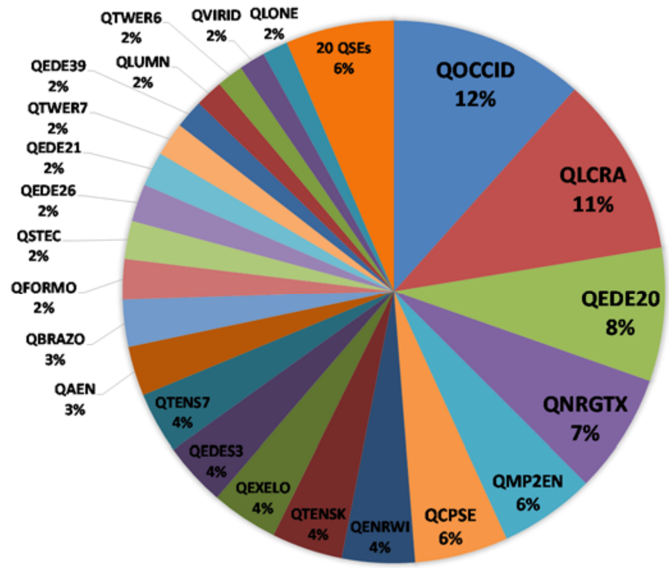
Figure 37: Ancillary Service Costs per MWh of Load



The average ancillary service cost per MWh of load increased to $1.60 per MWh in 2018 compared to the all-time low of $0.86 per MWh in 2017. Total ancillary service costs were 4.5% of the load-weighted average energy price in 2018, compared to 3.0 % in 2017, reversing the downward trend seen since 2015 when they went from 4.6 % and to 4.2 % in 2016.

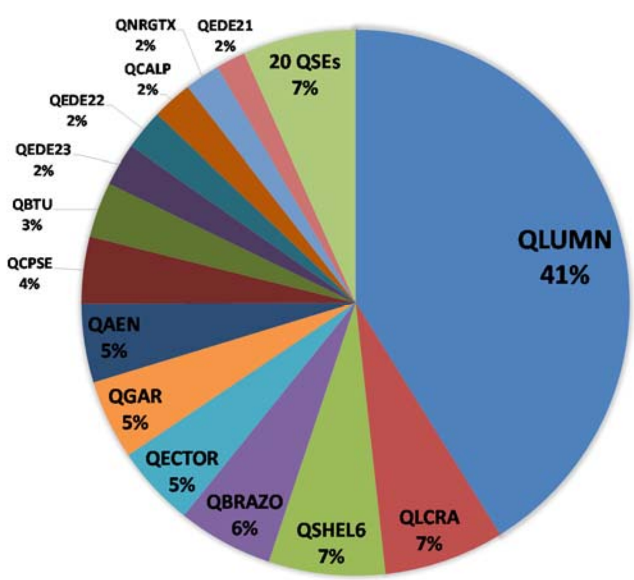
Responsive reserve service is the largest quantity purchased and typically the highest priced ancillary service product. Figure 38 below shows the share of the 2018 annual responsive reserve responsibility including both load and generation, displayed by QSE. During 2018, 43 different QSEs self-arranged or were awarded responsive reserves as part of the day-ahead market. The number of providers has been roughly the same for the past four years (45 in 2017, 42 in 2016, and 46 in 2015). There were no significant changes from 2017 in the largest providers or in the share of responsive reserve provided.

Figure 38: Responsive Reserve Providers



In contrast, Figure 39 below shows that the provision of non-spinning reserves is much more concentrated, with a single QSE (Luminant) bearing more than 40% the total responsibility. Luminant's 40% share of non-spin responsibility dropped from the 56% share it held in 2017 and 47% in 2016. The change in composition of Luminant's generation fleet, due to merger and retirements, likely explains the reduction. As Luminant's non-spin responsibility decreased in 2018, many other suppliers increased their share slightly. The lone exception was the addition of Shell as a large provider, with a 7% share of 2018 total requirements.

Figure 39: Non-Spinning Reserve Providers



The ongoing concentration in the supply of non-spinning reserve highlights the importance of modifying the ERCOT ancillary service market design to include real-time co-optimization of energy and ancillary services. Jointly optimizing all products in each interval will allow the market to substitute its procurements between units on an interval-by-interval basis to minimize costs and set efficient prices. Additionally, it will allow higher quality reserves (e.g., responsive reserves) to be economically substituted for lower quality reserves (e.g., non-spinning reserves), perhaps distributing the responsibility to provide among more entities.

Figure 40: Regulation Up Reserve Providers

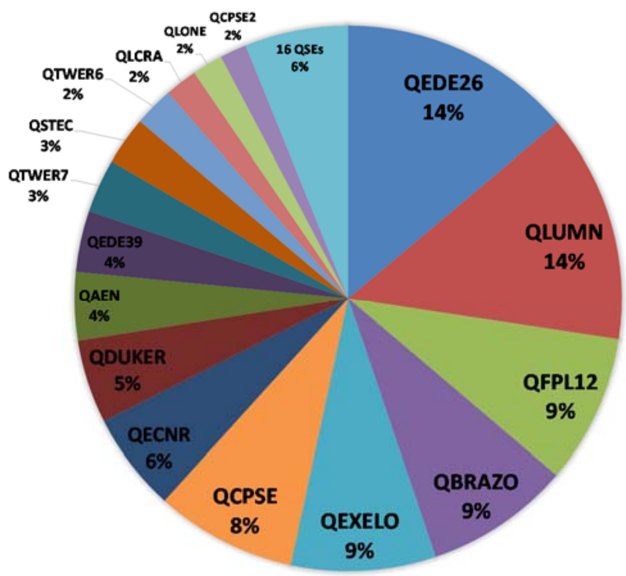
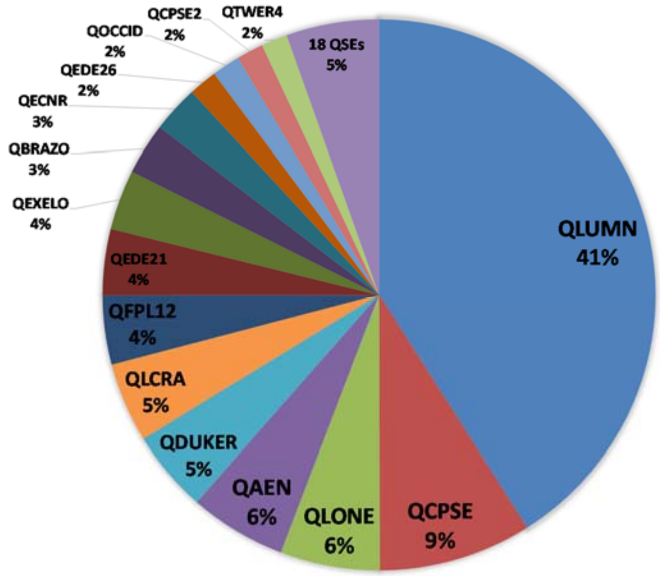


Figure 40 above shows the distribution for regulation up reserve service providers and Figure 41 shows the distribution for regulation down reserve providers. Figure 40 shows that regulation up was spread fairly evenly, similar to responsive reserve providers, while Figure 41 shows that that regulation down had similar concentration to non-spinning reserves in 2018. Again, Luminant had a dominant position in the provision of regulation down. Its 41% share of the regulation down responsibility in 2018 was higher than in recent years (25% in 2017 and 10% in 2016).

Figure 41: Regulation Down Reserve Providers



Ancillary service capacity is procured as part of the day-ahead market clearing. Between the time an ancillary service is procured and the time that it is needed, changes often occur that prompt a QSE to move all or part of its ancillary service responsibility from one unit to another. These changes may be due to a unit outage or to other changes in market conditions affecting unit commitment and dispatch. In short, QSEs with multiple units are continually reviewing and moving ancillary service requirements, presumably to improve the efficiency of ancillary service provision, at least from the QSE's perspective. Moving ancillary service responsibility is assumed to be in the QSE's self-interest. When all ancillary services are continually reviewed and adjusted in response to changing market conditions, the efficiencies will flow to all market participants and be greater than what can be achieved by QSEs acting individually.

The ERCOT market appropriately reflects the tradeoff between providing capacity for ancillary services versus providing energy in its co-optimized day-ahead market. Those same tradeoffs exist in real-time. Until comprehensive, market-wide co-optimization is implemented, the ERCOT market will continue to be subject to the choices of individual QSEs. These choices are likely to be in the QSE's best interest, and are not likely to lead to the most economic provision of energy and ancillary services for the market as a whole. Further, QSEs without large resource portfolios are effectively precluded from participating in ancillary service markets because of the replacement risk faced in having to rely on a supplemental ancillary services market (SASM). This replacement risk is substantial. Clearing prices for ancillary services procured in SASM are typically three to 40 times greater than annual average clearing prices from the day-ahead market.

ERCOT currently uses SASMs either to procure replacement ancillary service capacity when transmission constraints arise that make the capacity undeliverable, or when outages or limitations at a generating unit lead to failure to provide the ancillary service. A SASM may also be opened if ERCOT changes its ancillary service plan; this did not occur during 2018. A SASM was executed 21 times in 2018, with SASM awards providing 245 service-hours. SASMs were more frequent in 2018 than in 2017 when SASMs were executed 18 times, with SASM awards replacing 189 service-hours.

Figure 42 below provides the aggregate quantity of each service-hour that was procured via SASM. The volume of service-hours procured via SASM over the year (approximately 2,700 MWs of service-hours) is infinitesimal when compared to the total ancillary service requirement of nearly 43 million MWs of service-hours.

Figure 42: Ancillary Service Quantities Procured in SASM

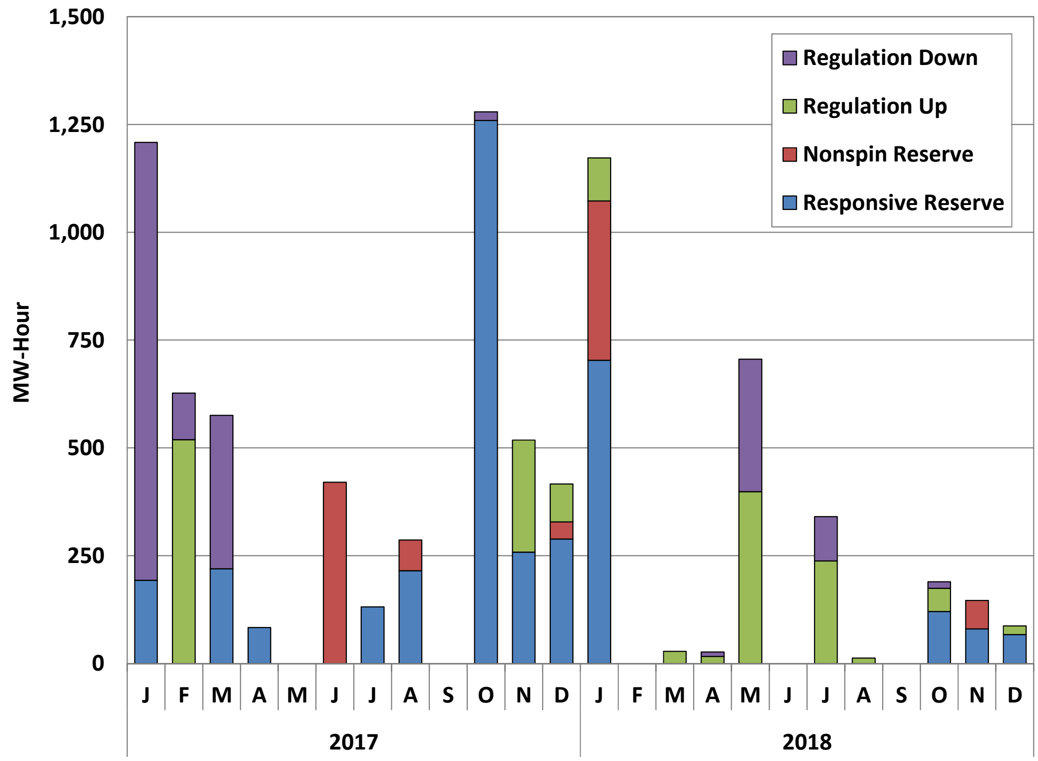
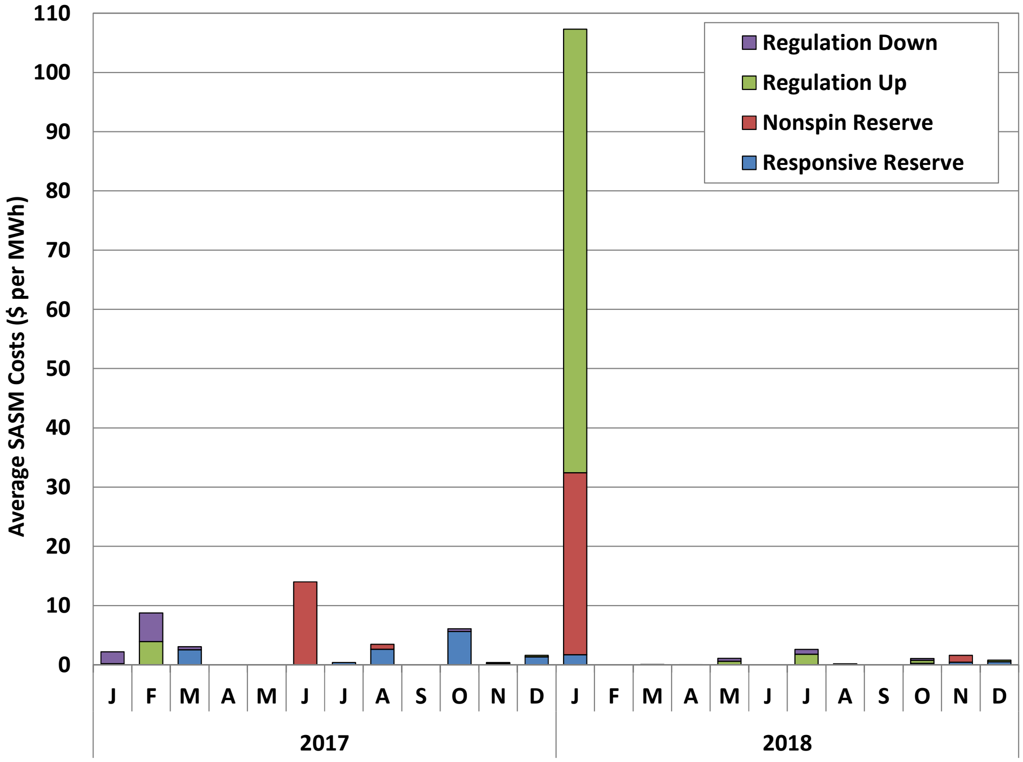


Figure 43 shows the average cost of the replacement ancillary services procured by SASM in 2018.

Figure 43: Average Costs of Procured SASM Ancillary Services



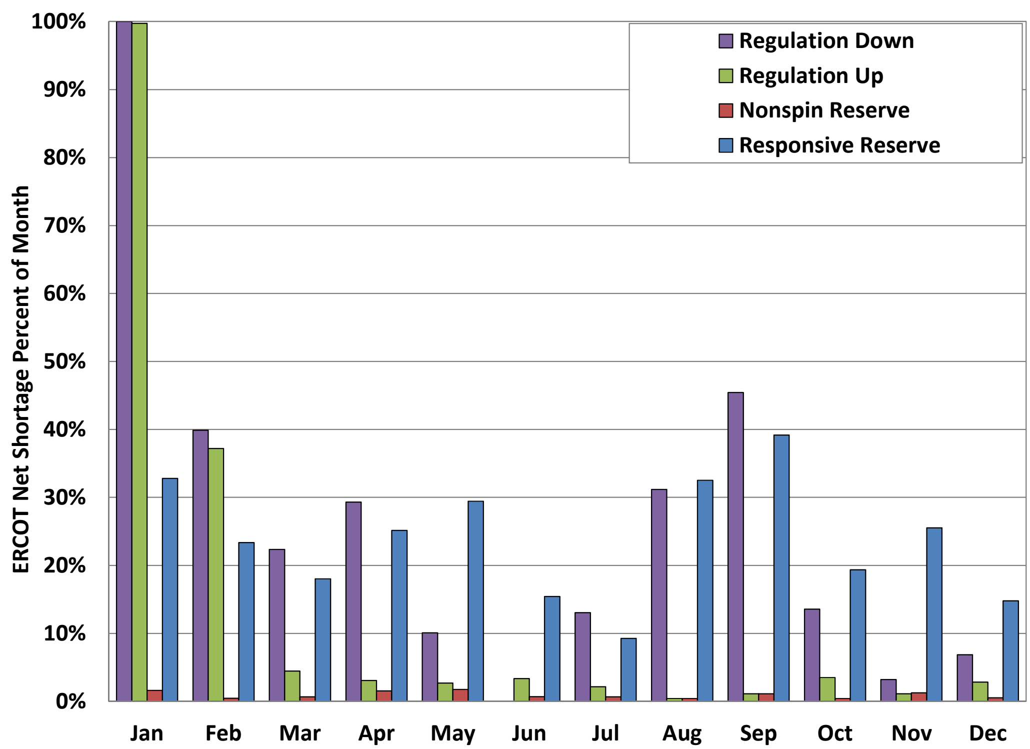
The opportunity exists for market participants to use the SASM process as a re-configuration market, or to move into or out of ancillary service positions awarded in the day-ahead market. SASMs were infrequent in 2018 largely because of the dearth of ancillary service offers typically available throughout the operating day, limiting re-configuration opportunities. The SASM procurement method, while offer-based, is inefficient and problematic and will be virtually eliminated when real-time co-optimization of energy and ancillary services is implemented.

Real-time co-optimization of energy and ancillary services will not require resources to estimate opportunity costs, will eliminate the need for the SASM mechanism, and allow ancillary services to be continually shifted to the most efficient provider. Because co-optimization allows the real-time market far more flexibility to procure energy and ancillary services from online resources, it will also reduce ERCOT's need to use RUC procedures to acquire ancillary services. The greatest benefit will be to effectively handle situations where entities that had day-ahead ancillary service awards were unable to fulfill that commitment, e.g. because of a generator forced outage. Thus, implementation of real-time co-optimization will provide benefits across the market in future years.

Because ancillary services are not currently co-optimized with energy in the SASM, potential suppliers are required to estimate opportunity costs rather than have the auction engine calculate those costs directly. As a result, resources that underestimate opportunity costs are inefficiently preferred over resources that overestimate opportunity costs. Further, the need to estimate the opportunity costs, which change constantly and significantly over time as the energy price changes, provides a strong disincentive to SASM participation, contributing to the observed lack of SASM offers. The paucity of SASM offers frequently leaves ERCOT with two choices in response to ancillary service un-deliverability or failure to provide: (1) use an out-of-market ancillary service procurement action with its inherent inefficiencies; or (2) operate with a deficiency of ancillary services with its inherent increased reliability risk.

In addition to its other weaknesses, a SASM is only useful for replacing ancillary services as part of a forward-looking view of the grid conditions. However, there are instances where the system is short ancillary services in real-time as per the resource details telemetered to ERCOT. Figure 44 depicts the percentage of hours in each month of 2018 where there was an ERCOT-wide shortage in the respective ancillary service. For this analysis, a shortage is defined as greater than 0.1 MW of obligation not being provided for at least 15 minutes out of an hour.

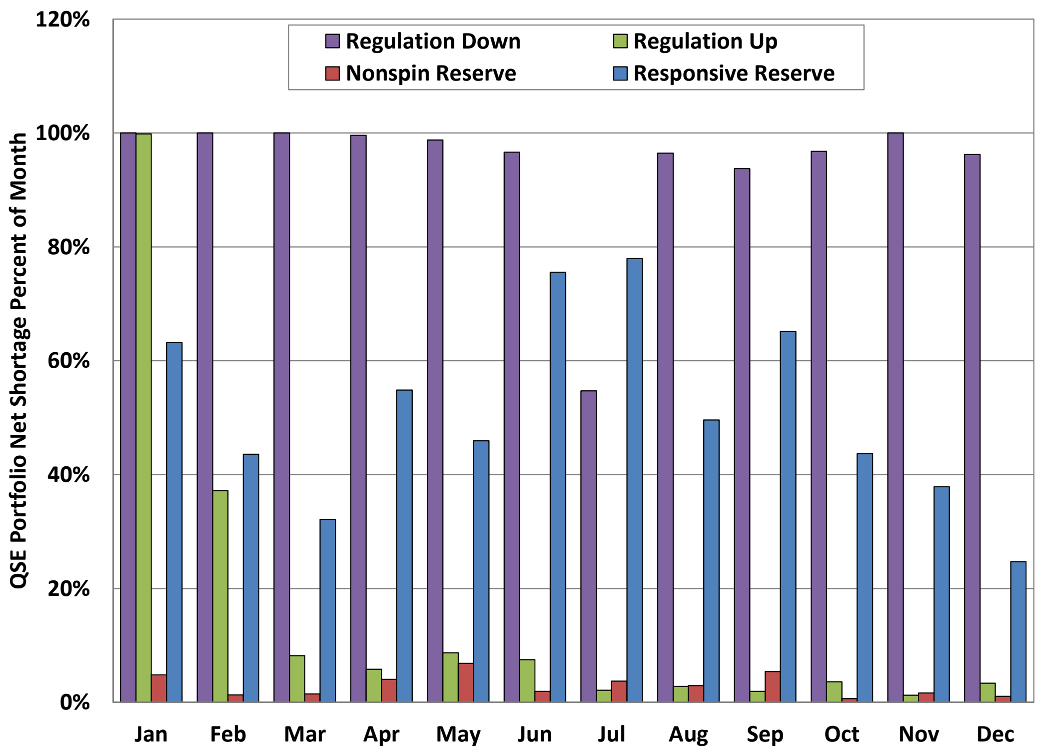
Figure 44: ERCOT-Wide Net Ancillary Service Shortages



This analysis shows that there was an ERCOT-wide shortage of both regulation down and regulation up services during nearly all of the hours in January. Again, this analysis is based on the telemetered status provided by the parties with the responsibility. Regulation up and down shortages continued into February in nearly 40% of the hours for both services. Shortages of responsive reserve occurred in more than 9% of the hours in every month, with the most frequent shortages occurring in September during almost 40% of the hours.

Figure 45 below shows the percentage of each month during which there was at least one QSE-portfolio that did not provide the full amount of ancillary service it had the responsibility to provide. As before, a shortage is defined as greater than 0.1 MW of obligation not being provided for at least 15 minutes out of an hour. QSE-specific shortages are much more frequent than ERCOT-wide shortages because a different QSE may be providing more ancillary services than they are obligated to provide, offsetting the deficit. This situation may occur if two QSEs do not register an ancillary service trade. As a result, one QSE would appear deficient, but ERCOT overall would not show a deficit.

Figure 45: QSE-Portfolio Net Ancillary Service Shortages



The seemingly pervasive deficiencies of QSEs meeting their ancillary service responsibilities is troubling from a market and possibly from a reliability perspective. Ancillary service awards are made from the day-ahead market and are for a certain amount of capacity for a one-hour duration. To date, there is not a mechanism in place to reduce payment for ancillary service awards in situations when the QSE has not fully met the award. While some explanations for the pervasive shortages shown in this analysis (such as the inability to register ancillary service trade agreements very close to real-time operations) would not cause a net ERCOT shortage of reserves and thus are of little reliability concern, the persistent net shortage shown in Figure 44 demonstrates the significance of this issue. One of the benefits of real-time co-optimization will be that responsibilities for providing ancillary services will be continually adjusted and updated in real-time, eliminating the need for SASMs and for registering individual trades.

III. Transmission Congestion and Congestion Revenue Rights

One of the most important functions of any electricity market is to manage the flows of power on the transmission network by not allowing additional power flow on transmission facilities that have reached their operating limits. Actions taken to ensure operating limits are not violated are called congestion management. The effect of congestion management is to change the output level of one or more generators to reduce the amount of electricity flowing on any transmission facility nearing its operating limit. This leads to higher costs. This increase in more expensive generation and decrease in less expensive generation results in different prices at different nodes. The decision about which generator(s) will vary its output is based on the generator's energy offer curve and how much of its output will flow across the overloaded transmission element. This leads to the dispatch of the most efficient generation to reliably serve demand while providing locational marginal pricing reflective of the actions taken to ensure system security.

The locational difference in prices produced by congestion can provide challenges to parties that have transacted in long term power contracts; namely, if the production point (for a seller) or consumption point (for a purchaser) is different from the contracted delivery point, the party is subject to the risk that the prices will be different when settled. Congestion Revenue Rights (CRR) markets enable parties to purchase the rights to those price differences in seasonal and monthly blocks, and thus achieve some level of price certainty.

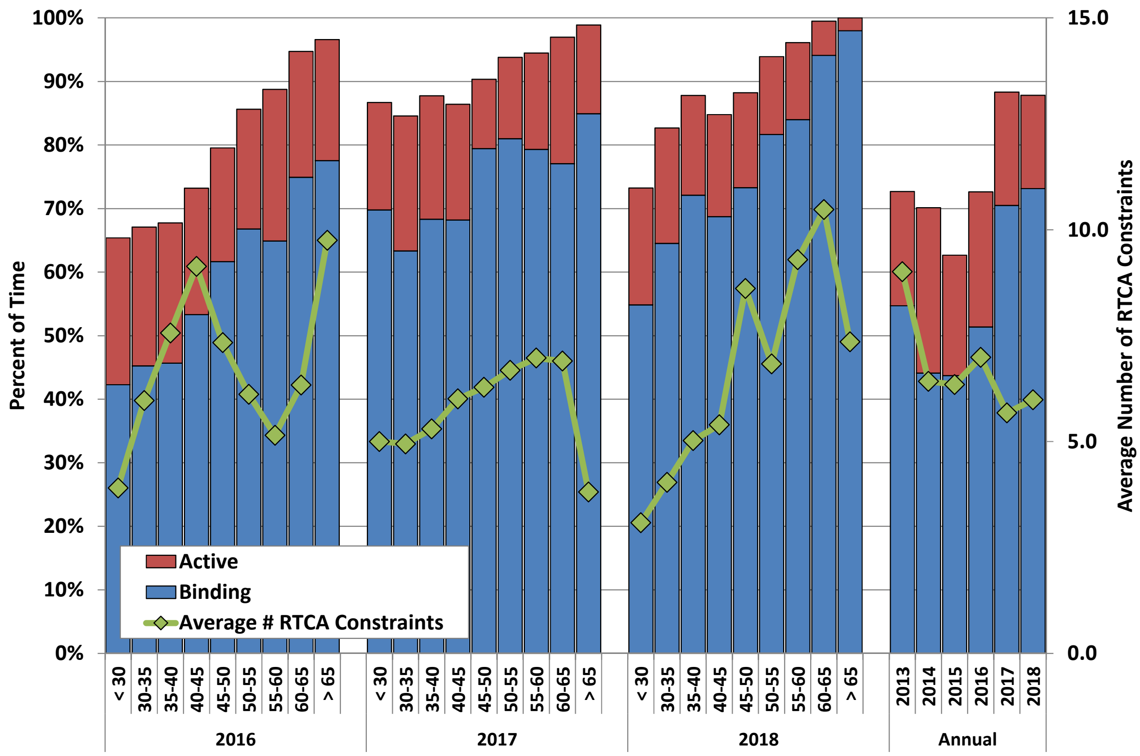
This section of the report summarizes transmission congestion in 2018, provides a review of the costs and frequency of transmission congestion in both the day-ahead and real-time markets, and concludes with a review of the activity in the CRR market.

A. Summary of Congestion

The total congestion costs experienced in the ERCOT real-time market in 2018 were $1.26 billion, a 30% increase from the 2017 value. A costly localized constraint in far west Texas was the primary cause of the increased congestion on the Yucca Switch to Basin 138 kV lines. This congestion was directly related to high loads associated with increased oil and gas production. Congestion occurred with roughly the same frequency in 2018 as in 2017, with only a slight increase in the number of binding intervals. All zones experienced increased congestion in 2018, though inter-zonal congestion decreased. The decrease in inter-zonal congestion can be attributed to the completion of the Houston transmission import project in April 2018.

Figure 46 provides a comparison of the amount of time transmission constraints were active and binding for various load levels from 2016 through 2018. This figure also indicates the average number of constraints in a Real-Time Contingency Analysis (RTCA) execution for each load level. RTCA is the process that evaluates the resulting flows on the transmission system under a large number of different contingency scenarios. A thermal constraint exists if the outage of a transmission element (contingency) results in a flow higher than the rating of a different element. Binding transmission constraints are those for which the dispatch levels of generating resources are altered in order to maintain transmission flows at reliable levels. The costs associated with this re-dispatch are the system's congestion value and are included in nodal prices. Active transmission constraints are those that the dispatch software evaluated, but did not require a re-dispatch of generation.

Figure 46: Frequency of Binding and Active Constraints



Binding constraints existed 73% of the time in 2018, a slight increase from 71% in 2017. Higher congestion costs in 2017 and 2018 are explained in part by the increased frequency of constraints. The pattern of constraints by load level was different in 2018 than in 2017. Constraints in 2018, similar to 2016, occurred more frequently during higher load levels. Real-time constraints were binding in 98% of intervals with load in excess of 65 GW.

Differences between 2018 and 2017 appear in the gradual increase for both the frequency of binding constraints and the average number of RTCA constraints as load levels increase. This difference is explained by the reduced amount of time with an active Generic Transmission Constraint (GTC), most noticeably at lower load levels. A GTC was active 33% of the time in 2018 compared to 43% in 2017. GTCs are not derived from RTCA, but rather are determined by off-line studies and their limits are typically determined prior to the operating day. GTCs are used to ensure that the generation dispatch does not violate a transient or voltage stability condition. Certain GTC limits are determined in real-time using Voltage Stability Assessment or

Transient Stability Assessment. Using these tools to continuously evaluate the North to Houston, Panhandle, Laredo, and the Rio Grande Valley Import limits provides a more accurate limit than what was could be determined as part of the day-ahead process.

Table 6 below shows the GTCs that were monitored in 2018.

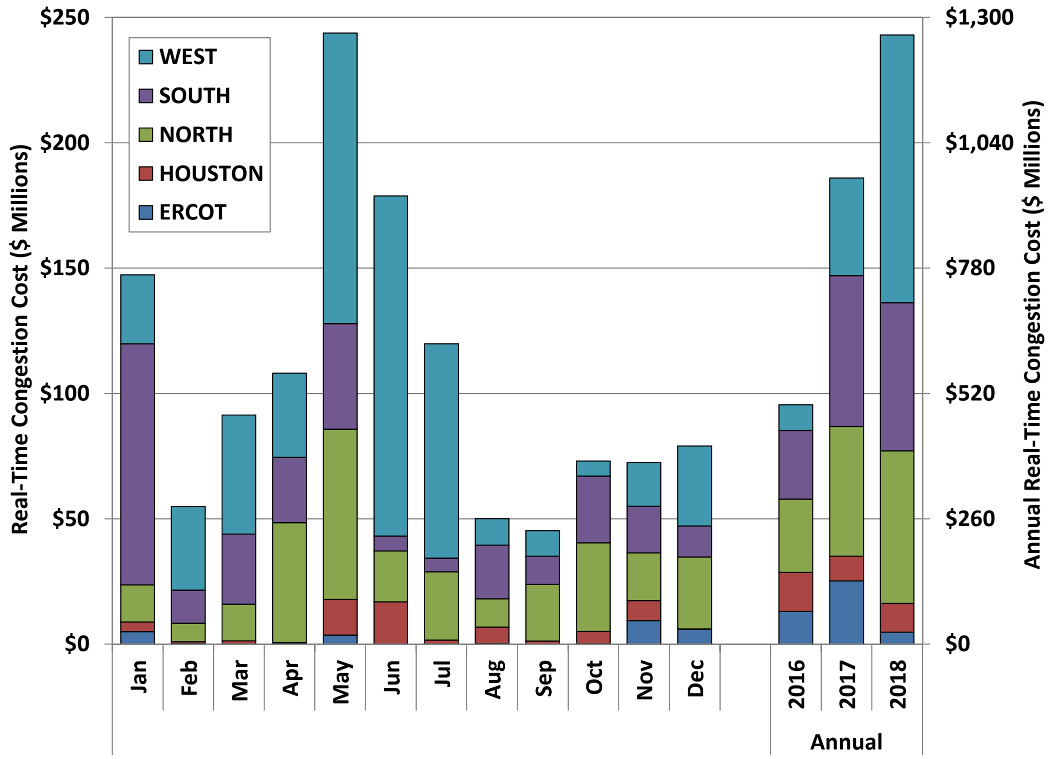
Table 6: Generic Transmission Constraints

Generic Transmission Constraint	Effective Date	# of Binding Intervals in 2018
North to Houston	December 1, 2010	7
Rio Grande Valley Import	December 1, 2010	92
Panhandle	July 31, 2015	14,660
Red Tap	August 29, 2016	25
North Edinburg - Lobo	August 24, 2017	0
Nelson Sharpe - Rio Hondo	October 30, 2017	1,900
East Texas	November 2, 2017	0
Treadwell	May 18, 2018	0
McCamey	March 26, 2018	0

With the exception of the North to Houston, Rio Grande Valley Import, and East Texas constraints, all GTCs resulted from issues identified during the generation interconnection process.

Figure 47 displays the amount of real-time congestion costs associated with each geographic zone, with the monthly values of 2018 preceding the annual values for the last three years. Costs associated with constraints that cross zonal boundaries (for example North to Houston) are shown in the "ERCOT" category.

Figure 47: Real-Time Congestion Costs



The months of February, August, and September exhibited the least amount of congestion costs, whereas the remaining months reflected much higher congestion. The 2018 monthly profile is relatively unexpected because shoulder months usually have the highest congestion costs. Shoulder months are when most transmission and generation outages for maintenance and upgrades occur. The increased congestion in January was due to cold weather conditions and higher than expected load conditions in the far west increased congestion costs in May, June and July.

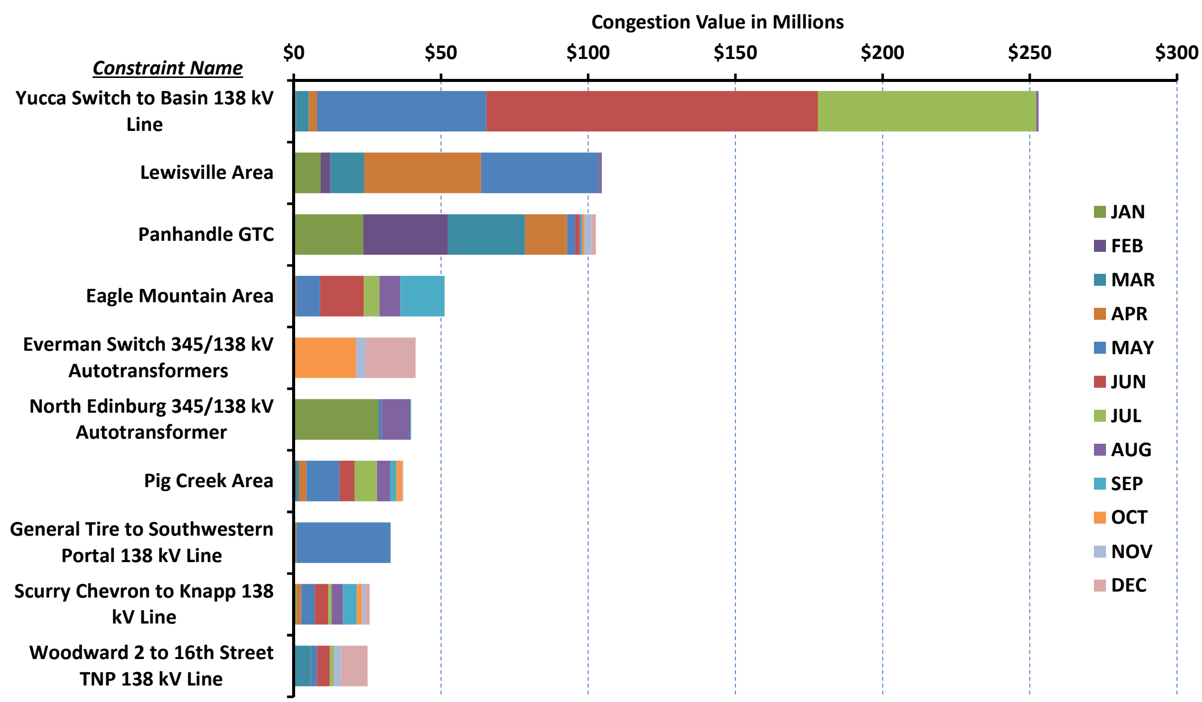
All zones experienced an increase in price impacts in 2018, whereas ERCOT cross-zonal congestion decreased. The North to Houston constraint was a significant contributor to ERCOT congestion in recent years, but the completion of the Houston transmission import project in April 2018 resulted in reduced congestion. The largest contributor to congestion costs in 2018 was the increased congestion in the West zone. This congestion was the result of the high load in the far west caused by oil and gas development activity in the Permian Basin. Transmission line outages related to the required inspection and repair of structures on facilities owned by Electric Transmission Texas (ETT) continued in 2018, and are expected to continue into the future. However, this activity had minimal impact on West zone congestion compared to the impacts of load growth in the Permian Basin.

