2017 State of the Market Report for the ERCOT Electricity Markets

POTOMAC ECONOMICS

Independent Market Monitor
for ERCOT

May 2018

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Executive Summary

As the Independent Market Monitor (IMM) for the Electric Reliability Council of Texas (ERCOT), Potomac Economics provides this report which reviews and evaluates the outcomes of the ERCOT wholesale electricity markets in 2017. It is submitted to the Public Utility Commission of Texas (PUCT) 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. This report also assesses the effectiveness of the Scarcity Pricing Mechanism (SPM) pursuant to the provisions of 16 Tex. Admin. Code (TAC) § 25.505(g).

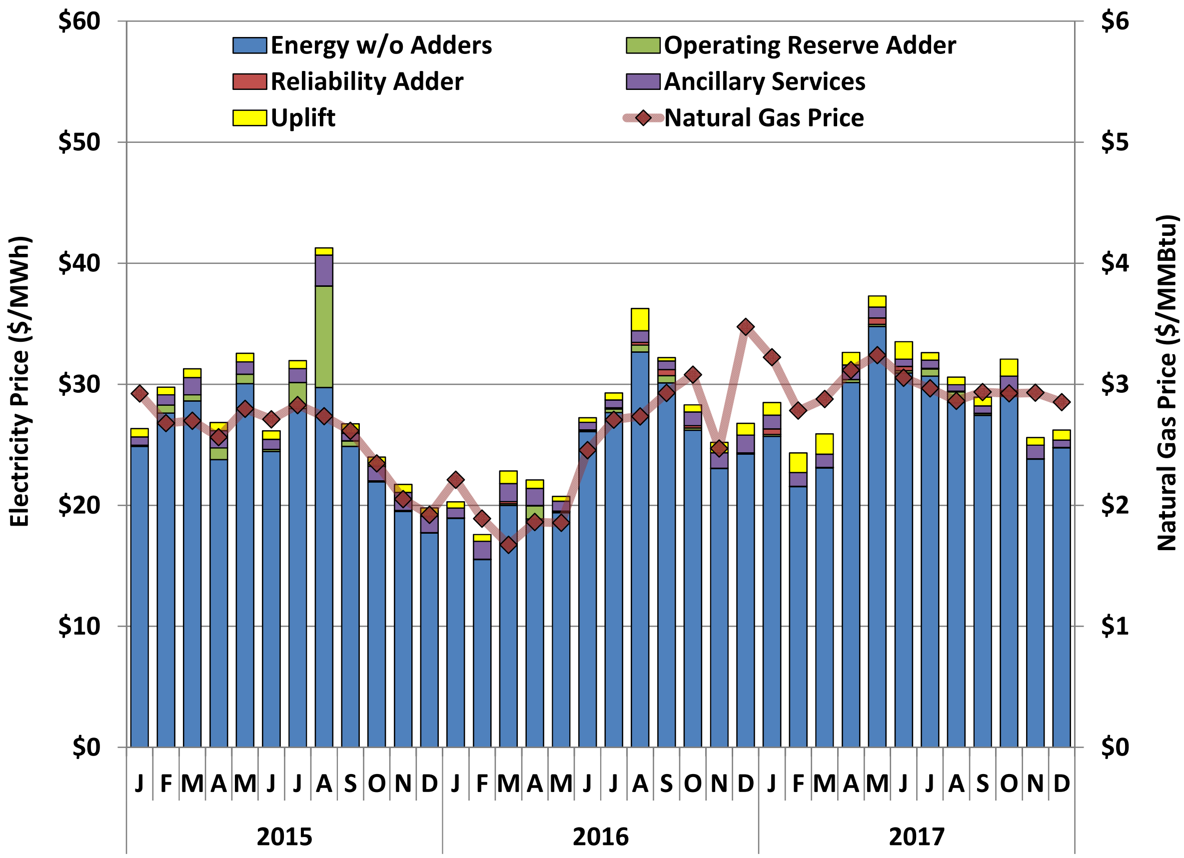
Overall, the ERCOT wholesale market performed competitively in 2017. Key findings and results from 2017 include the following:

- Higher natural gas prices led to higher energy prices in 2017:
  - The ERCOT-wide load-weighted average real-time energy price was $28.25 per MWh in 2017, a 14.7% increase from 2016.
  - The average price for natural gas was 22% higher in 2017 than in 2016, increasing from $2.45 per MMBtu in 2016 to $2.98 per MMBtu in 2017.
- Market conditions were rarely tight – real-time prices did not exceed $3,000 per MWh in 2017 and exceeded $1,000 per MWh for only 3.5 hours cumulatively for the year.
- The peak hour demand in ERCOT was 69,512 MW in 2017, a 2.2% decrease from the all-time hourly demand record of 71,110 MW set on August 11, 2016. However, average demand rose in 2017, increasing 1.9% from 2016.
- The total congestion costs experienced in the ERCOT real-time market in 2017 were $967 million, an increase of 95% from 2016. Three factors contributed to the substantial increase: 1) continued limitations on export capacity from the Panhandle, 2) planned outages associated with construction of the Houston Import Project, and 3) unusual operating conditions in the aftermath of Hurricane Harvey.
- Net revenues provided by the market during 2017 were less than the estimated amount necessary to support new greenfield generation investment, which is not a surprise given that planning reserves were above the minimum target and shortages were again rare in 2017. 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 spot market, real-time energy prices are very important because they set the expectations for prices in the day-ahead market and other forward markets where most transactions occur. Unless there are barriers preventing arbitrage of the prices between the spot and forward markets, the prices in the forward market should be directly related to the prices in the spot 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 to show them on a per MWh basis.[1]

Average All-in Price for Electricity in ERCOT



ERCOT developed two energy price adders that are designed to improve its real-time energy pricing when reserves become scarce or ERCOT takes out-of-market actions for reliability. To

1 For this analysis uplift includes: Reliability Unit Commitment (RUC) 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.

distinguish the effects of the energy price adders, the Operating Reserve Demand Curve Adder (operating reserve adder) and the Reliability Deployment Price Adder (reliability adder) are shown separate from the energy price. The operating reserve adder was implemented in mid-2014 to account for the shortage 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 depress the energy prices.

The largest component of the all-in price is the energy cost, which continues to be highly correlated with natural gas 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 22% increase in natural gas prices contributed to a 15% increase in ERCOT's average real-time energy prices. The all-in price in 2017 included small contributions from ERCOT's energy price adders – $0.24 per MWh from the operating reserve adder and $0.16 per MWh from the reliability adder.

Finally, the other classes of costs continue to be a small portion of the all-in electricity price – ancillary services costs were $0.87 per MWh, down from $1.03 per MWh in 2016 because of continued relatively low natural gas prices and lower ancillary service requirements. Uplift costs, including the ERCOT system administrative fee, accounted for $1.03 per MWh of the all-in electricity price, up from $0.74 per MWh in 2016.

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
ERCOT	**$53.23**	**$28.33**	**$33.71**	**$40.64**	**$26.77**	**$24.62**	**$28.25**
Houston	$52.40	$27.04	$33.63	$39.60	$26.91	$26.33	$31.81
North	$54.24	$27.57	$32.74	$40.05	$26.36	$23.84	$25.67
South	$54.32	$27.86	$33.88	$41.52	$27.18	$24.78	$29.38
West	$46.87	$38.24	$37.99	$43.58	$26.83	$22.05	$24.52
($/MMBtu)							
Natural Gas	$3.94	$2.71	$3.70	$4.32	$2.57	$2.45	$2.98

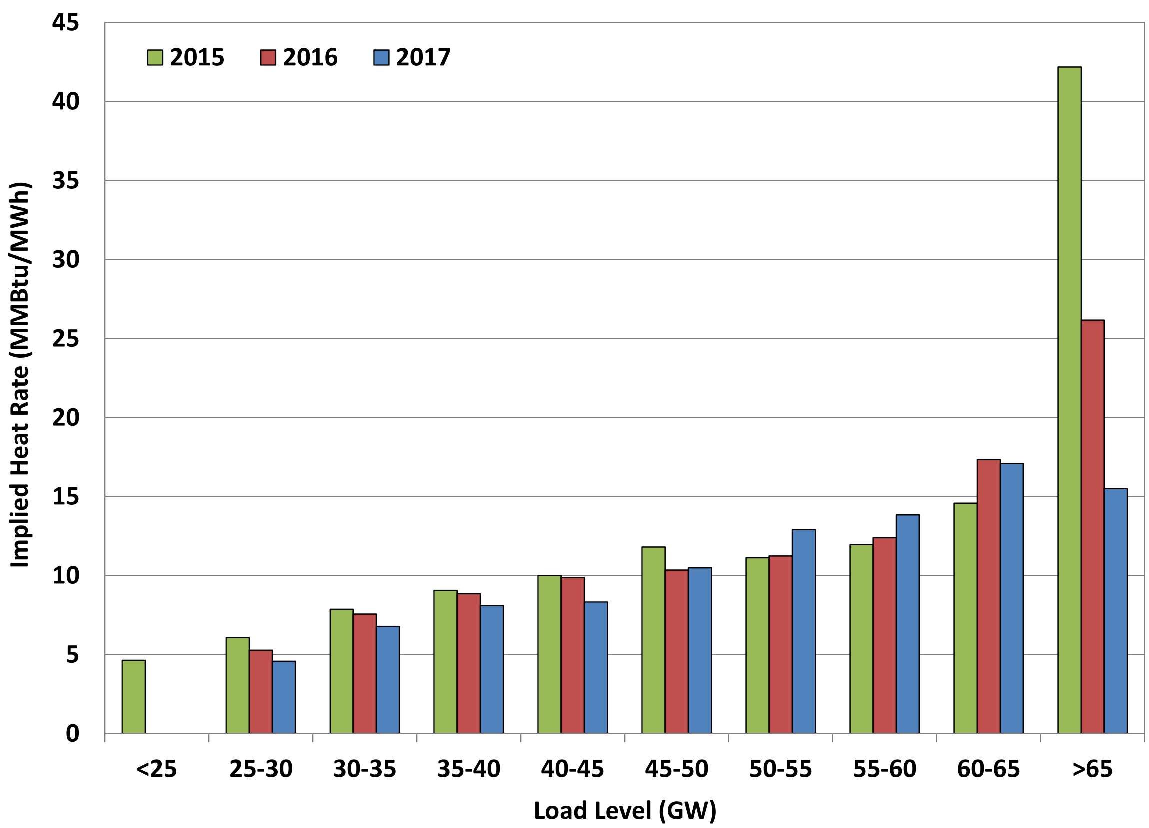
The table above provides the annual load-weighted average price for each zone for the past seven years. The difference in zonal prices in 2017 are directionally comparable to the prices in 2016. Constraints on the ability to import generation led to the Houston zone being the highest priced zone in 2017. Export limitations resulted in the West zone having the lowest price. However, price spreads were larger in 2017 because of higher natural gas prices and the increased impacts of transmission congestion.

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 surplus wind generation resulting from export limitations. Between 2012 and 2014, load growth because of higher oil and natural gas production activity resulted in localized import constraints and higher prices. Even with continued investment in transmission facilities, the continued entry of wind generation has led to export congestion and lower average prices since 2015.

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

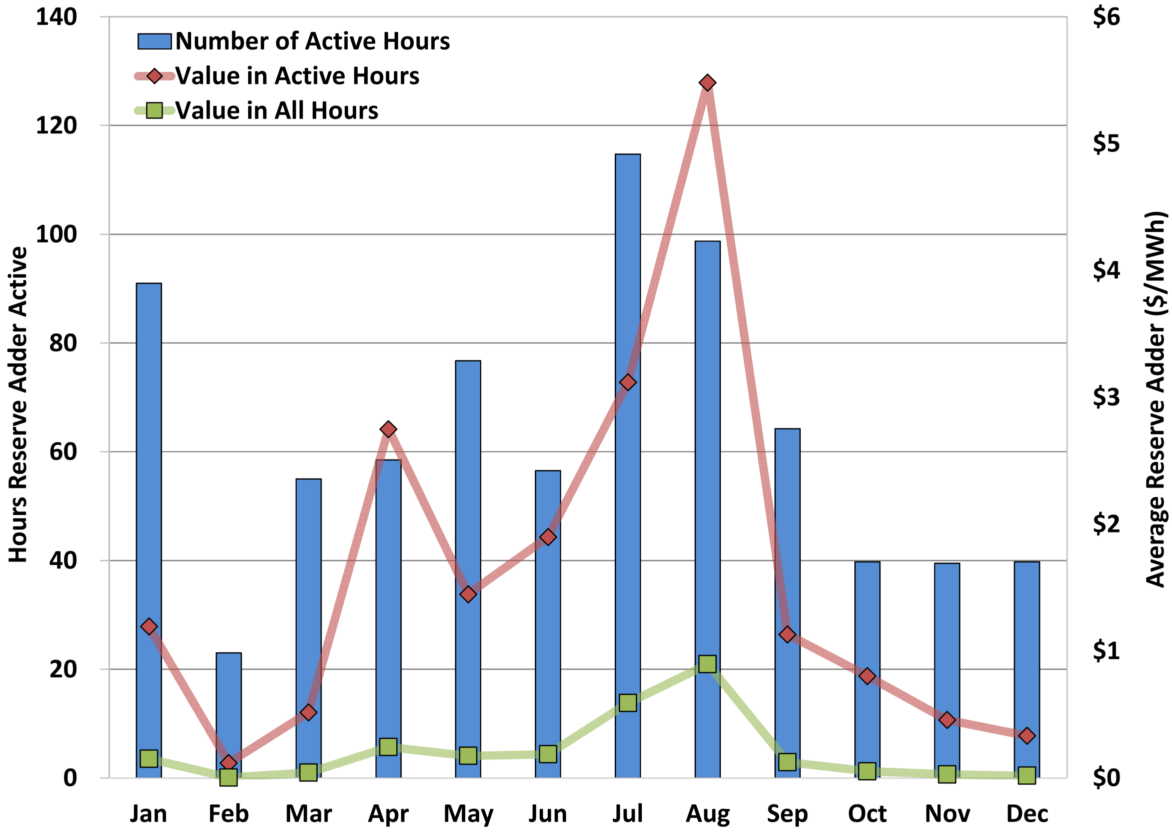


The figure shows the average implied heat rate at various system load levels from 2015 through 2017. In a well-performing market, a positive relationship between these two variables is expected because resources with higher marginal costs are dispatched to serve higher loads.