B. Real-Time Constraints

The review of real-time congestion begins with describing the areas with the highest financial impact from congestion. For this discussion, a congested area is determined by consolidating multiple real-time transmission constraints that are determined to be similar because of geographic proximity and constraint direction. There were 443 unique constraints that were binding at some point during 2018, with a median financial impact of approximately $177,000. In 2017, there were 399 unique constraints with a median financial impact of $235,000. The location of the constraints around load areas contributed to higher congestion costs in 2018.

Figure 48 displays the ten most costly real-time constraints as measured by congestion value.

Figure 48: Most Costly Real-Time Constraints



The constraint with the highest congestion cost in 2018 at $253 million was on a series of two 138 kV lines, the first connecting Yucca Switch and Gas Pad, and the second from Gas Pad and terminating at the Basin substation. The majority of the congestion value was generated on the line between Yucca Switch to Gas Pad. The Gas Pad to Basin segment only incurred $1 million of congestion cost. The congestion cost associated with Yucca Switch to Gas Pad in 2018 was almost double the most costly constraint in 2017. This constraint and the seventh most valued constraint on this list, the Pig Creek area, are discussed in detail later in the Permian Basin subsection.

The second most costly constraint in 2018 was the Lewisville area, which is north of Dallas-Fort Worth. The components of this constraint include the Lakepoint to Carrollton Northwest, the West TNP to TI TNP, and the Lewisville to Jones Street TNP 138 kV lines. The congestion costs for these constraints more than doubled from $40 million in 2017 to $105 million in 2018. This was due to recent transmission system additions enabling more export capability from the Panhandle constraint. The increased exports from the Panhandle led to more congestion in the DFW area. The retirement of five coal units at Big Brown and Monticello in east Texas also contributed to the increased congestion seen in this area.

The Panhandle constraint dropped from the most costly constraint in 2017 to third in 2018. The Panhandle constraint caused $102 million of congestion in 2018, a 30% decrease from $139 million in 2017. By the end of 2018, there was almost 5 GW of generation capacity in the Panhandle area, 90% of which was wind generation. The Panhandle GTC limit ranged in value from 2,100 MW to 4,300 MW during 2018. The GTC limit average was 3,500 MW, up by 400 MW from 3,100 MW in 2017, a 12% increase. This increase was attributable to the 345 kV installations and the two synchronous condensers, one located at Alibates and the other at Tule 345 kV substations in the Panhandle.

Maintenance activity performed by ETT on its transmission structures located in the Panhandle continued to reduce the Panhandle GTC limit, as work extended from 2017 into 2018. ETT continually monitors structures to find any additional damage and ETT has been providing updates to the market participants via the outage scheduler and ad hoc Congestion Management Working Group meetings. The average shadow price of the Panhandle GTC during binding intervals was $36 per MWh, a $2 per MWh increase from 2017, reflecting the difference between system-wide average price and negative prices from wind generation. The Panhandle GTC was active for 14% of the time in 2018, down from 16% in 2017.

Congestion in the Eagle Mountain area between Dallas and Fort-Worth was the fourth most costly constraint in 2018. ERCOT's *2017 Regional Transmission Plan* recommended transmission upgrades to this area to address the constraints of the Wagley Robertson to Blue Mound 138 kV line, the Wagley Robertson to Summerfield 138 kV line, and the Eagle Mountain to Morris Dido 138 kV line.[18] The same conditions that affected the Lewisville area also affected the Eagle Mountain area. In fact, the activation of constraints in the Lewisville area, Panhandle GTC, and the Eagle Mountain area all had the effect of dispatching wind output down and increasing the generation in the North.

The fifth, sixth, and eighth most costly constraints were due to planned outages or high loads in their respective areas. The Everman Switch 345/138 kV #1 and #2 autotransformers serve load

[18] ERCOT's *2017 Regional Transmission Plan* is available here: http://www.ercot.com/news/presentations/2017

into DFW, with congestion occurring October through December. The North Edinburg 345/138 kV autotransformer serves load into the Rio Valley, with congestion in months January and August. The eighth most costly constraint, the General Tire to Southwestern Portal 138 kV line, while associated with outages, was also affected by high wind generation. Congestion for this constraint occurred in May. The last two most costly constraints, Scurry Chevron to Knapp 138 kV line and Woodward 2 to 16th Street TNP 138 kV line, were due to issues with high wind output in the west and at times are increased by the ongoing transmission upgrades in the far west.

Permian Basin

The Permian Basin area in far west Texas has witnessed significant increases in load due to oil and gas development. The congestion related to this area was evident in the highest and seventh-highest valued constraints, Yucca to Basin 138 kV line and the Pig Creek area. The Gas Pad to Basin 138 kV line incurred $1 million of its share of Yucca to Basin 138 kV line constraint due to the shift in congestion from Yucca to Gas Pad 138 kV line after upgrades in the area were implemented. The upgrades were a result of an ERCOT Board of Directors approved project in October 2016 brought forward by AEP and Oncor to address the increasing load due to oil and gas activity. [19] While the project initially targeted summer 2019 for completion, the Yucca to Gas Pad 138 kV line portion of the transmission project was completed in July 2018. Construction of the project occurred while the lines were energized, indicating the importance of continuing to serve load in the area and limiting the impact of construction on the far west grid.

The Pig Creek area constraint consisted of three 138kV lines; Solstice to Pig Creek, Fort Stockton Plant to Linterna, and Linterna to Solstice. These constraints are in the far west region where transmission is sparse and congestion is in the opposite direction of the Yucca to Basin 138 kV line constraint. Roughly 900 MW of solar generation capacity is installed to the east of Pig Creek and was often limited by the Pig Creek area constraints in 2018. Congestion in this area was exacerbated by the operations of a small generator with output less than 10 MW. Under current requirements, the resource does not receive dispatch instructions from ERCOT. Further, because the small generator is paid a zonal price, the generator has no economic incentive to alter output to alleviate congestion.

Irresolvable Constraints

The shadow price of a constraint is the value at which economic dispatch results in profit-maximizing for the generators while also meeting demand at the lowest overall production cost. However, if the dispatch cannot resolve a reliability problem with the available generators, the shadow price would continue to increase as the economic dispatch sought a solution. In situations where there is no generation solution, the shadow price would theoretically rise to

[19] ..http://www.ercot.com/content/wcm/key_documents_lists/76340/11_AEPSC_and_Oncor_Barilla_Junction_Area_Improvement__RPG_Project.pdf

infinity. This would result in unreasonable prices and may even prevent the dispatch software from reaching a solution, therefore, the shadow price is capped.

Shadow price caps are based on a reviewed methodology,[20] and are intended to reflect the level of reduced reliability that occurs when a constraint is irresolvable. Prior to June 2018, the shadow price caps were $5,000 per MW for base-case (non-contingency) or voltage violations, $4,500 per MW for 345 kV constraints, $3,500 per MW for 138 kV, and $2,800 per MW for 69 kV thermal violations. GTCs are considered stability constraints either for voltage or transient conditions with a shadow price cap of $5,000 per MW. Effective June 20, 2018, the shadow price cap for base-case (non-contingency) or voltage violations, including the GTCs that track export capability, was changed to $9,251 per MW. The shadow price cap on these violations was raised to this level so that resources in the Panhandle and other areas with large impacts on a GTC would not be dispatched to levels that would jeopardize grid security, even in times of overall system shortage.

When ERCOT's dispatch software cannot find a dispatch combination to reduce the flows on the transmission element(s) of concern to a reliable level, the violated constraint will be priced at the shadow price cap. When this pricing becomes a chronic occurrence, a regional peaker net margin mechanism is applied such that once local price increases accumulate to a predefined threshold, the constraint is deemed irresolvable and the constraint's shadow price cap is re-evaluated. The shadow price cap is recalculated based upon the mitigated offer cap of existing resources and their ability to resolve the constraint.

20 ERCOT Business Practice Manual, Setting the Shadow Price Caps and Power Balance Penalties in Security Constrained Economic Dispatch (ERCOT Board Approved 6/12/18), available at http://www.ercot.com/mktrules/obd/obdlist.

Table 7: Irresolvable Constrained Elements

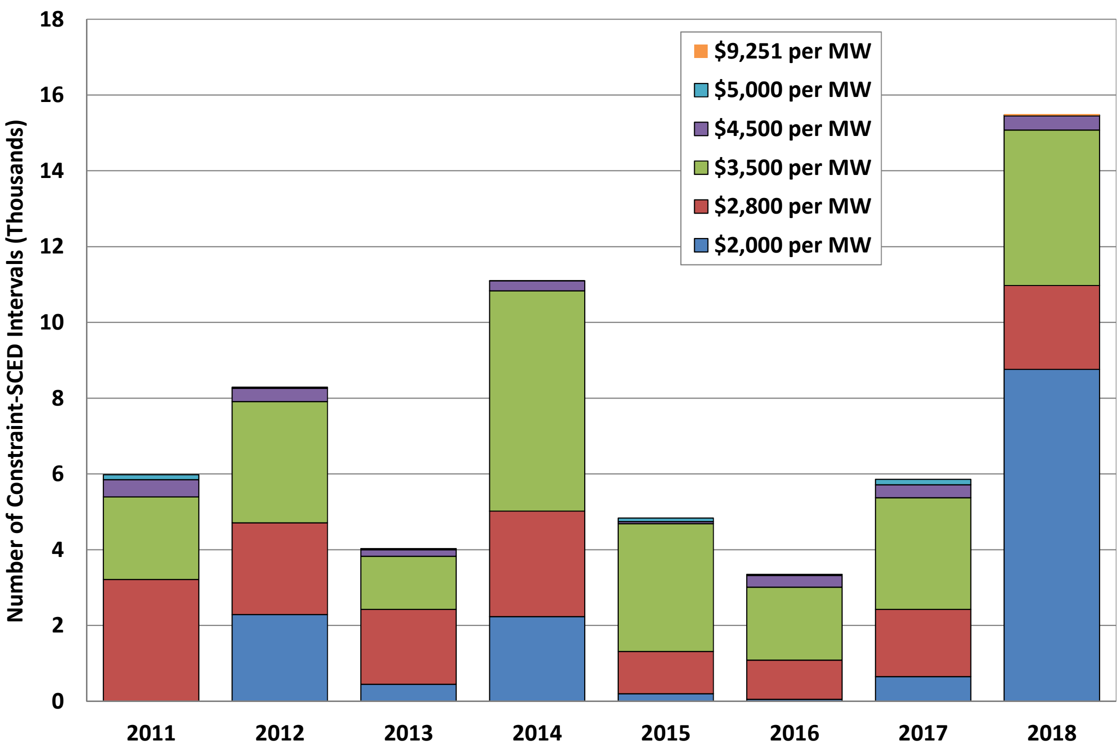
Irresolvable Element	Element Max Shadow Price	2018 Adjusted Max Shadow Price	Irresolvable Effective Date	Termination Date	Load Zone	# of Binding Intervals in 2018
Valley Import	$5,000	$2,000	1/1/12	-	South	92
Abilene Northwest to Ely Rea Tap 69 kV Line	$2,800	$2,000	9/26/14	1/30/18	West	0
Emma to Holt Switch 69 kV Line	$2,800	$2,000	10/27/14	-	West	0
San Angelo College Hills 138/69 kV Autotransformer	$3,500	$2,000	7/22/15	1/30/18	West	0
Fort Stockton Switch to Barilla 69 kV Line	$2,800	$2,800	1/1/18	-	West	2162
Moore to Hondo Creek Switching Station 138 kV Line	$3,500	$2,549	1/2/18	-	West	435
Wickett TNP to Winkler County 6 TNP 69 kV Line	$2,800	$2,000	4/9/18	-	West	480
Yucca Drive Switch – Gas Pad 138 kV line	$3,500	$2,000	5/4/18	-	West	9131
Yellow Jacket to Hext LCRA 69 kV line	$2,800	$2,000	5/18/18	-	West	541

As shown above in Table 7, nine elements were deemed irresolvable in 2018 and had a shadow price cap imposed according to the irresolvable constraint methodology. Two constraints, the Abilene Northwest to Ely Rea Tap 69 kV line and the San Angelo College Hills 138/69 kV Autotransformer, were deemed resolvable during ERCOT's annual review and were removed from the list. All irresolvable constraints are located in the West Load Zone with the exception of the Valley Import GTC which is located in the South Load Zone. The Fort Stockton Switch to Barilla 69 kV line constraint, located in far west Texas, was deemed irresolvable in January of 2018. The area was also impacted by the solar installations and Permian Basin load development.

While the constraint was deemed irresolvable, the shadow price cap was not lowered for 2018, so its irresolvable status had no impact. There is a future project planned to upgrade the 69 kV line to 138 kV identified in ERCOT's *2018 Constraints and Needs Report.*[21]

Because of shadow price caps, some constraints will be violated, as evidenced by the flow being greater than the limit of the constraint. In other words, it was not possible to resolve the constraint with the re-dispatch of available generation. Under these circumstances the shadow price will be equal to the designated maximum shadow price of the constraint. Figure 49 below shows the number of dispatch intervals a constraint reached its maximum shadow price for the years 2011 through 2018.

Figure 49: Frequency of Violated Constraints



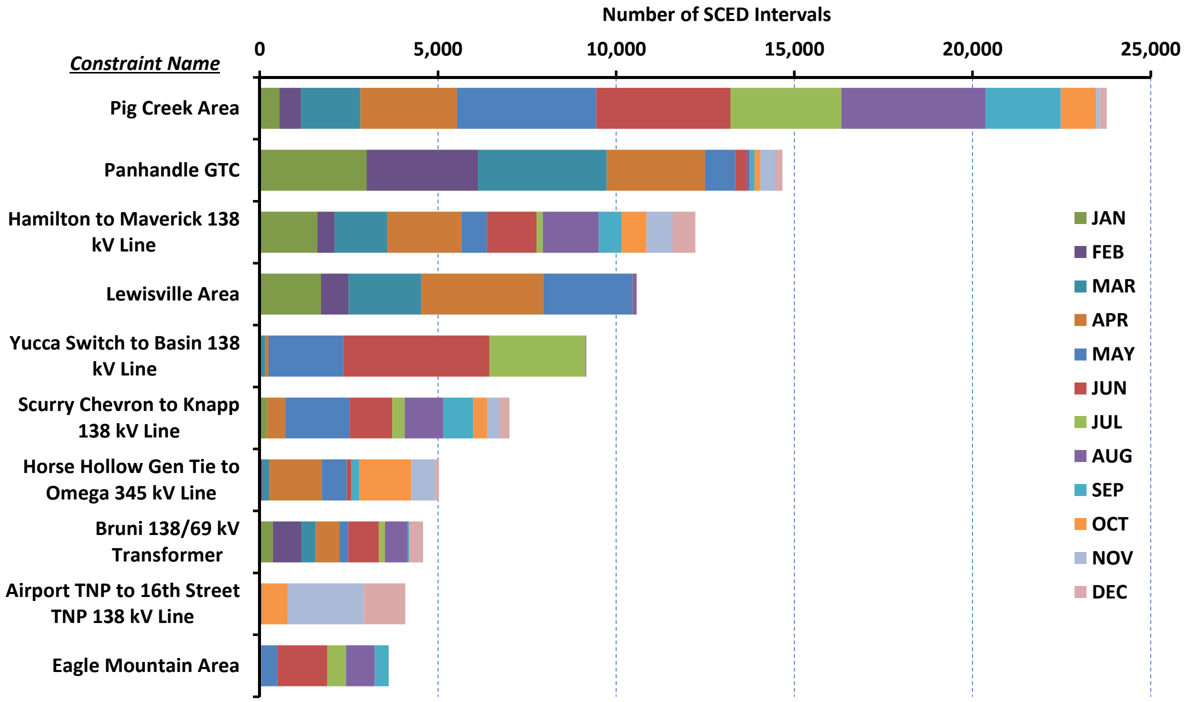
Constraints were at maximum shadow prices more frequently in 2018 than at any time under the Nodal market. The majority of the violated constraints occurring at the $2,000 per MW value were related to the Yucca Drive to Gas Pad 138 kV line irresolvable element. Violated constraints continued to occur in only a small fraction of all of the constraint-intervals, 8% in 2018, up from 3% in 2017.

Figure 50 below presents a slightly different set of real-time congested areas, showing the areas that were most frequently constrained in 2018.

---

[21] http://www.ercot.com/content/wcm/lists/144927/2018_Constraints_and_Needs_Report.pdf at 18.

Figure 50: Most Frequent Real-Time Constraints



All constraints listed in Figure 50 were frequently constrained due to high renewable output. Only the Pig Creek area (the most frequently congested constraint) was impacted by solar resources, whereas the rest of the constraints resulted from high output from wind resources. Six of the ten most frequently occurring constraints in 2018 were also among the ten most costly constraints including Pig Creek Area, Panhandle GTC, Lewisville area, Yucca Switch to Basin 138 kV line, Scurry Chevron to Knapp 138 kV line, and Eagle Mountain area.

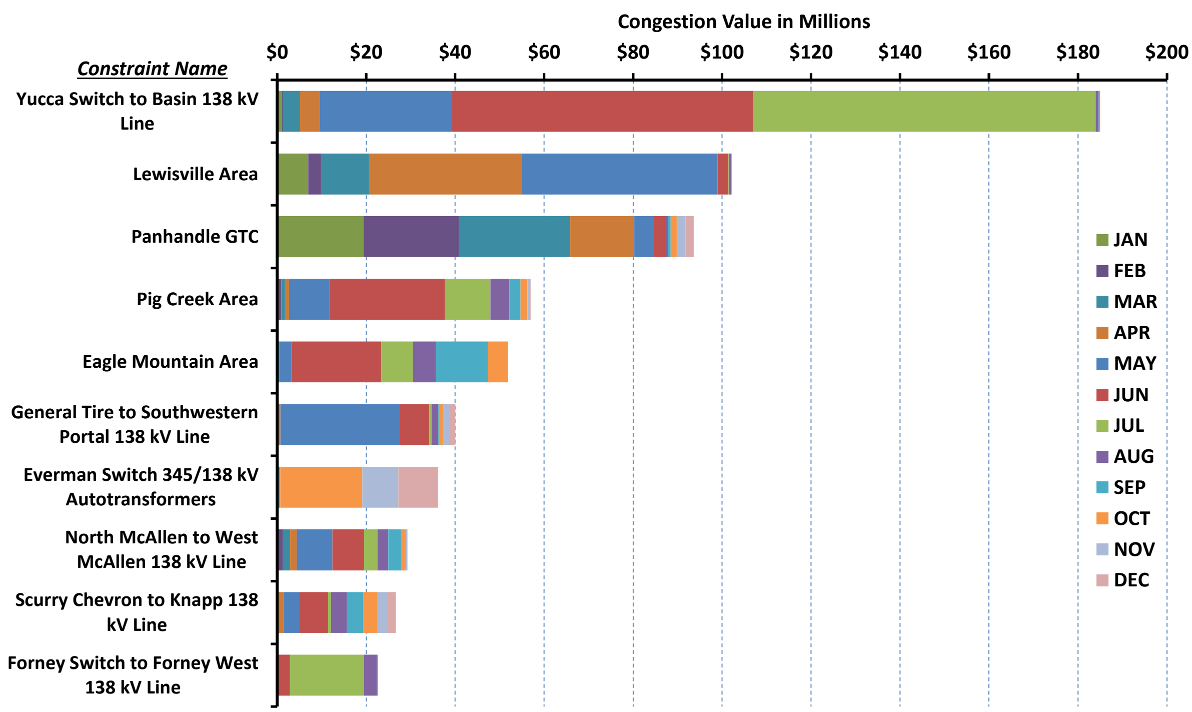
Five of these constraints were also in the top ten most frequent constraints in 2017, and three of the constraints, the Solstice to Pig Creek constraint, now called the Pig Creek area, Hamilton to Maverick 138 kV line, and the Lewisville area, occurred more frequently, about 80% of the time for each. The other two most frequent constraints in 2017, the Panhandle GTC and the Eagle Mountain area, were both less frequent in 2018.

The third, sixth, and seventh most frequently congested constraints, Hamilton to Maverick 138 kV line, Bruni 138/69 kV transformer, and Horse Hollow Gen Tie to Omega, respectively, are located in the South zone and were affected by high wind output. The ninth most frequently congested constraint, Airport TNP to 16th Street TNP 138 kV line, is located near the Pig Creek Area, and was congested because of high solar output flowing to the McCamey area post-contingency. This constraint is also adjacent to the tenth most valued constraint, the Woodward 2 to 16th Street TNP 138 kV line, but from the other direction. Only the ninth most frequent constraint experienced minimal congestion costs, as the generation that was re-dispatched to resolve the constraint was similarly priced. The rest of the most frequent constraints aggregated more than $10 million in congestion value.

C. Day-Ahead Constraints

This subsection provides a review of the transmission constraints from the day-ahead market. Figure 51 presents the ten most congested areas from the day-ahead market, ranked by their value. Eight of the constraints listed here were described in the previous subsection, Real-Time Constraints. To the extent the model of the transmission system used for the day-ahead market matches the real-time transmission system, and assuming market participants transact in the day-ahead market similarly to how they transact in real-time, the same transmission constraints are expected to appear in both markets.

Figure 51: Most Costly Day-Ahead Congested Areas



Since the start of the nodal market, it has been common for the day-ahead constraint list to contain many constraints that were unlikely to occur in real-time. However, for the second year in a row, the majority of the costliest day-ahead constraints in 2018 were also costly real-time constraints.

The Panhandle GTC incurred less congestion value in the day-ahead market than the real-time market as a result of less wind generation participating in the day-ahead market likely because of the uncertainty associated with predicting its output.

The General Tire to Southwestern Portal 138 kV line, located in Odessa, was the sixth most costly day-ahead constraint. The 138 kV line between North McAllen and West McAllen in the Valley was the eighth most costly day-ahead constraint. The tenth constraint, Forney Switch to Forney West 138 kV line is in the DFW area.

Figure 52: Day-Ahead Congestion Costs by Zone

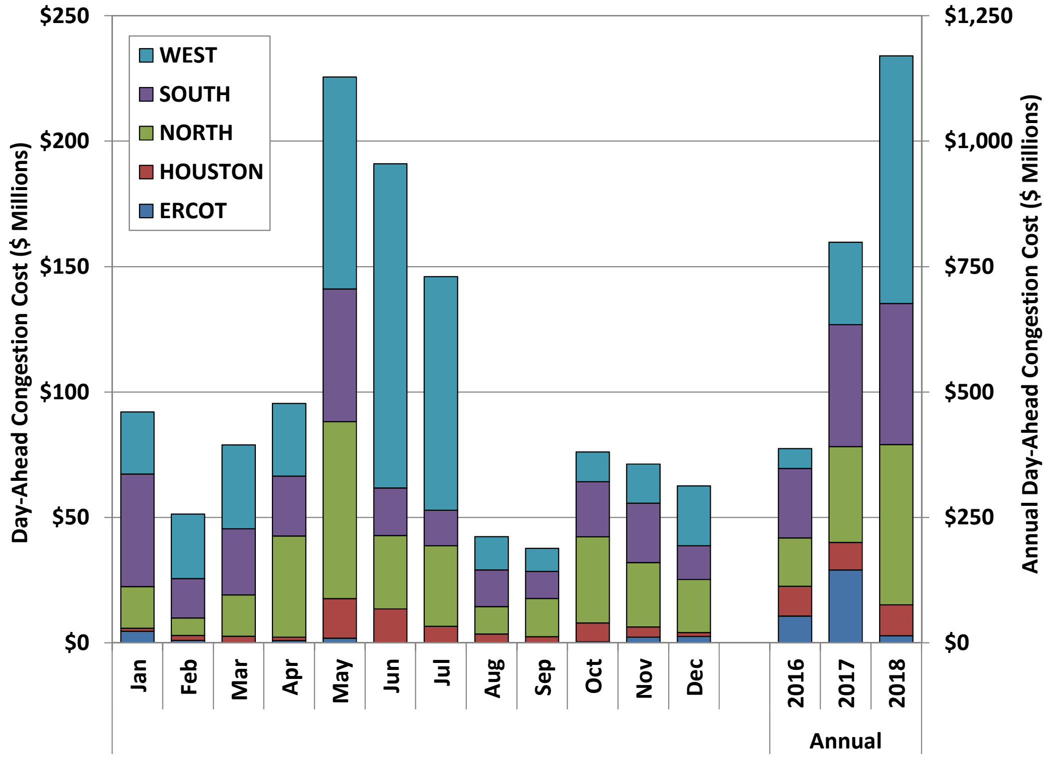


Figure 52 above presents day-ahead congestion costs by zone for the months of 2018 and annually for 2016 through 2018. Day-ahead congestion was higher in all zones in 2018 than it was 2017, while inter-zonal congestion was less. These outcomes are similar to those in the real-time market. The total day-ahead congestion costs in 2018 were almost a third higher than those costs in 2017; similar in magnitude to the increase seen in real-time congestion costs. The majority of West zone congestion was due to increased oil and gas development in far west Texas. The increased congestion costs during May, June and July were caused by high load growth in the west and were similar to the pattern of congestion costs observed in the real-time market.

D. Congestion Revenue Rights Market

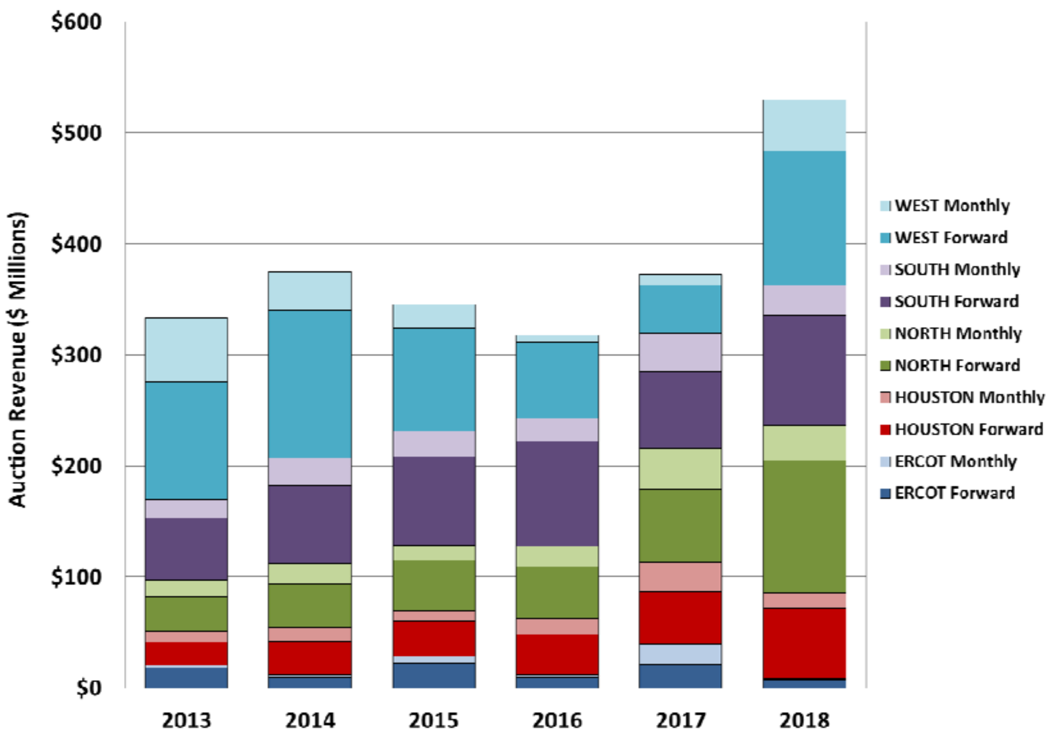
Congestion can be significant from an economic perspective, compelling the dispatch of higher-cost resources because power produced by lower-cost resources cannot be delivered because of transmission constraints. This causes different clearing prices for energy at different locations. Under the nodal market design, one means by which ERCOT market participants can hedge these price differences is by acquiring CRRs between any two settlement points.

CRRs may be acquired in semi-annual and monthly auctions while Pre-Assigned Congestion Revenue Rights (PCRRs) are allocated to certain participants (Non Opt-In Entities, or NOIEs) based on their historical patterns of transmission usage. Parties receiving PCRRs pay only a fraction of the auction value of a CRR between the same source and sink. Both CRRs and PCRRs entitle the holder to payments or charges that correspond to the difference in day-ahead locational prices of the source and sink.

CRR Costs and Auction Revenues

Figure 53 details the congestion cost as calculated by shadow price and flow on binding constraints in the CRR auctions. Note that this calculation, based on the binding constraint location, is similar to the calculation used earlier in this report to display the zonal location of real-time and day-ahead congestion costs and is different from the method used to determine CRR revenue allocation. The costs are broken down by the zonal location of the constraint and whether they were incurred in a monthly auction (Monthly) or a seasonal or annual auction (Forward).

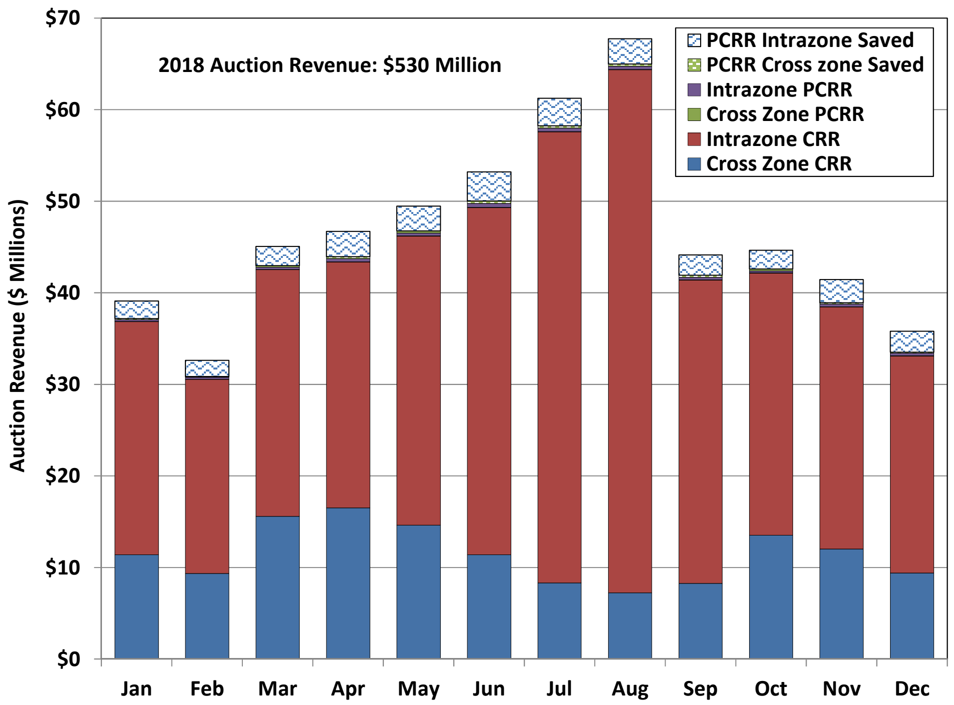
Figure 53: CRR Costs by Zone



Comparing the costs paid to acquire CRRs, as shown in Figure 53, to the trends of congestion costs seen in the real-time and the day-ahead markets, it is evident that CRR congestion, while still significantly lower than day-ahead and real-time congestion, increased by a greater percentage than congestion in either the day-ahead or real-time markets. Also, like the day-ahead and real-time markets, all zones saw increased congestion while inter-zonal congestion decreased.

Figure 54 summarizes the revenues collected by ERCOT in each month for all CRRs, including both auctioned and allocated. Also shown is the amount of discount provided to the PCRR recipients.

Figure 54: 2018 CRR Auction Revenue



CRR auction revenues are distributed to loads in one of two ways. Revenues from cross-zone CRRs are allocated to loads ERCOT wide. Revenues from CRRs that have the source and sink in the same geographic zone are allocated to loads within that zone. Allocating CRR auction revenues in this manner reduces the net cost for load purchases in heavily-congested areas, regardless of whether the congestion had raised prices or lowered prices in the area.

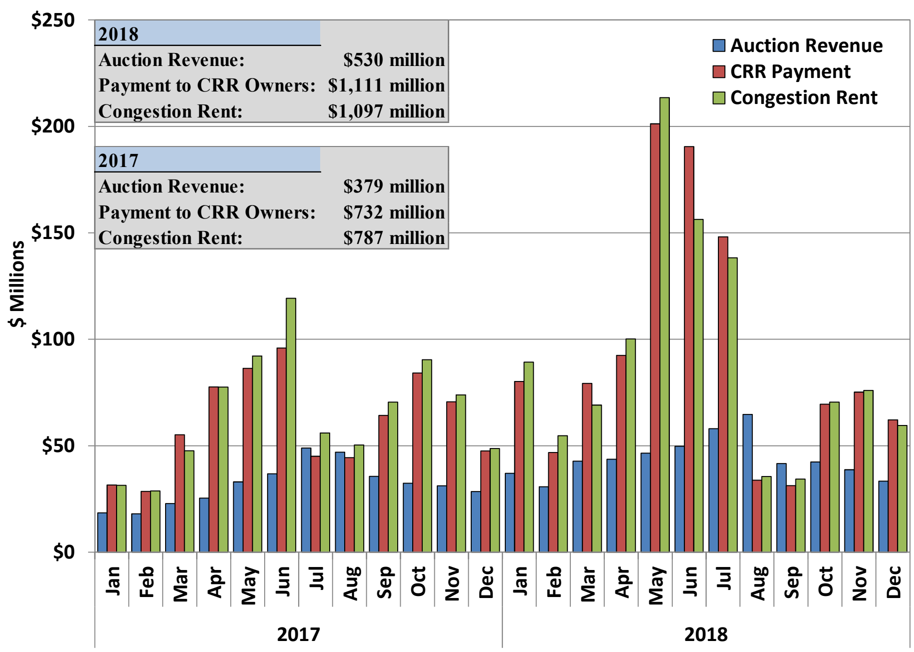
As previously discussed in this section, the only parties eligible to receive PCRRs are NOIEs, which pay only a fraction of the PCRR auction value. The difference between the auction value and the value charged to the purchaser is shown in Figure 54 as the PCRR discount. Even as the total amount of CRR auction revenue increased to $530 million in 2018 from $397 million in 2017, the total PCRR discount decreased from $50 million in 2016 to $31 million in 2018.

CRR Profitability

Although results for individual participants and specific CRRs varied, the aggregated results for the year and in most months show that participants paid much less for CRRs in 2018 than they received in payment from the day-ahead market. For the entire year of 2018, participants spent $530 million to procure CRRs and in aggregate received $1.1 billion. In general, this difference occurred because the substantial increase in congestion that occurred in 2018 was not foreseen by the market. The period of congestion that accounted for most of this difference was May, June and July when the CRR payments were $385 million higher than the congestion rent. This was mostly due to congestion in West Texas.

The next analysis of aggregated CRR positions adds day-ahead congestion rent to the picture. Day-ahead congestion rent is the difference between payments and charges of three-part offers, energy only offers, energy only bids, Point-to-Point (PTP) obligation bids, and PTP obligation bids linked to options in day-ahead market.[22] Day-ahead congestion rent creates the source of funds used to make payments to CRR owners. Figure 55 presents CRR auction revenues, payment to CRR owners, and congestion rent in 2017 and 2018, by month. Congestion rent for the year 2018 totaled $1,097 million and payment to CRR owners was $1,111 million.