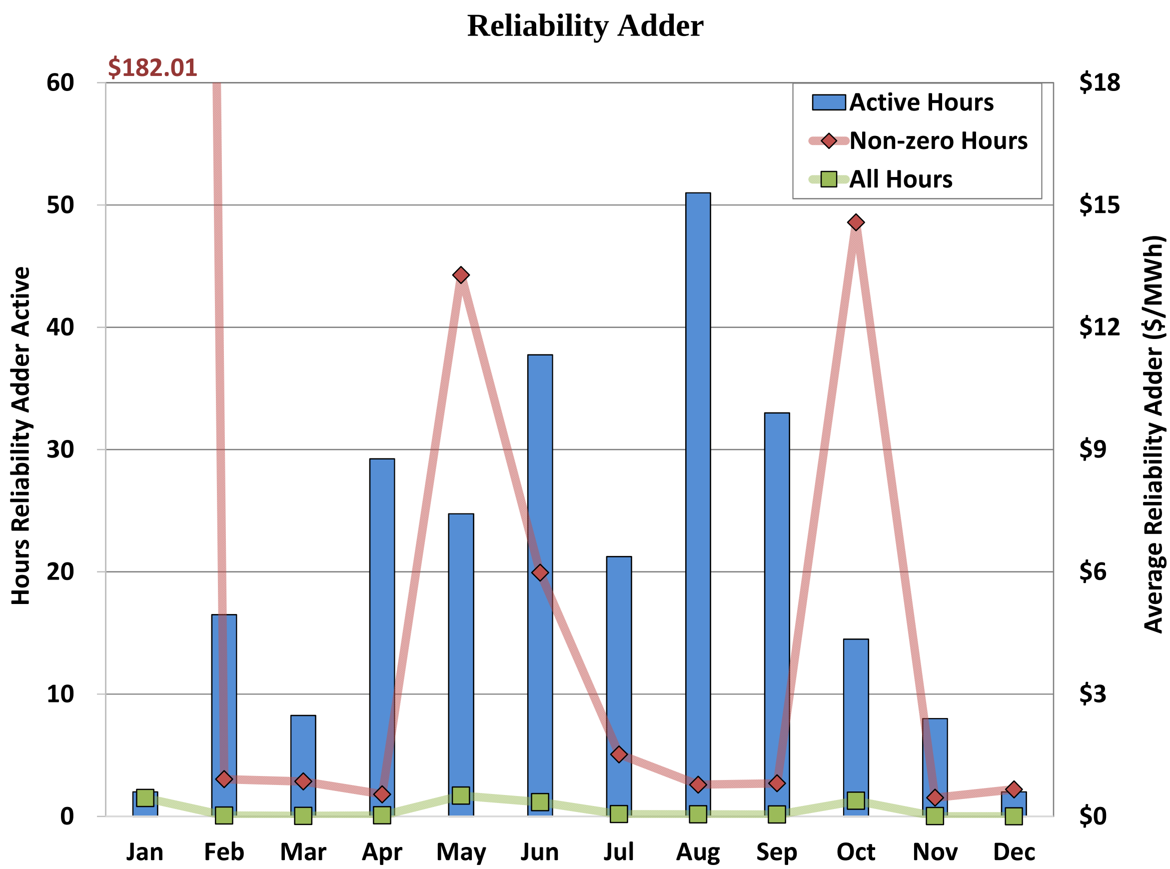
Energy Price Adders

As described above, the contributions of the energy price adders were relatively small in 2017. The first of the two adders, the operating reserve adder, is a shortage value intended to reflect the expected value of lost load (the loss of load probability, given online and offline reserve levels multiplied by the deemed value of lost load). The operating reserve adder had the largest impacts on prices during July and August. Overall, the operating reserve adder contributed $0.24 per MWh or less than 1% to the annual average real-time energy price because the system was rarely short of reserves.

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.



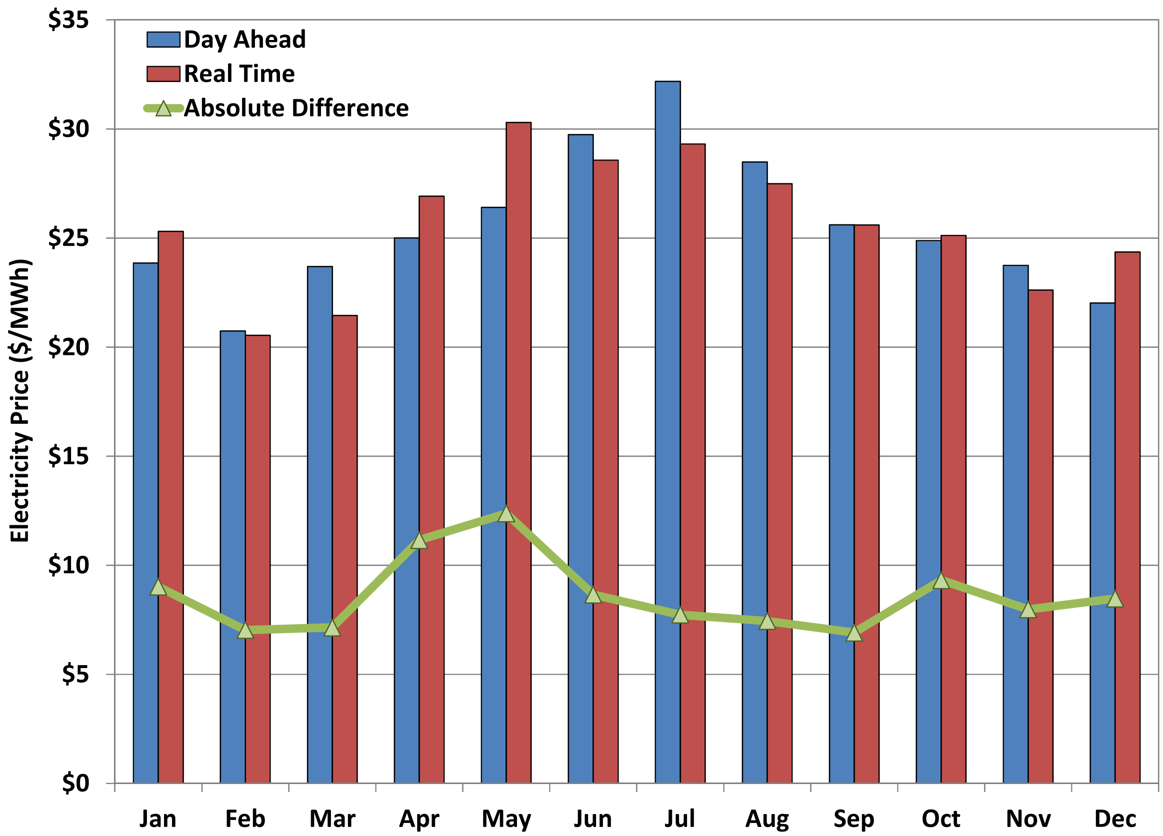
When the averaged across only the active hours, the largest price impacts of the reliability adder occurred in January. The reliability adder was non-zero for fewer than 250 hours, or less than 3% of the time in 2017, most of which occurred in August. The contribution from the reliability adder to the annual average real-time energy price was $0.16 per MWh. Like the operating reserve adder, it had very little overall effect on the market outcomes in 2017 because the supply conditions were rarely tight and ERCOT took fewer reliability actions in 2017.

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 volatility in the real-time market by purchasing in the day-ahead market. Finally, the day-ahead market plays a critical role in coordinating generator commitments. For all these reasons, the performance of the day-ahead market is essential.

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. The next figure shows the price convergence between the day-ahead and real-time markets in 2017.

Convergence Between Day-Ahead and Real-Time Energy Prices



Price convergence was very good in 2017; day-ahead and real-time prices both averaged $26 per MWh.[2] The average absolute difference between day-ahead and real-time prices was $8.60 per MWh in 2017 – a slight increase from $7.44 per MWh and $8.08 per MWh in 2016 and 2015, respectively.

This day-ahead premium 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

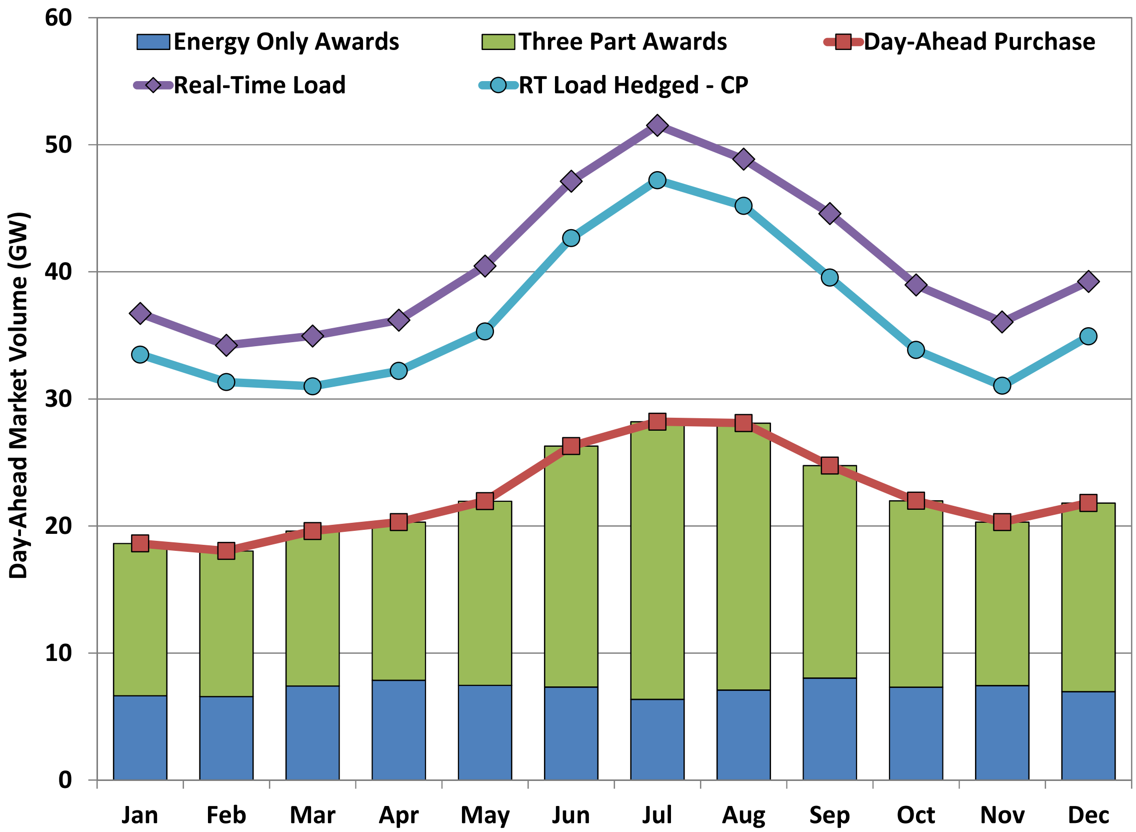
2 These values are simple averages, not load-weighted.

the highest premiums occurred during the summer months in 2017 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[3] 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 55% of real-time load in 2017, which was a slight increase compared to 53% in 2016.

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

3 A Point-to-Point obligation is a type of CRR that entitles the holder to be charged or to receive compensation and is evaluated in each CRR Auction and day-ahead market as the positive and negative power flows on all directional network elements created by the injection and withdrawal at the specified source and sink points of the quantity represented by the CRR bid or offer (MW).

participant to buy the network flow from one location to another.[4] When coupled with a self-scheduled generating resource, the PTP obligation allows a participant to service 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.

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 the figure above. Approximately 82% of QSEs’ real-time load was hedged in the day-ahead market. 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 amount of real-time load hedged increased to nearly 90%.

Ancillary Service Prices

Total requirements for ancillary services declined again in 2017, resulting in lower prices and lower total costs for ancillary services. Under the nodal market, ancillary services and energy are co-optimized in the day-ahead market. This means that market participants do not have to include expectations of forgone energy sales in ancillary service 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.

During the recent period of low energy prices, this correlation has not been apparent. Monthly average prices for responsive reserve service varied from $7 to $13 per MWh, with the highest price occurring in January. One possible explanation for this decoupling from day-ahead energy prices is that unit commitment patterns have changed because of high wind generation and less online capacity capable of providing reserves. This reduction in online capacity, especially in off-peak periods has led to higher prices for reserve prices in shoulder months.

The next table compares the average annual price for each ancillary service in 2017 with 2016. The changes in total requirements for ancillary services in 2017 led to concomitant changes in ancillary service prices. The average price for responsive reserve service decreased in 2017, as did the total requirements for the service. Reductions in the average price for non-spinning reserves is consistent with the reduced requirements for this product. Average prices for regulation up and down products increased in 2017 even though requirements for the two products both decreased slightly.

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

The prices for all of the ancillary service products remain modest in part because of the lack of shortages in 2017. When ERCOT experiences a shortage of operating reserves, real-time prices will rise to reflect the expected value of lost load embedded in the ORDC mechanism. The expectation of higher real-time prices will tend to drive up the day-ahead price for ancillary services. Hence, the lack of shortages contributed to the low average ancillary service prices shown in the table.

Average Annual Ancillary Service Prices by Service

	2016 ($/MWh)	2017 ($/MWh)
Responsive Reserve	$11.10	$9.77
Nonspin Reserve	$3.91	$3.18
Regulation Up	$8.20	$8.76
Regulation Down	$6.47	$7.48

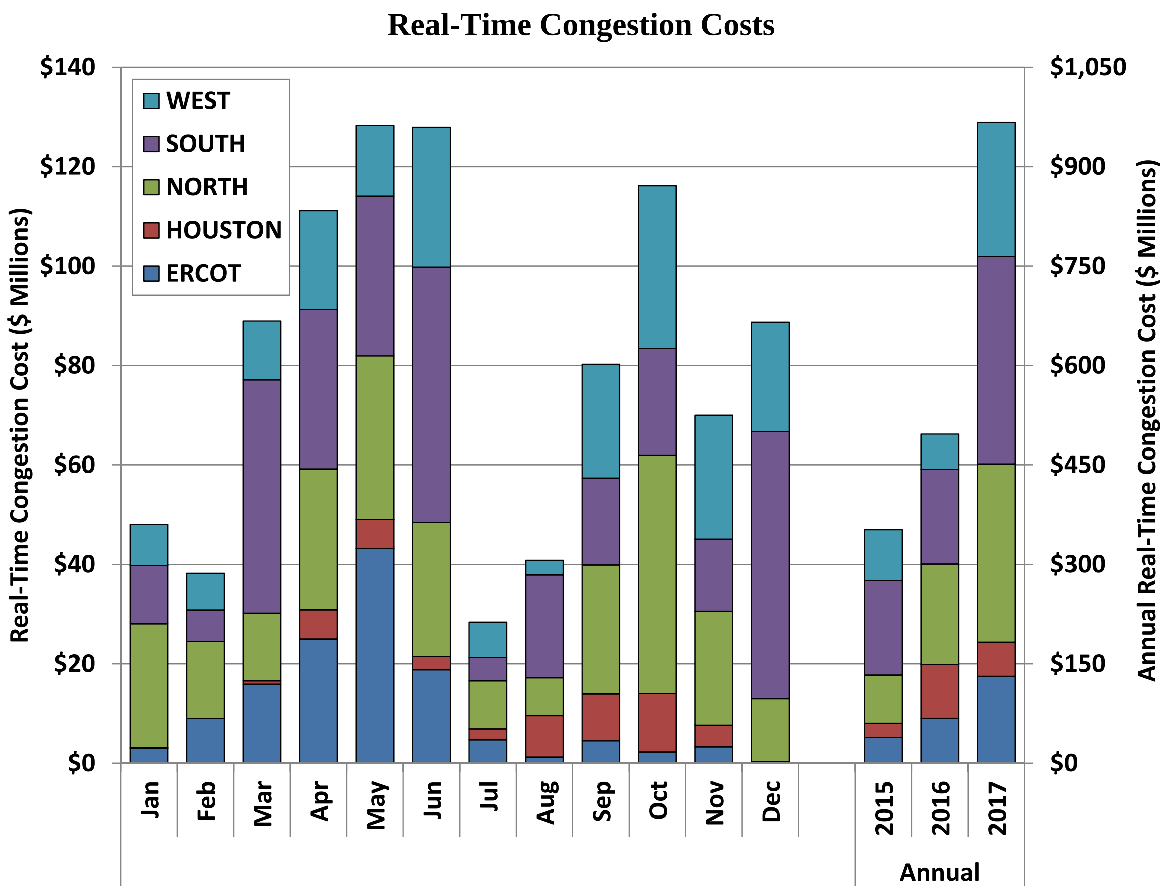
Transmission and Congestion

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 in the ERCOT real-time market in 2017 were $967 million, almost twice the congestions costs in 2016. Three factors contributed to the substantial increase: 1) continued limitations on export capacity from the Panhandle, 2) planned outages associated with construction of the Houston Import Project[5], and 3) unusual operating conditions in the aftermath of Hurricane Harvey. Congestion was more frequent in 2017, occurring in 16% more intervals than in 2016. All zones except for the Houston zone experienced increased congestion in 2017.

The next figure displays the amount of real-time congestion costs associated with each geographic zone, with the monthly values of 2017 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.

5 The Houston Import Project, which consists of the installation of a Limestone-Gibbons Creek-Zenith 345 kV double circuit line to meet reliability requirements for Houston load growth. The project was approved by the ERCOT Board of Directors on April 8, 2014



The months of January, February, July and August exhibited the least amount of congestion costs. The remaining months, typically the "shoulder months," reflected much higher congestion. This trend is expected because most transmission and generation outages for maintenance and upgrades occur during the shoulder months.