Figure 55: CRR Auction Revenue, Payments and Congestion Rent



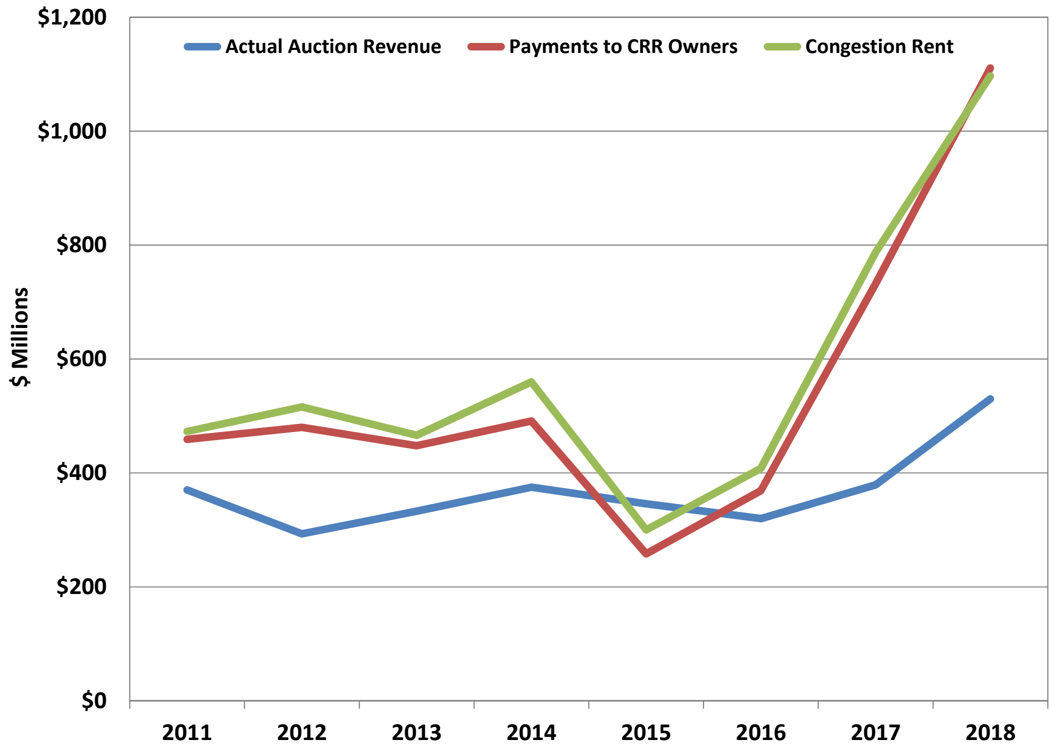
22 Under Protocol Section 7.9.3.1, Day-Ahead Market congestion rent is calculated as the sum of the following payments and charges: (a) The total of payments to all QSEs for cleared Day-Ahead Market energy offers, whether through Three-Part Supply Offers or through Day-Ahead Market Energy-Only Offer Curves, calculated under Section 4.6.2.1, Day-Ahead Energy Payment; (b) The total of charges to all QSEs for cleared Day-Ahead Market Energy Bids, calculated under Section 4.6.2.2, Day-Ahead Energy Charge; and (c) The total of charges or payments to all QSEs for PTP Obligation bids cleared in the Day-Ahead Market, calculated under Section 4.6.3, Settlement for PTP Obligations Bought in Day-Ahead Market. (d) The total of charges to all QSEs for PTP Obligation with Links to an Option bids cleared in the Day-Ahead Market, calculated under Section 4.6.3.

It is worth noting that because the CRR network model uses line ratings that are 90% of the expected lowest line ratings for the month, it is expected that CRRs would be somewhat undersold and that day ahead congestion rent would be higher than the payment to CRR owners. However, this was not the case in 2018 and payments to CRR owners were 101% of congestion rent. In 2017, this ratio was 93%. What appears to be an overpayment to CRR owners in 2018 was tied to shortfalls in real-time congestion rent, which were likely due to the inaccurate or inconsistent load distribution factors used in CRR auctions and day-ahead markets.

Figure 56 provides the annual history of these three CRR related values: auction revenues, reflecting the costs paid by owners to obtain the CRRs; Payments to CRR owners, reflecting the payments received by CRR owners; and day-ahead congestion rent, which is the funding source for most CRR payments. In 2018, like 2017, owners of CRRs in aggregate made a substantial profit on their CRR holdings. Payments to CRR owners in 2018 were almost double the total cost paid to acquire the CRRs. As discussed above, this was primarily due to unanticipated factors that led to significantly higher congestion in 2017. The figure shows that this was not the case in some recent years. In 2015, CRR owners were paid less than the total cost paid to obtain them. In 2016, it appears that CRR owners made a small profit, but the cost to obtain the CRRs reflects the discounted amounts that NOIEs paid to obtain PCRRs. Adding the NOIE discount to the auction revenue in 2016 would show CRRs, in aggregate, to be unprofitable.

Another item to note from these historical values is the relatively flat auction revenue. While congestion rent and payments to CRR owners has varied from less than $500 million to more than $1 billion per year, the costs paid to acquire CRRs varied in a much narrower range between $300 and $550 million per year since the start of the nodal market. This may imply that aggregate CRR profitability is less dependent on CRR owners making acquisition decisions based on sophisticated analysis, and more likely driven by the vagaries of annual transmission congestion patterns.

Figure 56: CRR History



CRR Funding Levels

The target value of a CRR is the megawatt amount of the CRR multiplied by the locational marginal price (LMP) of the sink of the CRR less the LMP of the source of the CRR. While the target value is paid to CRR account holders most of the time, there are two circumstances that cause ERCOT to pay less than the target value (i.e., CRRs are not fully funded). The first circumstance happens when the CRR is modeled on the day-ahead network and causes a flow on a transmission line that exceeds the line's limit. In other words, the transmission capability assumed in the CRR market is ultimately higher than in the day-ahead market, which can occur because of outages or other factors that reduce transfer capability. In this case, CRRs with a positive value that have a source or a sink located at a resource node settlement point are paid a lower amount than the target value.

The second circumstance occurs when there is not enough day-ahead congestion rent to pay all the CRRs at target (or derated, if applicable) value. In this case, all holders of positively valued CRRs receive a prorated shortfall charge such that the congestion revenue plus the shortfall charge can pay all CRRs at target or derated value. This shortfall charge has the effect of lowering the net amount paid to CRR account holders; however, if at the end of the month there is excess day-ahead congestion rent that has not been paid out to CRR account holders, the excess congestion rent can be used to make whole the CRR account holders that received shortfall charges. If there is not enough excess congestion rent from the month, the rolling CRR balancing fund can be drawn upon to fully pay CRR account holders that received shortfall charges.

Figure 57 shows the CRR balancing fund since the beginning of 2016. This year, due mostly to differences in the West Load Zone day-ahead and CRR load distribution factors, the CRR balancing fund was drawn on heavily in June and depleted entirely in July. While the balancing fund was once again up to its capped value of $10 million by August, another shortfall lowered the fund balance in November. At the end of the year the amount in the balancing fund was a little less than $8 million. In 2018, the total day-ahead surplus was $52 million, significantly lower than the surplus of $94 million in 2017, due largely to the shortfalls in June, July, and November. Because there was not enough day–ahead market surplus and CRR balancing fund to fully refund the shortfall charge for the CRR owners in July, for the first time since the creation of the balancing account the balancing fund went to $0 and the CRR owners received a shortfall charge, of $18.85 million, of which only $8.99 million was refunded at the end of July. From the perspective of the load, even though there was a shortfall month, the monthly CRR balancing account allocation to load was still a net positive in 2018 and resulted in a total amount of $65 million at the end of the year.

Figure 57: CRR Balancing Fund

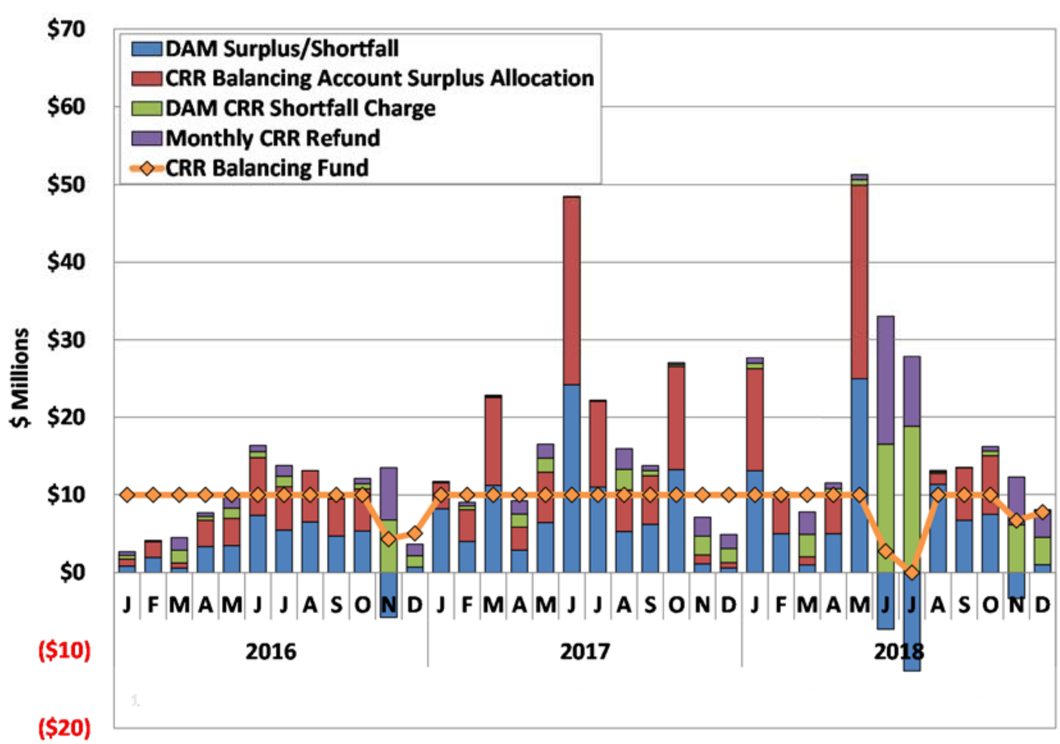
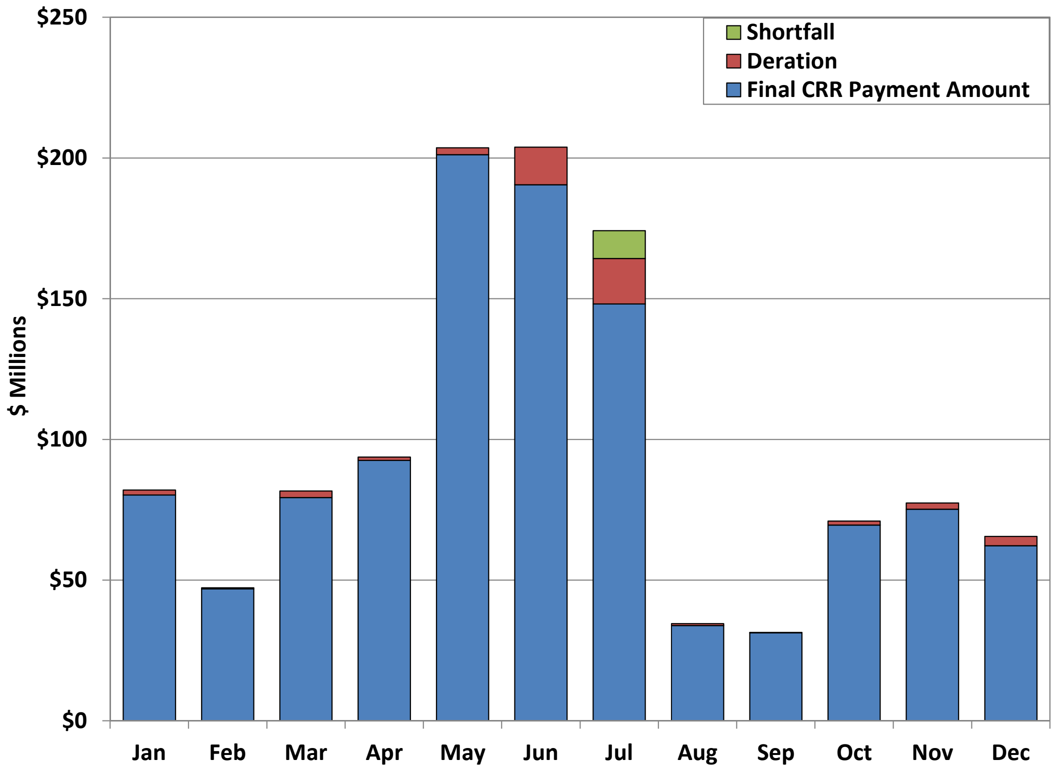


Figure 58 shows the amount of target payment, deration amount, and net shortfall charges (after make whole payments) for 2018. In 2018, the total target payment to CRRs was $1,166 million; however, there were $45 million of derations and $9 million of non-refunded shortfall charges resulting in a final payment to CRR account holders of $1,111 million. This final payment amount corresponds to a CRR funding percentage of 95%, slightly lower than 2017's funding percentage of 97%.

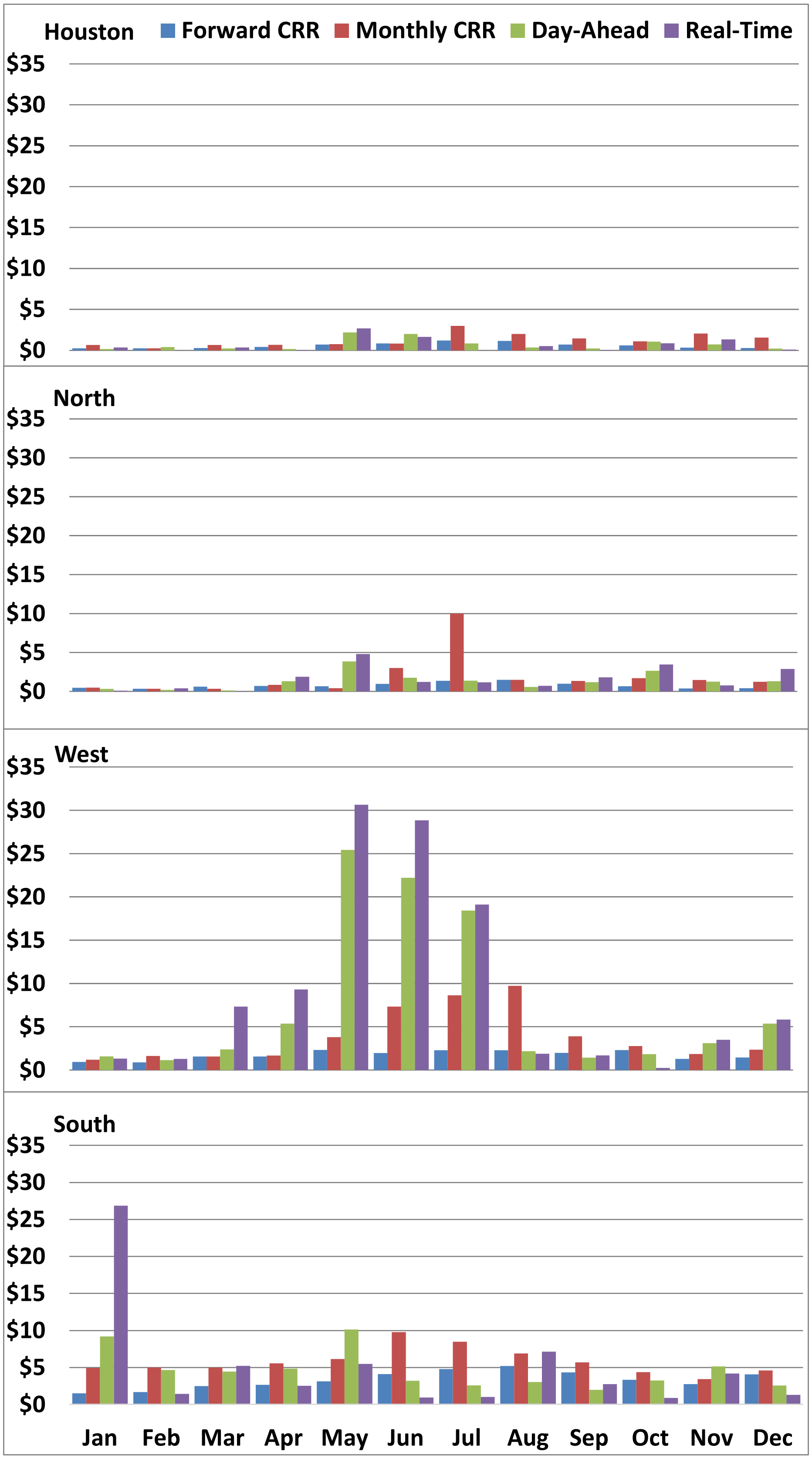
Figure 58: CRR Shortfalls and Derations



The last look at congestion examines the price spreads for each pair of hub and Load Zones in more detail. These price spreads are interesting as many loads may have contracts that hedge to the hub price and are thus exposed to the price differential between the hub and its corresponding Load Zone. Figure 59 presents the price spreads between all hub and Load Zones as valued at four separate points in time – at the average of the four semi-annual CRR auctions, monthly CRR auction, day-ahead and real-time.

Of note is the extreme divergence between the CRR price spreads for the West Load Zone and the day-ahead and real-time price spreads. The lagging nature of the CRR revenues can be seen by steady growth of the West zone monthly CRR spread from May to August, which continued one month after the day-ahead price spread dropped. The West Load Zone, due to the high congestion from March through July, had the highest hub to zone price spread in 2018, overtaking the South Load Zone.

Figure 59: Hub to Load Zone Price Spreads

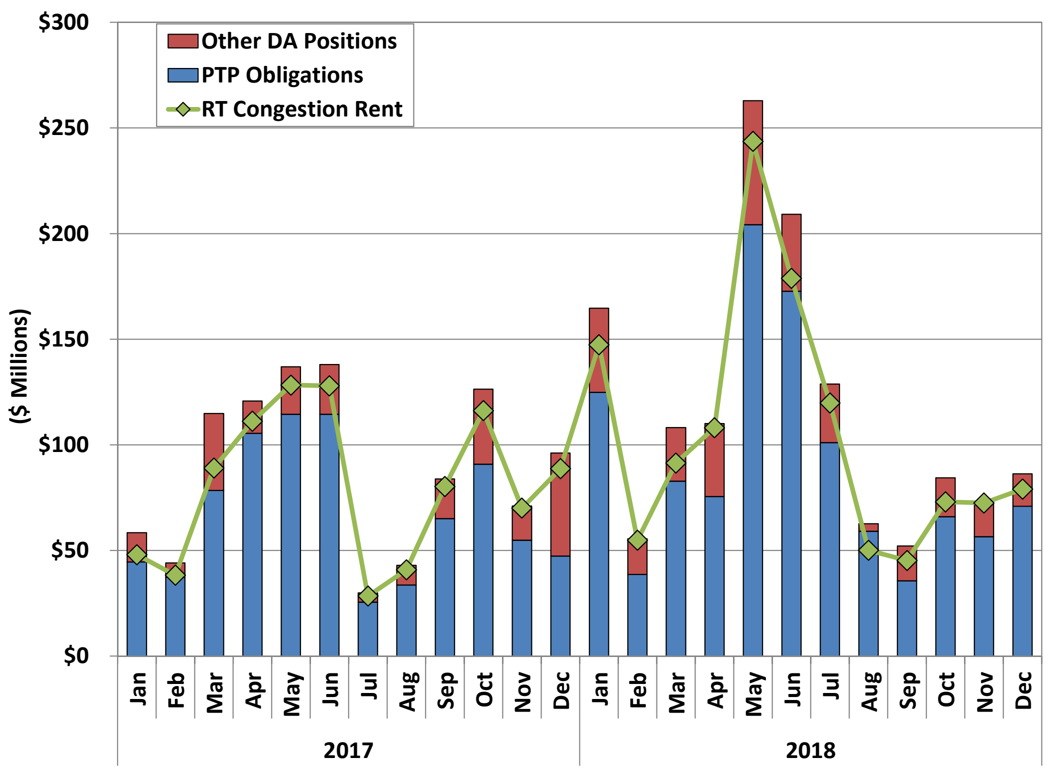


E. Revenue Sufficiency

In Figure 60, the combined payments to PTP obligation owners and effective payments to other day-ahead positions are compared to the total real-time congestion rent. For 2018, real-time congestion rent was $1,264 million, payments for PTP obligations (including those with links to CRR options) were $1,088 million and payments for other day-ahead positions were $310 million, resulting in a shortfall of approximately $134 million for the year.

By comparison, the real-time congestion rent was $967 million in 2017. Payments for PTP obligations and real-time CRRs were $812 million and payments for other day-ahead positions were $251 million, resulting in a shortfall of approximately $96 million for the year. This shortfall is paid for by charges to load.

Figure 60: Real-Time Congestion Rent and Payments



IV. Demand and Supply

This section reviews and analyzes the load patterns during 2018 and the existing generating capacity available to satisfy the load and operating reserve requirements. Specific analysis of the large quantity of installed wind generation is included, along with a discussion of the daily generation commitment characteristics. This section concludes with a review of the contributions from demand response resources.

A. ERCOT Load in 2018

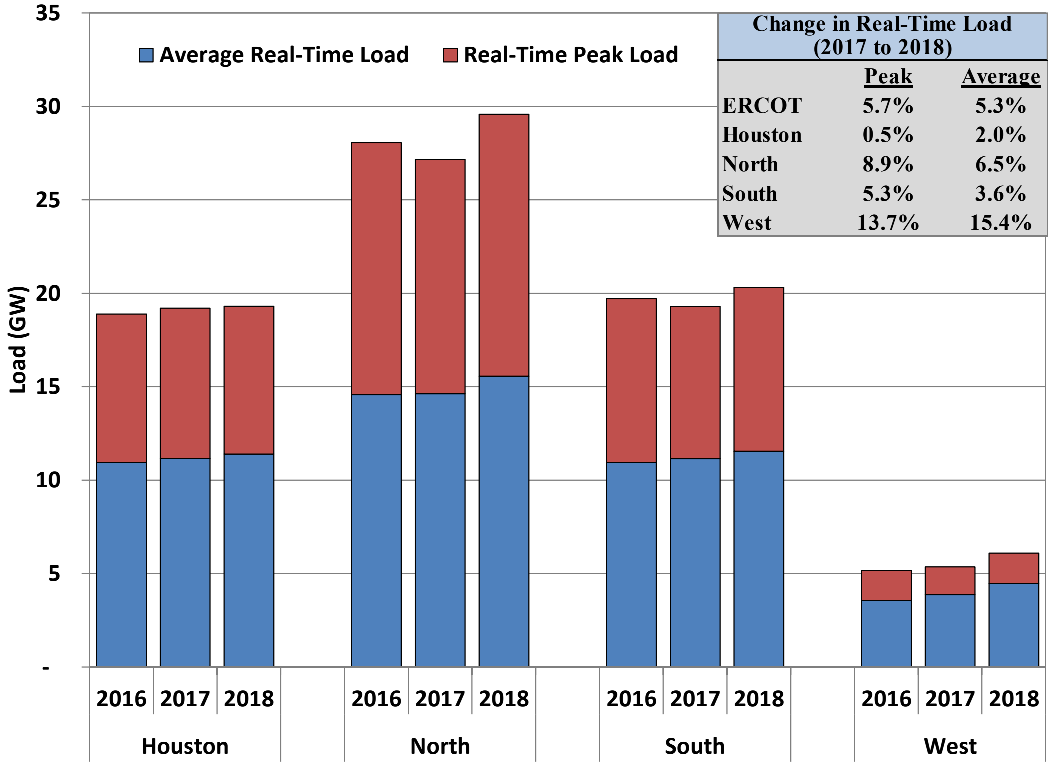
The changes in overall load levels from year to year can be shown by tracking the changes in average load levels. This metric tends to capture changes in load over a large portion of the hours during the year. Separately evaluating the changes in the load during the highest-demand hours of the year is also important. Significant changes in peak demand levels play a major role in assessing the need for new resources. The level of peak demand also affects the probability and frequency of shortage conditions (i.e., conditions where firm load is served but minimum operating reserves are not maintained). The expectation of resource adequacy is based on the value of electric service to customers and the harm or inconvenience to customers that can result from interruptions to that service. Hence, both of these dimensions of load during 2018 are examined in this subsection and summarized in Figure 61.

This figure shows peak load and average load in each of the ERCOT geographic zones from 2016 through 2018.[23] In each zone, as in most electrical systems, peak demand significantly exceeds average demand. The North zone is the largest zone (with about 36% of the total ERCOT load); the South and Houston zones are comparable (27%) while the West zone is the smallest (10% of the total ERCOT load).

Figure 61 also shows the annual non-coincident peak load for each zone. This is the highest load that occurred in a particular zone for one hour during the year; however, the peak can occur in different hours for different zones. As a result, the sum of the non-coincident peaks for the zones is greater than the annual ERCOT peak load.

23 For purposes of this analysis, Non-Opt In Entity (NOIE) Load Zones have been included with the proximate geographic zone.

Figure 61: Annual Load Statistics by Zone



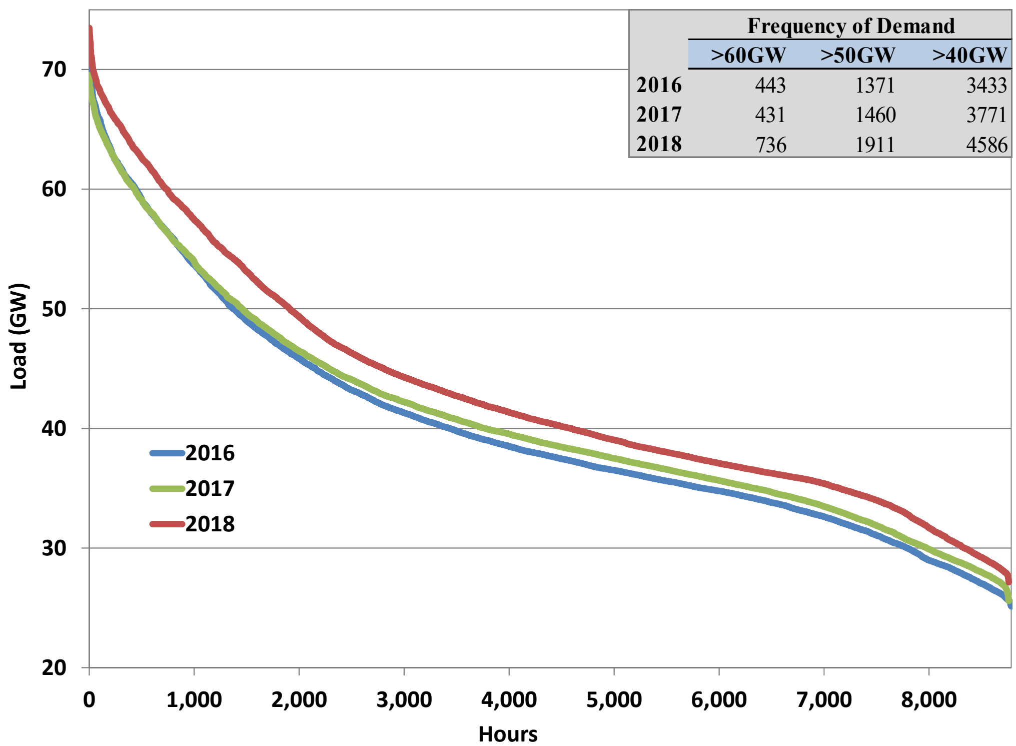
Change in Real-Time Load (2017 to 2018)		
	Peak	Average
ERCOT	5.7%	5.3%
Houston	0.5%	2.0%
North	8.9%	6.5%
South	5.3%	3.6%
West	13.7%	15.4%

Total ERCOT load in 2018 increased 5.3% from 2017 to total 376.4 TWh. This increase equates to approximately 2,200 MW per hour on average. All zones showed an increase in average real-time load in 2018 ranging from 2% increase in Houston to 15% increase in the West zone. Continuing robust oil and natural gas production activity in the West zone has been the driver for the high growth experienced recently. Weather impacts on load in 2018 were mixed across the zones. There were fewer annual cooling degree days across all zones. However, for the three summer months of June, July and August, there was a 12% increase from 2017 in the number of cooling degree days in Dallas. For the same time frame, Austin had a 6% increase and Houston was flat.

Summer conditions in 2018 produced a new record peak load of 73,473 MW on July 19, 2018, surpassing the previous ERCOT-wide coincident peak hourly demand record of 71,110 MW set on August 11, 2016. A new winter peak demand record of 65,915 MW was also set on January 17, 2018. All zones experienced varying increases in peak load ranging from 0.5% increase in Houston to more than 13% increase in the West zone, which continued to experience the highest percentage growth in peak load, due to continuing growth in oil and natural gas production.

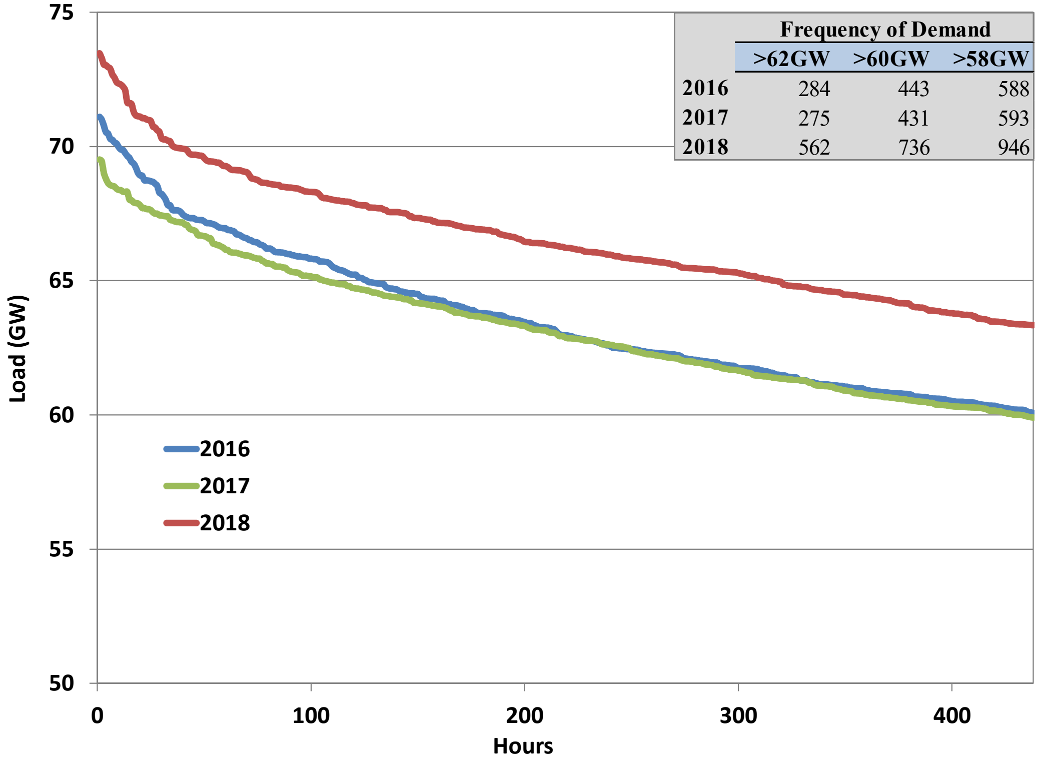
To provide a more detailed analysis of load at the hourly level, Figure 62 compares load duration curves for each year from 2016 through 2018. A load duration curve illustrates the number of hours (shown on the horizontal axis) that load exceeds a particular level (shown on the vertical axis). ERCOT has a fairly smooth load duration curve, typical of most electricity markets, with low to moderate electricity demand in most hours, and peak demand usually occurring during the late afternoon and early evening hours of days with exceptionally high temperatures. The load duration curve in 2018 was similar to 2017 and 2016, though slightly higher.

Figure 62: Load Duration Curve – All Hours



To better illustrate the differences in the highest-demand periods between years, Figure 63 below shows the load duration curve for the 5% of hours with the highest loads for the last three years. This figure also shows that the peak load in each year was significantly greater than the load at the 95th percentile of hourly load. Since 2011, the peak load has averaged 16% to 19% greater than the load at the 95th percentile. These load characteristics imply that a substantial amount of capacity – more than 10 GW – is needed to supply energy in less than 5% of the hours.

Figure 63: Load Duration Curve – Top 5% of Hours with Highest Load

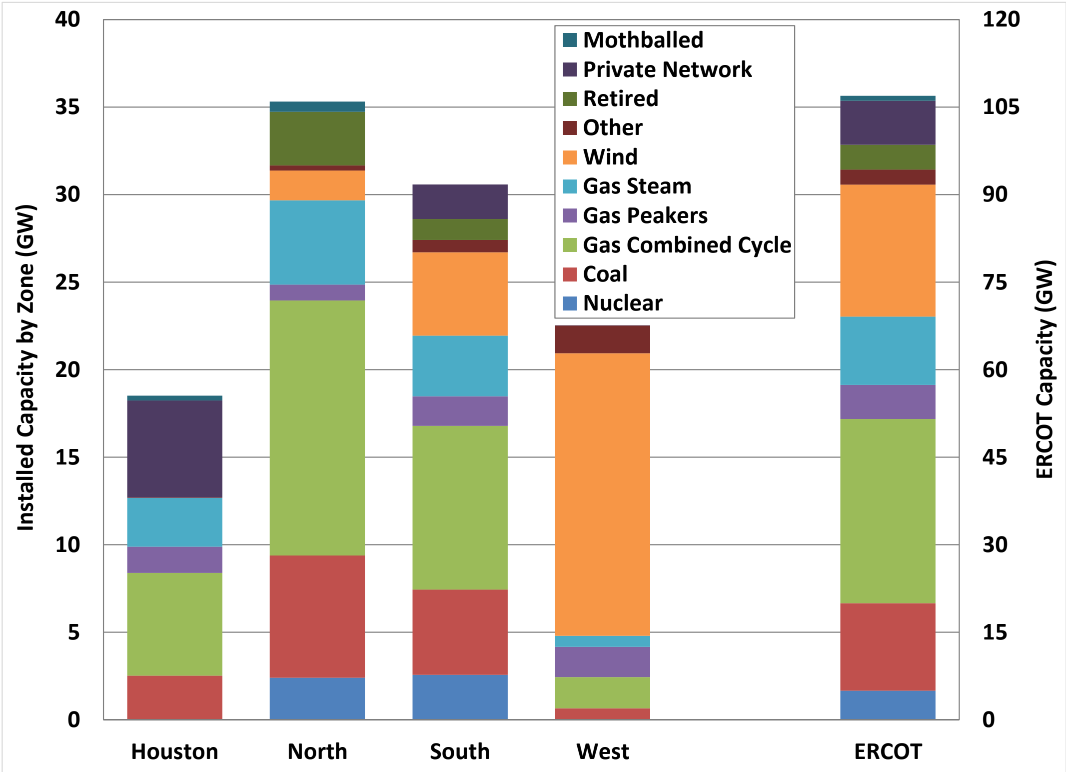


B. Generation Capacity in ERCOT

The generation mix in ERCOT is evaluated in this subsection. The distribution of capacity among the four ERCOT geographic zones is similar to the distribution of demand with the exception of the large amount of wind capacity in the West. In 2018, the North zone accounted for approximately 33% of capacity, the South zone 29%, the Houston zone 17%, and the West zone 21%. Excluding mothballed resources and including only the fraction of wind capacity available to reliably meet peak demand,[24] the North zone accounted for approximately 36% of capacity, the South zone 32%, the Houston zone 22%, and the West zone 10% in 2017. The installed generating capacity by type in each zone is shown in Figure 64.

24 The percentages of installed capacity to serve peak demand assume wind availability of 15% for non-coastal wind and 58% for coastal wind.

Figure 64: Installed Capacity by Technology for Each Zone



Approximately 3.8 GW of new generation resources came online in 2018, the bulk of which was multiple wind resources with total capacity of 2.2 GW, and an effective peak serving capacity of less than 520 MW. Approximately 150 MW of this wind capacity were re-powered resources. The remaining capacity additions were 670 MW of new combustion turbines and 880 MW of solar resources.

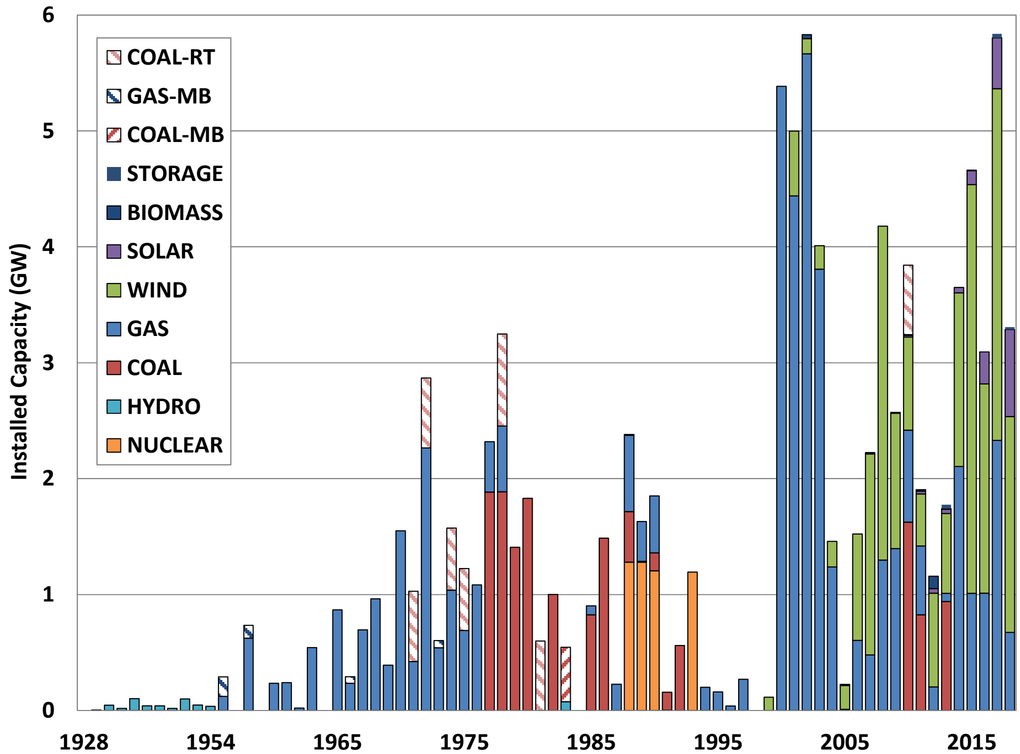
A total of nine coal generation resources, totaling 5.1 GW, were retired and permanently decommissioned in 2018. Luminant announced during the fall of 2017 that coal units at three locations, Monticello, Sandow, and Big Brown, would be retiring. All seven units, totaling approximately 4,300 MW, ceased operation in January and February of 2018. CPS Energy retired two coal units totaling 840 MW at the Calaveras location on December 31, 2018 (as noticed to ERCOT in 2013).