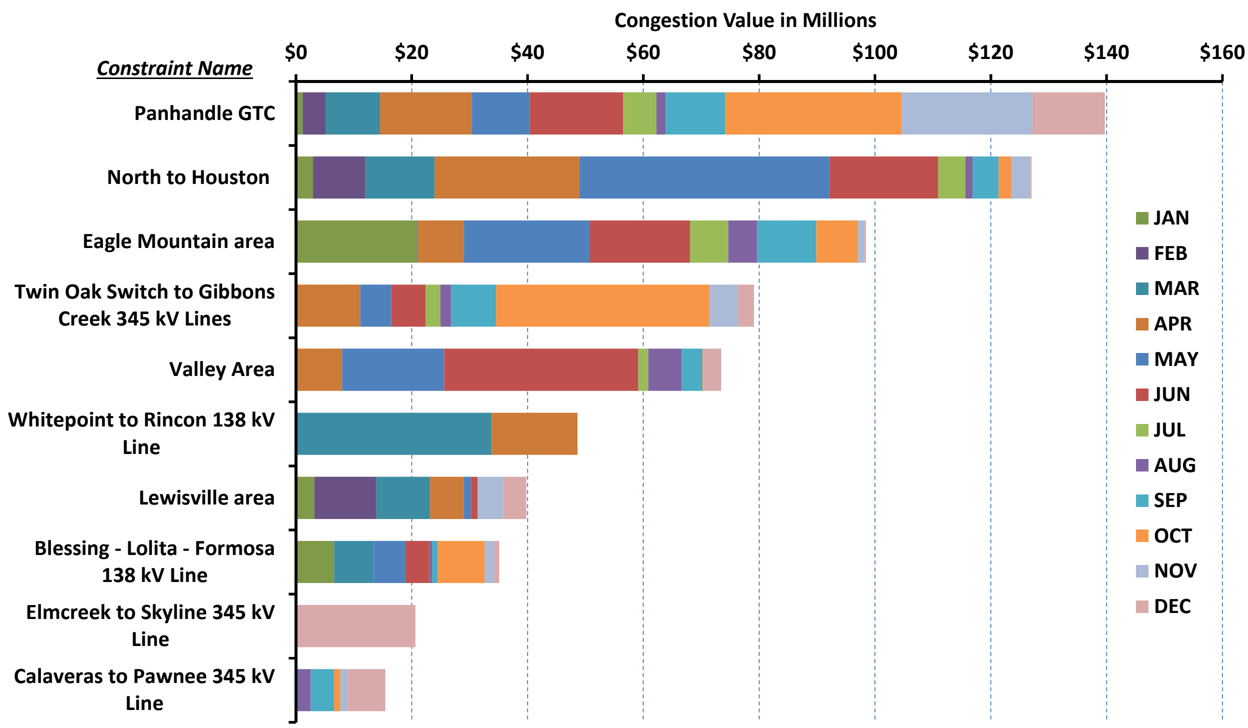
To better understand the main drivers of congestion in 2017, the next analysis summarizes the constraints that generated the highest congestion costs. For this discussion, the constraints groupings are determined by consolidating multiple real-time transmission constraints that are determined to be similar because of geographic proximity and constraint direction.

The figure below displays the ten most costly real-time constraints as measured by congestion value. The constraint with the highest congestion value in 2017 was the Panhandle Generic Transmission Constraint (GTC) at $139 million, a fivefold increase from 2016. By the end of 2017, there was almost 5 GW of generation capacity in the Panhandle area, 85% of which was wind generation. The highest GTC limit for the Panhandle was less than 4 GW, leading to frequent (16% of the intervals) and costly congestion in periods when wind output was high.

Congestion on the North to Houston constraint declined sharply after June due to the completion of a new 1200MW combined cycle generator located in Houston, combined with reduced load in

Houston as a result of the flooding damage caused by Hurricane Harvey. Lastly, the sizable congestion that occurred in September and October on the Twin Oak Switch to Gibbons Creek transmission path was largely caused by outages necessary to facilitate the construction of the Houston Import Project.

Most Costly Real-Time Constraints



Demand and Supply

Load in 2017

Total ERCOT load in 2017 increased 1.9% (approximately 780 MW per hour on average) to total 357.4 TWh. All zones showed an increase in average real-time load in 2017. The West zone saw the largest average load increase at 8.3%, which was likely due to continuing robust oil and natural gas production activity. Weather impacts on load in 2017 were mixed. Cooling degree days, a metric that is highly correlated with weather-related summer load, exhibited no change in Houston, decreased in Dallas and increased in Austin compared to 2016.

Summer conditions in 2017 produced load that peaked at 69,512 MW on July 28, short of the ERCOT-wide coincident peak hourly load record of 71,110 MW set on August 11, 2016. Further, demand did not ever exceed 70,000 MW in 2017, compared to five separate hours in 2016. The zones experienced varying changes in peak load. The West zone continued to experience the highest percentage growth in peak load, which was likely driven by continuing growth in oil and natural gas production.

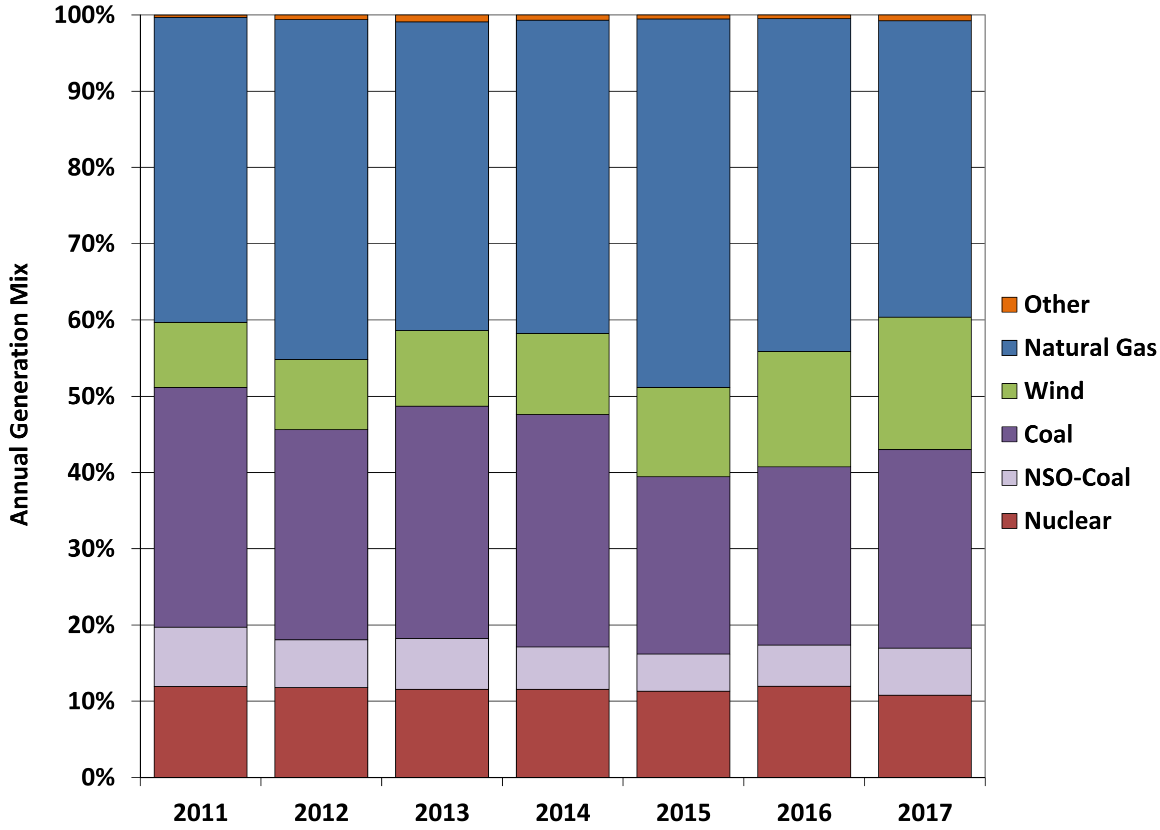
Generating Resources

Approximately 3.6 GW of new generation resources came online in 2017; the bulk of which was two new combined cycle natural gas units with total capacity of 2.2 GW. Wind additions totaled 1.1 GW with an effective peak serving capacity of less than 300 MW. The remaining capacity additions were 180 MW of new combustion turbines and 160 MW of solar. Fourteen generation resources totaling 1,222 MW, consisting primarily of aging natural gas generation, were retired in 2017.

Given these additions and retirements, shares of natural gas and coal capacity did not change significantly in 2017, representing 46% and 18% of installed capacity, respectively.

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 2017.

Annual Generation Mix



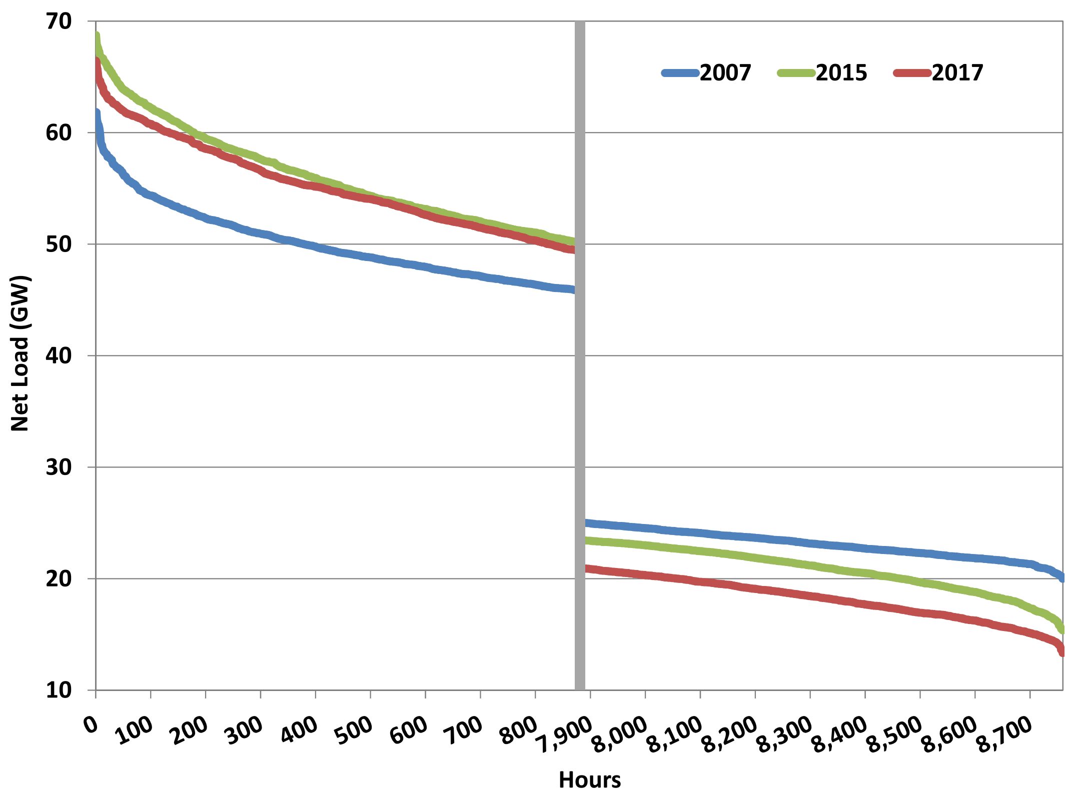
The generation share from wind has increased every year, reaching 17% of the annual generation requirement in 2017, up from 9% in 2011 and 15% in 2016. While the share of generation from coal had declined significantly from 2014 to 2015, its share has increased the last two years up to 32% in 2017. This figure separately shows the amount of energy produced from coal units scheduled to retire in 2018 (i.e., those that have submitted a Notification of Suspension of

Operations or NSO). These seven units have provided an average of 6% of the total annual generation output over the past seven years. As wind and coal output has increased, natural gas output declined from its high point of 48% in 2015 down to 39% in 2017. This trend should reverse, however, once the coal resources mentioned above retire.

Wind Output

ERCOT continued to set new records for peak wind output in 2017. On November 17, wind output set a new record at more than 16 GW, providing nearly 42% of the total load.[6] 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 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



Even with the increased development activity in the coastal area of the South zone, 73% 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. This profile 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

[6] Peak hourly wind generation was 16,035 MW on November 17, 2017, at 10:00 p.m.

low load hours of the year. Hence, 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 above), the difference between peak net load and the 95th percentile of net load has averaged 12.3 GW for the past three years. This means that 12.3 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), the minimum net load has dropped from approximately 20 GW in 2007 to below 13.3 GW in 2017, even with the sizable growth in annual load that has occurred. This continues to put operational pressure on the almost 25 GW of nuclear and coal generation that were in-service in 2017. Together with the decline in natural gas prices and average electricity price, this operational pressure has contributed to the recent retirement of more than 4 GW of coal.

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 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

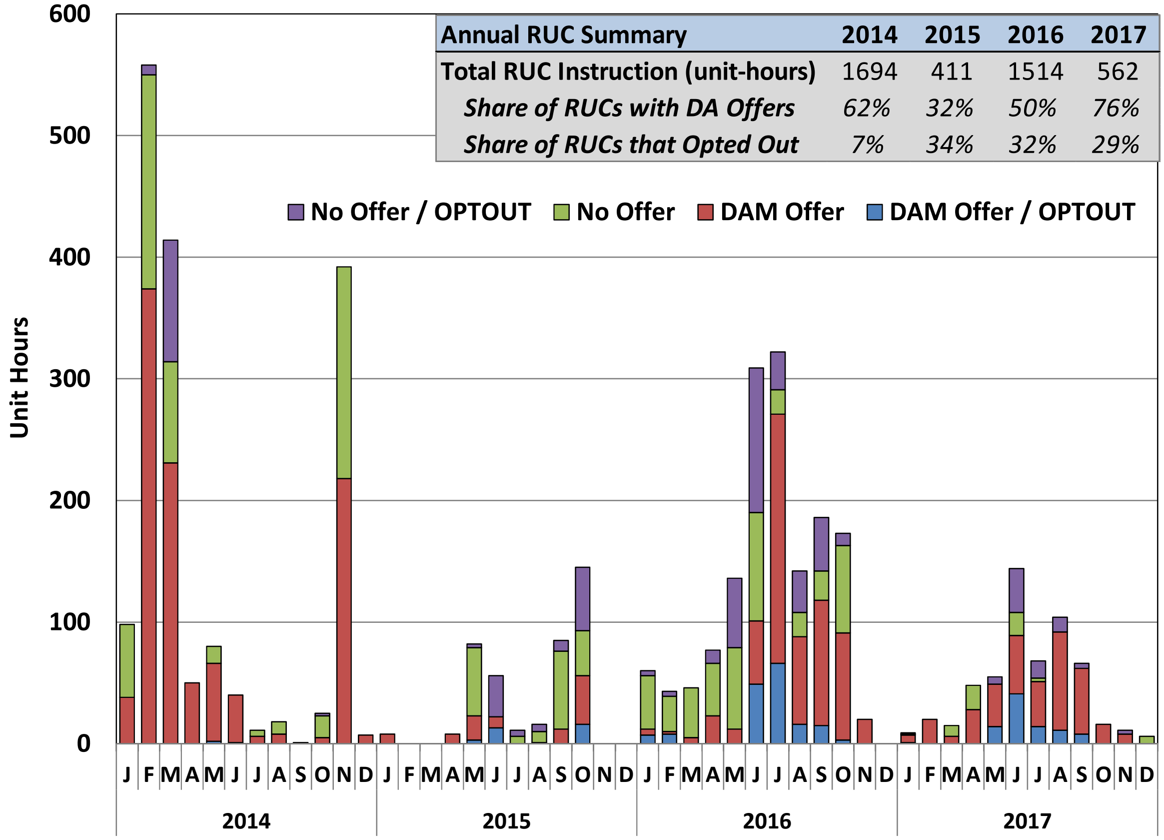
One of the important characteristics of any electricity market is the extent to which it results in the efficient commitment of generating resources. Under-commitment can cause apparent shortages in the real-time market and inefficiently high energy prices; while over-commitment can result in excessive 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 RUC process that 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 constraint may be either a thermal limit or a voltage concern. 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 2017 fell by 63% from 2016, despite the increase in congestion that occurred in 2017. Like 2016, most reliability commitments were made primarily to manage transmission constraints in 2017 (84% of unit-hours), including 7% to manage congestion in the aftermath of Hurricane Harvey. Only 13% of RUC instructions were made to ensure sufficient system-wide capacity and 2% for voltage support.

Having a day-ahead offer allows a generator to avoid revenue clawback associated with RUC instructions. Nonetheless, in 2017, only 76% of the generators receiving RUC instructions had day-ahead offers, a relatively low percentage considering the incentive to provide day-ahead offers inherent in the RUC claw-back rules. This low percentage was an increase from 2016 when the ratio was 50%. This may indicate that some of the reduction in the RUC activity in 2017 was due to a larger share of the units needed for reliability being committed through the day-ahead market.

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 2017 that chose to opt-out was 29%, similar to the 32% of generators that chose to opt-out in 2016.

During 2017, $1.2 million was clawed back from RUC units while only $0.5 million in make-whole payments were made to RUC units. All RUC make-whole payments in 2017 were collected from QSEs that were capacity short.

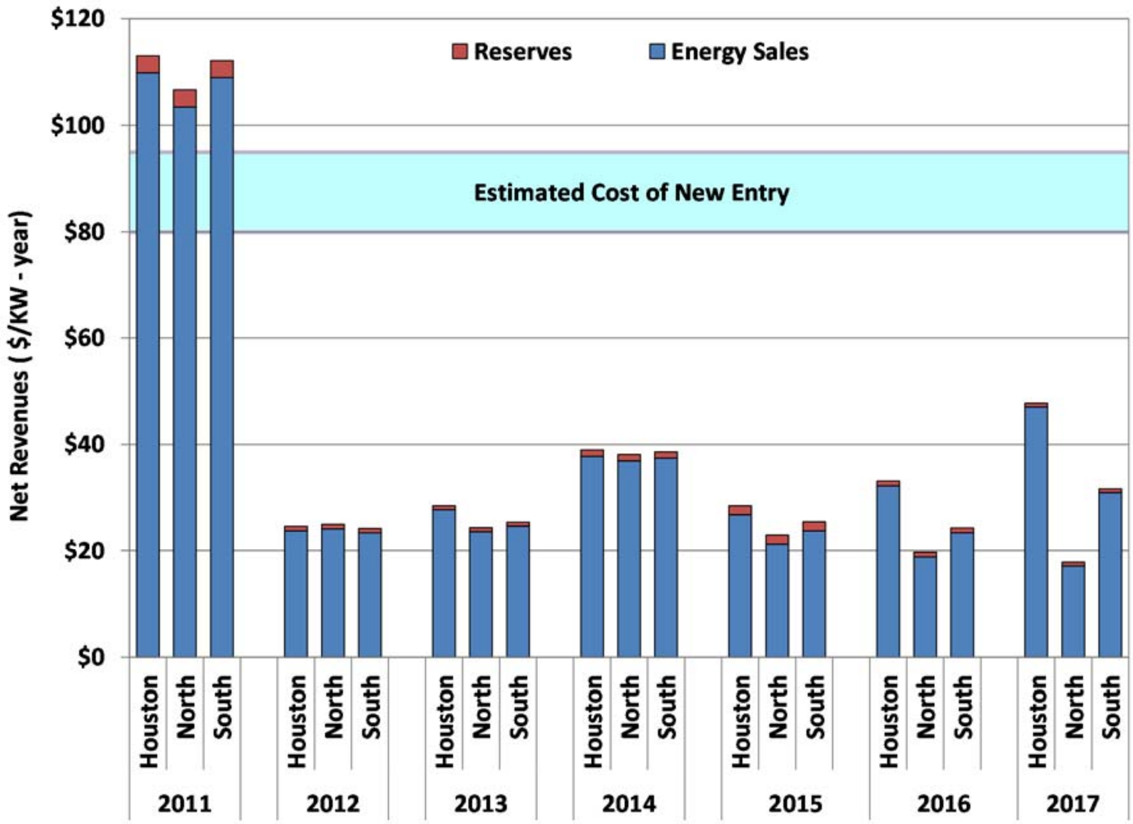
Resource Adequacy

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. The ERCOT market continued to provide net revenues well below the level needed to support new investment, ranging from less than $20 per kW-year in the North Zone to almost $48 per kW-year in Houston.

These results are consistent with continued surplus of capacity, which contributed to infrequent shortages over the past three years. In an energy-only market, shortages play a key role in delivering the net revenues needed to support new investment. Such shortages will tend to be clustered in years with little surplus capacity, unusually high load, or poor generator availability. Therefore, these results alone do not raise concerns regarding design or operation of ERCOT's ORDC mechanism for pricing shortages. Given the recent generation retirements and load growth, 2018 may well produce significantly more shortage pricing.

Given the low natural gas and resulting energy prices in 2017, 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. The generation-weighted average price for the four nuclear units in ERCOT (approximately 5 GW of capacity) was only $24.73 per MWh in 2017. This is similar to nuclear prices in 2016 and 2015, which were also lower than the ERCOT-wide prices in those years. Nuclear prices were $21.46 per MWh in 2016, down from $24.56 per MWh in 2015.

Assuming that operating costs of the nuclear units in ERCOT are similar to the U.S. average, it is likely that these units were not profitable in 2017, based on the fuel and operating and maintenance costs alone. Hence, it is unlikely that these nuclear units covered any capital costs that may have been incurred. However, unlike other regions with large amounts of nuclear generation, the four nuclear units in ERCOT are relatively new and owned by four entities with sizable load obligations. Although not profitable on a stand-alone basis, the nuclear units have substantial option value for the owners because they ensure that the cost of serving their load will not rise substantially if natural gas prices increase. Nonetheless, the economic pressure on these units raises resource adequacy issues that will need to be monitored.

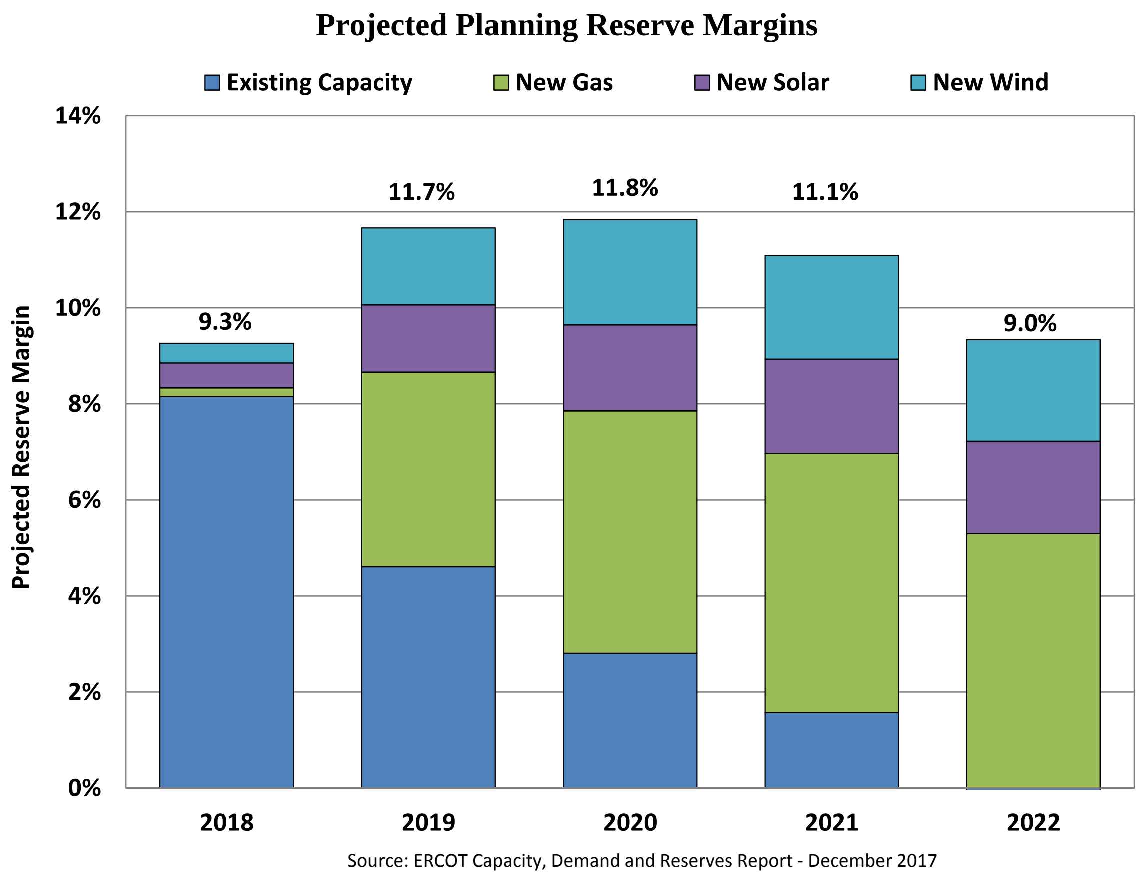
The generation weighted price of all coal and lignite units in ERCOT during 2017 was $26.32 per MWh, an increase from $23.98 per MWh in 2016. During 2015 and 2016 delivered coal costs in ERCOT were higher than natural gas prices at the Houston Ship Channel, resulting in reduced output for coal resources. With the increased natural gas prices in 2017, gas costs exceeded coal by nearly $0.40 per MMBtu. However, given coal units generally have higher heat rates and more expensive non-fuel operations and maintenance costs, economic pressure remains. During 2017, one coal unit was seasonally mothballed and Luminant declared its intention to retire seven other coal units in early 2018. The IMM reviewed each of these actions and found them to be supported by the unit-specific financials.

The figure below shows ERCOT's current projection of planning reserve margins and indicates that the region will have a 9.3% reserve margin heading into the summer of 2018. These projections are noticeably lower than those developed since May of last year.[7] The reduction was largely due to the approximately 5 GW of capacity taken offline by early 2018. The figure shows that ERCOT expects that the reserve margin will continue to be below the existing target level of 13.75% for the foreseeable future.[8]

---

7 See Report on the Capacity, Demand and Reserves (CDR) in the ERCOT Region (May 2, 2017); http://www.ercot.com/content/wcm/lists/114798/CapacityDemandandReserveReport-May2017.pdf

8 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 PUCT 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.



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.

However, because the surplus has now disappeared and shortages are likely to be more frequent in 2018, 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 closely. This response could cause planning reserve margins to exceed the forecast shown in the figure.

Analysis of Competitive Performance

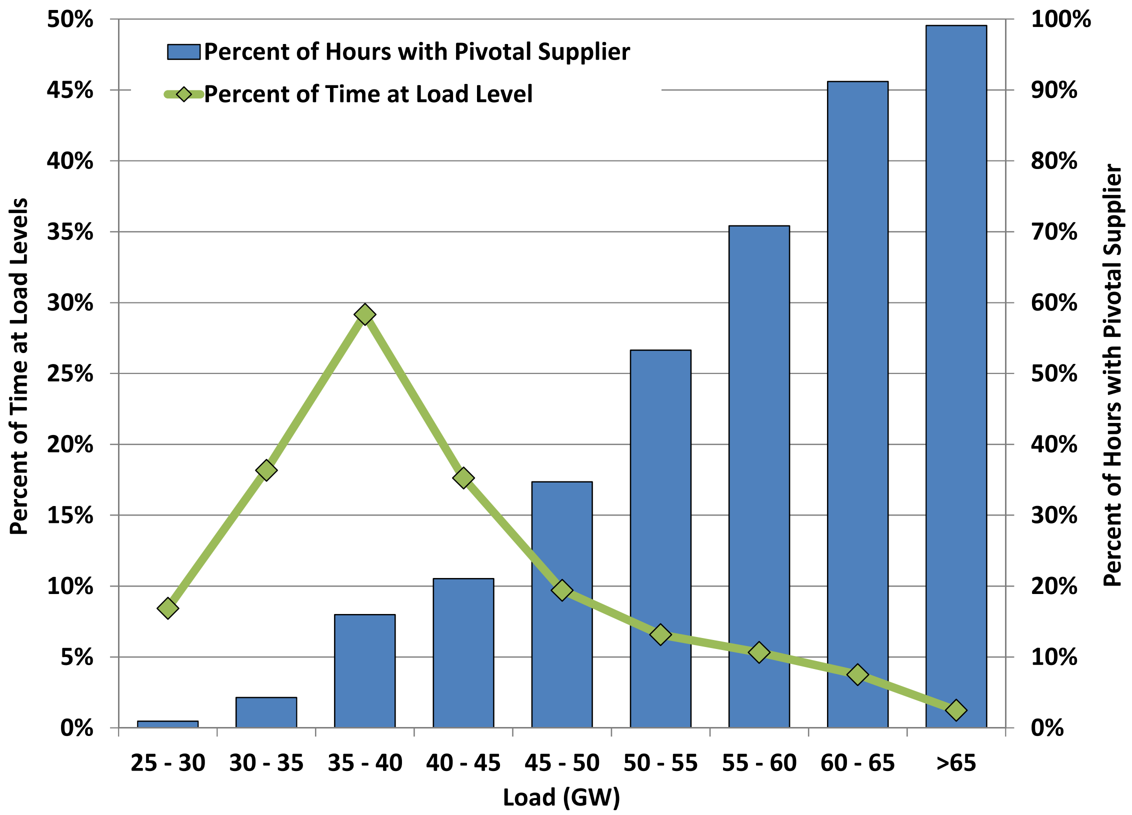
Market power is evaluated from two perspectives, structural (does market power exist) and behavioral (have attempts been made to exercise it).

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.

Pivotal Supplier Frequency by Load Level



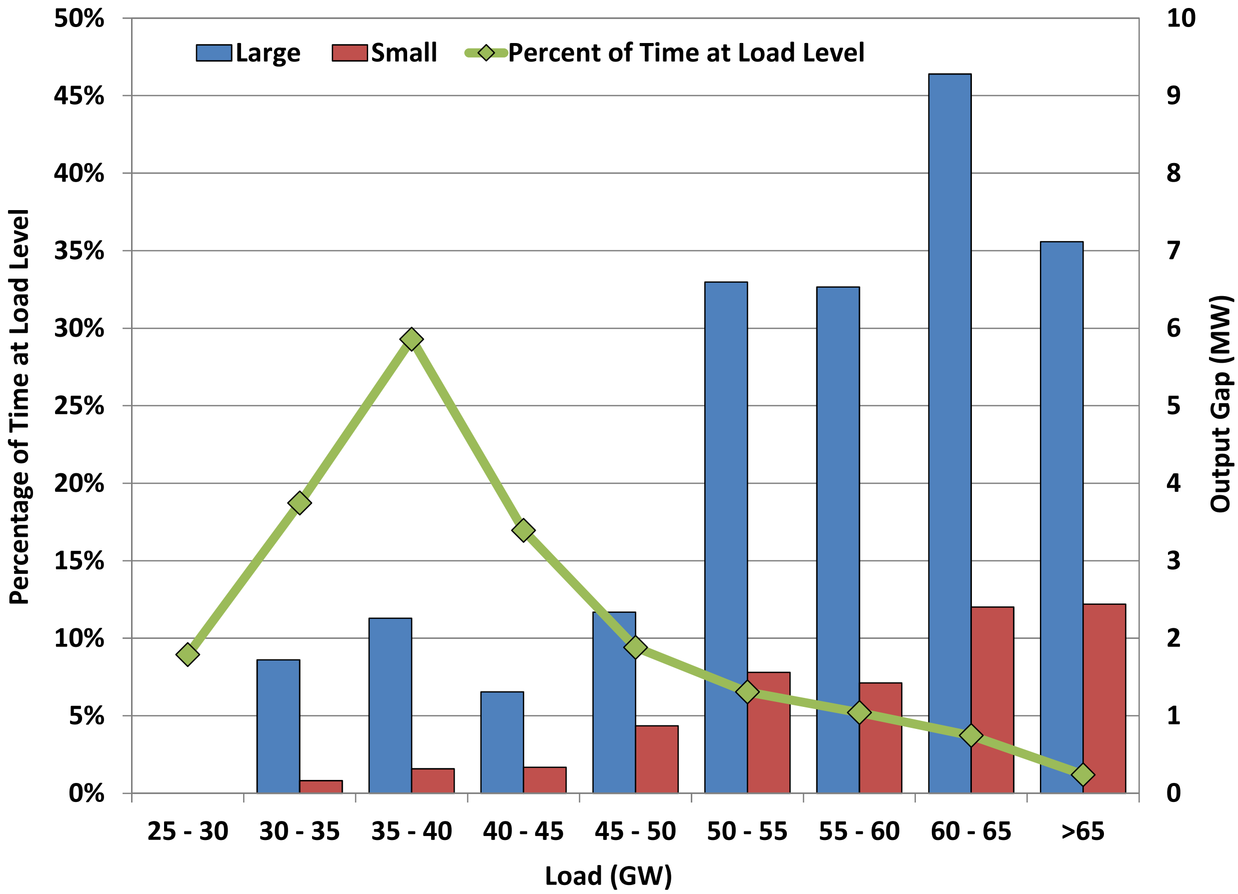
At loads greater than 65 GW there was a pivotal supplier 99% of the time. This is expected; at high load levels, the largest suppliers are more likely to be pivotal as other suppliers' resources are more fully utilized serving the load. There was a noticeable decrease in the percentage of time with a pivotal supplier at loads below 50GW in 2017. This led to a decrease in the pivotal supplier frequency to 24.5% of the time in 2017, down from 28.5% and 26% of all hours in 2016 and 2015, respectively. Even with the slight decrease, market power continues to be a potential concern in ERCOT and underscores the need for 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 margin of $30 per MWh is used for this analysis. To determine whether the output from a resource is economic to produce, the mitigated offer cap serves as a proxy for the marginal production cost of energy for each unit.