Figure 65 shows the age of generation resources in ERCOT shown as operational in the December 2018 Capacity, Demand, and Reserves Report (CDR),[25] as well as resources that came online but were not yet commercial. Private Use Network capacity contributions to the CDR were excluded. Seventy-five percent of the total coal capacity in ERCOT was at least thirty years old in 2018. After several years of limited generation additions in the late 1990's, there

25 ERCOT Capacity, Demand, and Reserves Report (Dec. 11, 2018), available at http://www.ercot.com/gridinfo/resource.

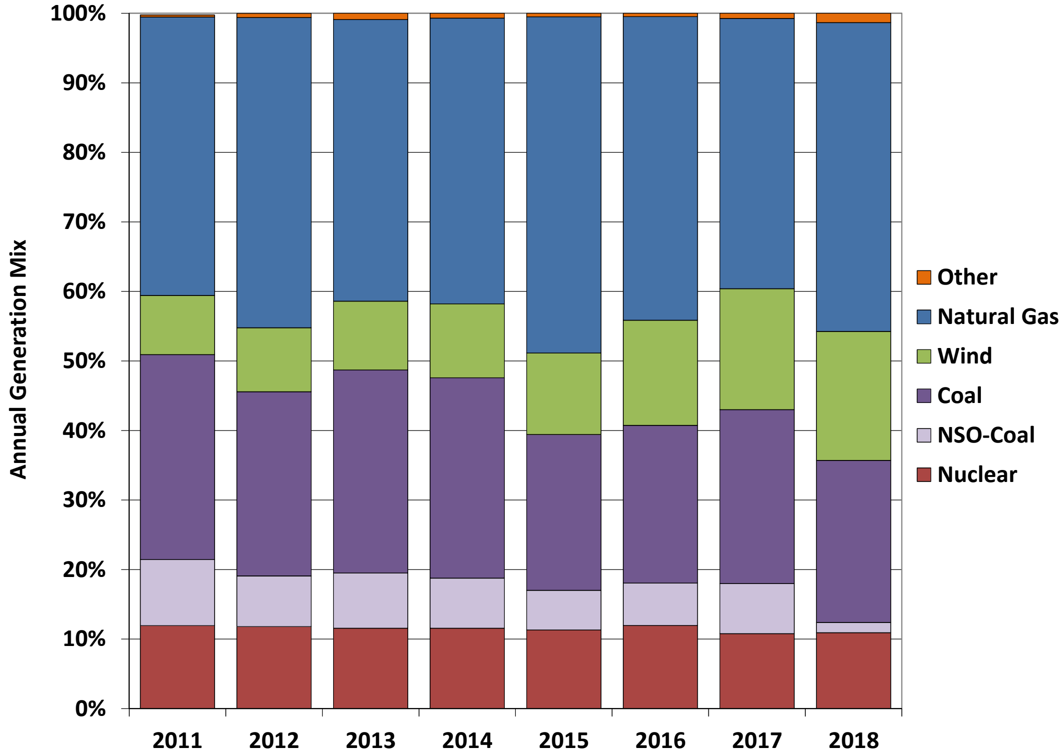
was more than 20 GW of combined cycle gas capacity added in the first five years after deregulation of the ERCOT generation market. Since the initial surge of combined cycle additions, there has been continued investment in that type of capacity, totaling another 10 GW over the past 15 years. The few new coal units added around 2010 were likely a response to high, and expected continued high natural gas prices. However, wind capacity has been the dominant technology for newly installed capacity since 2006.

Figure 65: Vintage of ERCOT Installed Capacity



The effects of coal retirements are evident in Figure 66, which shows the percentage of annual generation from each fuel type for the years since the implementation of the nodal market.

Figure 66: Annual Generation Mix

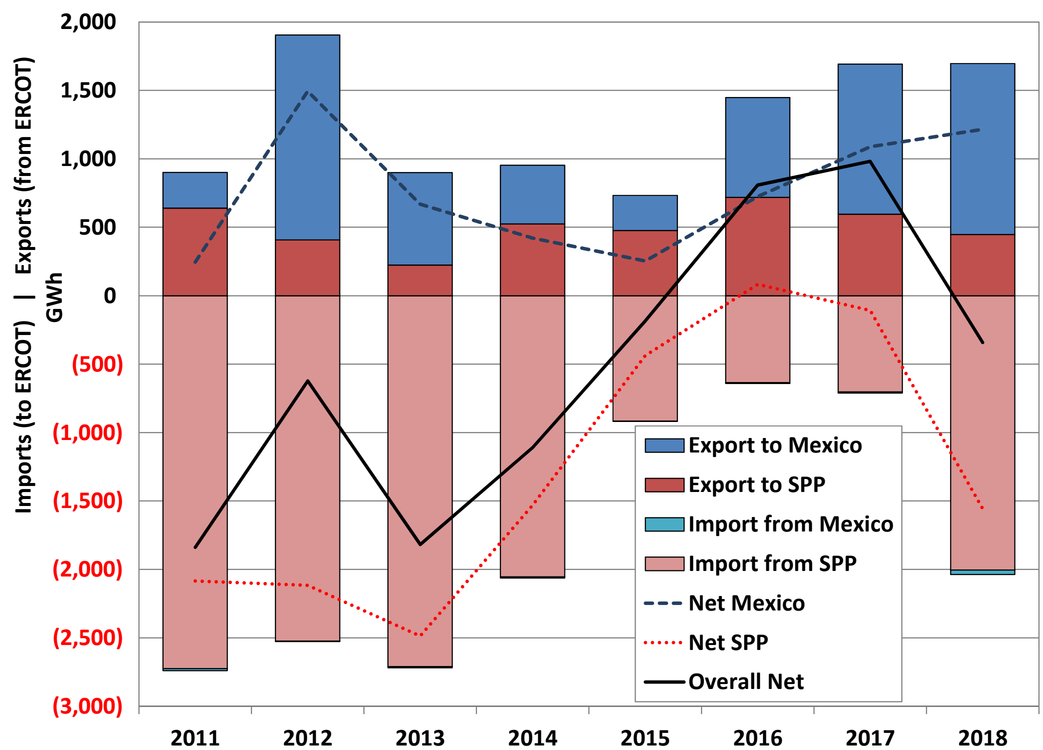


The generation share from wind has increased every year, reaching almost 19% of the annual generation requirement in 2018, up from 9% in 2011 and 17% in 2017. The share of generation from coal decreased to 25%, by far the lowest total since well before the advent of a competitive generation market in ERCOT. This figure separately shows the amount of energy provided from seven coal units that were retired at the start of 2018 and two coal units that were retired at the end of the year. The seven coal units provided an average of 7% of the total annual generation requirements from 2011 to 2017, and the two other units provided an additional 2%. In response to the reduction in coal generation, the share of natural gas generation increased again in 2018 to 44%, up from 39% in 2017.

While coal and lignite and nuclear plants operate primarily as base load units in ERCOT, the reliance on natural gas resources drives the high correlation between real-time energy prices and the price of natural gas fuel. This relationship was strengthened in 2018 with the retirement of 840 MW of coal. There were approximately 20 GW of coal and nuclear generation in ERCOT in 2018. Generally, when ERCOT load is above this level, natural gas resources will be on the margin and set the real-time energy spot price.

The ERCOT region is connected to other regions in North America via multiple asynchronous ties. Two ties, totaling 820 MW, connect ERCOT with the Southwest Power Pool (SPP) and three ties, totaling 430 MW, connect ERCOT with Comisión Federal de Electricidad (CFE) in Mexico. Transactions across the DC tie can be in either direction, into or out of ERCOT. These transactions can have the effect of increasing demand (exports) or increasing supply (imports). Figure 67 below shows the total energy transacted across the ties for each of the past several years.

Figure 67: Annual Energy Transacted Across DC Ties



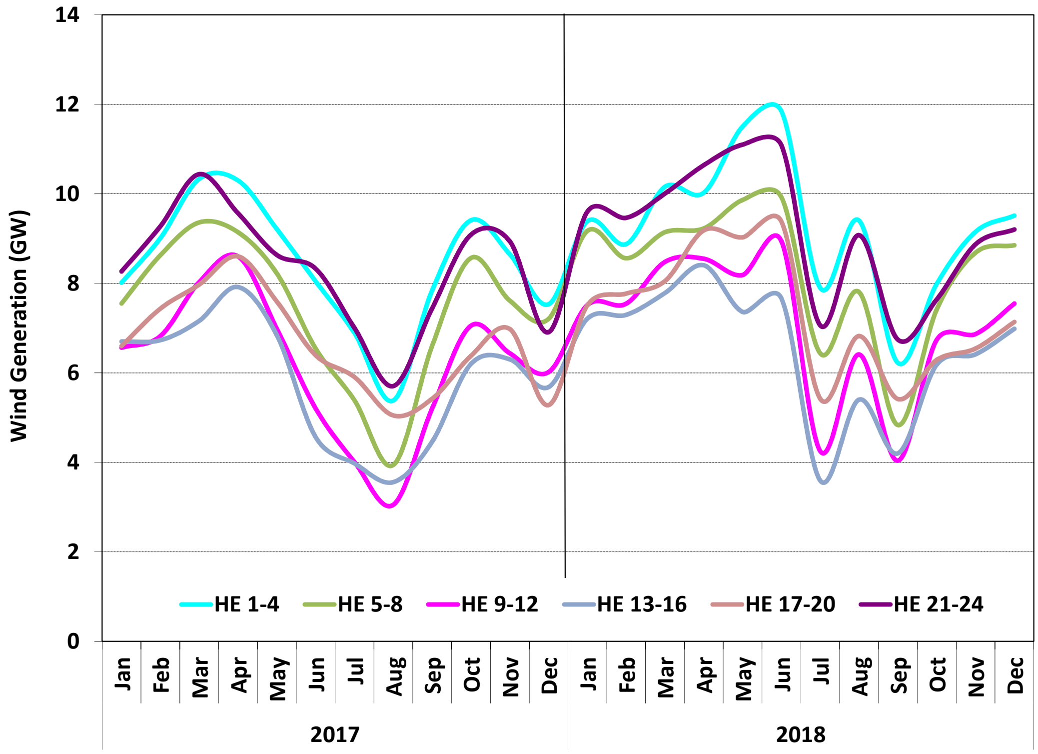
Between 2011 and 2014, far more energy was imported into the ERCOT market than exported into Mexico and SPP combined. In 2011, ERCOT was a net importer by 1,848 GWh, largely because of the high loads and tight conditions in ERCOT. Increased exports to Mexico led to decreased net imports in 2012, but return to previous levels in 2013. After 2013, there had been a trend of reduced imports from SPP and increased exports to Mexico because prices in ERCOT have remained relatively low. With the tightening supply in ERCOT and the resulting higher prices in 2018, that trend was reversed and imports into ERCOT exceeded exports into Mexico and SPP.

C. Wind Output in ERCOT

The amount of wind generation installed in ERCOT was approaching 22.5 GW by the end of 2018. Although the large majority of wind generation is located in the West zone, more than 4.5 GW of wind generation has been located in the South zone. Additionally, a private transmission line that went into service in late 2010 allows another nearly 1 GW of West zone wind to be delivered directly to the South zone. In 2007, wind generation in ERCOT was located in 14 counties; by 2018, there were 58 counties with wind generators serving ERCOT.

The average profile of wind production is negatively correlated with the load profile, with the highest wind production occurring during non-summer months, and predominately during off-peak hours. Figure 68 shows average wind production for each month in 2017 and 2018, with the average production in each month divided into four-hour blocks. Though the lowest wind output generally occurs during summer afternoons, there has been such a large amount of wind generation added in ERCOT that the average wind output during summer peak period now averages of 5.5 GW. This may be a small fraction of the total installed capacity but wind generation is now a non-trivial portion of generation supply, even at its lowest outputs.

Figure 68: Average Wind Production



ERCOT continued to set new records for peak wind output in 2018. On December 14, wind output exceeded 19 GW, setting the record for maximum output and on December 27, wind provided nearly 55% of the total load, also a new record.[26]

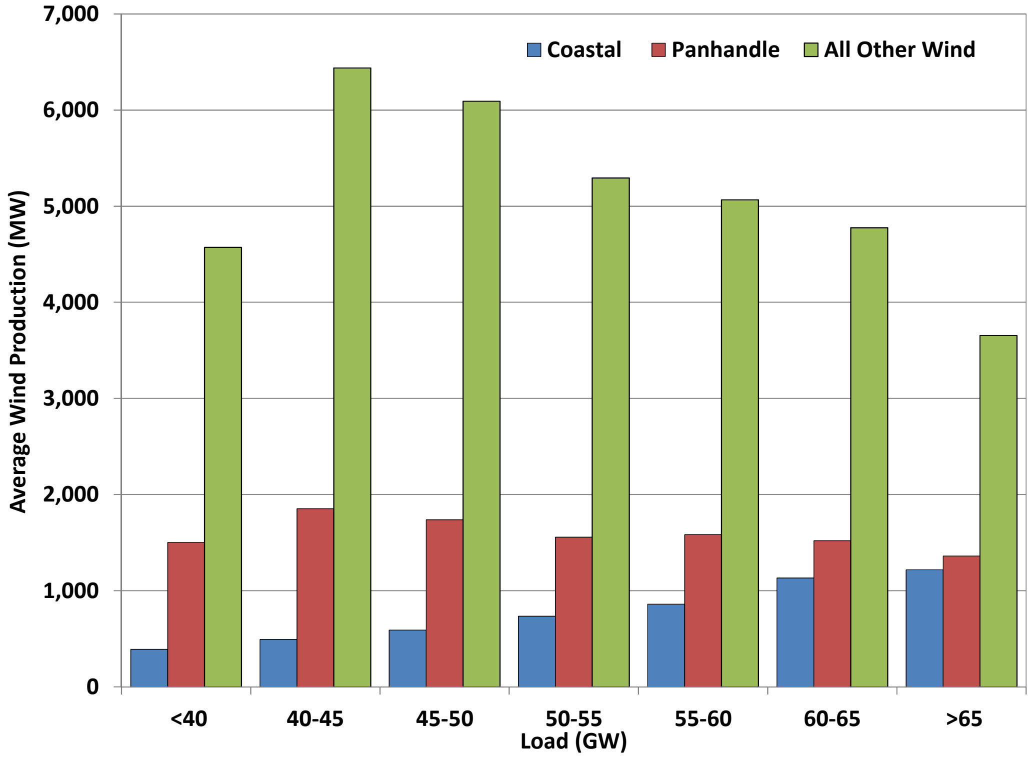
Examining wind generation in total masks the different wind profiles that exist for locations across ERCOT. Sites along the Gulf Coast of Texas are attractive due to the higher correlation of the wind resource in that location with electricity demand. However, average wind speeds along the coast are lower than at other Texas locations. More recently, the Texas Panhandle has attracted wind developer interest because of its abundant wind resources. The differences in

26 Peak hourly wind generation was 19,168 MW at 12:07 a.m. on December 14, 2018. Instantaneous wind penetration was 54.6% at 4:57 a.m. on December 27, 2018.

output for wind units located in the coastal area of the South zone and those located elsewhere in ERCOT are compared below.

Figure 69 shows data for the summer months of June through August, comparing the average output for wind generators located in the coastal region, the Panhandle and all other areas in ERCOT across various load levels.

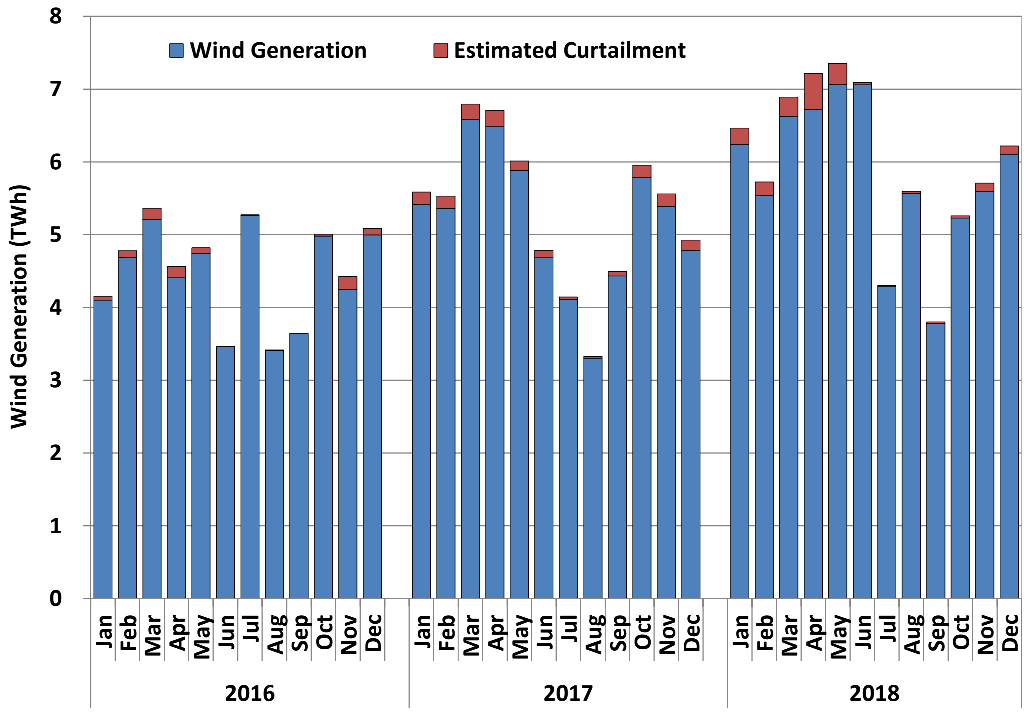
Figure 69: Summer Wind Production vs. Load



The typical profile for wind units not located along the coast or in the Panhandle is negatively correlated with peak electricity demand. However, output from wind generators located in the coastal area of the South zone is much more highly correlated with peak electricity demand. Panhandle wind shows a more stable output across the load levels.

Figure 70 shows the wind production and estimated curtailment quantities for each month of 2016 through 2018.

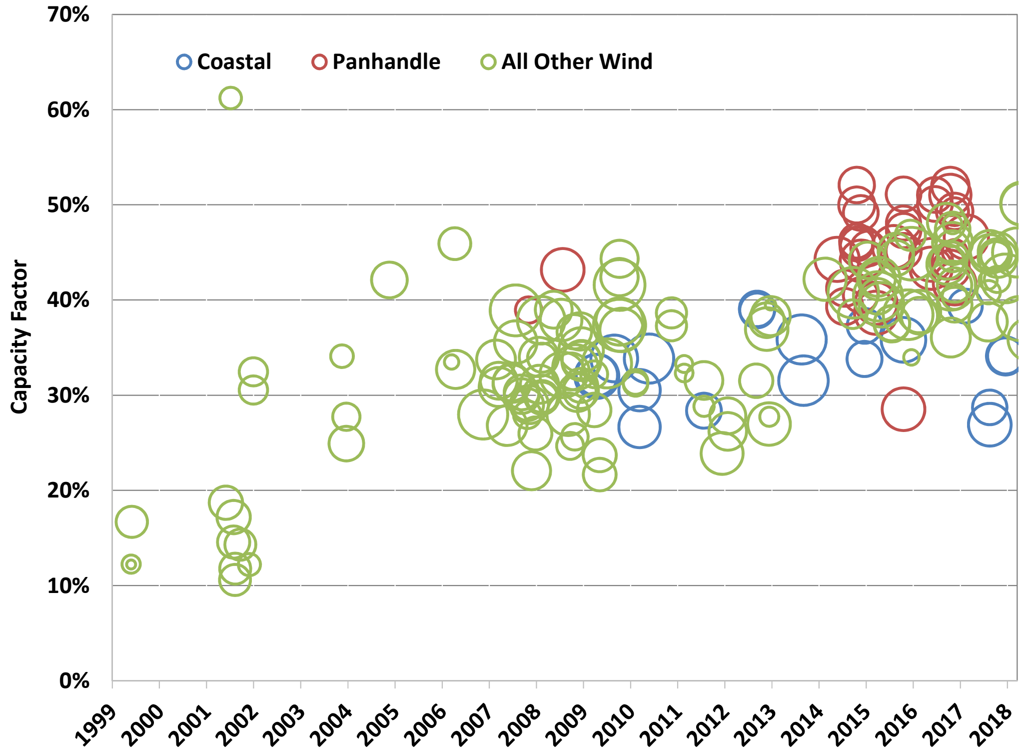
Figure 70: Wind Production and Curtailment



This figure reveals that the total production from wind resources continued to increase, while the quantity of curtailments also increased. The volume of wind actually produced in 2018 was estimated at 97.4% of the total available wind, continuing the small, but steady decline from 99.5% in 2014. As a comparison, in 2009, the year with the most wind curtailment, the amount of wind delivered was only 83%.

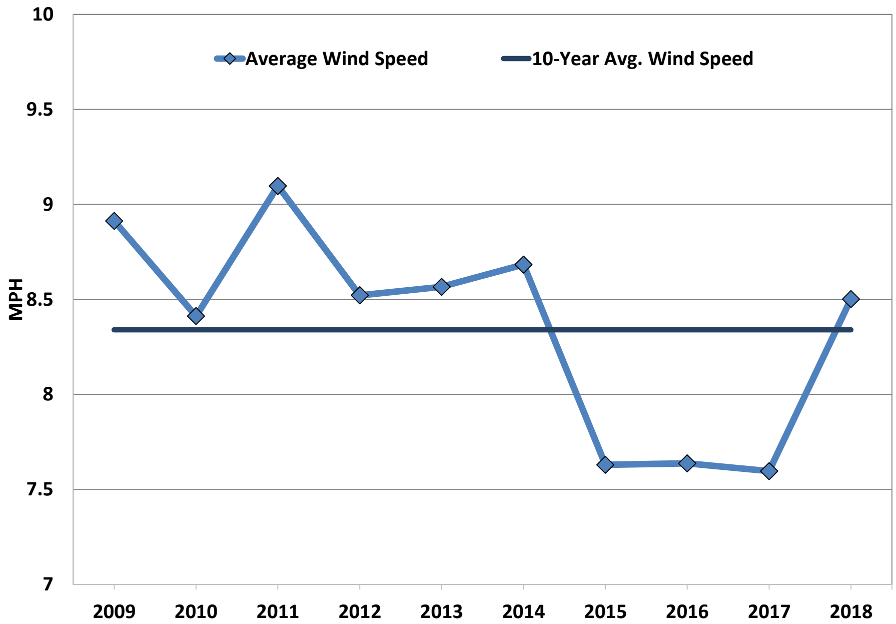
Figure 71 shows the capacity factor and relative size for wind generators by year installed. The chart also distinguishes wind generation units by location, with coastal units in blue and Panhandle resources in red, because of the different wind profiles for these regions. Capacity factors of coastal wind were consistent in 2018, and as previously described, output for coastal wind was generally more coincident with summer peak loads. Completion of the Competitive Renewable Energy Zones (CREZ) transmission lines has enabled more wind units to locate in the windier Panhandle and also facilitate more output from wind units in the west. Transmission maintenance for some 345 kV transmission lines had the effect of limiting output from some of the resources in the Panhandle, reducing their capacity factors. The figure shows a trend toward greater capacity factors for newer units.

Figure 71: Wind Generator Capacity Factor by Year Installed



As more wind generation capacity is installed in ERCOT, more energy from that capacity will be produced. However, the amount of energy produced will vary depending on actual wind speeds, which can vary from year to year. The next figure shows the average wind speed in ERCOT, as weighted by the locations of current installed wind generation. Figure 72 provides a means to compare wind speeds on an annual basis and indicates that the average wind speed in 2018 was slightly higher than the average over the past 10 years.

Figure 72: Historic Average Wind Speed



Net load is the total system load minus wind production. Figure 73 shows the net load duration curves for the years 2011, 2014, and 2018. Increasing wind output has important implications for the net load served by non-wind resources, reducing the energy available for them to serve, while not offering much contribution to serving peak load. This has important implications for resource adequacy in the ERCOT region as growth in peak demand requires additional resources to be added, but the energy available to be served is reduced.

Figure 73: Net Load Duration Curves

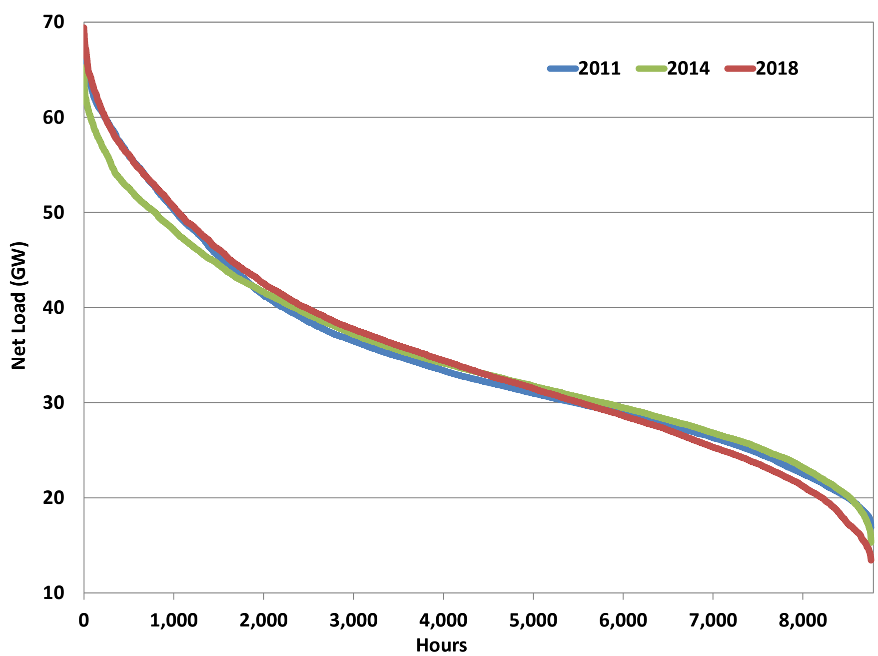
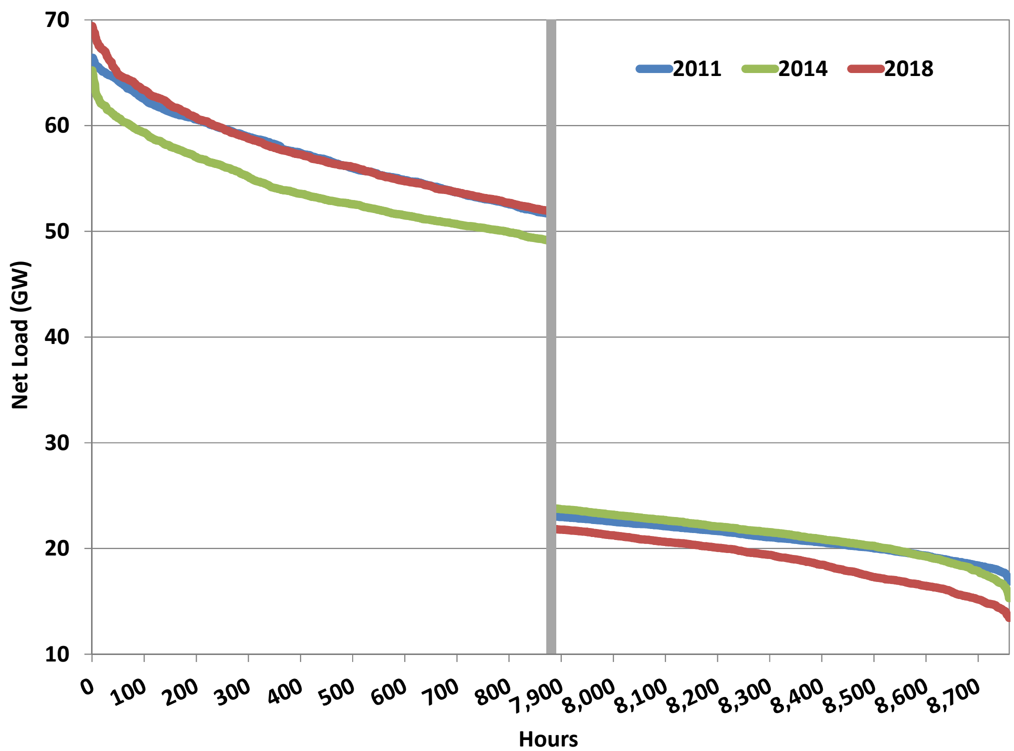


Figure 74 shows net load in the highest and lowest hours. Even with the increased development activity in the coastal area of the South zone, 71% of the wind resources in the ERCOT region are located in West Texas. The wind profiles in this area are such that most of the wind production occurs during off-peak hours or other times of low system demand. These profiles result in only modest reductions of the net load relative to the actual load during the highest demand hours, but much larger reductions in the net load in the other hours of the year. Wind generation erodes the total load available to be served by base load coal units, while doing very little to reduce the amount of capacity necessary to reliably serve peak load.

In the hours with the highest net load (left side of the figure below), the difference between peak net load and the 95th percentile of net load has averaged 12.1 GW the past three years. This means that 12.1 GW of non-wind capacity is needed to serve load less than 440 hours per year.

Figure 74: Top and Bottom Deciles (Hours) of Net Load

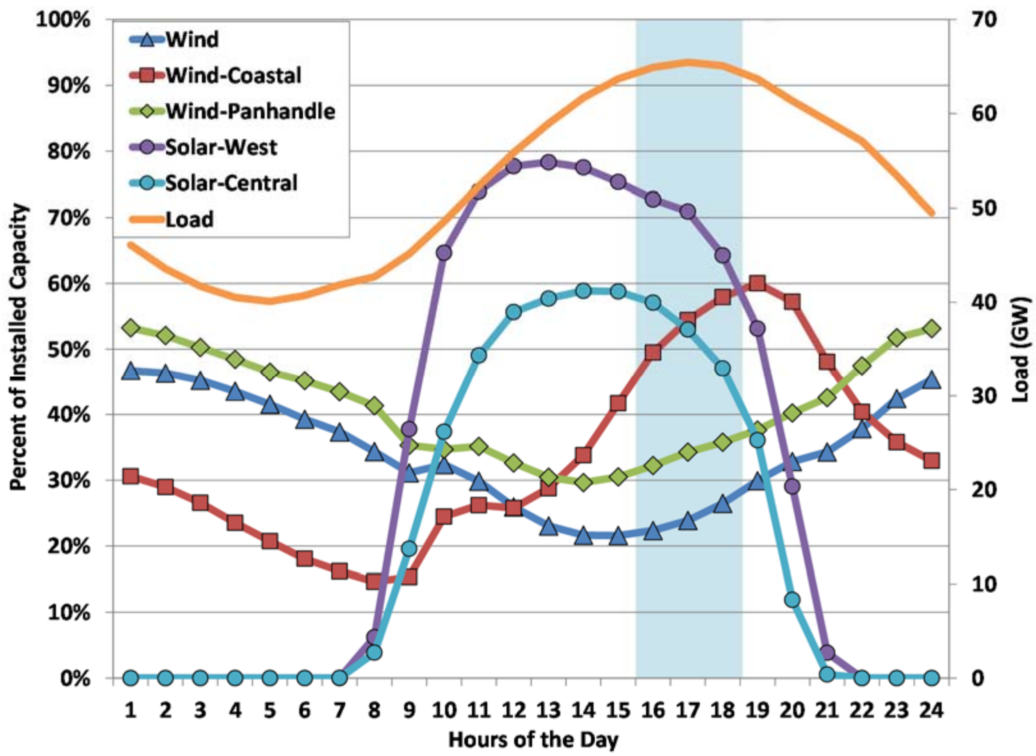


In the hours with the lowest net load (right side of the figure above), the minimum net load has dropped from approximately 20 GW in 2007 to below 13.4 GW in 2018, even with the sizable growth in annual load that has occurred. This trend has put operational pressure on the almost 20 GW of nuclear and coal generation that were in-service in 2018. This operational pressure was certainly one of the contributors to the recent retirement of more than 5 GW of coal in 2018.

Thus, although the peak net load and reserve margin requirements are projected to continue to increase and create an increasing need for non-wind capacity to satisfy ERCOT's reliability requirements, the non-wind fleet can expect to operate for fewer hours as wind penetration increases. This outlook further reinforces the importance of efficient energy pricing during peak demand conditions and other times of system stress, particularly in the context of the ERCOT energy-only market design.

The growing numbers of solar generation facilities in ERCOT have an expected generation profile highly correlated with peak summer loads. Figure 75 compares average summertime (June through August) hourly loads with observed output from solar resources, both in the West and Central zones, as well as wind resources. Generation output is expressed as a ratio of actual output divided by installed capacity.

Figure 75: Summer Renewable Production



This figure shows that solar production as a percentage of installed capacity is the highest in the early afternoon, approaching 80% for facilities located in the west and 60% for those in central Texas, and producing almost 70% and 55%, respectively, of its installed capacity during peak load hours. Even though the amount of solar generation is currently much less than the amount of wind generation, large amounts of solar anticipated to be installed in ERCOT will be a large contributor to serving peak demand.

The contrast between coastal wind and all other wind is also clearly displayed in Figure 75. Coastal wind produced 50 to 60% of its installed capacity during summer peak hours. Output from Panhandle wind and all other wind (primarily West zone) was less than 40% during summer peak hours.

D. Demand Response Capability

Demand response is a term that broadly refers to actions that can be taken by end users of electricity to reduce load in response to instructions from ERCOT or in response to certain market or system conditions. The ERCOT market allows participants with demand-response capability to provide energy and reserves in a manner similar to a generating resource. The ERCOT Protocols allow for loads to actively participate in the ERCOT-administered markets as load resources. A second way that loads may participate is through ERCOT-dispatched reliability programs, including Emergency Response Service (ERS) and legislatively-mandated demand response programs administered by the transmission and distribution utilities in their energy efficiency programs. Additionally, loads may self-dispatch by adjusting consumption in response to energy prices or by reducing consumption during specific hours to lower transmission charges.

Reserve Markets

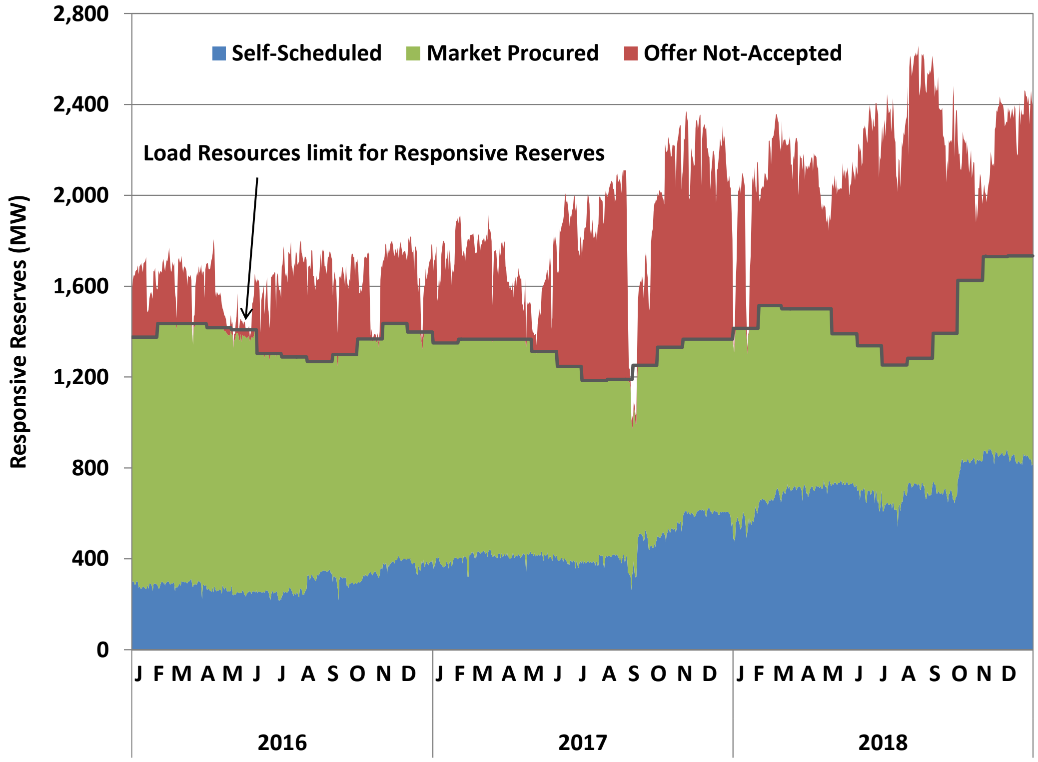
ERCOT allows qualified load resources to offer responsive reserves into the day-ahead ancillary services markets. Tripping load has the effect of increasing system frequency and can be a very effective mechanism for maintaining system frequency at 60Hz. Load resources providing responsive reserves have high set under-frequency relay equipment, which enables the load to be automatically tripped when the system frequency falls below 59.7 Hz. These events typically occur only a few times each year. As of December 2018, approximately 5,064 MW of qualified load resources were capable of providing responsive reserve service, an increase of approximately 350 MW during 2018.