The figure above shows the output gap levels, separately showing the results aggregated for the five largest suppliers (those with greater than five percent of ERCOT installed capacity) and all other suppliers (i.e., the small category).[9]

These results show that potential economic withholding levels were extremely low for the largest suppliers and small suppliers alike in 2017. 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 2017.

9 In the second step of the dispatch, the after-mitigation offer curve is used to determine dispatch instructions and locational prices. The output gap at Step 2 showed very small quantities of capacity that would be considered part of this output gap.

Recommendations

Overall, we find that the ERCOT market performed well in 2017. However, we have identified and recommended a number of potential improvements to the ERCOT markets. We make seven recommendations in this report, all but one we have previously recommended. These recommendations are categorized by principle objective: a) to improve the operation of the ERCOT system and its resources; and b) to improve price formation in ERCOT's energy and ancillary services markets. We describe each recommendation below and the benefits that each would provide. For recommendations repeated from prior reports, we discuss the status of progress made to evaluate or implement the recommendation.

Improving Real-Time Operations and Resource Performance

One of the primary functions of the wholesale markets is to coordinate the operations of all resources to satisfy the system's needs at the lowest cost. The recommendations in this section are principally intended to improve the operation of the ERCOT markets, but in doing so will also improve ERCOT's prices and performance incentives. Many of the recommendations were considered over the past year, which we describe in the status section for each recommendation.

1. **Implement real-time co-optimization of energy and ancillary services.**

Substantial benefits can be achieved by implementing real-time co-optimization of energy and ancillary services. First, jointly optimizing all products in each interval allows ancillary service responsibilities to be continually adjusted in response to changing market conditions. The efficiencies of this continual adjustment would flow to all market participants and would be greater than what can be achieved by QSEs acting individually. The continual, optimal system-wide allocation of resources between providing energy and providing reserves will lower the cost of satisfying both requirements. Additionally, it will ensure that energy is produced in locations where it may be most valuable.

The second benefit of real-time co-optimization will be improved 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. In a co-optimized system, the real-time market will determine every five minutes whether a shortage of either energy or any class of reserves exists and set prices accordingly. By reallocating reserves and energy in an optimal manner, the system often has access to more reserves. Thus, a system without co-optimization may perceive and price shortages that could be eliminated by allocating resources optimally.

Additionally, under a co-optimized system, a demand curve would be established for each type of reserve (potentially including locational reserve products in the future). Currently, capacity providing responsive or regulating reserves is not available to be converted into energy at any price. With co-optimization, when it is economic to release reserves to provide energy, the value of these reserve shortages will be reflected efficiently in the energy and reserve prices. This is especially important in ERCOT because pricing during shortage conditions is key for the success of ERCOT's energy-only market.

Other economic benefits would 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). For all of these reasons, implementing real-time co-optimization of energy and ancillary services is our highest priority recommendation.

Status: In September 2013, the PUCT initiated a project to consider the feasibility of implementing real-time co-optimization.[10] After initial investigation including a draft whitepaper by ERCOT, the project was temporarily put on hold to consider whether a Multi-Interval Real-Time Market (MIRTM) should be pursued first or in conjunction with real-time co-optimization. In early 2017, the PUCT provided direction to ERCOT to restart the evaluation of implementing real-time co-optimization.[11] The PUCT created a project to "assess price-formation rules in ERCOT's energy-only market and led multiple workshops on scarcity pricing and other price-formation issues in ERCOT's energy-only market in 2017.[12] The IMM filed comments detailing the benefits of real-time co-optimization in Project No. 47199.[13]

At the open meeting on December 14, 2017, the PUCT approved ERCOT's proposed plan, created in conjunction with Commission Staff and the IMM, to assess the benefits of the potential implementation of real-time co-optimization and marginal losses in the ERCOT wholesale electricity market in Project No. 47199. The IMM has developed software to estimate the benefits of co-optimization by simulating it in historic periods and will conduct this simulation for 2017 using publicly available data. The IMM expects to submit our results to the PUCT in June 2018. In coordination with ERCOT, the IMM intends to make the software, input data, and results available to all market participants to facilitate transparency and understanding of the analytic approach and results.

10 *See* PUCT Project No. 41837, *PUCT Review of Real-Time Co-Optimization in the ERCOT Region.*

11 *Id.*, ERCOT Letter to Chairman and Commissioners (Apr. 27, 2017), responding to Commissioner direction at the April 13, 2017 Open Meeting directing ERCOT "to restart the evaluation of the potential implementation of the co-optimization of energy and operating reserves in the real-time market."

12 *See* PUCT Project No. 47199, *Project to Assess Price-Formation Rules in ERCOT's Energy-Only Market.*

13 Comments of Potomac Economics at 2, 10 (Sept. 15, 2017); IMM Reply Comments at 2-5 (Dec. 22, 2017).

2. 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.

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 in 2016 and 2017.[14]

Load curtailment to avoid transmission charges may be resulting in price distortion 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 curtailments 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. This trend continued in 2017, with significant load curtailments on peak load days in June, August and September when real-time prices were less than $100 per MWh.

14 *See* ERCOT, 2017 Annual Report of Demand Response in the ERCOT Region (Mar. 2018) at 7, available at http://www.ercot.com/services/programs/load.

Status: The PUCT made no changes to the ERS program or transmission service rates in 2017.[15] The IMM discussed the importance of the 4CP allocation mechanism as part of its recommendations in Project No. 47199.[16]

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

The real-time market relies primarily on two classes of resources: online resources and offline resources that can start quickly. 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 decommitment of peaking 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.

Status: 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 this time.[17] The finding of insufficient benefits is not surprising given the current 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 and ever increasing quantities of intermittent renewable resources, the benefits of improving the short-term commitment process will grow. In addition, it is likely much less costly to develop a process to

15 The PUCT considered changes to transmission service rates in Project No. 46393, but changes to the 4CP allocation method were not part of that project. *See* PUCT Project No. 46393, *Rulemaking Proceeding to Repeal and Replace 16 Texas Administrative Code § 25.192, Relation to Transmission Service Rates*. The PUCT ultimately opted not to pursue changes to 16 TAC § 25.192 at the February 15, 2018 Open Meeting.

16 *See* Comments of Potomac Economics at 8 (Sept. 15, 2017.

17 *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).

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 as part of its recommendations in Project No. 47199.[18]

4. Price congestion at all generator 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 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.

Status: This is a new recommendation.

Improving Price Formation in the ERCOT Market

5. Consider including marginal losses in ERCOT locational marginal prices.
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.

[18] See Comments of Potomac Economics at 3 (Sept. 15, 2017).

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.

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: The IMM filed comments detailing the benefits of marginal losses in the price-formation Project No. 47199.[19] At the open meeting on December 14, 2017, the PUCT approved ERCOT's proposed plan, created in conjunction with Commission Staff and the IMM, to assess the benefits of the potential implementation of marginal losses in the ERCOT wholesale electricity market in Project No. 47199. ERCOT will model a future year case with average transmission losses and separately with marginal transmission losses. ERCOT is expected to provide the resulting benefits assessment of marginal losses in June 2018.

6. 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: In the absence of a comprehensive redesign of ancillary services, multiple incremental modifications have been and are being considered. Two proposed changes pertinent to this recommendation are NPRR848 and NPRR815. NPRR848, as submitted, would modify the clearing process for responsive reserve service in accord with this recommendation. It remains tabled in the stakeholder process. NPRR815, which was approved in December 2017 and scheduled for implementation in mid-2018, would: 1) increase the allowable percentage of responsive reserve service that load resources may provide from 50% to 60%, and 2) specify the

19 Comments of Potomac Economics at 2 (Sept. 15, 2017); IMM Reply Comments at 5-7 (Dec. 22, 2017).

minimum amount of primary frequency response (generator provided) as 1150 MW. These changes are in a helpful direction and we will monitor their effects.

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 recommendations in Project No. 41799, 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. This proposal would require real-time co-optimization as part of its implementation so it could not be introduced in the near term.[20] We are prepared to work with ERCOT and market participants to evaluate this proposal or others to address this recommendation.

20 See Comments of Potomac Economics at 2, 8-10 (Sept. 15, 2017).

I. REVIEW OF REAL-TIME MARKET OUTCOMES

Although only a small share of the power produced in ERCOT is transacted in the spot market, real-time energy prices are very important because they 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 spot and forward markets, the prices in the forward market should be directly related to the prices in the spot market (i.e., the spot 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. This section evaluates and summarizes electricity prices in the real-time market during 2017.

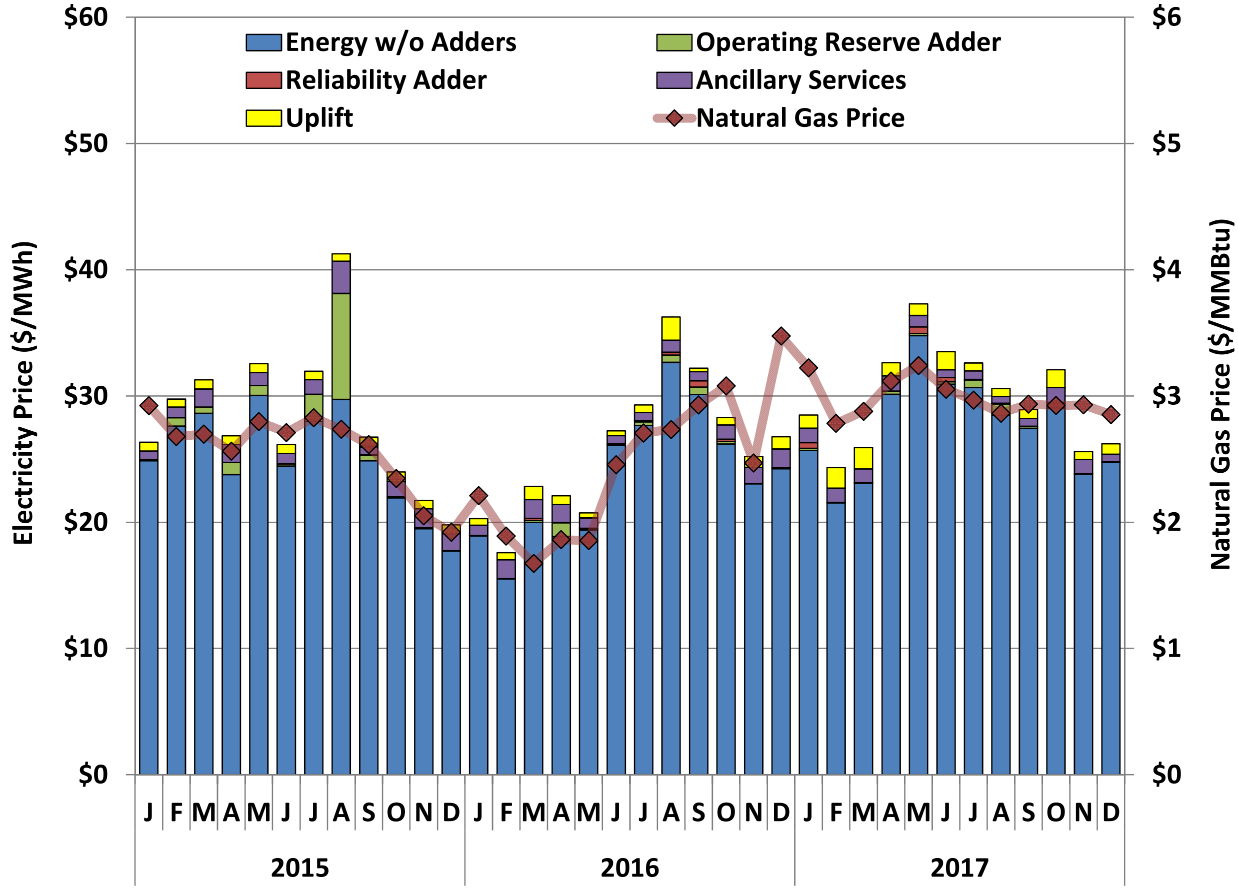
A. Real-Time Market Prices

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 2015 through 2017. 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 to show them on a per MWh basis.[21] ERCOT developed two energy price adders that are designed to improve its real-time energy pricing when conditions warrant or when ERCOT takes out-of-market actions for reliability. To distinguish the effects of the energy price adders, the Operating Reserve Demand Curve Adder (operating reserve adder) and the Reliability Deployment Price Adder (reliability adder) are shown separate 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 a separate price run of SCED, 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.

21 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.

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 2017 was $2.98 per MMBtu, up approximately 22% from the 2016 average price of $2.45 per MMBtu. ERCOT average real-time energy prices increased nearly 15%, increasing from $24.62 per MWh in 2016 to $28.25 per MWh in 2017.

The average real-time energy price in 2017 included small contributions from ERCOT's energy price adders: $0.24 per MWh from the operating reserve adder and $0.16 per MWh from the reliability adder. These values were similar to the levels in 2016; $0.27 and $0.13 per MWh, for reserve and reliability adder, respectively. The highest monthly average operating reserve adder for 2017 occurred in August, while the highest monthly average reliability adder occurred in May.

Other cost categories continue to be a small portion of the all-in electricity price. Ancillary services costs were $0.87 per MWh in 2017, down from $1.03 per MWh in 2016 because of continued low natural gas prices and lower ancillary service requirements.