On June 1, 2015, ERCOT began procuring a variable amount of responsive reserve service based on season and time of day. ERCOT established equivalency ratios at this time, to better ascertain the amount of primary frequency response expected from the procurement of responsive reserves. In 2016, the first full year with variable procurement, the quantity of megawatts offered but not accepted by load resources increased. During 2016, there were no system-wide manual deployments of load resources providing responsive reserves. There was, however, one automatic deployment of 927 MW of frequency responsive load on May 1, 2016.

In 2018, the total amount of responsive reserves procured by ERCOT varied between 2,300 MW and 3,200 MW per hour. During 2018, there was one automatic deployment of 546 MW load resources providing responsive reserve service, which occurred on April 21st.

Figure 76 below shows the average amount of responsive reserves provided from load resources on a daily basis for the past three years.

Figure 76: Daily Average of Responsive Reserves Provided by Load Resources



Prior to June 1, 2018, load resources were limited to providing a maximum of 50% of responsive reserves. The implementation of NPRR815 allows load resources to now provide up to 60% of the responsive reserve obligation, and also requires that at least 1,150 MW of responsive reserves be provided from generation resources.[27] The quantity of load providing responsive reserve offers submitted exceeded the limit for most of 2018. One exception is when real-time prices are expected to be high. Because load resources provide capacity by reducing consumption, they have to be consuming energy to be eligible to provide the service. During periods of expected high prices, for example brief periods in January 2018, the price paid for energy can exceed the value received from providing responsive reserves. Reduced offer quantities observed during the spring and fall months may reflect the lack of availability of load resources due to annual maintenance at some of the larger load resource facilities. ERCOT Protocols also permit load resources to provide non-spinning reserves and regulation services, but for a variety of reasons, load resources have participated only minimally in providing these services.

Reliability Programs

There are two main reliability programs in which demand can participate in ERCOT, ERS and load management programs offered by the transmission and distribution utilities. The ERS

[27] *See* NPRR815: *Revise the Limitation of Load Resources Providing Responsive Reserve (RRS)* Service (http://www.ercot.com/mktrules/issues/NPRR815).

program is defined by a Commission rule enacted in March 2012, which set a program budget of $50 million.[28] The program was modified from a pay-as-bid auction to a clearing price auction in 2014, providing a clearer incentive to load to submit offers based on the costs to curtail, including opportunity cost. In 2016, the procurement for ERS shifted from four time periods per contract term to six time periods per contract term. The additional time periods were created to separate the higher risk times of early morning and early evening from the overnight and weekend hours. The time and capacity-weighted average price for ERS over the contract periods from February 2018 through January 2019 was $6.72 per MWh, similar to the outcome of $6.86 per MWh as the previous program year. For the first time since the inception of the program, this price was lower than the average price of $9.20 per MWh paid for non-spinning reserves in 2018. The average price for non-spinning reserves in 2017 was much lower at $3.18 per MWh. ERS was not deployed in either year.

Beyond ERS, there were slightly more than 250 MW of load participating in load management programs administered by transmission and distribution utilities in 2018.[29] Energy efficiency and peak load reduction programs are required by statute and Commission rule and most commonly take the form of load management, where participants allow electricity to selected appliances (typically air conditioners) to be curtailed.[30] These programs administered by transmission and distribution utilities may be deployed by ERCOT during a Level 2 Energy Emergency Alert (EEA).

Self-dispatch

In addition to active participation in the ERCOT market and ERCOT-dispatched reliability programs, loads in ERCOT can observe system conditions and reduce consumption accordingly. This response comes in two main forms. The first is by participating in programs administered by competitive retailers or third parties to provide shared benefits of load reduction with end-use customers. The second is through actions taken to avoid the allocation of transmission costs. Of these two methods, the more significant impacts are related to actions taken to avoid incurring transmission costs that are charged to certain classes of customers based on their usage at system peak.

---

28 *See* 16 TAC § 25.507.

29 *See* ERCOT 2018 Annual Report of Demand Response in the ERCOT Region (Mar. 2019) at 7, available at http://www.ercot.com/services/programs/load.

30 *See* PUCT Project 45675, *2016 Energy Efficiency Plans and Reports Pursuant to 16 TAC §25.181(n)*; SB 7. Section 39.905(a)(2) (http://www.capitol.state.tx.us/tlodocs/76R/billtext/html/SB00007F.htm).

For decades, transmission costs have been allocated on the basis of load contribution to the highest 15-minute system demand during each of the four months from June through September. This allocation mechanism is routinely referred to as four coincident peak, or 4CP. By reducing demand during peak periods, load entities seek to reduce their share of transmission charges. Transmission costs have doubled since 2012, increasing an already substantial incentive to reduce load during probable peak intervals in the summer.[31] ERCOT estimates that as much as 1700 MW of load were actively pursuing reduction during the 4CP intervals in 2018, an increase of about 200 MW from 2017.[32]

Voluntary load reductions to avoid transmission charges may be distorting prices during peak demand periods because the response is targeting peak demand rather than responding to wholesale prices. This was readily apparent in 2016 as there were significant load reductions corresponding to peak load days in June, July and September when real-time prices on those days were in the range of $25 to $40 per MWh. The trend continued in 2017, with significant reductions on peak load days in June, August and September when real-time prices were less than $100 per MWh. Even with higher prices in 2018, reductions were observed during June, July, and August at times with wholesale prices less than $40 per MWh.

Two factors in the ERCOT market continue to advance appropriate pricing actions taken by load in the real-time energy market. First, the initial phase of "Loads in SCED" was implemented in 2014, allowing controllable loads that can respond to 5-minute dispatch instructions to specify the price at which they no longer wish to consume. Although an important step, there are currently no loads qualified to participate in real-time dispatch. Second, the reliability adder, discussed in more detail in Section I: Review of Real-Time Market Outcomes, performs a separate pricing run of the dispatch software to account for the amount of load deployed, including ERS. Proposed changes to the calculation method of the reliability adder were discussed in 2018 in NPRR904, *Revisions to Real-Time On-Line Reliability Deployment Price Adder for ERCOT-Directed Actions Related to DC Ties*.[33]

---

31 See PUCT Docket No. 47777, Commission Staff's Application to Set 2018 Wholesale Transmission Service Charges for the Electric Reliability Council of Texas, Final Order (Mar. 29, 2018); PUCT Docket No. 46604, Commission Staff's Application to Set 2017 Wholesale Transmission Service Charges for the Electric Reliability Council of Texas, Final Order (Mar. 30, 2017); PUCT Docket No. 45382, Commission Staff's Application to Set 2016 Wholesale Transmission Service Charges for the Electric Reliability Council of Texas, Final Order (Mar. 25, 2016).

32 See ERCOT, 2018 Annual Report of Demand Response in the ERCOT Region (Mar. 2019) at 7, available at http://www.ercot.com/services/programs/load.

33 The primary flaw identified in the calculation method of the Real-Time On-Line Reliability Deployment Price Adder was that LDL relaxation made the price adders higher, even when the RUC-instructed Resource is being dispatched above LDL in the pricing run. The price adders fluctuated based on interval-to-interval changes in the system, including changes for Resources that were not RUC-instructed. HDL and LDL relaxation of Resources that were not RUC-instructed was intended to avoid ramp limitations that could exaggerate the pricing impacts of the out-of-market action.

V. Reliability Commitments

One of the important characteristics of any electricity market is the extent to which it results in the efficient commitment of generating resources. Under-commitment can cause apparent shortages in the real-time market and inefficiently high energy prices, while over-commitment can result in excessive start-up costs, uplift charges, and inefficiently low energy prices.

The ERCOT market does not include a mandatory centralized unit commitment process. The decision to start-up or shut-down a generator is made by the market participant. ERCOT's day-ahead market informs these decisions, but is only financially binding. That is, when a generator's offer to sell is selected (cleared) in the day-ahead market there is no corresponding requirement to actually start that unit. The generator will be financially responsible for providing the amount of capacity and energy cleared in the day-ahead market whether or not the unit operates. This decentralized commitment depends on clear price signals to ensure an efficient combination of units are online and available for dispatch. ERCOT, in its role as reliability coordinator, has the responsibility to commit units it deems necessary to ensure the reliable operation of the grid. Gaps exist between what individual resources, in aggregate, view as economic commitment and what ERCOT views as necessary to ensure the reliability of the region. In the event of these gaps, ERCOT uses its discretion to commit additional units to ensure reliability.

This section describes the evolution of rules and procedures regarding reliability unit commitments (RUC), and the outcomes of RUC activity in 2018. It also describes the Current Operating Plan data submitted by Qualified Scheduling Entities (QSEs) and used by ERCOT to determine the need for a RUC. Price mitigation that occurs during RUC and local congestion is described and the section concludes with a discussion of the Reliability Must Run (RMR) process revisions in ERCOT in 2018.

A. History of RUC-Related Protocol Changes

The RUC process has undergone several modifications since the nodal market began in 2010. The following changes were implemented in an effort to improve the commitment process and market outcomes associated with RUC. In March 2012, an offer floor was put in place for energy above the Low-Sustained Limit (LSL) for units committed through RUC.[34] Initially, the RUC offer floor was set at the system-wide offer cap. The RUC offer floor was subsequently

[34] *See* NPRR435, *Requirements for Energy Offer Curves in the Real Time SCED for Generation Resources Committed in RUC*, implemented on March 1, 2012.

adjusted to $1,000 per MWh[35] and then to the current offer floor of $1,500 per MWh.[36] Resources committed through the RUC process receive a make-whole payment and forfeit market revenues through a "clawback" provision. Beginning on January 7, 2014, resources committed through the RUC process could forfeit the make-whole payments and waive the clawback charges, effectively self-committing and accepting the market risks associated with that decision.[37] This buyback or "opt-out" mechanism for RUC initially required a resource to update its Current Operating Plan (COP) before the close of the adjustment period for the first hour of a RUC.

On June 25, 2015, ERCOT automated the RUC offer floor of $1,500 per MWh and implemented the Real-Time On-Line Reliability Deployment Adder (reliability adder).[38] ERCOT systems now automatically set the energy offer floor at $1,500 per MWh when a resource properly telemeters a status indicating it has received a RUC instruction. The reliability adder, as discussed more in Section I: Review of Real-Time Market Outcomes, captures the impact of reliability deployments such as RUC on energy prices.

The RUC process was modified again in 2017. On June 1, 2017, ERCOT began using a telemetered snapshot at the start of each RUC instruction block as the trigger to calculate the reliability adder. This was an improvement over the previous calculation trigger, which required the QSE to accurately telemeter an ONRUC status.[39] Resources now have the ability to opt-out of RUC instructions after the close of the adjustment period, because the opt-out decision is no longer communicated via the COP.

Resources are also now permitted to opt out of RUC instructions via real-time telemetry; opting out of a RUC instruction is available for resources that telemeter ONOPTOUT during the first dispatchable interval within the first RUC-hour of the commitment block instruction. During 2018, approximately 54% of RUC instructions were given after the close of the adjustment period compared to 28% of instructions issued after the close of the adjustment period in 2017.

---

35 NPRR568, *Real-Time Reserve Price Adder Based on Operating Reserve Demand Curve*, implemented on June 1, 2014.

36 NPRR626, *Reliability Deployment Price Adder*, partially-implemented to update the RUC offer floor on October 1, 2014.

37 NPRR416, *Creation of the RUC Resource Buyback Provision* (formerly "Removal of the RUC Clawback Charge for Resources Other than RMR Units"), as modified by NPRR575, *Clarification of the RUC Resource Buy-Back Provision for Ancillary Services*.

38 NPRR626, *Reliability Deployment Price Adder* (Formerly "ORDC Price Reversal Mitigation Enhancements").

39 NPRR744, *RUC Trigger for the Reliability Deployment Price Adder and Alignment with RUC Settlement*, implemented on June 1, 2017.

The most significant change to the RUC process in 2018 was the approval and implementation of NPRR864, *RUC Modifications to Consider Market-Based Solutions*, which modified the RUC engine to consider fast-start generators (start time of 1 hour or less) as self-committed for future hours.[40] The culmination of these RUC changes has enabled ERCOT to defer supplementary commitment decisions, allowing market participants full opportunity to make their own unit commitment

Several other proposed Protocol revisions were initiated in 2018, including a Board-approved modification that will allow ERCOT systems to adjust the configuration of a self-committed combined cycle unit, to increase available capacity.[41] Further, market participants considered a proposal in 2018 that would address the situation of a QSE receiving a RUC instruction for a unit that had received a DAM award but electing to not actually operate. In this situation the resource should not receive RUC settlement treatment.[42] And finally, market participants also began considering a proposal to address the settlement of switchable generation resources receiving an instruction to switch from a non-ERCOT Control Area to the ERCOT Control Area.[43]

B. RUC Outcomes

ERCOT continually assesses the adequacy of market participants' resource commitment decisions using the RUC process, which executes both on a day-ahead and hour-ahead basis. Additional resources may be determined to be needed for two reasons – to satisfy the total forecasted demand, or to make a specific generator available resolve a transmission constraint. The transmission constraint may be either a thermal limit or a voltage concern.

The number of RUC instructions in 2018 increased somewhat from 2017. The 642 unit-hours of RUC instructions in 2018 represent a 14% increase from the 562 unit-hours in 2017. These 2018 RUC instructions were geographically diverse as well, with 16% to generators in the South zone in a variety of locations: San Antonio, Corpus Christi, and the Rio Grande Valley, 2% were to generators in the Houston zone, 24% were to generators in the North zone, and the remaining 58% were to generators in the West zone.

As in 2017, most reliability commitments in 2018 were made primarily to manage transmission constraints (80% of unit-hours). Only 19% of RUC instructions were made to ensure sufficient

40 NPRR864, *RUC Modifications to Consider Market-Based Solutions*, implemented on October 26, 2018; see also NPRR875, *Clarification for the Implementation of NPRR864, RUC Modifications to Consider Market-Based Solutions*.

41 NPRR884, *Adjustments to Pricing and Settlement for Reliability Unit Commitments (RUCs) of On-Line Combined Cycle Generation Resources*, approved on December 11, 2018.

42 NPRR910, *Clarify Treatment of RUC Resource that has a Day-Ahead Market Three-Part Supply Award*.

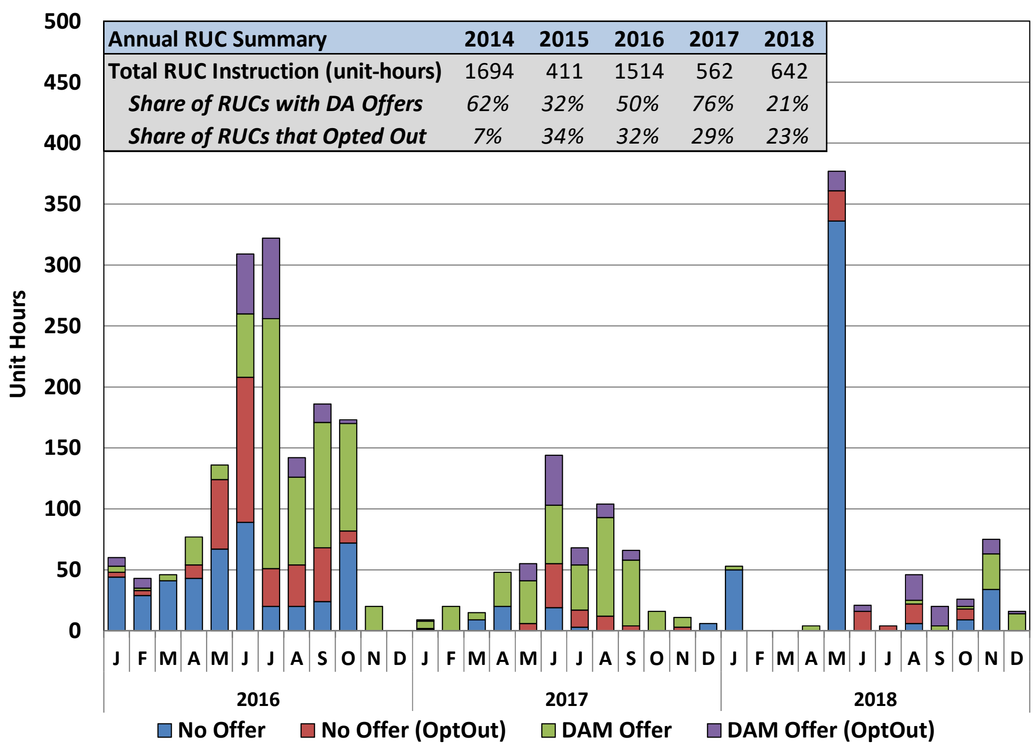
43 NPRR912, *Settlement of Switchable Generation Resources (SWGRs) Instructed to Switch to ERCOT*.

system-wide capacity and 1% were for voltage support. Most of the RUC instructions in 2018 occurred during a single month. Congestion in the Permian Basin area of west Texas related to higher than expected load, planned and forced outages led to 377 unit-hours of RUC instructions during May.

RUC activity in previous years has been the result of specific, typically short-term system needs. Looking back at the primary drivers for RUC activity in previous years no single or long-standing cause is apparent. In 2017, construction of the Houston transmission import project and the necessity to maintain stability during outages immediately following Hurricane Harvey in 2017 were primary drivers. The high amount of RUC activity in 2016 was due to localized transmission congestion for units located in Houston and the Valley. In 2015, RUCs were most frequent in the fall because of congestion in Dallas and the Valley. RUC activity in 2014 was concentrated during cold weather events in February and March and in response to transmission outages in March and November.

Figure 77 below shows RUC activity by month for 2016 through 2018, indicating the volume of generators receiving a RUC instruction that had offers in the day-ahead market or chose to opt-out of the RUC instruction. The monthly data shows no consistent pattern of RUC activity over the past three years. Annual summaries are also provided going back to 2014, the year with the highest amount of RUC activity.

Figure 77: Day-Ahead Market Activity of Generators Receiving a RUC



Annual RUC Summary	2014	2015	2016	2017	2018
Total RUC Instruction (unit-hours)	1694	411	1514	562	642
Share of RUCs with DA Offers	*62%*	*32%*	*50%*	*76%*	*21%*
Share of RUCs that Opted Out	*7%*	*34%*	*32%*	*29%*	*23%*

A unit receiving a RUC instruction is guaranteed payment of its start-up and minimum energy costs (RUC make-whole payment). However, if the energy payments received by a unit operating under a RUC instruction exceed its operating costs, payment to that generator is reduced (RUC clawback charge). Generators without offers submitted to the day-ahead market forfeit all excess revenues, whereas generators with day-ahead offers forfeit only 50% of excess revenues. Given this incentive to have offers submitted in the day-ahead market, it is somewhat surprising that all units do not submit day-ahead offers. In 2018, only 21% of the total RUC unit-hours had day-ahead offers. This is lower than expected, considering the incentive to provide day-ahead offers inherent in the RUC claw-back rules. The percentage of RUC unit-hours with day-ahead offers was much higher in 2017 and 2016 at 76% and 50%, respectively. The very low value in 2018 may be explained by the large number of fast starting generators receiving RUC instructions, primarily during May due to higher than expected load, planned and forced outages. Sixty percent of the unit-hour instructions in 2018 were for fast starting generators, whereas, since 2014 a more typical share has been 15%. It is not unusual for the decision to commit fast starting units to be made in real-time.

Since January 2014, a generator receiving a RUC instruction has had the choice to "opt out," meaning it forgoes all RUC make-whole payments in return for not being subject to RUC clawback charges. The percentage of generators receiving RUC instructions in 2018 that chose to opt-out was 23%, similar to the 29% of generators that chose to opt-out in 2017.

There were 494 hours in which units were settled as RUC in 2018. The amount of time with non-zero reliability adders was much less (233 hours). There were 120 unit-hours of RUC instructions to ensure system-wide adequacy in 2018, representing 19% of the total RUC-hour instructions. This is an increase from 73 unit-hours in 2017 (13% of the total) and, 33 unit-hours (2% of the total) in 2016. The increased occurrence of RUC instructions for system-wide adequacy is not unexpected given ERCOT's decreasing levels of reserve capacity.

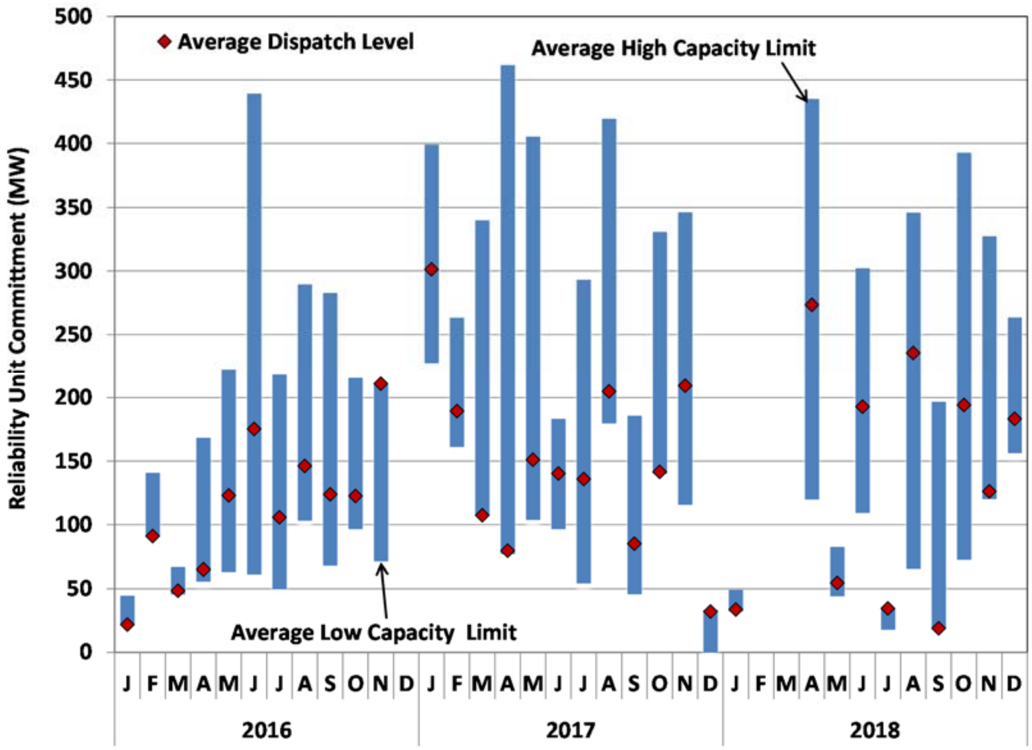
Table 8 lists the generation resources that received the most RUC instruction in 2018 and includes the total hours each unit was settled as ONRUC and the number of hours in which the unit opted-out. The units highlighted in gray are the ones that frequently received RUC instructions in 2017. Frequent RUC instructions were issued to the Permian Basin CTs in May due to localized transmission congestion related to high loads and line outages. Although there were 327 unit-hours of RUC instructions issued to the Permian CTs, these occurred in only 262 hours. In many hours more than one CT was necessary.

Table 8: Most Frequent Reliability Unit Commitments

Resource	Location	Unit-RUC Hours	Unit OPTOUT Hours	Average LSL during Dispatchable Hours	Average LDL during Dispatchable Hours	Average Dispatch during Dispatchable Hours	Average HSL during Dispatchable Hours
Permian CT 3	Far West	104	0	41	41	45	67
Permian CT 4	Far West	103	0	41	41	45	66
Permian CT 5	Far West	76	7	41	41	49	68
Permian CT 1	Far West	44	4	41	41	47	65
Silas Ray CC1	Valley	34	0	45	45	46	54
Lake Hubbard Unit 2A	DFW	16	11	45	229	255	465
Silas Ray 10	Valley	22	4	18	18	21	40
Handley Unit 4	DFW	15	4	120	147	169	435
Mountain Creek Unit 6	DFW	0	16	15	49	53	122
WA Parish G2	Houston	0	15	54	95	107	161
Stryker Unit 2	DFW	13	2	47	66	72	502
Handley Unit 5	DFW	0	13	120	277	314	432
Olinger 3	DFW	0	12	25	25	25	138
Nueces Bay CC1	Corpus Christi	2	8	157	185	201	383
Duke CC1	Valley	8	2	177	187	197	235

The next analysis compares the average dispatched output of the reliability-committed units, including those that opted-out, with the average operational limits of the units. Figure 78 shows that the monthly average dispatch of units receiving RUC instructions has rarely been close to the average high capacity limit. This figure shows that the average quantity dispatched exceeded 100 MW only six months in 2018, and that no RUC activity occurred in February and March. In April the average dispatch level was more than 200 MW because of mitigation of the ONRUC resource. The other months of June, August, and October were close to 200 MW due to RUC resources choosing to opt-out and thus not having a $1500 per MWh offer floor.

Figure 78: Reliability Unit Commitment Capacity



Units committed for RUC in 2018 showed a slight decrease in the dispatch level compared to prior years. In 27% of intervals with RUC resources, one or more resources were dispatched above their low dispatchable limit (LDL), whereas in prior years, resources receiving a RUC were infrequently dispatched above LDL. This higher dispatch level indicates that most units receive RUC instructions to resolve local constraints, and that these local constraints are non-competitive. As a result, units are paid based on their mitigated offer caps. It is rare for a generator receiving a RUC instruction to be dispatched above LDL with its offer at or above the $1,500 per MWh offer floor. In 2018, this occurred in only 2% of the intervals with an ONRUC resource.

When a unit is committed for RUC, the unit will receive a make-whole payment if the real-time revenues received are less than the costs incurred to commit the unit. These costs can be based on generic values or unit-specific verifiable costs. Of the 33 different resources that received a RUC instruction in 2018, 30 resources had approved unit-specific verifiable costs for start-up costs and minimum load costs. Those 30 resources represent 95% of total RUC-instructed megawatt-hours in 2018.

Table 9 displays the total annual amounts of make-whole payments and claw-back charges attributable to RUCs since 2011. There are two sources of funding for RUC make-whole payments. The first is from QSEs that do not provide enough capacity to meet their obligations. If there are remaining RUC make-whole funds required after contributions from any capacity short QSEs, any remaining RUC make-whole funding will be uplifted to all QSEs on a load-ratio share.

Table 9: RUC Settlement

	Claw-Back from Generator in millions	Make-Whole to Generator in millions
2011	$8.5	$27.8
2012	$0.3	$0.4
2013	$1.1	$2.9
2014	$2.8	$3.8
2015	$0.3	$0.5
2016	$1.4	$1.2
2017	$1.2	$0.5
2018	$3.1	$0.6

As stated above, if real-time revenues received by a RUC resource exceed the operating costs incurred by the unit, then excess revenues are clawed-back and returned to QSEs representing load. During 2018, more than $3.1 million was clawed-back from RUC units while only $0.6 million in make-whole payments were made to RUC units. RUC make-whole payments in 2018 were collected almost exclusively from QSEs that were capacity short. The amount of make-whole that was uplifted was de minimis. The magnitude of both the claw-back and make-whole amounts are very small compared to the size of the ERCOT real-time energy market.

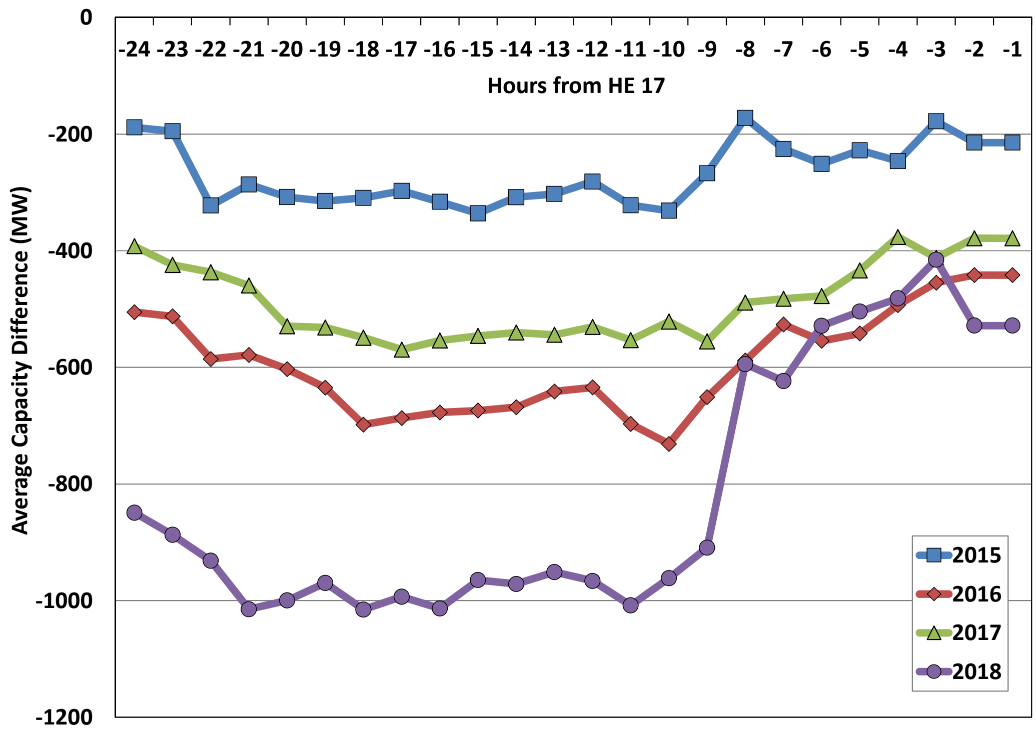
C. QSE Operation Planning

One of the important characteristics of any electricity market is the extent to which it results in the efficient commitment of generating resources. Under-commitment can cause shortages in the real-time market and inefficiently high energy prices; while over-commitment can result in excessive start-up costs, uplift charges, and inefficiently low energy prices. The COP is the mechanism used by QSEs to communicate the expected status of their resources to ERCOT. After aggregating COP information about the amount of capacity that QSEs expect to be on line every hour, ERCOT then evaluates any potential locational or system-wide capacity deficiency. If such a deficiency is identified, ERCOT will issue a RUC instruction to ameliorate the shortfall. The accuracy of COP information greatly affects ERCOT's ability to effectively perform supplemental commitment using the RUC process.

COPs are updated an ongoing basis by QSEs, providing multiple views of their expectations for a particular operating hour. Presumably, QSE expectations about which units will be online in a particular hour are most accurate for the COP submitted just before the operating hour. Figure 79 shows the average difference between the actual online unit capacity in the peak hour and the amount of capacity planned to be online in the peak hour as submitted each of the 24 hours leading up to the close of the adjustment period. This data is derived from COP submissions and averaged for hour ending 17 in the months of July and August, for each year 2015 through 2018. The figure shows that the amount of capacity needed to serve peak load has increasingly been committed in the final hour. In 2015 about 200 MW of capacity was committed in the last hour before real-time, and in 2016, the amount increased to more than 420 MW, with even larger deficiencies seen in the hours leading up to real-time. The increase in self-committed capacity seen during 2017 may have been a reaction to the increased RUC activity observed in 2016. However nearly 400MW remained to be committed in the last hour during 2017. In 2018, the average amount of capacity committed in the last hour exceeded 500 MW. Compared to the average load of 66 GW in hour 17, the amount of last minute self-commitment is just less than one percent of total capacity needs.

As previously described, only a small portion of total RUC instructions were issued to ensure system-wide capacity sufficiency. This is testament to the restraint exhibited by ERCOT operators to allow market participants to make their own commitment decisions with regard to the nearly 500 MW of close-to real-time capacity commitments. The fact that there is nearly 5,000 MW of fast starting generators controlled by multiple market participants highlights the complexity of these decisions and suggests that improvements to these close-to-real-time commitments may be warranted.

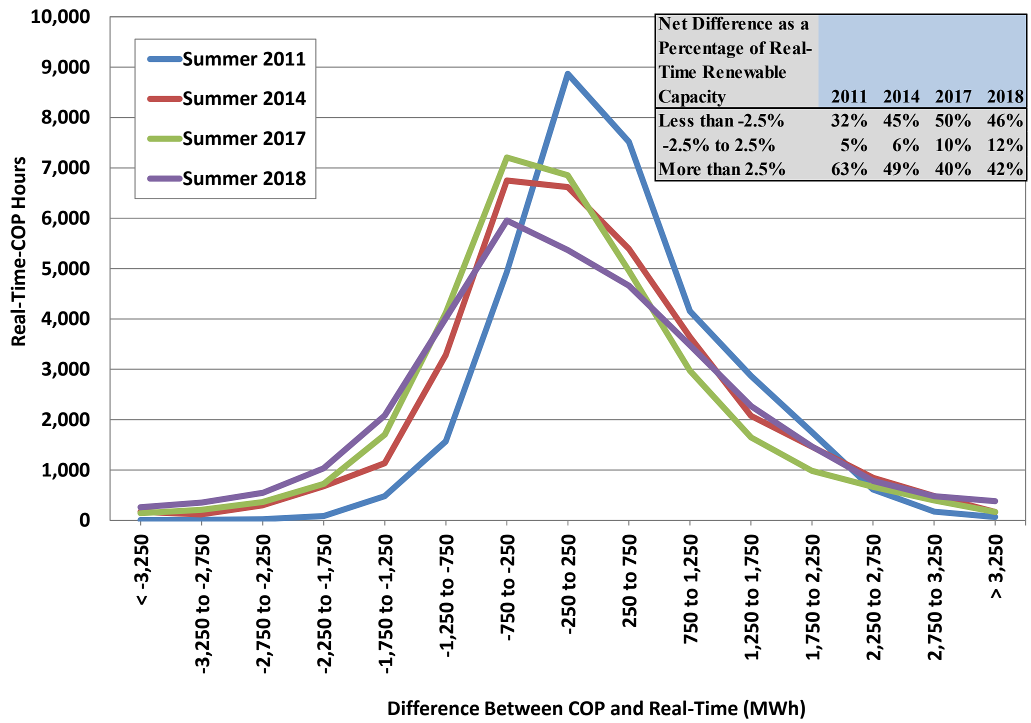
Figure 79: Capacity Commitment Timing – July and August Hour Ending 17



The next set of analyses quantify the amount of difference between the aggregated capacity commitments as described by all the COP submissions, and the actual capacity commitments observed in real-time. These analyses are limited to the peak hours of 7 through 22 for the summer months of June, July, and August. Multiple COP submissions provide data for each of the hours being evaluated, and there can be large variations in unit commitment expectations reflected in those multiple COPs, even for the same operating hour. Because unit commitment decisions for renewable resources are influenced by much different factors than other types of resources, the first analysis focuses only on renewable resources.

Figure 80 summarizes the difference between the amount of capacity expected to be online as shown in various views of the COP and capacity that actually was online in real-time for renewable resources. Renewable resources for this analysis are wind, solar, biomass, renewable, and power storage resources. For the on-peak hours analyzed, there were hours in which the difference between the various COP representations and actual capacity in real-time were as large as 3,250 MW, both positive (real-time capacity exceeded COP) and negative (real-time capacity was less than COP). The table in Figure 80 shows the aggregate renewable difference divided by the aggregate renewable capacity seen in the relevant operating hour to determine a percentage error.

Figure 80: Real-Time to COP Comparisons for Renewable Capacity

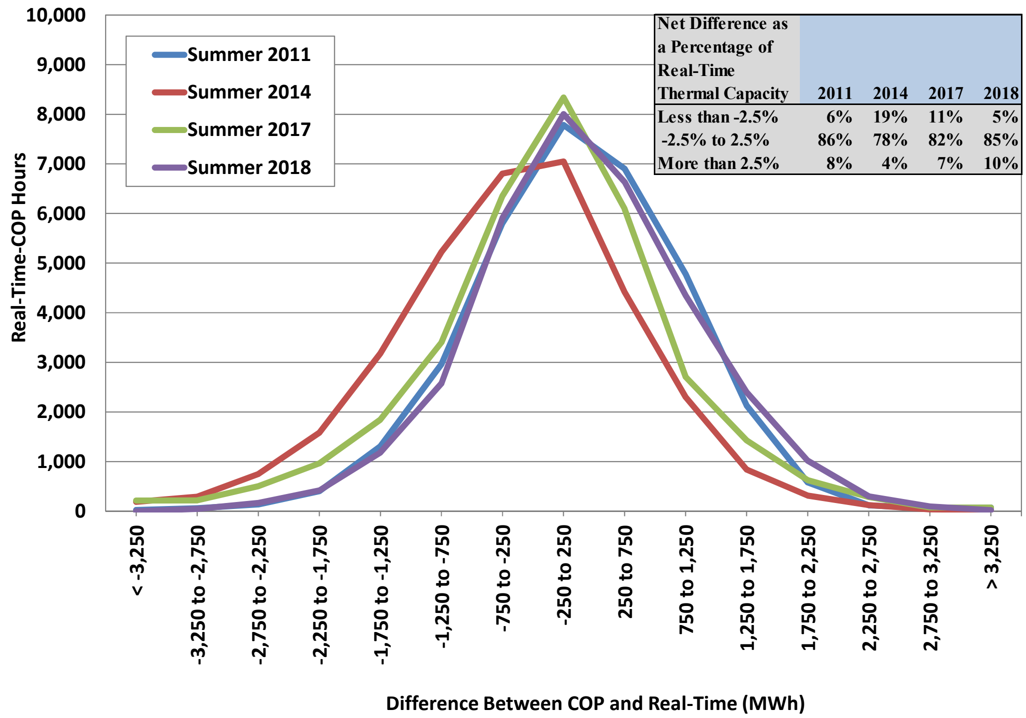


In 2018, only 12% of the aggregate renewable capacity shown in the COP was within 2.5% of the capacity seen in real-time by the same resources. The translation of the 2.5% threshold in renewable capacity could be a difference of as much as a 300 MW error. This small value is not indicative of any concern with specific market participants, because COP estimates for the preponderance of renewable capacity is determined based on an ERCOT developed forecast. Indeed, the almost symmetrical distribution of the errors observed in 2018 is evidence of ERCOT's capacity forecast being relatively unbiased.

This is in contrast to the values shown for 2011 when the real-time capacity for renewables tended to be higher than the 2.5% error 63% of the time: that is, the tendency was for more renewable resources to be operating in real-time than had been indicated via the COPs. The increase in low-error occurrences can be explained by continuous improvements that have been made to forecasting procedures by both ERCOT and their vendors. But even with these improvements, the COP submitted just before the operating hour were within 2.5% of the real-time MWh capacity only 16% of the time in 2018.

Figure 81 summarizes the same analysis for thermal capacity. Summer 2018 experienced zero hours where there was a net difference of less than 3,250 MWh between the COP capacity and thermal real-time capacity during the on-peak hours of hour ending 7 to hour ending 22.

Figure 81: Real-Time to COP Comparisons for Thermal Capacity



Net Difference as a Percentage of Real-Time Thermal Capacity	2011	2014	2017	2018
Less than -2.5%	6%	19%	11%	5%
-2.5% to 2.5%	86%	78%	82%	85%
More than 2.5%	8%	4%	7%	10%

The table in Figure 81 shows thermal resources were much more likely to indicate capacity within 2.5% of the respective operating hour at 85% of the summer on-peak hours. The last COP values for the on-peak hours ending 7 to 22 were within 2.5% of the real-time MWh capacity 94% of the time in 2018. The translation of the 2.5% threshold in thermal capacity could be a difference of as much as a 1,400 MW error. In contrast to the renewable capacity outside of the 2.5% band, this analysis indicates a bias for thermal resources to under report capacity in the COP at 10% versus over reporting capacity at 5% in 2018.

Specific to the last COP submitted for hours ending 7 to 22 in 2018, there was a 10% occurrence of the real-time-last-COP hours where a resource deviated by more than 10% of the HSL reported in the COP. Explanation for COP under-reporting includes resources that were currently on startup or shutdown before and after their operating periods, additional capacity available from power augmentation not being shown in the COP value, non-spin resources changing from OFFNS codes to an ON code, resources coming online responding to market activity, or combined cycle resources increasing their configuration size. Conversely, explanations for over reporting capacity would include resource derations, unexpected unit outages, offline due to market activity, transitioning from ON to OFFNS codes for non-spin service, or combined cycle resources decreasing their configuration size.

Figure 82: Standard Deviations of Real-Time to COP Capacity Differences

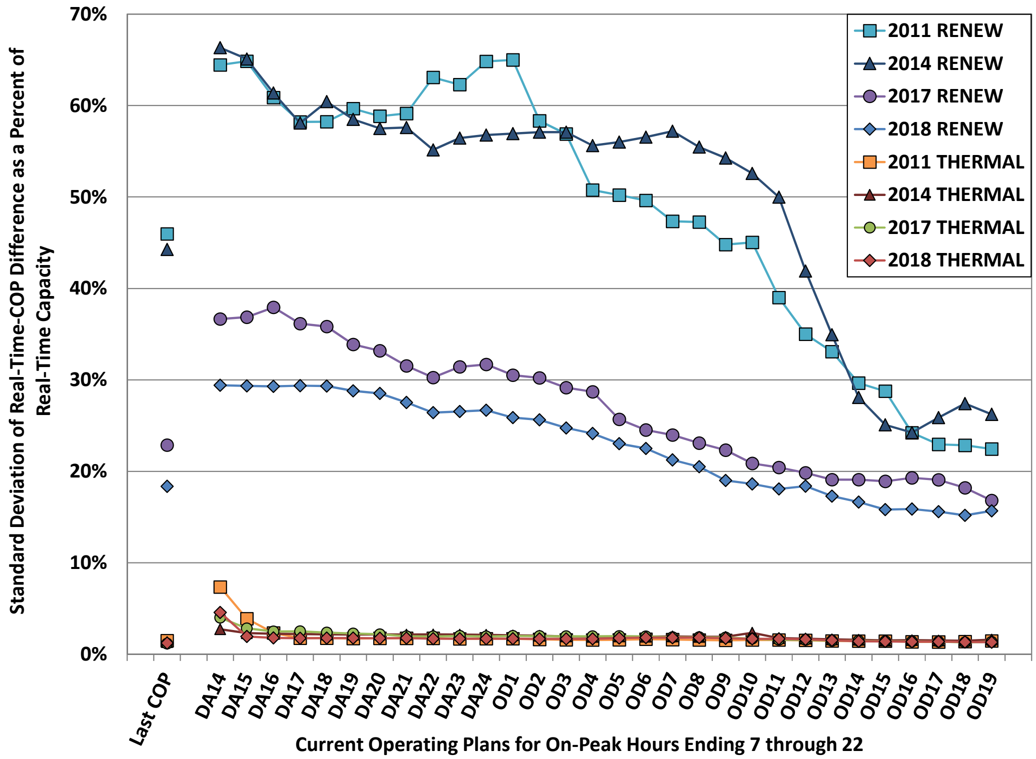


Figure 82 shows the error between COP and real-time dispatch capacity and how the error decreases as COP submittals get closer to real-time. Data for all relevant COPs is included, meaning that a COP submitted day-ahead at 1400 (DA14) includes data for all operating hours, 7 to 22. However, the COP submitted at 1900 on the operating day (OD19) only contains data for operating hour ending 21. The last COP is the last relevant COP after the adjustment period for all hours ending 7 through 22. The standard deviation for the OD19 is lower than the last COP standard deviation as there is minimal variability in net load for hour ending 21, whereas many large net load ramping hours would be included in the last COP. While the chart shows a larger disparity in the standard deviations based on percentages for thermal and renewable, the standard deviation in MW are between the bandwidth of 1,000 MW to 1,250 MW for both.

These previous analyses indicate the wide degree of uncertainty facing ERCOT operators as they fulfill their responsibilities to ensure sufficient capacity, in the right location, is available to meet real-time requirements. The current mechanism allowing resources to OPTOUT of a RUC instruction, which allows resources to forego the guaranty of start-up costs in exchange for the opportunity to keep all profits from energy revenues, provides the incentive for resources to defer their own commitment decision to the last possible moment at no risk. If the unit is required, ERCOT will issue a RUC instruction, which the resource will then chose to OPTOUT of. Having the ability to convert an ERCOT-issued commitment instruction into a self-commitment is expected to become more problematic with smaller reserve margins and less system flexibility.

D. Mitigation

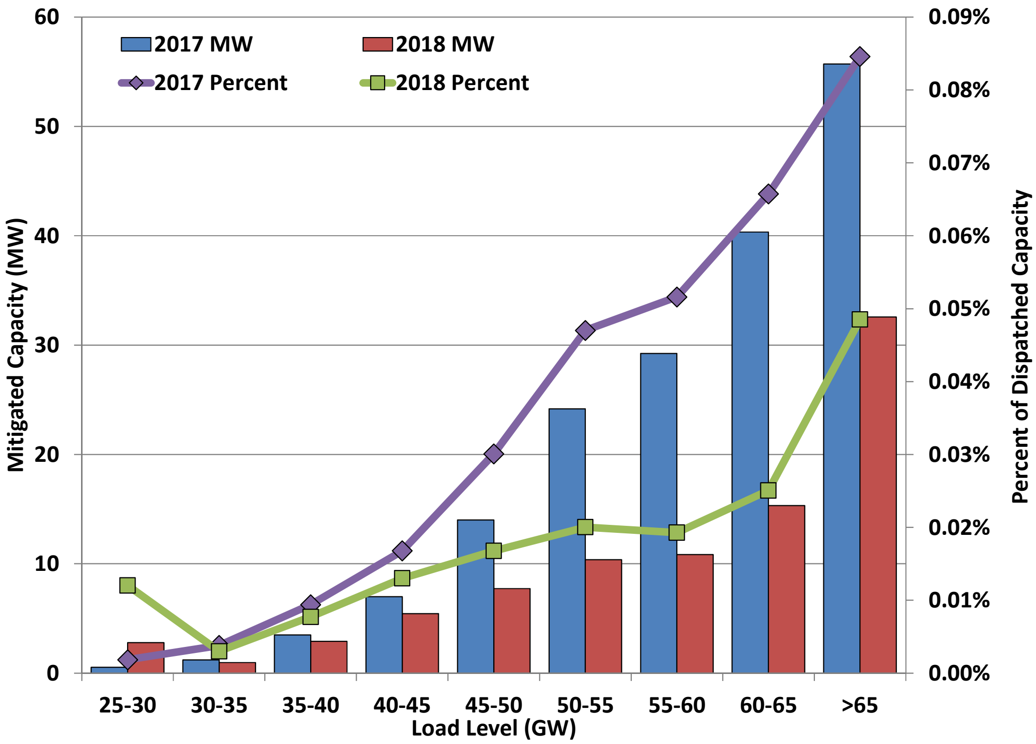
In situations where competitive forces are not sufficient, it is necessary to mitigate prices to a level that approximates competitive outcomes. ERCOT's real-time market includes a mechanism to mitigate prices for resources that are required to resolve a transmission constraint. Mitigation applies whether the unit is self-committed or receives a RUC instruction. Units typically receive a RUC instruction to resolve transmission constraints and as such are typically required to resolve a transmission constraint, and therefore mitigated. As shown previously in Figure 78, units that received a RUC instruction were frequently dispatched above their low operating limits in 2018. This higher dispatch was most often the result of the RUC units being dispatched based on their mitigated price, not based on the RUC offer floor of $1,500 per MWh.

ERCOT's dispatch software includes an automatic, two-step price mitigation process. In the first step, the dispatch software calculates output levels (Base Points) and associated locational marginal prices using the participants' offer curves and considers only the transmission constraints that have been deemed competitive. These "reference prices" at each generator location are compared with that generator's mitigated offer cap, and the higher of the two is used to formulate the offer curve to be used for that generator in the second step in the dispatch process. The resulting mitigated offer curve is used by the dispatch software to determine the final output levels for each generator, taking all transmission constraints into consideration.

This approach is intended to limit the ability of a generator to raise prices in the event of a transmission constraint that requires its output to resolve. In this subsection, the quantity of mitigated capacity in 2018 is analyzed. Although executing at all times, the automatic price mitigation aspect of the two-step dispatch process only has the potential to have an effect when a non-competitive transmission constraint is active. Since the introduction of an impact test in 2013 to determine whether units are relieving or contributing to a transmission constraint, only the relieving units are subject to mitigation. This change significantly reduces the amount of capacity subject to mitigation.

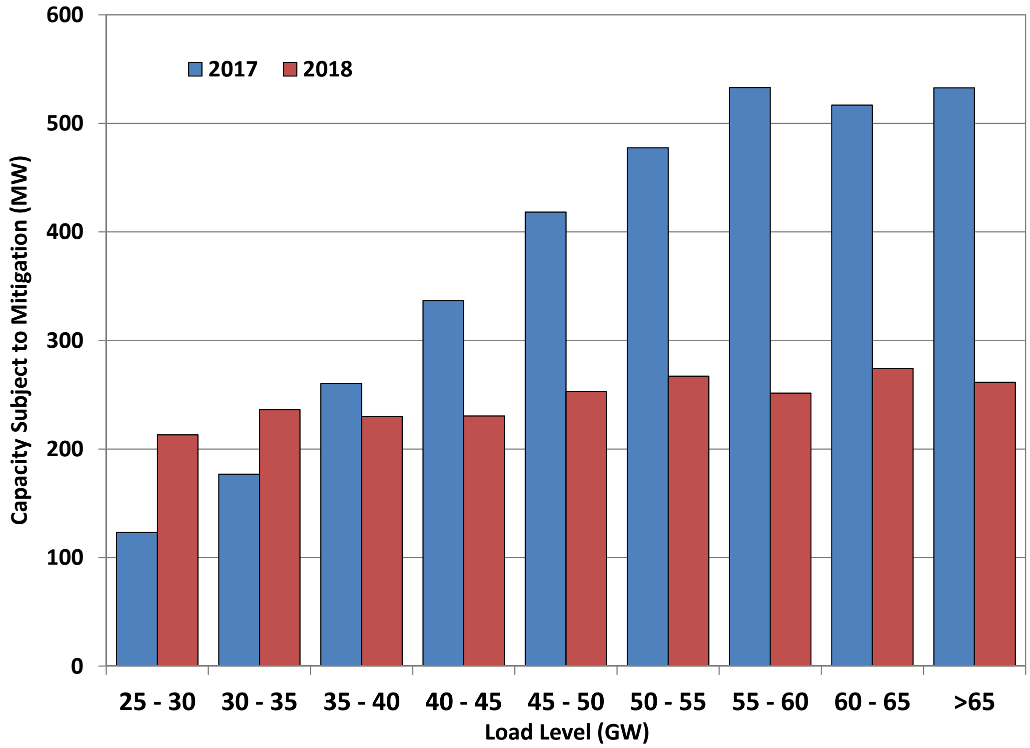
The analysis shown in Figure 83 computes the percentage of capacity, on average, that is actually mitigated during each dispatch interval. The results are provided by load level.

Figure 83: Mitigated Capacity by Load Level



The amount of mitigation in 2018 was generally lower than in 2017. This is somewhat unexpected given the similar frequency, although much higher cost of congestion. The reduction may be explained by the overall higher costs in 2018. Another factor may be the separation in natural gas prices between Houston Ship Channel and Waha fuel price indices. To the extent some generator offers (and costs) were based on very low Waha gas prices, there would be no need to mitigate them based on Houston Ship Channel prices. Only the amount of capacity that could be dispatched within one interval is counted as mitigated for the purpose of this analysis. The next analysis computes the total capacity subject to mitigation, by comparing a generator's mitigated and unmitigated (as submitted) offer curves and determining the point at which they diverge. The difference between the total unit capacity and the capacity at the point the curves diverge is calculated for all units and aggregated by load level. The results are shown in Figure 84.

Figure 84: Capacity Subject to Mitigation



As in the prior analysis, the amount of capacity subject to mitigation in 2018 was lower than 2017 in all but the lowest load levels. As described previously, the reduction may be explained by the overall higher costs in 2018 and the separation in natural gas prices between Houston Ship Channel and Waha. It is important to note that this measure includes all capacity above the point at which a unit's offers become mitigated, without regard for whether that capacity was actually required to serve load.

E. Reliability Must Run and Must Run Alternative

A total of nine generation resources provided a Notification of Suspension of Operations (NSO) with suspension dates in 2018, accounting for approximately 4,273 MW of the capacity being retired or mothballed during the year.[44] ERCOT determined that the units were not necessary to support ERCOT transmission system reliability, and as a result no new RMR contracts were

44 Spencer 4 & 5 on January 3, 2018; Monticello Units 1, 2, and 3 on January 4, 2018; Sandow Units 4 and 5 on January 11, 2018; and Big Brown Units 1 and 2 on February 12, 2018.

awarded in 2018.[45] However, review of the RMR process remained active throughout the 2018, building off the momentum gathered in previous years.

Several proposed RMR-related Protocol revisions initiated in 2017 were discussed throughout 2018. The first proposal is a reevaluation of the process for determining the Mitigated Offer Cap for RMR resources, previously contemplated in NPRR784.[46] The proposal would allow the RMR resource to be dispatched but be priced above other resources that solve the same constraint. The second proposed revision would clarify that operations and maintenance (O&M) costs are to be updated and submitted to ERCOT every three months, consistent with the schedule for provision of updated budgets for RMR resources, and would clarify the requirement for variable O&M costs submissions to include all variable costs incurred by the RMR resource for up to a ten year historical period.[47] Both NPRRs were pending at the close of 2018.

A proposal submitted in 2018 now allows third-party evaluation of submitted budget items, changes to the standby payment as cost information changes, and a final reconciliation intended to ensure that RMR payments are as accurate as possible.[48] This Protocol change was approved and includes a requirement for ERCOT to issue a miscellaneous invoice to reconcile final RMR costs no later than 30 days after the Real-Time Market True-Up Statement is issued for the termination date of the RMR agreement.

Finally, ERCOT began the process of addressing certain Commission-approved requirements for RMR and must-run alternative (MRA) services in 2018. In 2017, the Commission adopted amendments to 16 TAC §25.502[49] to adjust the notice requirements and complaint timeline applicable to suspending a resource's operation. The Commission also gave ERCOT the

---

45 The last RMR contract was executed in 2016, for Greens Bayou 5, a 371 MW natural gas steam unit built in 1973 and located in Houston. On March 29, 2016, NRG submitted an NSO indicating that Greens Bayou 5 would be mothballed indefinitely beginning June 27, 2016. On May 27, 2016, ERCOT made a final determination that Greens Bayou 5 was necessary for RMR service. The Greens Bayou 5 RMR agreement was effective June 2, 2016 for a term of 25 months and a budgeted cost of $58.1 million, plus the opportunity for up to 10% more as an availability incentive. ERCOT initially determined that Greens Bayou 5 was needed for transmission system stability in the Houston region during the summers of 2016 and 2017 until the Houston Import Project transmission upgrade was completed. However, following changes to the RMR study parameters and an earlier than expected completion of new generation in Houston, ERCOT provided NRG, the owner of Greens Bayou 5, with notice of termination of the RMR Agreement on February 27, 2017. The RMR contract was cancelled effective May 29, 2017. The total cost paid to the NRG for the Greens Bayou RMR contract was approximately $22 million, and the unit was never operated during the term of the contract. On December 5, 2017, NRG submitted a Notification of Change of Generation Resource Designation for Greens Bayou 5, declaring the unit permanently decommissioned as of December 31, 2017.

46 NPRR826, *Mitigated Offer Caps for RMR Resources*.

47 NPRR838, *Updated O&M Cost for RMR Resources*.

48 NPRR845, *RMR Process and Agreement Revisions*, approved on October 9, 2018.

49 The amendments to §25.502 relating to pricing safeguards in markets operated by ERCOT became effective on January 1, 2018.

discretion to decline to enter into an RMR agreement based on the economic value of lost load, require ERCOT approval of RMR and MRA agreements and require refunds in some instances for capital expenditures related to those agreements. The first NPRR to incorporate these rule changes into the ERCOT Protocols was approved on August 7, 2018.[50]

ERCOT proposed additional NPRRs in 2018 to further comply with the changes to 16 TAC §25.502, specifically NPRR885 and NPRR896, which in tandem provide an appropriate Protocol framework for MRA evaluation, contracting, processes and settlement. Posted on July 3, 2018, NPRR885 proposes new Protocol language to address numerous issues related to the solicitation and operation of MRA service. For example, this NPRR defines MRA Service, clarifies MRA-eligible Resource types, clarifies MRA procedures, codifies MRA submission process, sets processes for Demand Response MRA measurement and verification, defines MRA settlements, and creates a Standard Form MRA Agreement.[51] Taken in conjunction with NPRR885, NPRR896 outlines the process ERCOT will use to evaluate the cost-effectiveness of procuring RMR or MRA service.[52] Both of these NPRRs were pending at the close of 2018 as well.

[50] NPRR862, *Updates to Address Revisions under PUCT Project No. 46369.*

[51] NPRR885, *Must-Run Alternative (MRA) Details and Revisions Resulting from PUCT Project No. 46369, Rulemaking Relating to Reliability Must-Run Service*

[52] NPRR896, *RMR and MRA Alternative Evaluation Process.*

VI. Resource Adequacy

One of the primary functions of the wholesale electricity market is to provide economic signals that will facilitate the investment needed to maintain a set of resources that are adequate to satisfy the system's needs. This section begins with an evaluation of these economic signals by estimating the "net revenue" resources received from ERCOT real-time and ancillary services markets and providing comparisons to other markets. Next, the effectiveness of the Scarcity Pricing Mechanism is reviewed. The current estimate of planning reserve margins for ERCOT are presented, followed by a description of the factors necessary to ensure resource adequacy in an energy-only market design.

A. Net Revenue Analysis

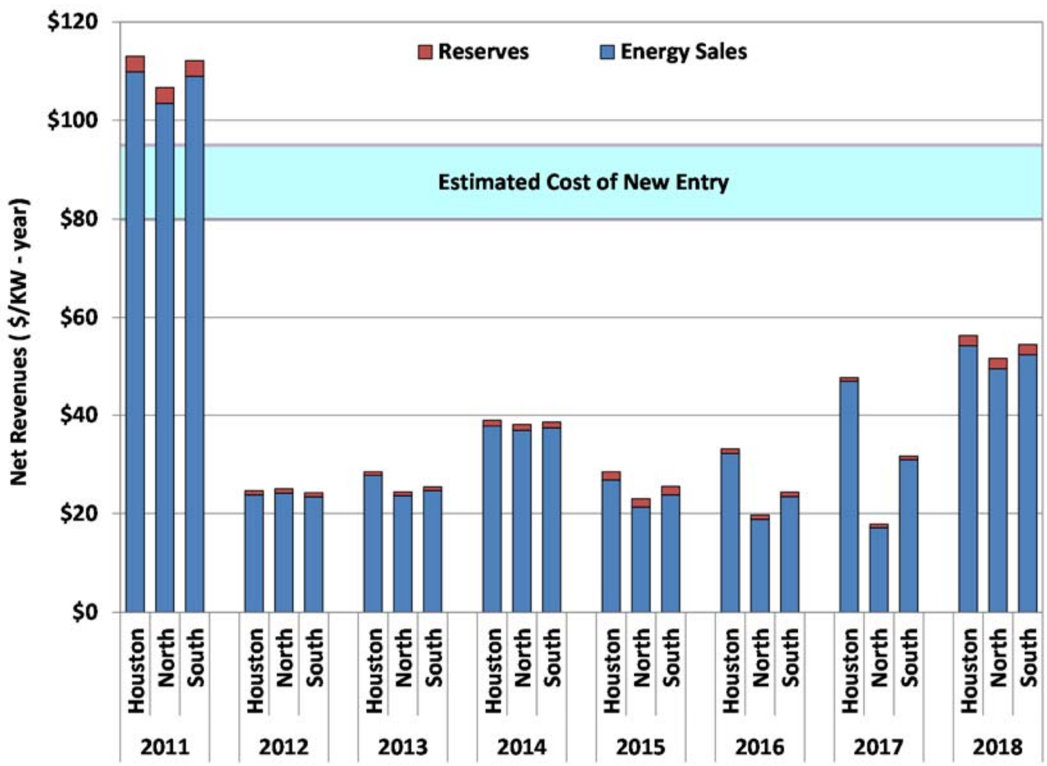
Net revenue is calculated by determining the total revenue that could have been earned by a generating unit less its variable production costs. Put another way, it is the revenue in excess of short-run operating costs that is available to recover a unit's fixed and capital costs, including a return on the investment. In ERCOT's energy-only market, the net revenues from the real-time energy and ancillary services markets alone provide the economic signals that inform suppliers' decisions to invest in new generation or retire existing generation. To the extent that revenues are available through the day-ahead market or other forward bilateral contract markets, these revenues are ultimately derived from the expected real-time energy and ancillary service prices. Although most suppliers are likely to receive the bulk of their revenues through bilateral contracts, the spot prices produced in the real-time energy market should drive bilateral energy prices over time and thus are appropriate to use for this evaluation. It is important to note that this net revenue calculation is a look back at the estimated contribution based on actual market outcomes. Suppliers will typically base investment decisions on expectations of future electricity prices. Although expectations of future prices should be informed by history, they will also factor in the likelihood of shortage pricing conditions that could be very different than what actually occurred.

The energy net revenues are computed based on the generation-weighted settlement point prices from the real-time energy market. Weighting the energy values in this way facilitates comparisons between geographic zones, but will mask what could be very high values for a specific generator location. This analysis does not consider any payments for potential RUC actions. The analysis relies on simplifying assumptions that can lead to over-estimates of the profitability of operating in the wholesale market. Start-up costs and minimum running times are not accounted for in the net revenue analysis. Ramping restrictions, which can prevent generators from profiting during brief price spikes, are also excluded. But despite these limitations, the net revenue analysis provides a useful summary of signals for investment in the wholesale market.

For purposes of this analysis, the following assumptions were used for natural gas units: heat rates of 7 MMBtu per MWh for a combined cycle unit, 10.5 MMBtu per MWh for a combustion turbine, and $4 per MWh in variable operating and maintenance costs. A total outage rate (planned and forced) of 10% was assumed for each technology. Net revenue is calculated by assuming the unit will produce energy in any hour for which it is profitable and by assuming it will be available to sell reserves and regulation (combined cycle units only) in all other hours.

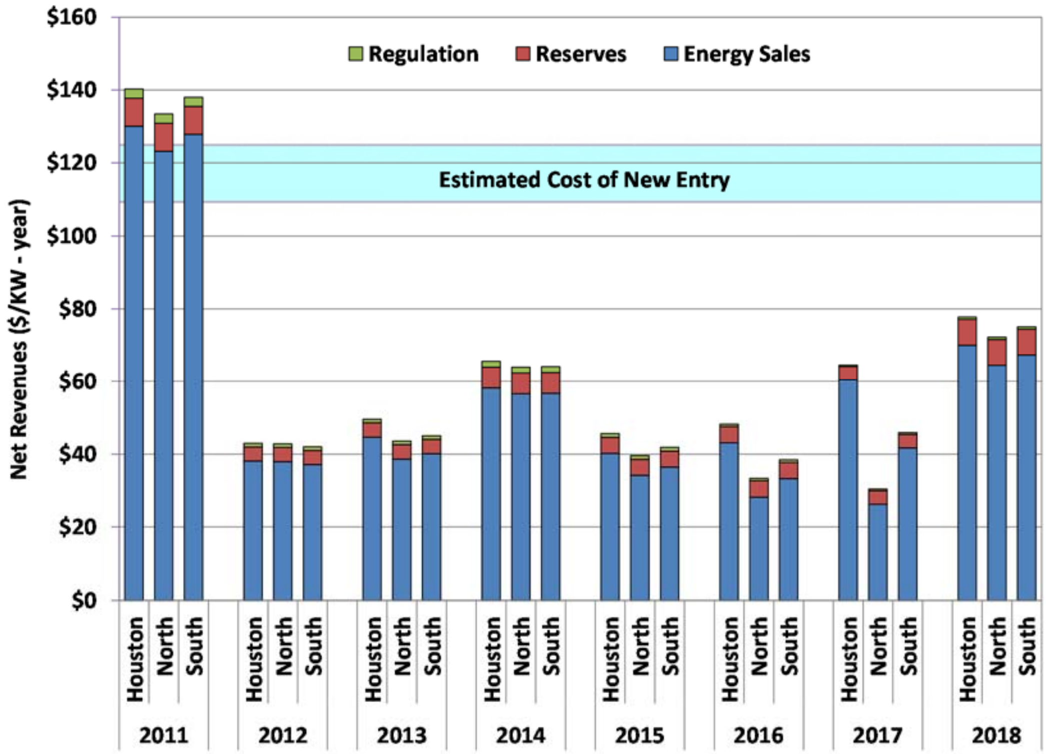
The next two figures provide an historical perspective of the net revenues available to support investment in a new natural gas combustion turbine (Figure 85) and combined cycle generation (Figure 86), selected to represent the marginal new supply that may enter when new resources are needed. Values for the West zone are excluded because historically lower energy prices make it a less attractive location to site natural gas generation. The figure also shows the estimated "cost of new entry," which represents the revenues needed to break even on the investment.

Figure 85: Combustion Turbine Net Revenues



Based on estimates of investment costs for new units, the net revenue required to satisfy the annual fixed costs (including capital carrying costs) of a new combustion turbine unit ranges from $80 to $95 per kW-year. Although higher overall in 2018 than any year since 2011, the ERCOT market continued to provide net revenues well below the level needed to support new investment, ranging from below $52 per kW-year in the North Zone to more than $56 per kW-year in Houston.

Figure 86: Combined Cycle Net Revenues

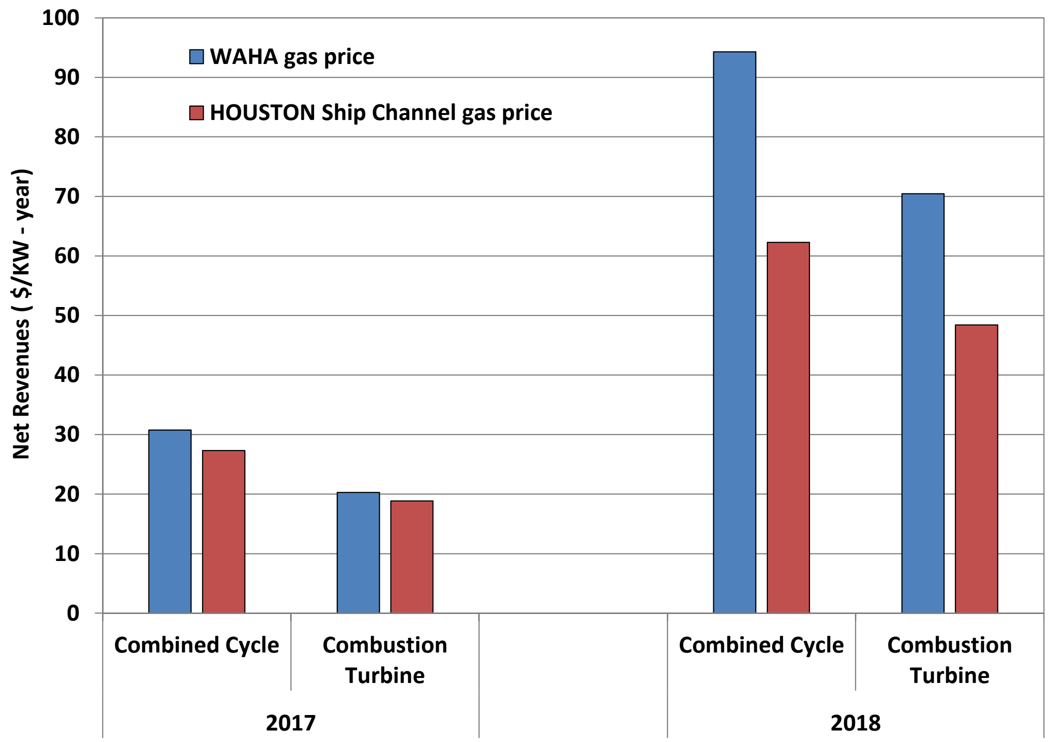


For a new combined cycle natural gas unit, the estimate of net revenue requirement is approximately $110 to $125 per kW-year. The net revenues in 2018 for a new combined cycle unit were calculated to be approximately $72 to $78 per kW-year, depending on the zone. Again, these values are well below the estimated cost of new combined cycle generation.

These results are consistent with a shrinking surplus of capacity, which contributed to more frequent shortages in 2018 compared to recent years. In an energy-only market, shortages play a key role in delivering the net revenues an investor needs to recover its investment. Such shortages will tend to be clustered in years with unusually high load or poor generator availability. The results in 2018 do not by themselves raise substantial concern regarding design or operation of ERCOT's Operating Reserve Demand Curve (ORDC) mechanism for pricing shortages. Given the recent generation retirements and continued load growth, 2018 was in fact a year with significantly more occurrences of shortage pricing, with that trend expected to continue in 2019.

A new trend evident in 2018 was the growing separation in natural gas prices between the Waha and Houston locations. Increased drilling activity in the Permian Basin of far west Texas has produced a glut of natural gas and consequently, much lower prices at the Waha location. Figure 87 below provides a comparison of net revenue for both types of natural gas units assuming Houston and Waha gas prices. With Waha gas prices, the net revenues of gas units located in the West zone are higher than the other locations. Even though the load price in the West is higher than the corresponding prices in the other three zones, the generation weighted Settlement Point Price in the West is much lower.

Figure 87: West Zone Net Revenues



Given the continuing effects of relatively low natural gas prices, the economic viability of existing coal and nuclear units was evaluated. Non-shortage prices, which have been substantially affected by the prevailing natural gas prices, determine the vast majority of net revenues received by these base load units. As previously described, the load-weighted ERCOT-wide average energy price in 2018 was $35.63 per MWh. The generation-weighted average price for the four nuclear units in ERCOT (approximately 5 GW of capacity) was lower at $29.00 per MWh.

According to data published by the Nuclear Energy Institute, the total generating cost for nuclear energy in the U.S. was $33.50 per MWh in 2017.[53] The generation-weighted average price in 2017 for the four nuclear units in ERCOT was less than $25.00 per MWh.
Assuming that operating costs in ERCOT are similar to the U.S. average, and that nuclear operating costs have either continued their multi-year decline or held stable in 2018, it is likely that the nuclear units in ERCOT are marginally profitable. Unlike other regions with large amounts of nuclear generation, the four nuclear units in ERCOT are relatively new and owned by

53 https://www.nei.org/resources/reports-briefs/nuclear-costs-in-context

four entities with sizable load obligations. Even if these units are not profitable on a stand-alone basis, the nuclear units have substantial option value for the owners because they ensure that their cost of serving their load will not rise substantially if natural gas prices increase. With the apparent reduction in nuclear operating costs, the economic pressure on these units seems to be easing.

The generation-weighted price of all coal and lignite units in ERCOT during 2018 was $33.31 per MWh, an increase from $26.32 per MWh in 2017. Although specific unit costs may vary, index prices for Powder River Basin coal delivered to ERCOT were approximately $2.65 per MMBtu in 2018; remaining at 2017 (and 2015) levels after decreasing to $2.51 per MMBtu in 2016. At these average prices coal units in ERCOT are likely receiving just enough revenue to cover operating costs.

During 2015 and 2016, delivered coal costs in ERCOT were higher than natural gas prices at the Houston Ship Channel, resulting in reduced market share for coal generation. With the increased natural gas prices in 2017 and 2018, the spread between coal and natural gas increased to nearly $0.60 per MMBtu. However, given coal units generally have higher heat rates and more expensive non-fuel (O&M) costs than combined-cycle natural gas units, economic pressure remain for coal units. Luminant retired seven coal units in early 2018, CPS Energy retired two coal units at the end of the year. [54] Additionally, Texas Municipal Power Authority also announced the continued mothballing of the Gibbons Creek coal unit late in the year. The IMM reviewed each of these actions and found them to be supported by the unit specific financials.

54 Monticello Units 1, 2, and 3, totaling 1,865 MW, retired on January 4, 2018; Sandow Units 4 and 5, totaling approximately 1,200 MW, retired on January 11, 2018; and Big Brown Units 1 and 2, totaling 1,208 MW, retired on February 12, 2018. Deely Units 1 and 2, totaling 840 MW, retired on December 31, 2018.

Table 10 displays the calculated output-weighted price by generation type.