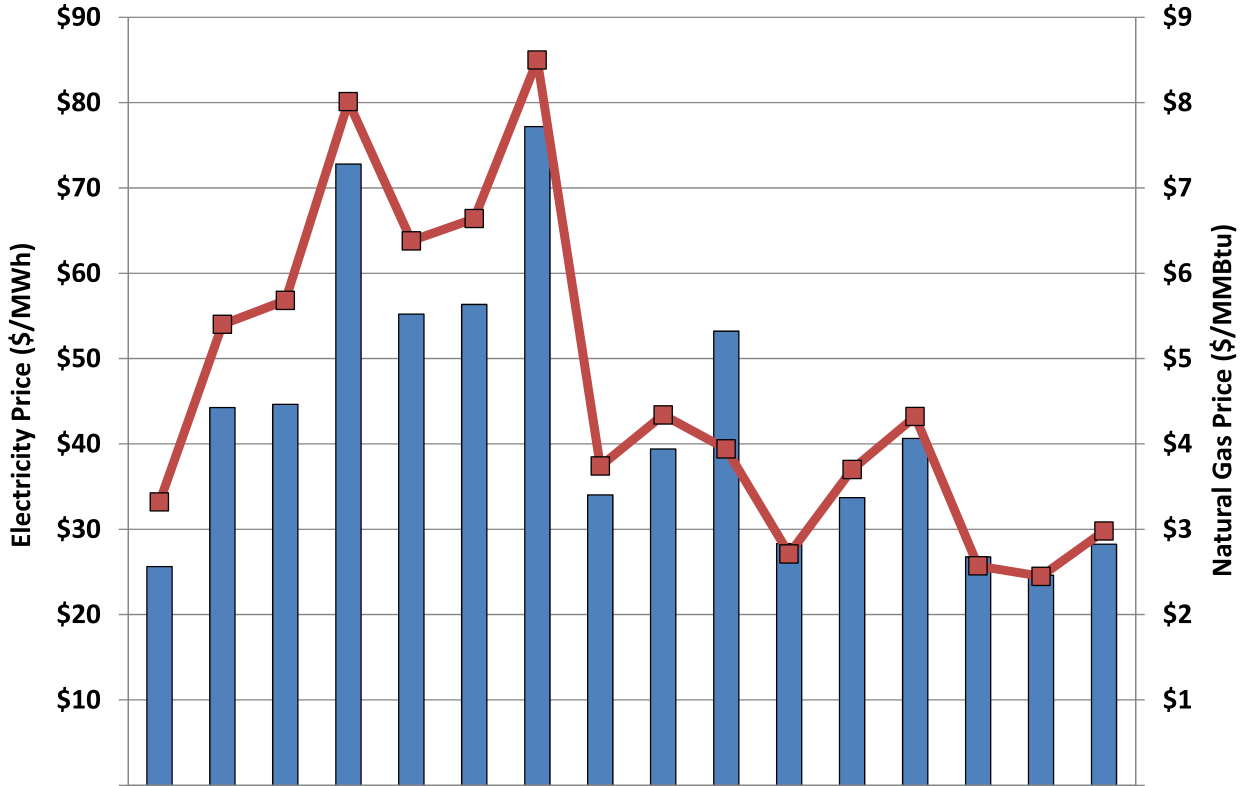
Uplift costs accounted for $1.03 per MWh of the all-in electricity price in 2017, up from $0.84 per MWh in 2016. In the context of providing the total cost of serving load in ERCOT, these values include both the ERCOT system administrative fee and the program costs for Emergency Response Service (ERS), which are assessed to all loads. The total amount of uplifted costs in 2017 was approximately $365 million. There are many costs included as uplift, but the largest components are the ERCOT system administrative fee ($199 million or $0.56 per MWh), ERS program costs ($50 million or $0.14 per MWh) and the revenue neutrality allocation (RENA), which totaled $96 million or $0.27 per MWh in 2017.

Virtually all of the increase in uplift costs in 2017 was due to the increase in RENA. Specifically, RENA was $28 million ($0.08 per MWh) in 2016 and increased to $96 million ($0.27 per MWh) in 2017. Several factors can contribute to RENA uplift, including 1) setting a price floor in the real-time market at -$251; 2) settlement of day-ahead PTP obligations linked to options; 3) manual corrections that occur when the clearing price of PTP obligations in the day-ahead market is higher than the submitted bid price; 4) inconsistency between day-ahead and real-time market during market clearing; and 5) not including private network load when calculating Load Zone prices.

More detailed studies show that almost all the RENA uplift occurred in market hours when there was transmission congestion. The two factors contributing most to RENA uplift in 2017 were the settlement of day-ahead PTP obligations linked to options and not including private network load when calculating Load Zones prices. The amount of RENA uplift associated with not including private network load in Load Zone prices is estimated to have exceeded $40 million in 2017. These impacts were addressed in late 2017 with the implementation of NPRR831.

Figure 2 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 2017.

Figure 2: 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
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
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

Like Figure 1, Figure 2 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. A noticeable exception occurred in 2011, when energy prices were affected by scarcity conditions.

Energy prices vary across the ERCOT market because of congestion costs that are incurred as power is delivered over the network. Figure 3 shows the monthly load-weighted average prices in the four geographic ERCOT zones during 2017 and 2016. 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 3: Average Real-Time Energy Market Prices by Zone

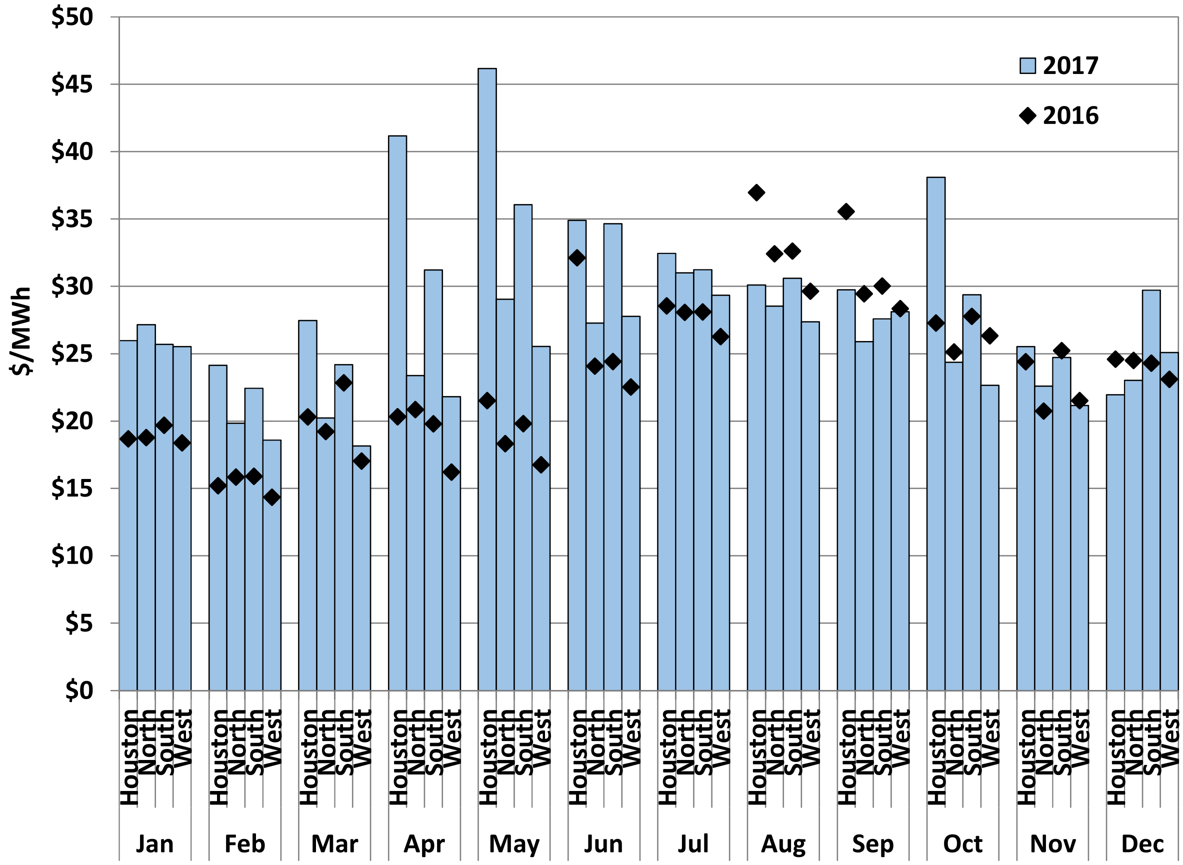


Table 1 provides the annual load-weighted average price for each zone for the past seven years, and includes the annual average natural gas price for reference.

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

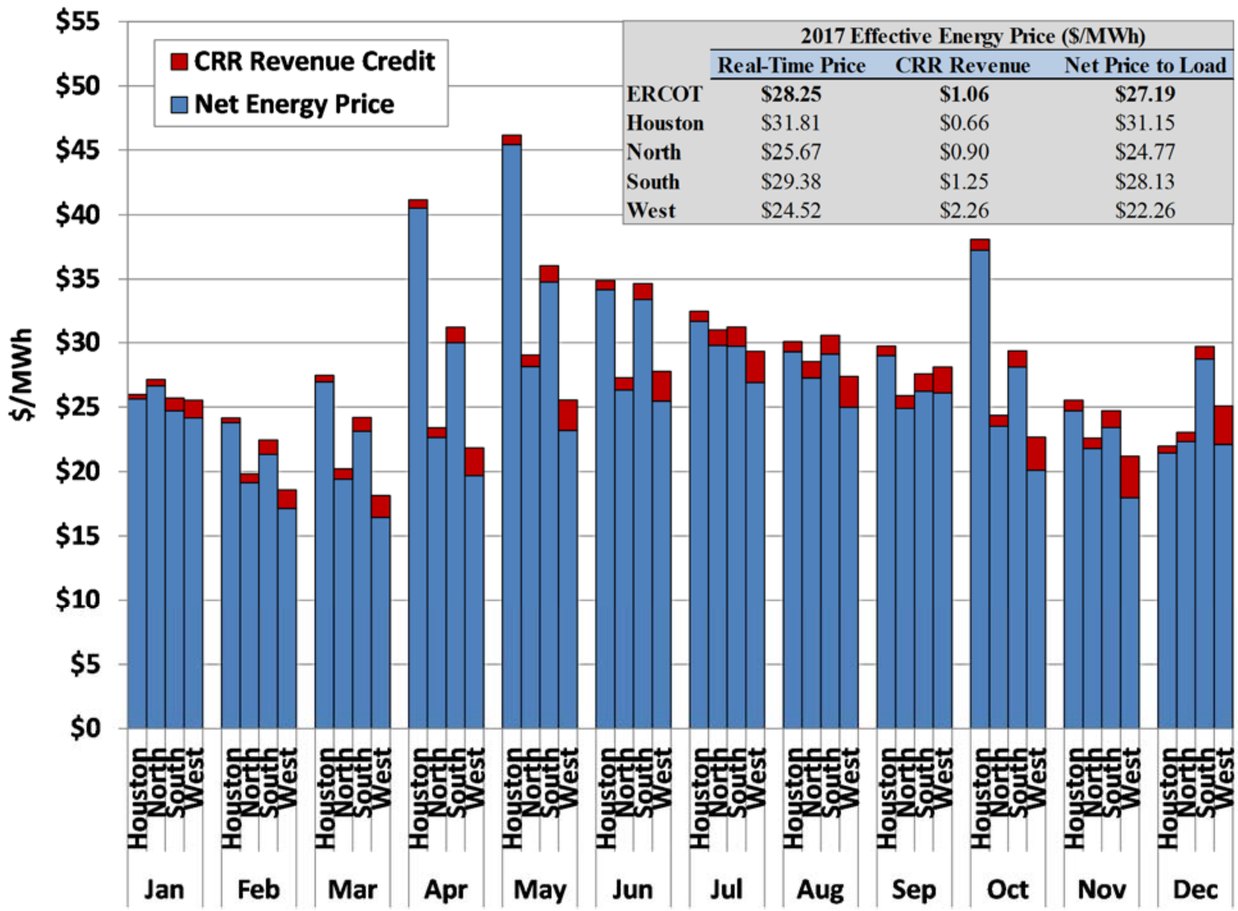
($/MWh)	2011	2012	2013	2014	2015	2016	2017
ERCOT	**$53.23**	**$28.33**	**$33.71**	**$40.64**	**$26.77**	**$24.62**	**$28.25**
Houston	$52.40	$27.04	$33.63	$39.60	$26.91	$26.33	$31.81
North	$54.24	$27.57	$32.74	$40.05	$26.36	$23.84	$25.67
South	$54.32	$27.86	$33.88	$41.52	$27.18	$24.78	$29.38
West	$46.87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

The difference in zonal prices in 2017 are directionally comparable to the prices in 2016. Constraints on the ability to import generation led to the Houston zone being the highest priced zone in 2017. Export limitations resulted in the West zone having the lowest price. However, price spreads were larger in 2017 because of higher natural gas prices and the increased impacts of transmission congestion.

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 with continued investment in transmission facilities, the amount of wind generation additions have meant export limitations and resulting lower prices since 2015.

Another factor influencing zonal price differences is Congestion Revenue Right (CRR) auction revenue distributions. They 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 4 below shows the effect that this reduction has on a monthly basis, by zone.

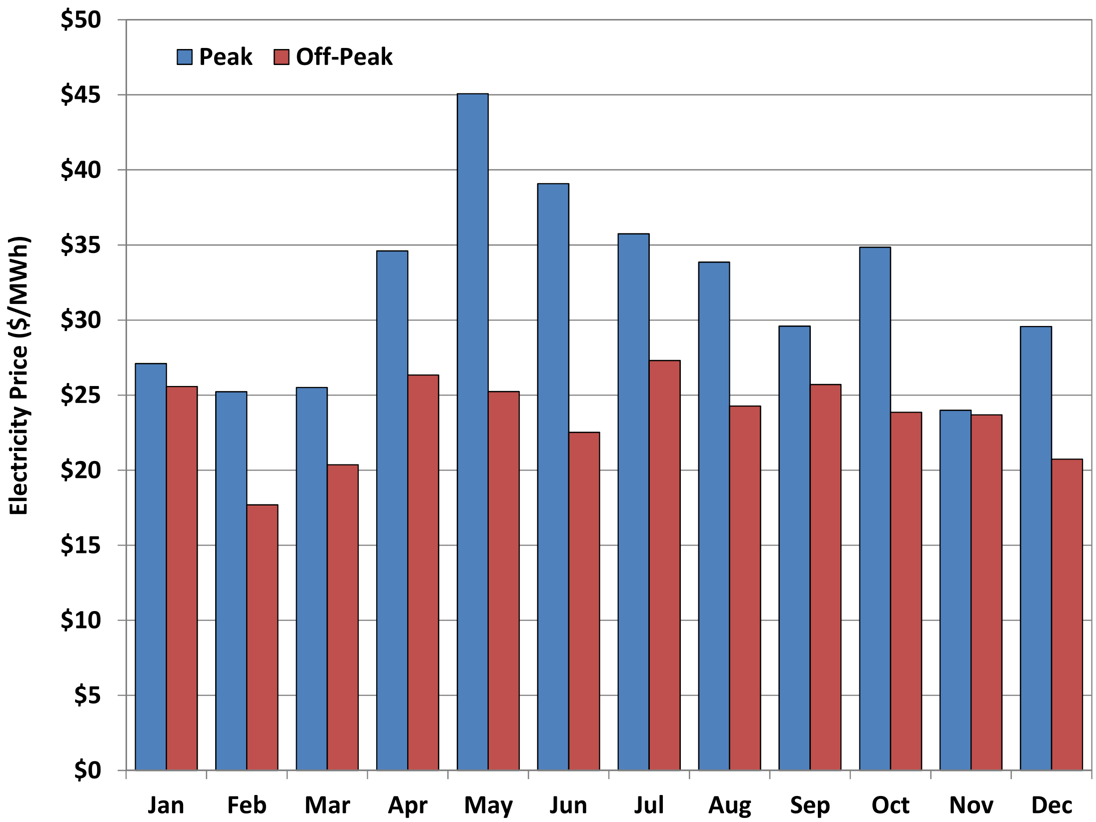
Figure 4: Effective Real-Time Energy Market Prices



With the CRR auction revenue offset included, the ERCOT-wide load-weighted average price rose by $3.48 per MWh to $27.19 per MWh in 2017 compared to $23.71 per MWh in 2016. Focusing on zonal differences, a smaller credit in Houston relative to the ERCOT-wide CRR auction revenue credit and a larger credit in the West again resulted in the net price difference between the two zones being even higher in 2017.