Table 10: Settlement Point Price by Fuel Type

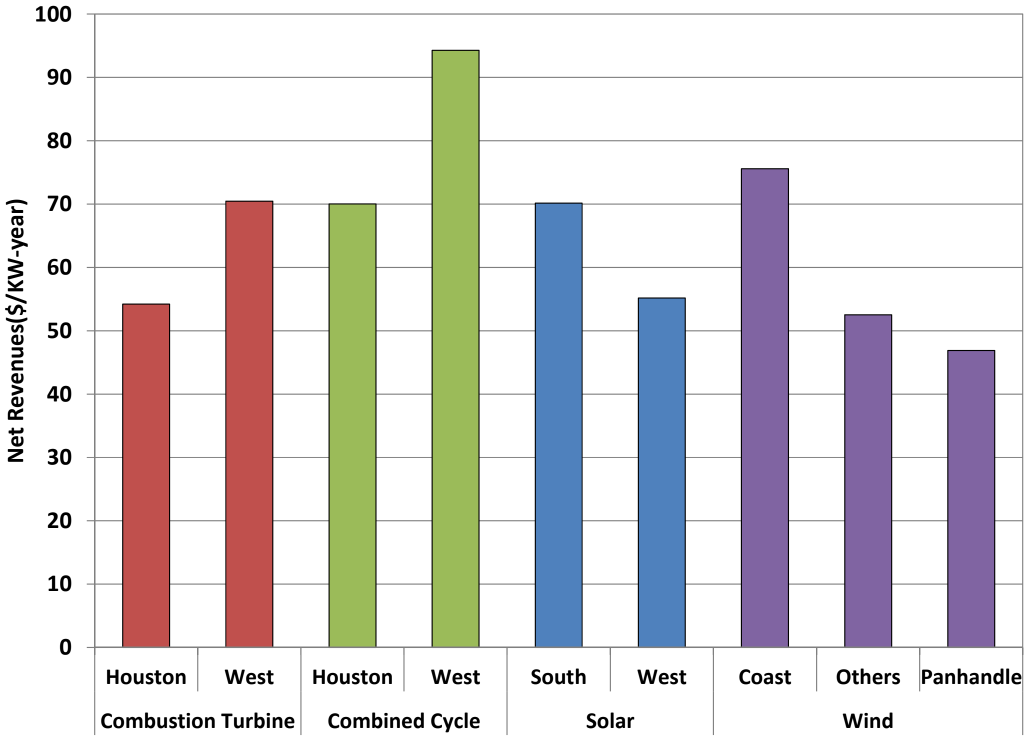
Generation Type	Output-Weighted Price
Coal	$33.31
Combined Cycle	$35.53
Gas Peakers	$71.64
Gas Steam	$66.09
Hydro	$34.40
Nuclear	$29.00
Power Storage	$103.19
Private Network	$34.41
Renewable	$39.84
Solar	$35.37
Wind	$19.26

Although electricity prices at generator locations in the West zone are lower than in other areas, natural gas prices at Waha have been low enough that profitability would be higher for generators in the West compared to other locations. This is evident in Figure 88 which compares the net revenues for gas generation (combustion turbine and combined cycle) located in the Houston zone paying Houston Ship Channel gas prices, with the same technologies located in the West zone and paying Waha natural gas prices. Although higher than in other locations, net revenues for generators in the West were not sufficient to support new build in 2018.

The current low natural gas prices at Waha are a result of the apparent high profitability of oil production in the Permian Basin. Natural gas is produced as an associated product and there is insufficient transport capacity for the current amounts of natural gas being produced. New transportation projects have been identified and are currently being worked. However, it is unclear how much longer the large basis difference in natural gas prices will continue to exist.

Also shown in Figure 88 are the net revenues for wind and solar generation at multiple locations. As the cost to install wind or solar does not vary much by location, the profitability of those resources are chiefly determined by the available natural resource and the prevailing price to be received. For both wind and solar generators, profitability in 2018 appears to have been inversely related to the available natural resource. Net revenues for wind and solar in some locations were equal to gas technologies in 2018.

Figure 88: Net Revenues by Generation Resource Type



These results indicate that during 2018, the ERCOT markets would not have provided sufficient revenues to support profitable investment in any of the types of generation technology evaluated. As detailed in Figure 65, very little non-renewable capacity was added in 2018. This seems inconsistent with the increased levels of scarcity pricing present in the ERCOT market in 2018 compared to recent years. There continues to be robust investment in renewable generation additions. It is unclear how much of that is in anticipation of the expiration of federal tax credits and how much will continue in the long term.

The backward-looking net revenue analysis provides insight into the potential profitability of new generation, and therefore the ability of the ERCOT market to attract sufficient new resources to meet growing customer demands for electricity. However, resource investment decisions are driven by multiple factors.

First, resource investments are driven primarily by forward price expectations. Historical net revenue analyses do not provide a view of the future pricing expectations that will spur new investment. Suppliers will develop their own view of future expected revenue and given the level to which prices will rise under shortage conditions, small differences in expectations about the frequency of shortage pricing can greatly influence revenue expectations. Real-time prices in the summer of 2018 did not reach the expectations as set by forward prices. Forward prices for summer of 2019 are currently lower than they were at a similar time for 2018, even though the installed capacity margins are smaller.

Second, this analysis does not account for bilateral contracts. The only revenues considered in the net revenue calculation are those that came directly from the ERCOT real-time energy and ancillary services markets in a specific year. Some developers may have bilateral contracts for unit output that would provide more revenue than the ERCOT market did in 2018. Given the level to which prices will rise under shortage conditions, buyers may enter bilateral contracts to hedge against high shortage pricing.

Third, net revenues in any one year may be higher or lower than an investor would require over the long term. In 2018, shortages were more common than in recent years, but still much less frequent than would be expected over the long term. Shortage revenues play a pivotal role in motivating investment in an energy-only market like ERCOT. Hence, in some years shortage pricing will be frequent and net revenues may substantially exceed the cost of entry, while in most others it will be less frequent and net revenue will be less than the cost of entry.

Finally, the costs of new entry used in this report are generic for the U.S. and reflective of the costs of a new unit on an undeveloped greenfield site. They have been reduced somewhat to reflect the lower costs of construction in Texas. However, companies may have opportunities to build generation at much lower cost than these estimates; either by having access to lower cost equipment, or by adding the new unit to an existing site, or some combination of both. Financing structures and costs can vary greatly between suppliers and may be improved to be lower than the generic financing costs assumed in the net revenue analysis.

To provide additional context for the net revenue results presented in this subsection, the net revenue in the ERCOT market for two types of natural gas generation technologies are compared with the net revenue that those technologies could expect in other wholesale markets.

Figure 89 provides a comparison of net revenues for a hypothetical combustion turbine with an assumed heat rate of 10,500 MMBtu per MWh installed in ERCOT, Midcontinent ISO (MISO), New York ISO (NYISO), and the PJM Interconnection. Net revenues for two locations in both ERCOT and NYISO are provided to highlight the variation in value that can exist even within the same market. Additionally, Figure 89 includes estimated total net revenues for a combustion turbine generator located in Southwest Power Pool (SPP) and California ISO (CAISO), shown without the component values.

Figure 89: Combustion Turbine Net Revenue Comparison Between Markets

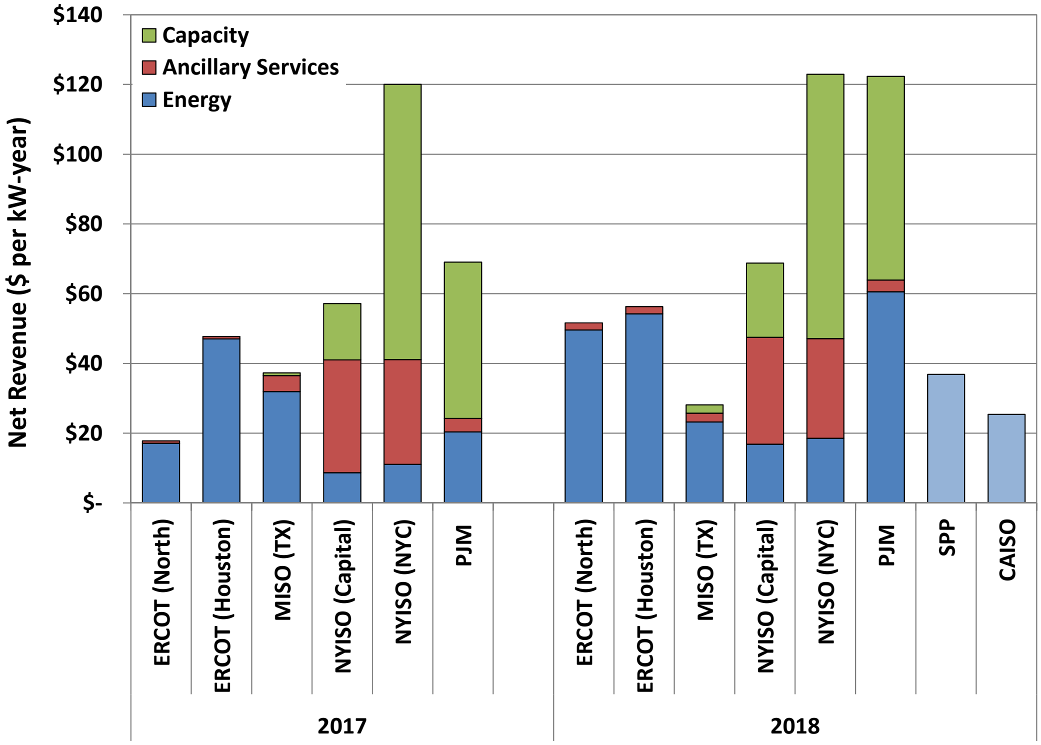
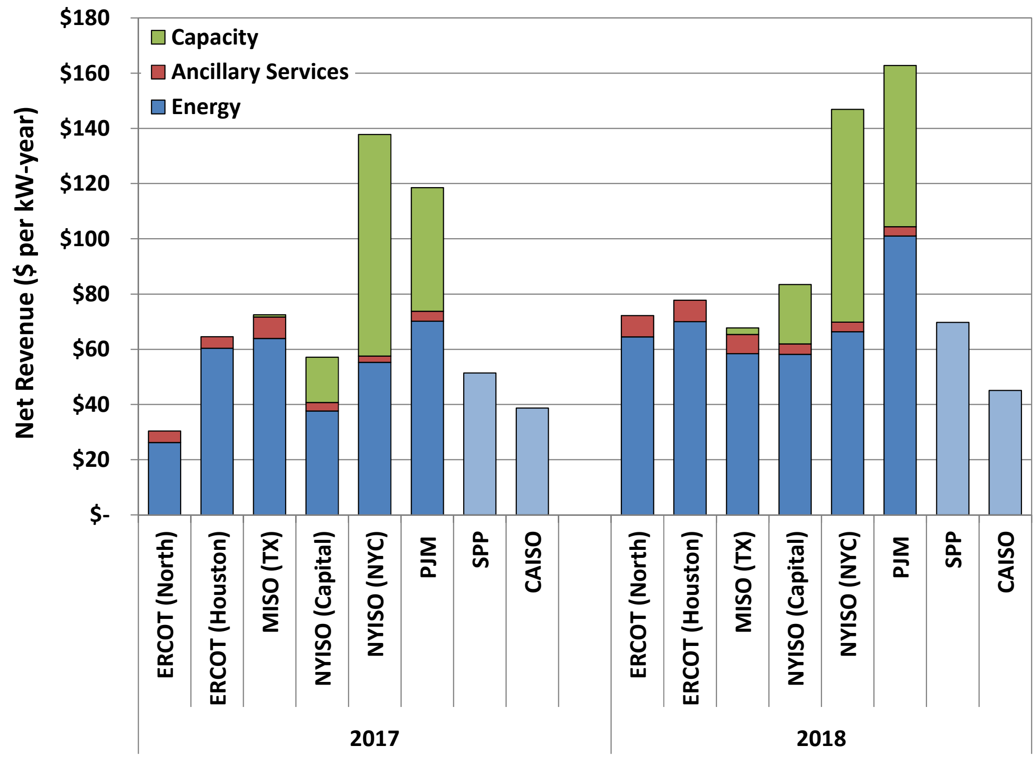


Figure 90 provides the net revenues for a hypothetical combined cycle unit with an assumed heat rate of 7,000 MMBtu per MWh installed in ERCOT, MISO, NYISO, and PJM. Both Figure 89 and Figure 90 display estimates of net revenue from energy, reserves and regulation, and capacity. ERCOT does not have a capacity market, and thus, does not have any net revenue from capacity sales. Figure 90 also includes estimated total net revenues for a combined cycle generator located in SPP and CAISO, shown without the component values.

Figure 90: Combined Cycle Net Revenue Comparison Between Markets



Both figures indicate a general increase in net revenues across all markets. The exception to this trend was MISO's TX zone. Most other markets also have sufficient installed reserves, typically a result of low or no load growth. The two figures also show that capacity revenues in NYISO and PJM provide a meaningful portion of the net revenues for new resources. In ERCOT, these revenues will be provided through its shortage pricing, which is evaluated in the next section. Because shortage pricing has not been frequent in the past two years, the net revenues in ERCOT have tended to be lower than those in RTO areas with functional capacity markets despite the fact that ERCOTs capacity margin is much lower than any of the other markets shown.

B. Effectiveness of the Scarcity Pricing Mechanism

The Commission adopted rules in 2006 that define the parameters of an energy-only market. In accordance with the IMM's charge to conduct an annual review,[55] this subsection assesses the Scarcity Pricing Mechanism (SPM) in 2018 under ERCOT's energy-only market structure.

Revisions to 16 TAC § 25.505 were adopted in 2012 that specified a series of increases to the ERCOT system-wide offer cap. The last step went into effect on June 1, 2015, increasing the system-wide offer cap to $9,000 per MWh. As shown in Figure 21 on page 23, there have been very brief periods when energy prices rose to the cap since the system-wide offer cap was

55 *See* 16 TAC § 25.505(g)(6)(D).

increased to greater than $3,000 per MWh, with prices hitting $9,000 for the first time on January 23, 2018.

The SPM includes a provision termed the Peaker Net Margin (PNM) that is designed to provide a fail-safe pricing measure, which if exceeded would cause the system-wide offer cap to be reduced. If the PNM for a year reaches a cumulative total of $315,000 per MW, the system-wide offer cap is then reduced to the higher of $2,000 per MWh or 50 times the daily natural gas price index.[56] PNM also serves as a simplified measure of the annual net revenue of a hypothetical peaking unit.[57]

Market outcomes in 2018 indicate that net revenues provided by the ERCOT real-time market were not at a level sufficient to support new natural-gas fueled generation. It appears that market outcomes in 2018 were reflective of the specific conditions experienced rather than any flaw with the scarcity pricing mechanism. ERCOT's energy-only market design is intended to produce very high energy prices (including the operating reserve adder) during times of supply scarcity, i.e., when available capacity is insufficient to meet both total demand for electricity and required operating reserves. This structure creates a very strong incentive for generators to be available at the times when they are expected to be needed most.

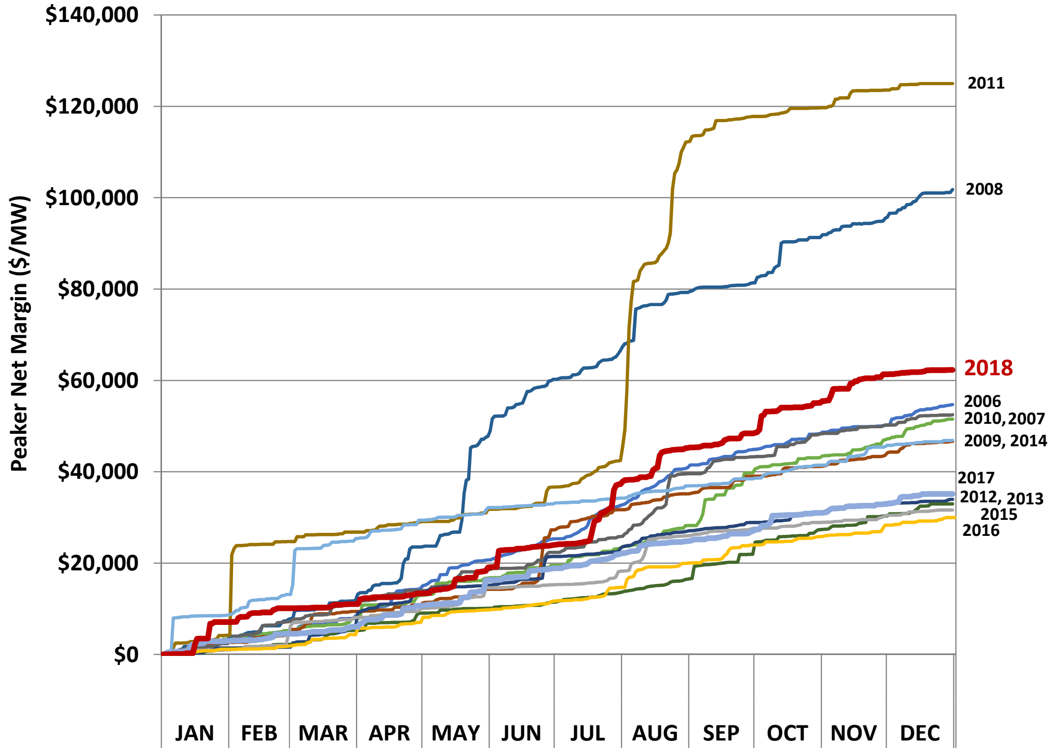
The benefits from the incentives for generator availability in ERCOT's energy-only market design are somewhat thwarted by the lack of a priced-based mechanism to determine when to convert reserve capacity into energy in real-time. In an energy-only market that depends on scarcity pricing signals to provide incentive for proper levels of investment, it is imperative the scarcity pricing reflects actual scarcity rather than the inefficient assignment of reserve capacity. Real-time co-optimization will improve the accuracy of the shortage pricing mechanism.

Figure 91 shows the cumulative PNM results for each year since the creating of the scarcity pricing mechanism and shows that PNM in 2018 was higher than in recent years, but still not approaching the levels reached in 2008 and 2011. Considering the purpose of PNM was initially defined to provide a "circuit breaker" trigger for lowering the system-wide offer cap, PNM has not approached levels that would dictate a needed reduction in the system wide offer cap, even in a relatively high year like 2018.

56 The threshold established in the initial Rule was $300,000 per MW-year. For 2014 and each subsequent year, ERCOT shall set the PNM threshold at three times the cost of new entry of new generation plants. The current threshold is based on the most recent version of an Other Binding Document entitled "System-Wide Offer Cap and Scarcity Pricing Mechanism Methodology."

57 The proxy combustion turbine in the Peaker Net Margin calculation assumes a heat rate of 10 MMBtu per MWh and includes no other variable operating costs or startup costs.

Figure 91: Peaker Net Margin



As with net revenues, the PNM is expected to be less than the cost of new entry in most years. Concerns with the SPM under the zonal market design were addressed in every State of the Market Report produced during that period.[58] The implementation of the nodal market design, which included a power balance penalty curve, created the opportunity for real-time energy prices to systematically reflect the value of reduced reliability imposed under shortage conditions, regardless of submitted offers.

In 2013, the Commission took another step toward improving resource adequacy signals by directing ERCOT to implement the ORDC. As discussed in Section I: Review of Real-Time Market Outcomes, ORDC is a shortage pricing mechanism that reflects the loss of load probability at varying levels of operating reserves multiplied by the value of lost load. In the short time it has been in effect, ORDC has had a modest impact on real-time prices, even in 2018 when reduced installed reserves led to higher expectations of shortage pricing.

In the spring of 2018, a change to the ORDC was implemented at the behest of the Commission via Other Binding Document Revision Request (OBDRR) 002, which modified the *Methodology for Implementing Operating Reserve Demand Curve (ORDC) to Calculate Real-Time Reserve*

58 The zonal market design was not the problem per se, rather its reliance on high-priced offers to set high prices during periods of shortage was of concern.

Price Adder.[59] This OBDRR revised the on-line and off-line capacity reserves for resources on-line during a reliability unit commitment (RUC) instruction. With this change, ONRUC capacity is now excluded from the reserves used to determine the ORDC price adder, increasing the value of the ORDC price adder and resulting in potentially higher real-time energy prices. This was a relatively minor change to the ORDC implemented prior to the summer of 2018 in order to improve pricing signals under tight conditions.

A more profound change to the ORDC was contemplated by the Commission in the latter half of 2018 when reviewing summer market performance. The Commission considered proposals modifying various defining aspects of the curves used to determine the ORDC price adder including shifting the loss of load probability (LOLP) portion of the curve, as well as modifying the minimum contingency level ("X") and the value of lost load (VOLL).[60] The LOLP portion of the curves used to determine the ORDC price adder has typically been constructed using normal probability distributions which are defined by two factors, MU and SIGMA. For each of the twenty-four curves currently used (six time of day blocks and four seasons), these factors are calculated by taking the average of historical differences between expected and actual operating reserves ("MU") and the standard deviation in those values ("SIGMA"). The Commission deliberated the merits of these proposed ORDC changes, as well as implementation of marginal losses and real-time co-optimization, through the close of 2018.[61]

Regardless of the ORDC parameter values, the fact that responsive and regulating reserves are forced to be maintained (held behind the High Ancillary Service Limit (HASL)) under the current market design will continue to be problematic until real-time co-optimization is implemented. By not allowing reserves to be converting to energy at any price, reserves will not gradually decrease, which means that prices will not gradually increase. Reserves will be sufficient until they are not, which means that the contribution from the operating reserve adder will be relatively small until reserves are deficient and the contribution from the adder is quite large. Jointly optimizing all products will improve the utilization of ERCOT resources, ensure that shortage pricing only occurs when the system is actually short after fully utilizing its resources, and establish prices for each product that efficiently reflect its reliability value without the use of administrative caps and adders.

59 *See* OBDRR002, *ORDC OBD Revisions for PUCT Project 47199*, effective May 31, 2018.

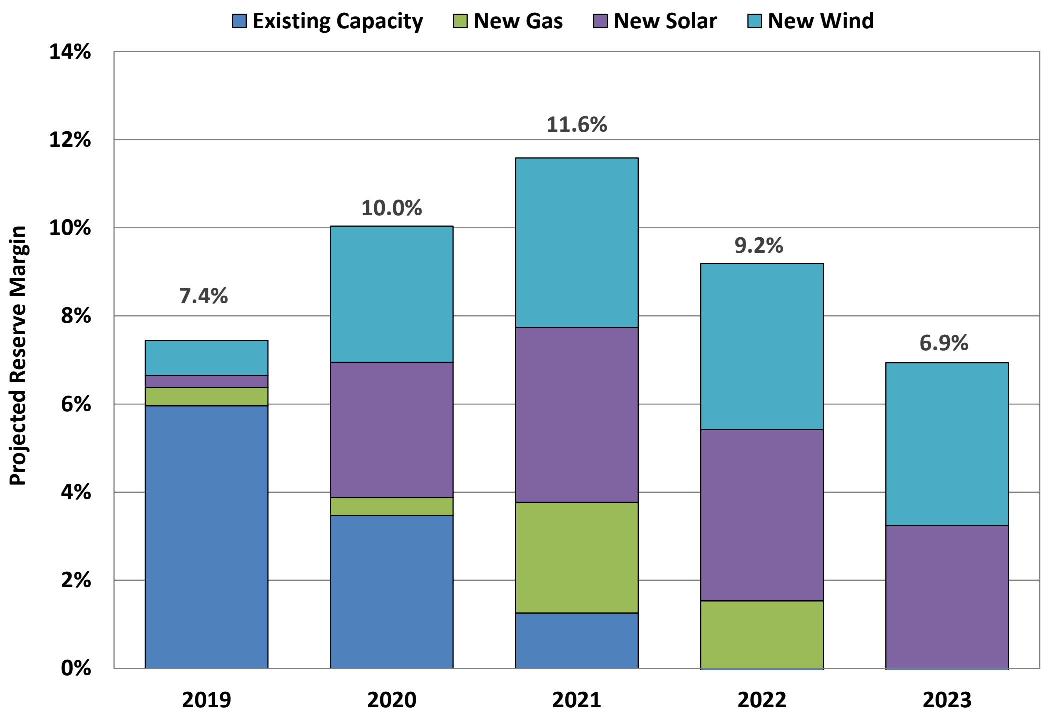
60 PUCT Project No. 48551, *Review of Summer 2018 ERCOT Market Performance*.

61 On January 17, 2019, the Commission approved a phased-process to implement changes to the ORDC and directed ERCOT to implement a .25 standard deviation shift in the LOLP calculation using a single blended ORDC curve, with a second step of .25 in the spring of 2020. The ORDC changes were approved by the ERCOT Board of Directors at its February 12, 2019 meeting, implemented on March 1, 2019 via OBDRR011, ORDC OBD Revisions for PUCT Project 48551. The Commission also instructed ERCOT to proceed with the implementation of real-time co-optimization, but chose not to direct the implementation of marginal losses.

C. Planning Reserve Margin

The prior subsection discusses and evaluates the economic signals produced by the ERCOT markets to facilitate efficient decisions by suppliers to maintain an adequate base of resources. This subsection summarizes and discusses the current level of capacity in ERCOT, as well as the long-term need for capacity in ERCOT. Figure 92 below shows ERCOT's current projection of planning reserve margins.

Figure 92: Projected Planning Reserve Margins



Source: ERCOT Capacity, Demand and Reserves Report, December 2018 with Gibbons Creek capacity removed

Figure 92 indicates that Texas is heading into the summer months of 2019 with a historically low reserve margin of 7.4%, just over half of ERCOT's previously stated reserve margin goal of 13.75%.[62] Also evident is that all currently operating capacity will be insufficient to serve forecasted peak load in 2022. These reserve margin projections are even lower than those

62 A target planning reserve margin of 13.75% was approved by the ERCOT Board of Directors in November 2010, based on a one in ten loss of load expectation (LOLE). The Commission directed ERCOT to evaluate planning reserve margins based on an assessment of the Economically Optimal Reserve Margin (EORM) and the Market Equilibrium Reserve Margin (MERM). *See* PUCT Project No. 42303, ERCOT Letter to Commissioners (Oct. 24, 2016). On December 12, 2017, ERCOT published its "Study Process and Methodology Manual: Estimating Economically Optimum and Market Equilibrium Reserve Margins" as part of its ongoing reporting initiative.

developed in December 2018,[63] due in large part to higher expected seasonal peak loads, additional delays and cancellations of planned projects, and the indefinite mothballing of 470 MW at the Gibbons Creek coal unit.[64] The reserve margin is expected to continue to be below existing target levels for the foreseeable future.

Installed reserve margins for summer of 2018 were also historically low. What seem like very low reserves may just be the new normal. Given the overall size of the system and projected growth, a more robust reserve margin may no longer be required to cover load forecast errors and mitigate generator availability risks. For example, a 15% reserve margin for a 40 GW system would provide 6 GW of installed reserves, whereas for a 70 GW system the installed reserves would be 10.5 GW. It is not clear that the magnitude of forecasting errors are directly proportional to the size of peak load. Further, the size of generators has not increased and there are more of them, which means generator availability risks are proportionally smaller. With smaller, more distributed generation technologies playing an increasingly important role in ERCOT, the risk associated with generator outages should decrease.

On October 12, 2018, ERCOT filed a draft report with the Commission titled "*Estimation of the Market Equilibrium and Economically Optimal Reserve Margins for the ERCOT Region*."[65] The report estimates the Market Equilibrium Reserve Margin (MERM) and Economically Optimal Reserve Margin (EORM) for ERCOT's wholesale electric market with projected system conditions for 2022. In the draft report, The Brattle Group (Brattle) estimated a MERM of 10.25% under projected 2022 market conditions. Brattle further estimated the EORM as 9.0%, based on the risk-neutral, probability-weighted-average cost of 57,000 simulations. The draft report stated that estimated societal costs are relatively flat with respect to reserve margin near the minimum, with only modest variation between reserve margins of 7% and 11%. Brattle stated that the economic optimum occurs at the reserve margin that minimizes societal costs net of all supply costs and the lost value from any disruptions in electric service. Finally, Brattle

63 See Report on the Capacity, Demand and Reserves in the ERCOT Region (December 11, 2018); http://www.ercot.com/content/wcm/lists/167023/CapacityDemandandReservesReport-Dec2018.pdf; the 2019 summer reserve margin was projected to be 8.1%, a reduction of 2.9 percentage points from the May 2018 CDR report.

64 On December 21, 2018, ERCOT received an NSO from the City of Garland for Gibbons Creek (GIBCRK_GIB_CRG1) indicating that the Resource will be mothballed indefinitely effective June 1, 2019. Gibbons Creek is a 470 MW coal unit located in Grimes County (20 miles southeast of College Station) and owned by the Texas Municipal Power Agency (TMPA), which is an organization jointly owned by four municipalities – the cities of Garland, Denton, Bryan and Greenville.

65 *See Estimation of the Market Equilibrium and Economically Optimal Reserve Margins for the ERCOT Region* (Oct. 12, 2018); http://www.ercot.com/content/wcm/lists/143980/10.12.2018_ERCOT_MERM_Report_Final_Draft.pdf.

stated that its analysis shows that the market equilibrium of 10.25% is greater than the economically optimal level of capacity by 1.25%.[66]

This current projection of planning reserve margins is consistent with the economic signals produced by the market in recent years, which are themselves the product of the sustained capacity surpluses that have existed in ERCOT. Hence these results demonstrate that the market is functioning properly. Less efficient, uneconomic units are retiring in times of relatively low prices. Of the eleven generation units retired or mothballed since the May 2017 CDR, eight of those units (totaling approximately 4,500 MW) were coal units.[67] The IMM views the decisions to retire the coal units to be justified based on the operating history and estimated costs of continued operations. Similar to the forces that have led to the retirement of less efficient natural gas fueled steam units, the retirement of older, less efficient coal units is an expected market outcome. The retirement of uneconomic generation should not be viewed as failure to provide resource adequacy. In fact, facilitating efficient decisions by generators to retire uneconomic units is nearly as important as facilitating efficient decisions to invest in new resources.

With expectations for future natural gas prices to remain relatively low, the economic pressure on coal units in ERCOT is not expected to subside any time soon. AEP publicly announced in September 2018 that the 650 MW Oklaunion coal unit will retire by September 30, 2020. The plant is located on the Texas side of the border between the two states and feeds power into the Southwest Power Pool (SPP) and ERCOT, and accounts for 4.4% of ERCOT's summer coal capacity. Its retirement would leave ERCOT with 24 operational coal units. The effects of this proposed retirement have not been factored in to current projections because AEP has not formally filed a notice to suspend operations with ERCOT.

Even with low prices, there continues to be high interest in the ERCOT market from generation developers as evidenced by the amount of capacity under consideration for interconnection. At the end of 2018, there was more capacity in the various stages of interconnection evaluation than at the beginning of the year. However, the composition of that capacity had changed with much more solar and wind generation and reduced amounts of natural gas generation.

Because the surplus has now disappeared and shortages are likely to be even more frequent in 2019, the economic signals could change rapidly. These short-term market outcomes and price signals, as well as investors' response to these economic signals, will be monitored. This

66 ERCOT posted an updated version of the report on February 21, 2019. The revised report (http://www.ercot.com/content/wcm/lists/167026/2018_12_20_ERCOT_MERM_Report_Final.pdf) does not materially differ from the previous version. The revisions were generally intended only to clarify or to provide additional explanatory detail.

67 Monticello Units 1, 2, and 3, totaling 1,865 MW, to be retired on January 4, 2018; Sandow Units 4 and 5, totaling approximately 1,200 MW, to be retired on January 11, 2018; Big Brown Units 1 and 2, totaling 1,208 MW, to be retired on February 12, 2018; Gibbons Creek, a 470 MW unit mothballed indefinitely effective June 1, 2019.

response could cause the planning reserve margins to exceed the forecast shown in Figure 92 above.

D. Ensuring Resource Adequacy

One of the primary goals of an efficient and effective electricity market is to ensure that, over the long term, there is an adequate supply of resources to meet customer demand plus any required installed or planning reserves. In a region like ERCOT, where customer requirements for electricity have been and are expected to continue to increase, even with growing demand response efforts, maintaining adequate supply requires capacity additions. To incent these additions the market design must provide revenues such that the marginal resource receives revenues sufficient to make that resource economic. In this context, "economic" includes both a return of and on capital investment.

Generators earn revenues from three sources: energy prices during non-shortage, energy prices during shortage and capacity payments. The capacity payments generators receive in ERCOT are related to the provision of ancillary services. Ancillary service payments are a small contributor to generator revenues, approximately $7 per kW-year. Setting ancillary service payments aside, generator revenue in ERCOT is overwhelmingly derived from energy prices under both shortage and non-shortage conditions. Bilateral contracts should be priced based on expectations for these values. If real-time prices are expected to be $30 per MWh, it would not be reasonable to contract to sell at a price much below that level, or inversely, contract to buy at a price far above that value.

Energy pricing under non-shortage conditions should be the same regardless of whether payments for capacity exist. In ERCOT, with no capacity payments available, the amount a generator may receive from energy pricing under shortage conditions must be large enough to provide the necessary incentives for new capacity additions. This will occur when energy prices are allowed to rise substantially during times when the available supply is insufficient to simultaneously meet both energy and minimum operating reserve requirements.

Ideally, energy and reserve prices during shortages should reflect the diminished system reliability under these conditions, which is equal to the increased probability of "losing" (not serving) load times the value of the lost load. Allowing energy prices to rise during shortages mirrors the outcome expected if loads were able to actively specify the quantity of electricity they wanted and the price they would be willing to pay. The energy-only market design relies exclusively on these relatively infrequent occurrences of high prices to provide the appropriate price signal for demand response and new investment, when required. In this way, energy-only markets can provide price signals that will sustain a portfolio of resources to be used in real-time to satisfy the needs of the system. However, this portfolio may not include enough capacity to meet a specified target quantity of planning reserves.

Faced with reduced levels of generation development activity coupled with increasing loads that resulted in falling planning reserve margins, the Commission has devoted considerable effort deliberating issues related to resource adequacy. Over the long term, a co-optimized energy and operating reserve market will provide more accurate shortage pricing. Planning reserves should continue to be monitored to determine whether shortage pricing alone is leading to the desired level of planning reserves.

VII. Analysis of Competitive Performance

In this section, market power is evaluated from two perspectives: structural (does market power exist) and behavioral (have attempts been made to exercise it). Market structure is examined by using a pivotal supplier analysis that indicates the frequency with which a supplier was pivotal at higher load levels. This section also includes a summary of the Voluntary Mitigation Plans in effect during 2018. Market participant conduct is evaluated by reviewing measures of physical and economic withholding. These withholding patterns are further examined relative to the level of demand and the size of each supplier's portfolio. Based on these analyses, we find the overall performance of the ERCOT wholesale market to be competitive in 2018.

A. Structural Market Power Indicators

The market structure is analyzed by using the Residual Demand Index (RDI). The RDI is used to measure the percentage of load that cannot be served without the resources of the largest supplier, assuming that the market could call upon all committed and quick-start capacity owned by other suppliers.[68] When the RDI is greater than zero, the largest supplier is pivotal (i.e., its resources are needed to satisfy the market demand). When the RDI is less than zero, no single supplier's resources are required to serve the load if the resources of its competitors are available.

The RDI is a useful structural indicator of potential market power, but it is important to recognize its limitations. As a structural indicator, the RDI does not illuminate actual supplier behavior to indicate whether a supplier may have exercised market power, or whether it would have been profitable for a pivotal supplier to exercise market power. Nonetheless, the RDI does identify conditions under which a supplier could raise prices significantly by withholding resources.

Figure 93 shows the ramp-constrained RDI, calculated at the QSE level, relative to load for all hours in 2018. The occurrences of a pivotal supplier are not limited to just the high load summer period. This analysis indicated the existence of a pivotal supplier for some fraction of time at load levels as low as 35 GW. The trend line indicates a strong positive relationship between load and the RDI.

68 For the purpose of this analysis, "quick-start" includes off-line combustion turbines that are flagged as on-line in the current operating plan with a planned generation level of 0 MW that ERCOT has identified as capable of starting-up and reaching full output after receiving a dispatch instruction from the real-time energy market.

Figure 93: Residual Demand Index

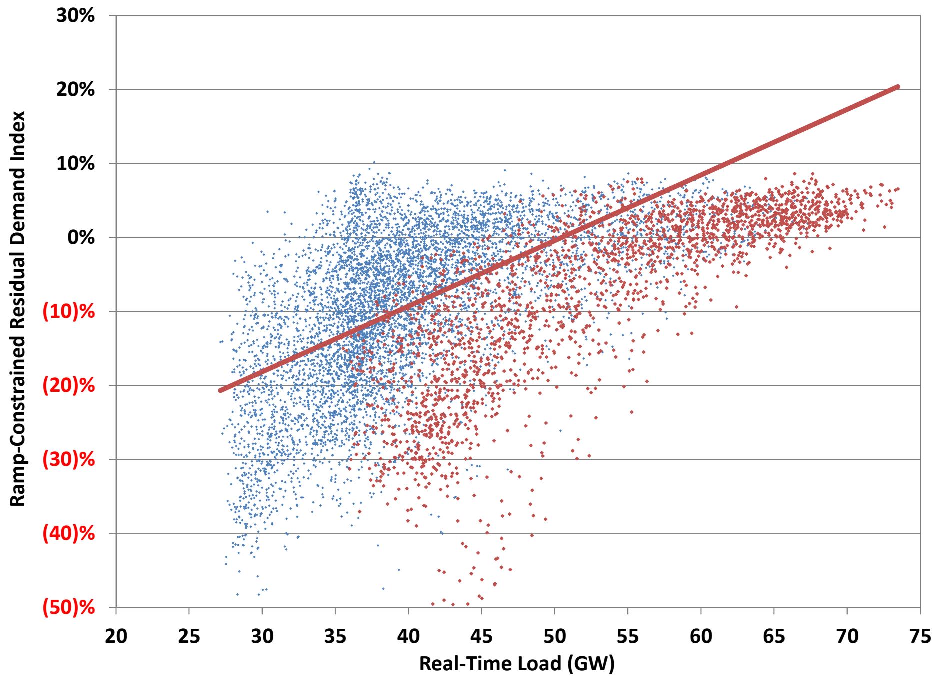
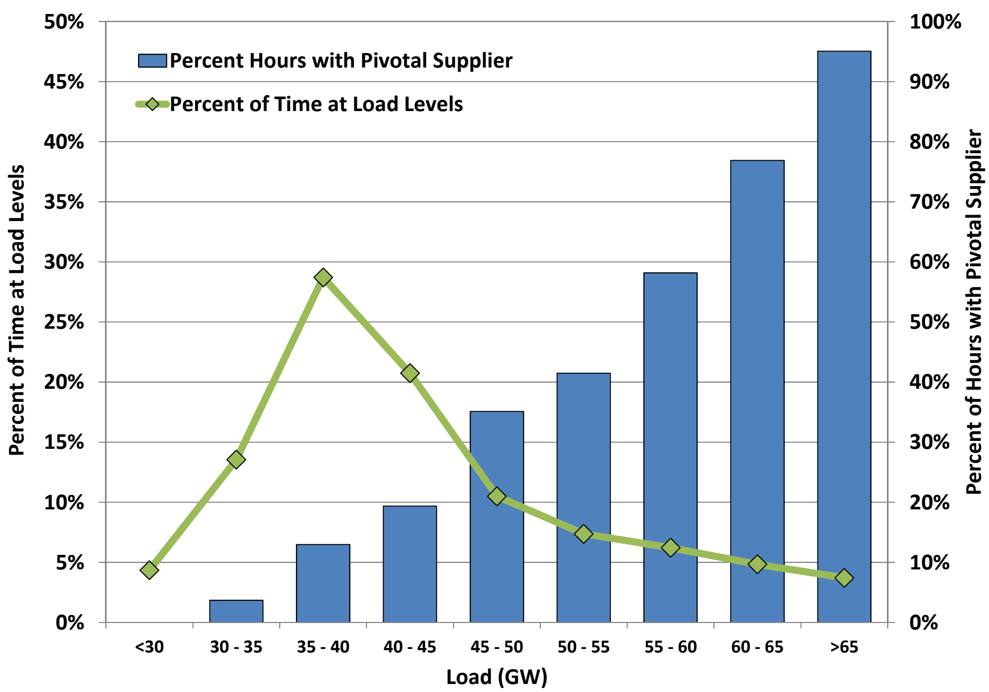


Figure 94 below summarizes the results of the RDI analysis by displaying the percentage of time at each load level there was a pivotal supplier. The figure also displays the percentage of time each load level occurs.

Figure 94: Pivotal Supplier Frequency by Load Level



At loads greater than 65 GW there was a pivotal supplier 95% of the time. This is expected because at high load levels, the largest suppliers are more likely to be pivotal as other suppliers' resources are more fully utilized serving the load. Pivotal suppliers existed 30% of all hours in 2018, which was more frequent than in 2017 when pivotal suppliers existed in 25% of all hours. The increase was due to the greater number of high load (>65GW) hours in 2018. Over the past several years, as generation supply ownership in ERCOT has become less concentrated, the fraction of time with a pivotal supplier has decreased from 75-80% of the time down to 25-30% seen in recent years. Even with this reduction, market power continues to be a potential concern in ERCOT, requiring effective mitigation measures to address it.

Inferences regarding market power cannot be made solely from pivotal supplier data. Bilateral and other financial contract obligations can affect a supplier's potential market power. For example, a small supplier selling energy only in the real-time energy market may have a much greater incentive to exercise market power than a large supplier with substantial long-term sales contracts. The RDI measure shown in the previous figures does not consider the contractual position of the supplier, which can increase a supplier's incentive to exercise market power compared to the load-adjusted capacity assumption made in this analysis.

It should be noted that the analysis above evaluates the structure of the entire ERCOT market. In general, local market power in narrower areas that can become isolated by transmission constraints raise more substantial competitive concerns. As more fully discussed in Section V, Reliability Commitments, this local market power is addressed through: (a) structural tests that determine "non-competitive" constraints that can create local market power; and (b) the application of limits on offer prices in these areas.

Voluntary Mitigation Plans

Voluntary Mitigation Plans (VMPs) existed for four market participants at various times in 2018. Generation owners are motivated to enter into VMPs because adherence to a plan approved by the Commission constitutes an absolute defense against an allegation of market power abuse through economic withholding with respect to behaviors addressed by the plan. This increased regulatory certainty afforded to a generation owner regarding its energy offers in the ERCOT real-time market must be balanced by appropriate protections against a potential abuse of market power in violation of PURA §39.157(a) and 16 TAC §25.503(g)(7).

VMPs should promote competitive outcomes and prevent abuse of market power through economic withholding in the ERCOT real-time energy market. The same restrictions are not required in forward energy markets (e.g., the ERCOT day-ahead market) because the prices in forward energy markets are derived from expectations for real-time energy prices. Because forward energy markets are voluntary and the market rules do not inhibit arbitrage between the forward energy markets and the real-time energy market, competitive outcomes in the real-time energy market serve to discipline the potential abuse of market power in the forward energy markets.

By the end of 2018, only Calpine and NRG had active and approved VMPs. Luminant terminated its previously-approved VMP in 2018, while Exelon's was terminated automatically, for the reasons outlined below.

Calpine's VMP was approved in March of 2013.[69] Because its generation fleet consists entirely of natural gas fueled combined cycle units, the details of the Calpine plan are somewhat different than the others. Calpine may offer up to 10% of the dispatchable capacity of its portfolio at prices up to $500 per MWh. Additionally, Calpine may offer up to 5% of the dispatchable capacity of its portfolio at prices no higher than the system-wide offer cap. When approved, the amount of capacity covered by these provisions was approximately 500 MW. With additions to Calpine's generation fleet made since the VMP was approved, its current amount of offer flexibility has increased to approximately 700 MW. Calpine's VMP shall remain in effect from the date it was approved by the Commission until terminated by the Executive Director of the Commission or Calpine.

NRG's plan, initially approved in June 2012 and modified in May 2014,[70] allows the company to offer some of its capacity at prices up to the system-wide offer cap. Specifically, up to 12% of the difference between the high sustained limit and the low sustained limit – the dispatchable capacity – for each natural gas unit (5% for each coal or lignite unit) may be offered no higher than the greater of $500 per MWh or 50 times the natural gas price. Additionally, up to 3% of the dispatchable capacity for each natural gas unit may be offered no higher than the system-wide offer cap. The amount of capacity covered by these provisions is approximately 500 MW. NRG's VMP shall remain in effect from the date it was approved by the Commission until terminated by the Executive Director of the Commission or by NRG.

Luminant received approval from the Commission for a VMP in May 2015.[71] The Luminant plan was similar in many respects to the NRG plan. Under the VMP, Luminant was permitted to offer a maximum of 12% of the dispatchable capacity for its natural gas units (5% for coal/lignite units) at prices up to $500 per MWh and offer a maximum of 3% of the dispatchable capacity for

---

69 PUCT Docket No. 40545, *Petition of Calpine Corporation for Approval of Voluntary Mitigation Plan*, Order (Mar. 28, 2013).

70 PUCT Docket No. 40488, *Request for Approval of a Voluntary Mitigation Plan for NRG Companies Pursuant to PURA § 15.023(f) and P.U.C. Subst. R. 25.504(e)*, Order (Jul. 13, 2012); PUCT Docket No. 42611, *Request for Approval of an Amended Voluntary Mitigation Plan for NRG Companies*, Order (Jul. 11, 2014).

71 PUCT Docket No. 44635, *Request for Approval of a Voluntary Mitigation Plan for Luminant Companies Pursuant to PURA § 15.023(f) and P.U.C. Subst. R. 25.504(e)*, Order Approving VMP Settlement (May 22, 2015).

natural gas units up to the system-wide offer cap. When approved, the amount of capacity covered by these provisions was approximately 500 MW. With the acquisition of three combined cycle units, the amount of offer flexibility had increased to approximately 900 MW. In addition, the plan contained a maximum offer for the approximately 1,000 MW of quick-start qualified combustion turbines owned by Luminant based on unit-specific verifiable costs and index prices for fuel and emissions. Luminant terminated its VMP on April 9, 2018 as a result of its merger with Dynegy, Inc.[72]

Approved on August 31, 2017,[73] Exelon's VMP provided for up to 12% but no more than 40 MW of dispatchable capacity from non-quick start natural gas units to be offered no higher than $500 per MWh or fifty times the fuel index price defined in the VMP. Up to 3% of the difference between the high sustained limit and the low sustained limit could have been offered at prices up to and including the system-wide offer cap. The amount of capacity covered by these provisions was slightly less than 600 MW. Exelon's VMP contained an automatic termination provision upon the earlier of three years from the date of the Commission's August 31, 2017 Order, or the day Exelon's Installed Generation Capacity dropped below 5% of the total ERCOT Installed Generation Capacity. On April 17, 2018, sufficient assets of Exelon's subsidiary, ExGen Texas Power, were transferred in a Chapter 11 bankruptcy process to drop Exelon below 5% of the total ERCOT Installed Generation Capacity. The VMP was automatically terminated, effective April 17, 2018.

The amount of offer flexibility afforded by the VMPs is small when compared to the offer flexibility that small participants – those with less than 5% of total ERCOT capacity – are granted under 16 TAC § 25.504(c). Although 5% of total ERCOT capacity may seem relatively trivial, the potential market impacts of a market participant whose size is just under the 5% threshold choosing to exercise flexibility and offering a significant portion of their fleet at very high prices can be large.

Currently, the 5% "small fish" threshold is just under 4,000 MW. The combined amount of capacity afforded offer flexibility under the VMPs granted to Calpine and NRG totals less than 1,200 MW of capacity.

---

72 *See Application of Luminant Power Generation LLC, Big Brown Power Company LLC, Comanche Peak Power Company LLC, La Frontera Holdings LLC, Oak Grove Management Company LLC, and Sandow Power Company Under Section § 39.158 of the Public Utility Regulatory Act*, Docket No. 47801 (Nov. 22, 2017); on April 9, 2018, Luminant filed a letter with the Commission terminating its VMP upon closing of the proposed transaction approved by the Commission in Finding of Fact No. 36 of the Order in Docket No. 47801.

73 PUCT Docket No. 47378*, Request for Approval of a Voluntary Mitigation Plan for Exelon Generation Company, LLC*, Order (Aug. 31, 2017).

Allowing small amounts of high-priced offers is intended to accommodate potential legitimate fluctuations in marginal cost that may exceed the base offer caps, such as operational risks, short-term fluctuations in fuel costs or availability, or other factors. However, the remaining VMPs contain a requirement that these offers, if offered in any hour of an operating day, must be offered in the same price and quantity pair for all hours of the operating day. This provision, along with the quantity limitations, significantly reduces the potential that the VMPs will allow market power to be exercised.

The final key elements in the VMPs are the termination provisions. The approved VMPs may be terminated by the Executive Director of the Commission with three business days' notice, subject to ratification by the Commission. PURA defines market power abuses as "practices by persons possessing market power that are unreasonably discriminatory or tend to unreasonably restrict, impair, or reduce the level of competition."[74] The exercise of market power may not rise to the level of an abuse of market power if the actions in question do not unreasonably impair competition, which would typically involve profitably raising prices significantly above the competitive level for a significant period of time. Thus, although the offer thresholds provided in the VMPs are designed to promote competitive market outcomes, the short termination provision provides additional assurance that any unintended consequences associated with the potential exercise of market power can be addressed in a timely manner rather than persisting and rising to the level of an abuse of market power.

B. Evaluation of Supplier Conduct

The previous subsection presented a structural analysis that supports inferences about potential market power. This subsection provides the results of evaluating actual participant conduct to assess whether market participants have attempted to exercise market power through physical or economic withholding. First, unit deratings and forced outages are examined to detect physical withholding. This is followed by an evaluation of the "output gap," used to detect economic withholding.

In a single-price auction like the real-time energy market, suppliers may attempt to exercise market power by withholding resources. The purpose of withholding is to cause more expensive resources to set higher market clearing prices, allowing the supplier to profit on its other sales in the real-time energy market. Because forward prices will generally be highly correlated with spot prices, price increases in the real-time energy market can also increase a supplier's profits in the bilateral energy market. This strategy is profitable only if the withholding firm's incremental profit as a result of higher price is greater than the lost profit from the foregone sales of its withheld capacity.

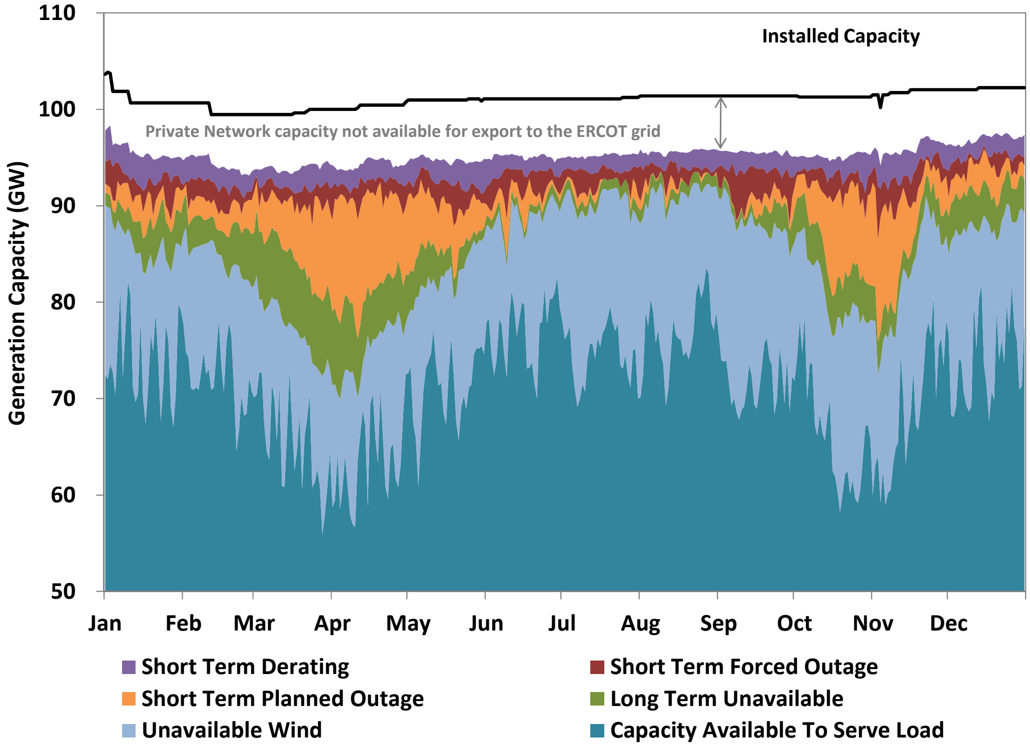
[74] PURA § 39.157(a).

Generation Outages and Deratings

At any given time, some portion of installed capacity is unavailable because of generator outages and deratings. Because of limitations in outage data, the outage type must be inferred. The outage type can be inferred by cross-referencing unit status information communicated to ERCOT with scheduled outage submissions. If there is a corresponding scheduled outage, the unit is considered to be on a planned outage. If not, it is considered to be a forced outage. The derated capacity is defined as the difference between the summertime maximum capacity of a generating resource and its actual capability as communicated to ERCOT on a continuous basis. It is very common for generating capacity to be partially derated because the resource cannot achieve its installed capacity level because of technical or environmental factors (e.g., component equipment failures or ambient temperature conditions). Wind generators rarely produce at the installed capacity rating because of variations in available wind input. Because such a large portion of derated capacity is related to wind generation it is shown separately in the following evaluation of long-term and short-term deratings.

Figure 95 shows a breakdown of total installed capacity for ERCOT on a daily basis during 2018. This analysis includes all in-service and switchable capacity. From the total installed capacity the following are subtracted: (a) capacity from private networks not available for export to the ERCOT grid; (b) wind capacity not available because of the lack of wind input; (c) short-term deratings; (d) short-term planned outages; (e) short-term forced outages; and (e) long-term outages and deratings greater than 30 days. What remains is the capacity available to serve load.

Figure 95: Reductions in Installed Capacity



Short-term outages and deratings of non-wind generators fluctuated between 2 and 20 GW, as shown in Figure 95, while wind unavailability varied between 4 and 20 GW. Short-term planned outages were largest in the shoulder months of April and October, while smallest during the summer months, consistent with expectations. Short-term forced outages and deratings had no discernable seasonal pattern, occurring throughout the year, also consistent with expectations.

The quantity of long-term (greater than 30 days) unavailable capacity, peaked in April at more than 7 GW, with almost all capacity returned to service in anticipation of tighter conditions during the summer of 2018.

The next analysis focuses specifically on short-term planned outages and forced outages and deratings of non-wind units because these classes of outages and deratings are the most likely to be used to physically withhold units in an attempt to raise prices. Figure 96 shows the average magnitude of the outages and deratings lasting less than 30 days for the year and for each month during 2018.

Figure 96: Short-Term Outages and Deratings

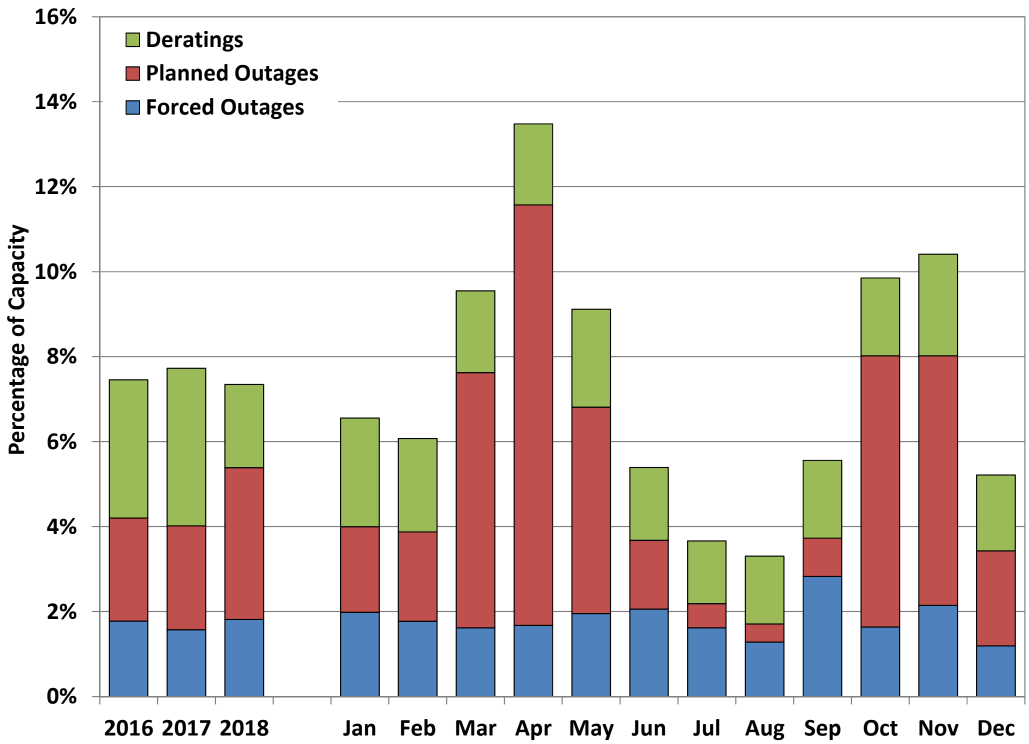
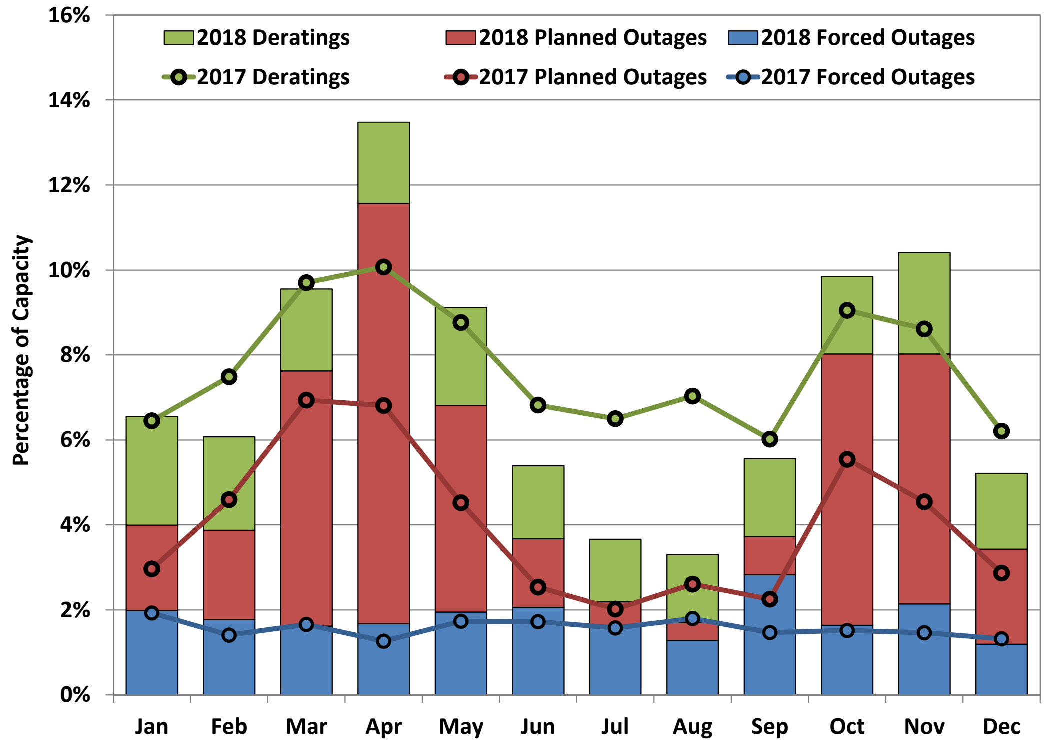


Figure 96 shows that total short-term deratings and outages were almost 14% of installed capacity in April and dropped to less than 4% during July and August. Most of this fluctuation was due to planned outages. The amount of capacity unavailable during 2018 averaged 7.3% of installed capacity, a slight decrease from 7.7% experienced in 2017 and 7.5% in 2016. The amount of planned outages was noticeably higher in 2018. This may be explained by the heightened expectations for shortages during the summer and generators taking outage time to ensure higher availability. The decrease in the amount of deratings may also be similarly explained by generators operating in modes that would allow them to maximize generation.

The following Figure 97 provides a comparison of the monthly outage and derating values for 2018 and 2017.

Figure 97: Derating, Planned Outages and Forced Outages



It shows the consistency of forced outages, with the interesting exception of lower forced outages in August 2018 followed by a higher forced outage rate in September. This pattern would seem to imply that generator operators were able to somehow defer the impacts of unexpected equipment limitations through August. However, those actions may have come at the cost of higher outage rates in September. The significant increase in planned outages scheduled during Spring and Fall 2018 may be an indicator of intense preparation for what was expected to be a summer with very tight operating conditions. The large reduction in deratings across all months of 2018 would also seem to indicate that generators were intent on making every bit of capability available to the market. It is unclear as to whether the low outage rates during August and the low level of derations were a direct result of increased planned maintence activities, or merely good fortune.

Evaluation of Potential Physical Withholding

Physical withholding occurs when a participant makes resources unavailable for dispatch that are otherwise physically capable of providing energy and that are economic at prevailing market prices. A plant operator can withhold by either by derating a unit or declaring the unit as forced out of service. Because generator deratings and forced outages are unavoidable, the goal of the analysis in this subsection is to differentiate justifiable deratings and outages from physical withholding. Physical withholding is tested for by examining deratings and outage data to ascertain whether the data are correlated with conditions under which physical withholding would likely be most profitable.

The RDI results shown in Figure 93 and Figure 94 indicate that the potential for market power abuse rises at higher load levels as the frequency of positive RDI values increases. Hence, if physical withholding is occurring, one would expect to see increased deratings and outages at the highest load levels. Conversely, because competitive prices increase as load increases, deratings and outages in a market performing competitively will tend to decrease as load approaches peak levels. Suppliers that lack market power will take actions to maximize the availability of their resources because their output is generally most profitable in peak periods.

Figure 98 shows the average relationship of short-term deratings and forced outages as a percentage of total installed capacity to real-time load levels for large and small suppliers during summer months. Portfolio size is important in determining whether individual suppliers have incentives to withhold available resources. Hence, the patterns of outages and deratings of large suppliers can be usefully evaluated by comparing them to the small suppliers' patterns.

Long-term deratings are not included in this analysis because they are unlikely to constitute physical withholding given the cost of such withholding. Wind and private network resources are also excluded from this analysis because of the high variation in the availability of these classes of resources. The large supplier category includes the five largest suppliers in ERCOT. The small supplier category includes the remaining suppliers.

Figure 98: Outages and Deratings by Load Level and Participant Size, June-August

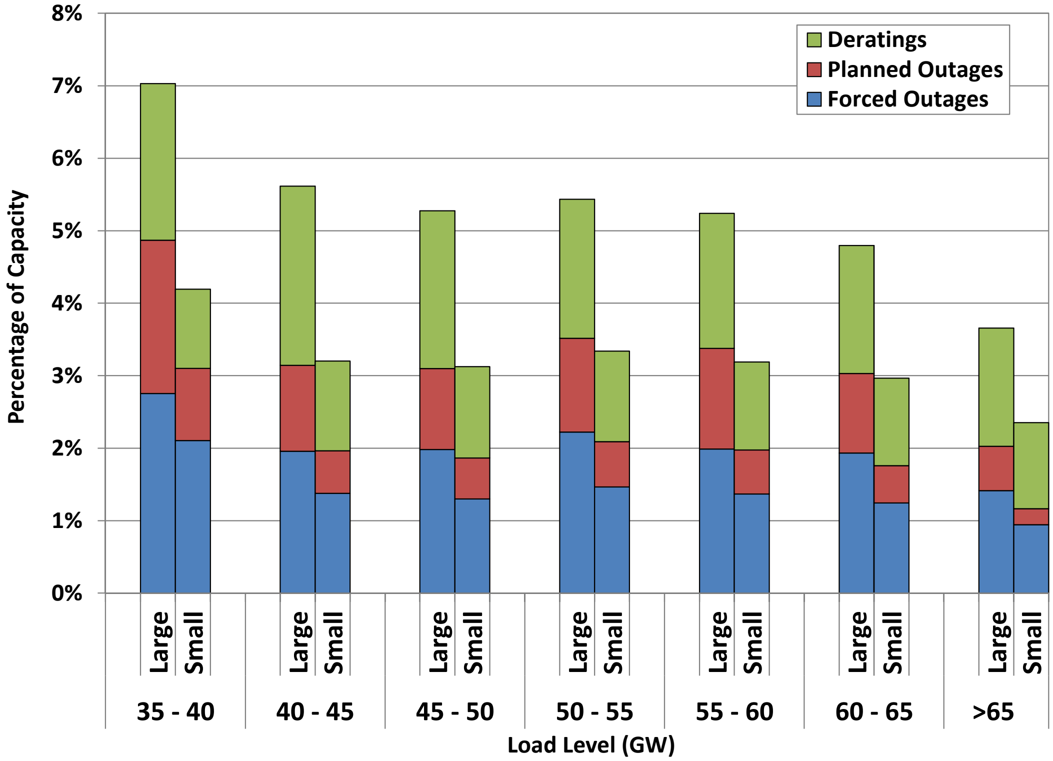


Figure 98 suggests that as demand for electricity increases, all market participants tend to make more capacity available to the market by scheduling planned outages during low load periods. Because small participants have less incentive to physically withhold capacity, the outage rates for small suppliers serves as a good benchmark for competitive behavior expected from the larger suppliers. Outage rates for large suppliers at all load levels exceeded those for small suppliers, but are not at levels that raise competitiveness concerns. Outages rates for small suppliers were historically low in 2018. Small suppliers have the most incentive to ensure generator availability because each unit in their fleet makes up a larger percentage of the total, which means that any outage has the potential for larger financial impacts. Again, it is not clear whether the low outage rates seen in 2018 were the result of improved scheduled maintence procedures, or merely good fortune.

Evaluation of Potential Economic Withholding

To complement the prior analysis of physical withholding, this subsection evaluates potential economic withholding by calculating an "output gap." The output gap is the quantity of energy that is not being produced by online resources even though the output is economic to produce by a substantial margin given the real-time energy price. A participant can economically withhold resources, as measured by the output gap, by raising its energy offers so as not to be dispatched.

A resource is evaluated for inclusion in the output gap when it is committed and producing at less than full output. Energy not produced from a committed resource is included in the output gap if the real-time energy price exceeds that unit's mitigated offer cap by at least $30 per MWh.[75] The mitigated offer cap serves as a proxy for the marginal production cost of energy from that resource.

Before presenting the results of the output gap analysis, a description of ERCOT's two-step dispatch software is required. In the first step, the dispatch software calculates output levels (base points) and associated locational marginal prices using the participants' offer curves and only considering transmission constraints that have been deemed competitive. These "reference prices" at each generator location are compared with the generator's mitigated offer cap, and the higher of the two is used to formulate the offer curve for that generator during the second step in the dispatch process. The resulting mitigated offer curve is used by the dispatch software to determine the final output levels for each generator, taking all transmission constraints into consideration.

If a market participant has sufficient market power, it might raise its offer in such a way to increase the reference price in the first step. Although in the second step the offer appears to be mitigated, the market participant has still influenced the market price. This output gap is measured by the difference between the capacity level on a generator's original offer curve at the first step reference price and the capacity level on the generator's cost curve at the first step reference price. However, this output gap is only indicative because no output instructions are sent based on the first step. The output gap is only used to screen whether a market participant is withholding in a manner that may influence the reference price.

The results of the analysis shown below in Figure 99 indicate that only very small amounts of capacity would be considered part of the first step output gap. The amount of capacity shown reflects the average for only the intervals in which an output gap was detected. If averaged over all intervals the amounts are de minimis.

75 Prior to 2015, the State of the Market report used $50 for the output gap margin. With the reduction in average prices, the gap was reduced.

Figure 99: Incremental Output Gap by Load Level and Participant Size – Step 1

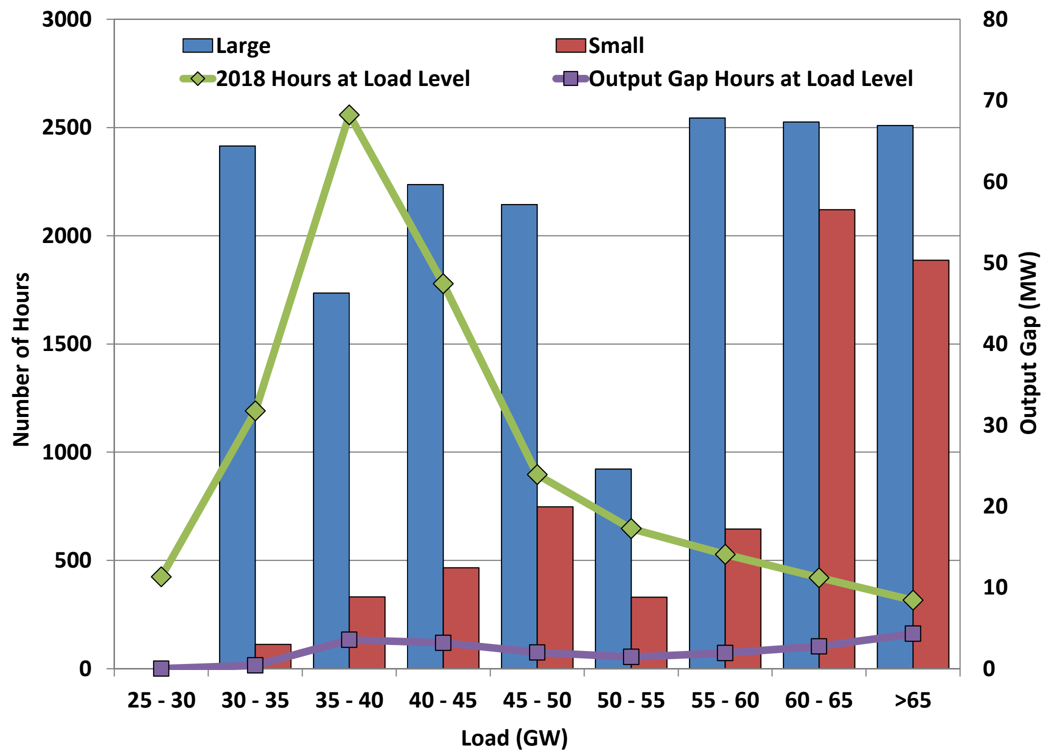
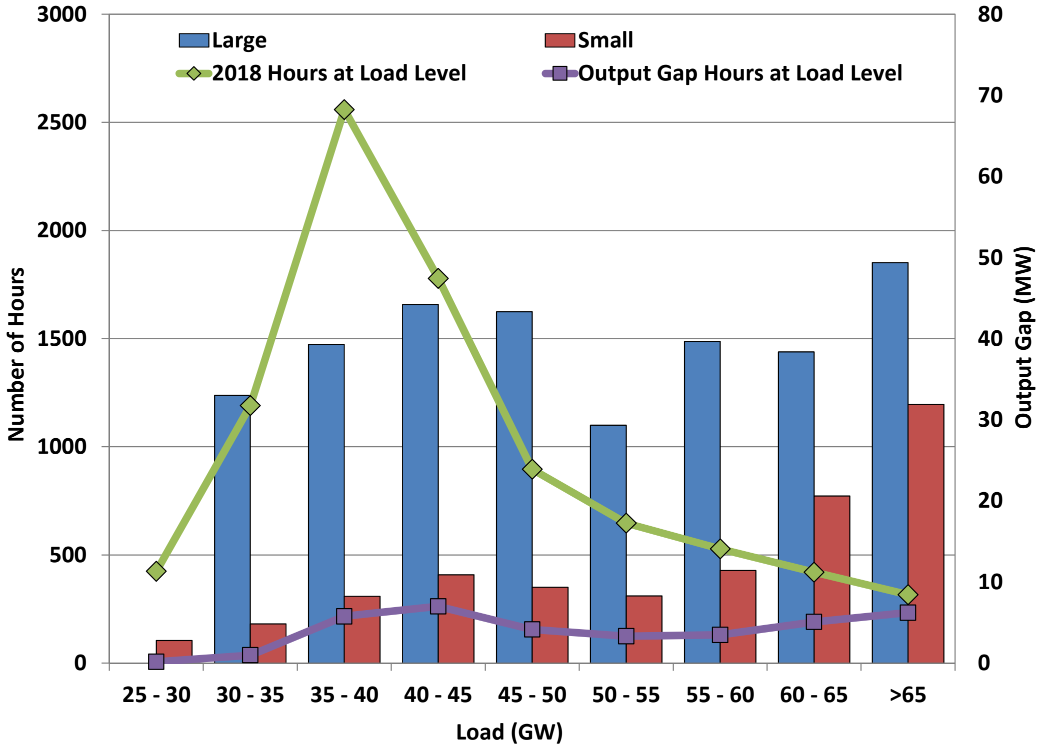


Figure 99 above shows the ultimate output gap levels, measured by the difference between a unit's operating level and the output level had the unit been competitively offered to the market. In the second step of the dispatch, the after-mitigation offer curve is used to determine dispatch instructions and locational prices. As previously illustrated, even though the offer curve is mitigated, there is still the potential for the mitigated offer curve to be increased. This increase would be the result of a high first-step reference price being influenced by a market participant raising prices. The percentage of a step 1 output gap hours evident in 2018 was 8%.

Similar to the previous analysis, Figure 100 also shows very small quantities of capacity that would be considered part of this output gap, with only 15% of the hours in 2018 that exhibited an output gap.

Figure 100: Incremental Output Gap by Load Level and Participant Size – Step 2



These results show that potential economic withholding levels were extremely low in 2018. Output gaps of the largest suppliers are routinely monitored individually and were found to be consistently low across all load levels. These results, together with our evaluation of the market outcomes presented in this report, allow us to conclude that the ERCOT market performed competitively in 2018.