Real-time energy prices not only vary by location, they vary by time of day. Figure 5: Peak and Off-Peak Pricing shows the load-weighted average real-time prices in ERCOT for the categories of Peak and Off-Peak for each month in 2017. 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 5: Peak and Off-Peak Pricing

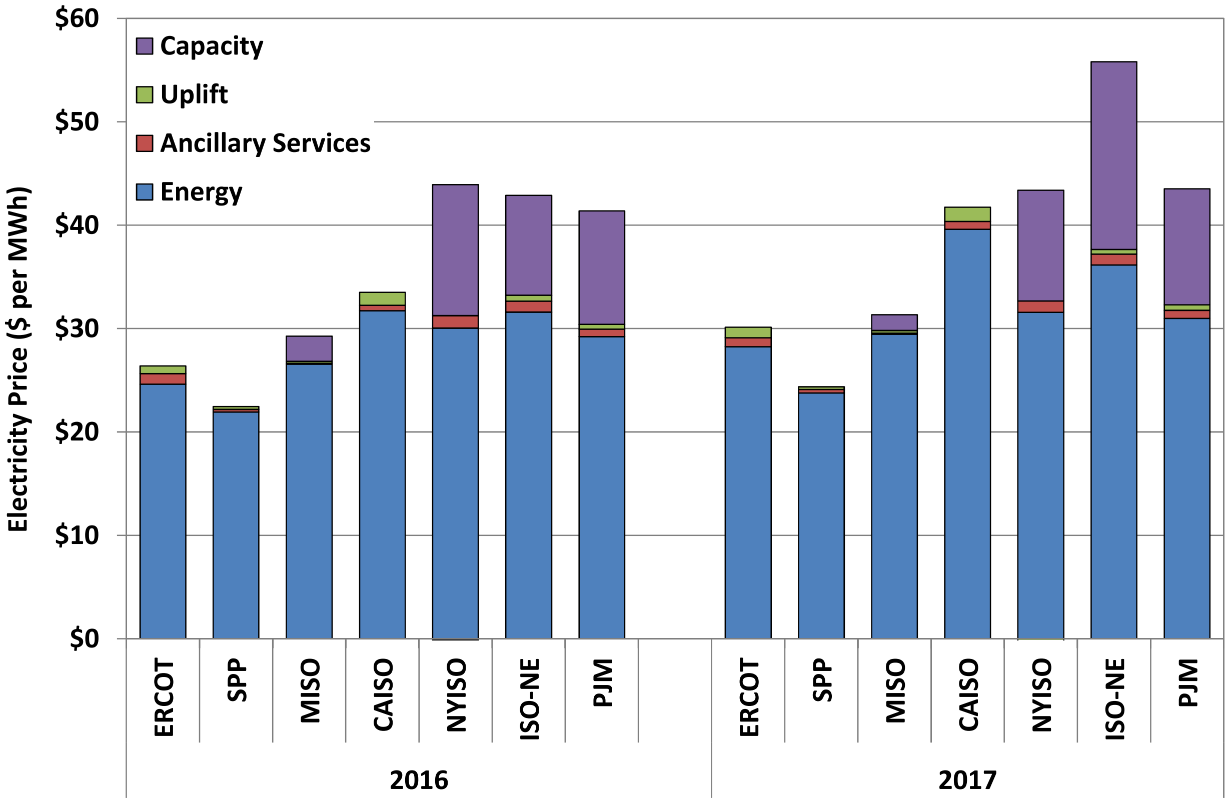


As would be expected, Peak hours were higher priced than Off-Peak hours for every month in 2017. The monthly difference ranged from a minimum of $0.31 per MWh in November to a maximum of $19.84 per MWh in May. The average difference between monthly Peak and Off-Peak pricing was $8.41 per MWh.

To provide additional perspective on the outcomes in the ERCOT market, Figure 6 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, New York ISO, ISO New England, and the Pennsylvania-New Jersey-Maryland (PJM) Interconnection.

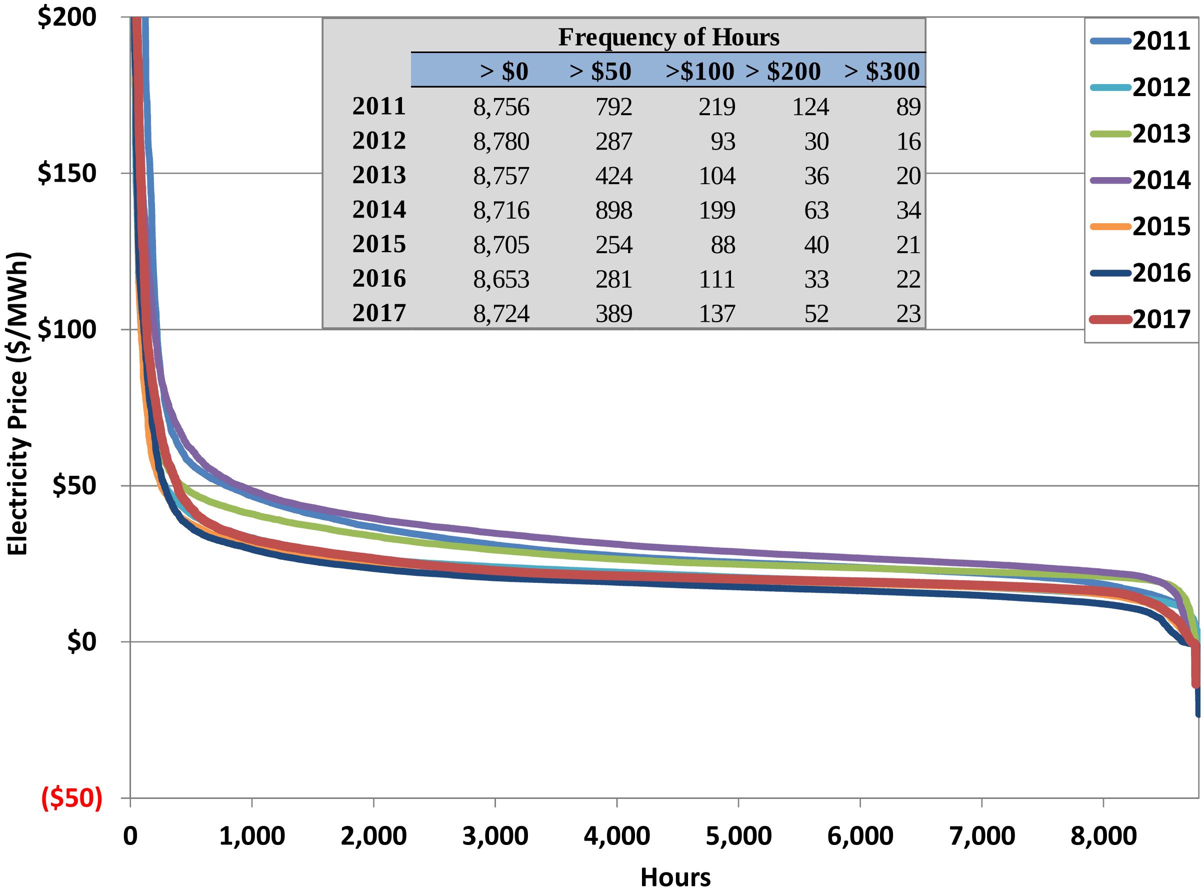
Figure 6: Comparison of All-in Prices Across Markets



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. Figure 6 shows that, with the exception of a small decrease in NYISO, all-in prices were higher across U.S. markets in 2017. Modest increases in natural gas prices across the United States led to small increases to the energy component of electricity prices. The exceptions were CaISO and ISO-NE, which had much larger increases to the energy component. ISO-NE also had a sizable increase to the capacity component.

Figure 7 below shows price duration curves for the ERCOT energy market in each year from 2011 to 2017. 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). The prices in this figure are the hourly ERCOT average prices derived by load weighting the zonal settlement point prices.

Figure 7: ERCOT Price Duration Curve

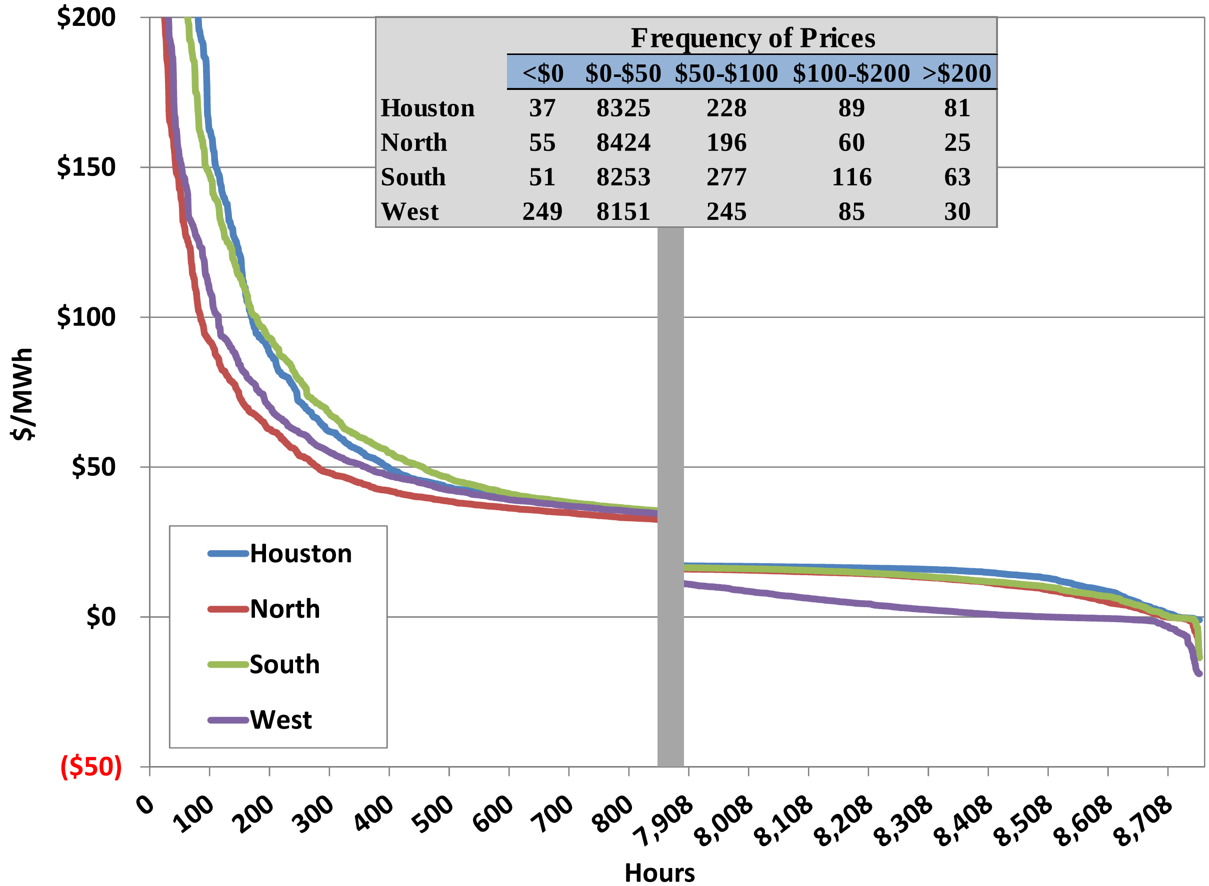


Frequency of Hours	> $0	> $50	>$100	> $200	> $300
2011	8,756	792	219	124	89
2012	8,780	287	93	30	16
2013	8,757	424	104	36	20
2014	8,716	898	199	63	34
2015	8,705	254	88	40	21
2016	8,653	281	111	33	22
2017	8,724	389	137	52	23

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. In 2017, there were 36 hours with ERCOT-wide prices at or below zero.

To more closely examine the variation in zonal real-time energy prices, Figure 8 shows the top and bottom 10% of the hourly average price duration curve in 2017 for the four zones.

Figure 8: Zonal Price Duration Curves



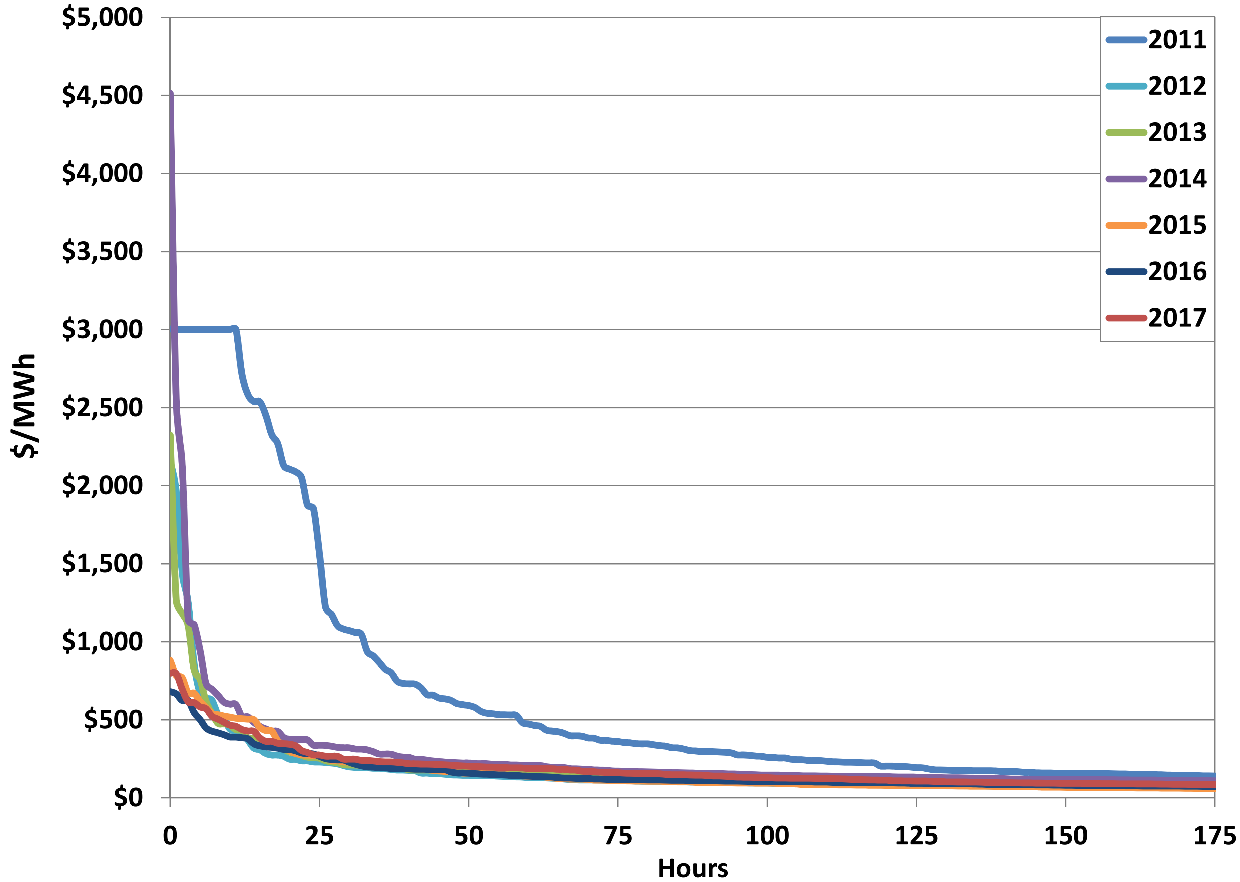
Frequency of Prices	<$0	$0-$50	$50-$100	$100-$200	>$200
Houston	37	8325	228	89	81
North	55	8424	196	60	25
South	51	8253	277	116	63
West	249	8151	245	85	30

Compared to the other zones, low prices in the West Zone were much lower in 2017. This can be explained by the increased occurrences of transmission constraints limiting exports of low-priced wind generation.

The higher frequency of prices greater than $50 per MWh in the Houston and South zones is explained by North to Houston congestion, which continued to have high impacts in 2017. More details about the transmission constraints influencing zonal energy prices are provided in Section III: Transmission Congestion and Congestion Revenue Rights.

To see where the prices during 2017 diverged from prior years, Figure 9 compares prices for the highest-priced 2% of hours in each year. Energy prices for the top 100 hours of 2011 were significantly higher, while all subsequent years have followed an almost identical pattern. 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.

Figure 9: 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 Table 2. 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.

Table 2: Number and Impacts of Price Spikes on Average Real-Time Energy Prices

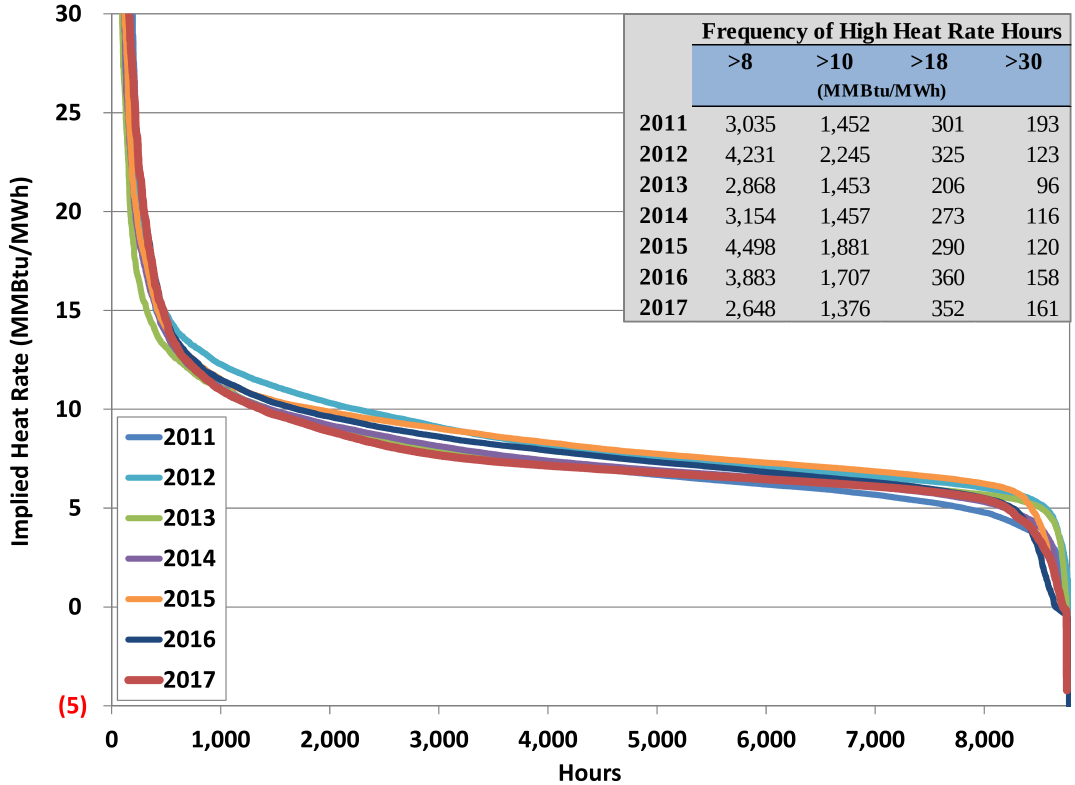
	Average Number of Spikes Per Month	Magnitude (per MWh)	Price Impact
2011	83	$14.09	48%
2012	94	$3.63	16%
2013	54	$3.43	12%
2014	74	$5.28	16%
2015	89	$3.35	16%
2016	99	$3.53	19%
2017	87	$4.33	20%

The overall impact of price spikes in 2017 was $4.33 per MWh. This result is generally consistent with the pricing impact of price spikes in past years. Of this price spike impact, $0.19 per MWh was due to the effects of the operating reserve adder and $0.13 per MWh was due to the effects of the reliability adder.

B. Real-Time Prices Adjusted for Fuel Price Changes

Although real-time electricity prices are driven to a large extent 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 10 and Figure 11 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 would be implied if natural gas was always on the margin.

Figure 10: Implied Heat Rate Duration Curve – All Hours

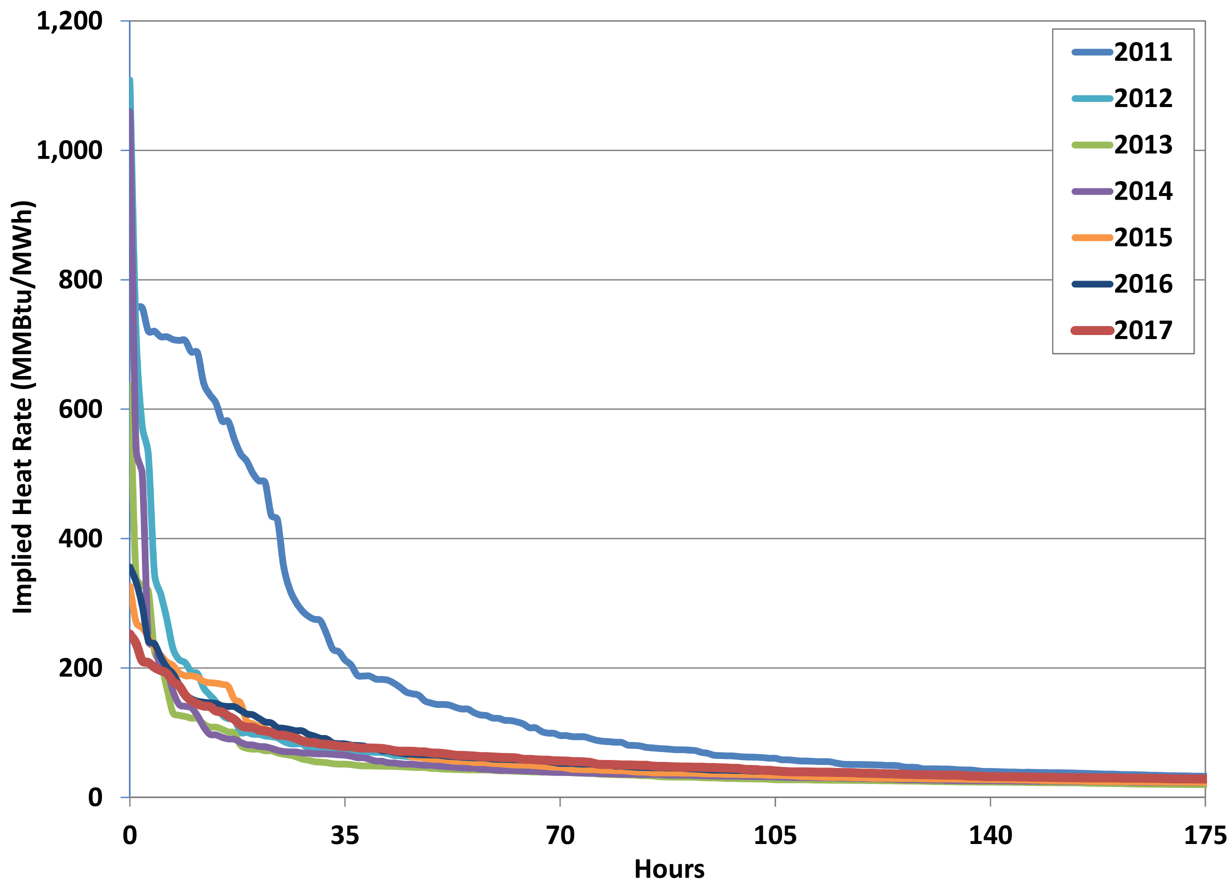


	>8	>10	>18	>30
	(MMBtu/MWh)			
2011	3,035	1,452	301	193
2012	4,231	2,245	325	123
2013	2,868	1,453	206	96
2014	3,154	1,457	273	116
2015	4,498	1,881	290	120
2016	3,883	1,707	360	158
2017	2,648	1,376	352	161

Implied system-wide heat rates for most hours of the year have been dropping since 2015, as evidenced by the decrease in the number of hours with an implied heat rate of greater than 8 MMBtu/MWh. This decrease can be explained by improvements in the efficiency of the ERCOT generation fleet, including the growing influence of wind generation.

Figure 11 shows the implied marginal heat rates for the top 2% of hours for years 2011 through 2017. The implied heat rate duration curve for the top 2% of hours in 2017 closely resembles that for 2016. Among all years presented, 2011 remains an outlier.

Figure 11: Implied Heat Rate Duration Curve – Top 2% of Hours



To further illustrate these differences, Figure 12 shows the implied marginal heat rates on a monthly basis in each of the ERCOT zones for 2016 and 2017. This figure is the fuel price-adjusted version of Figure 3 in the prior subsection, Real-Time Market Prices. Implied heat rates in 2017 were lower in all zones in 2017 as compared to 2016, with the largest drop in the West zone.

Figure 12: Monthly Average Implied Heat Rates

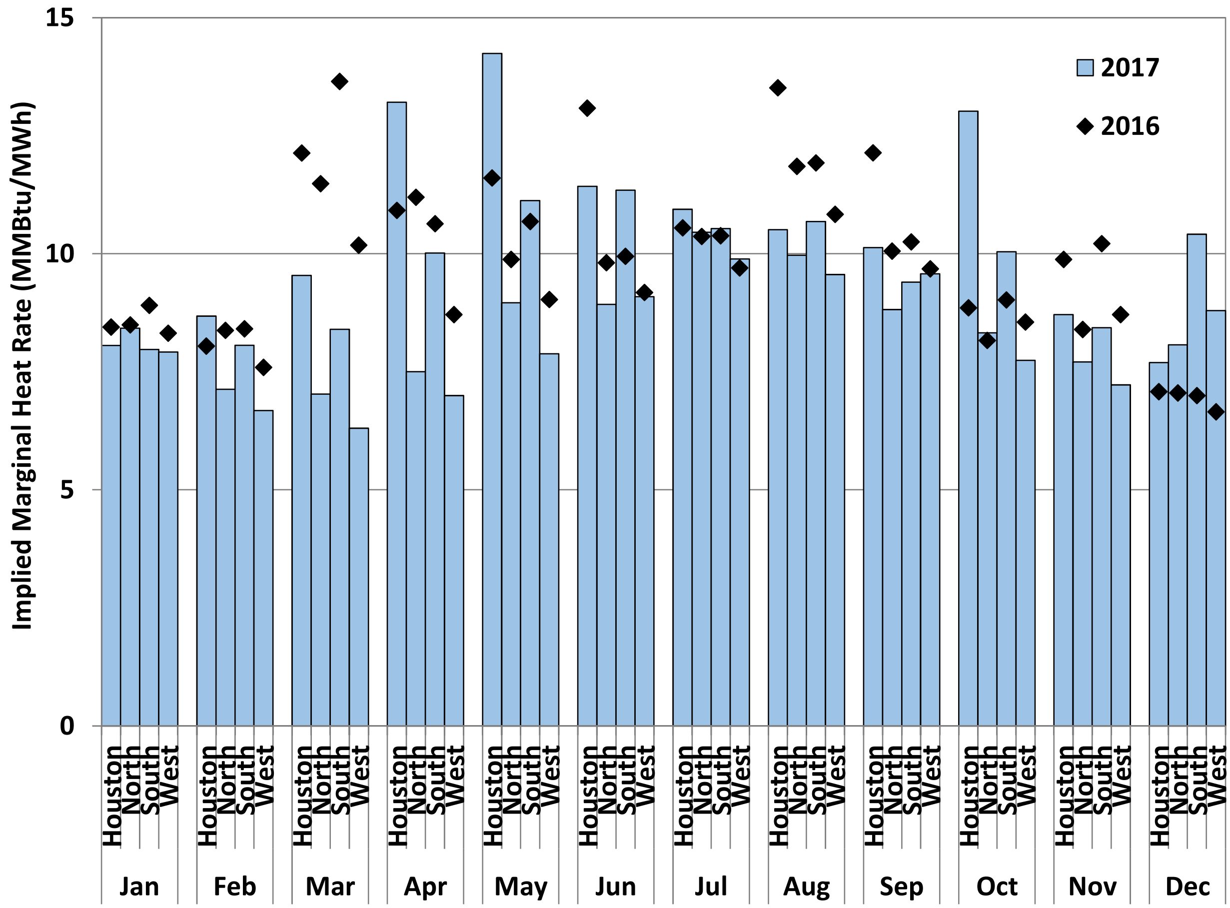


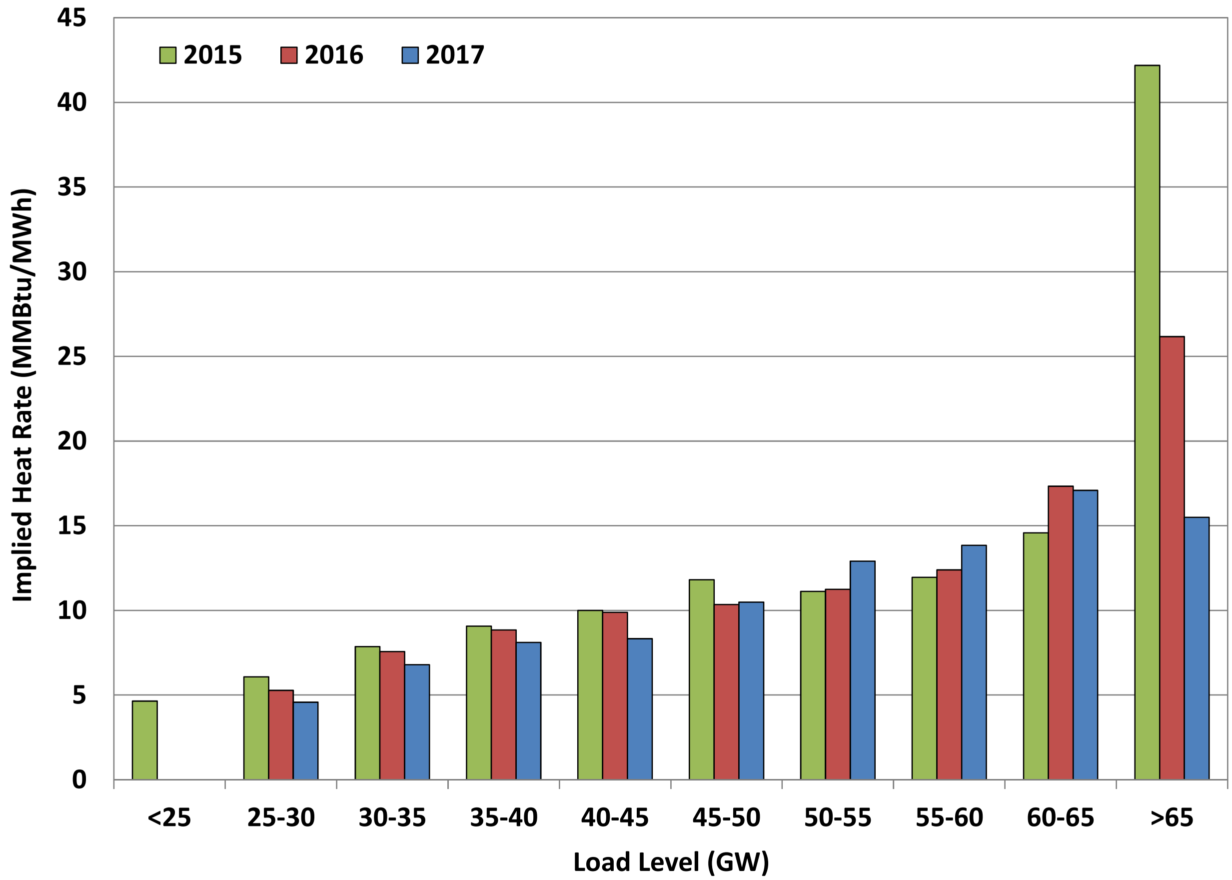
Table 3 displays the annual average implied heat rates by zone for 2011 through 2017. Adjusting for natural gas price influence, Table 3 shows that the annual, system-wide average implied heat rate decreased in 2017 compared to 2016. Zonal variations in the implied heat rate were greater in 2017 because of the increased influence of transmission congestion.

Table 3: Average Implied Heat Rates by Zone

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

The examination of implied heat rates from the real-time energy market concludes by evaluating them at various load levels. Figure 13 below provides the average implied heat rate at various system load levels for years 2015 through 2017.

Figure 13: Implied Heat Rate and Load Relationship

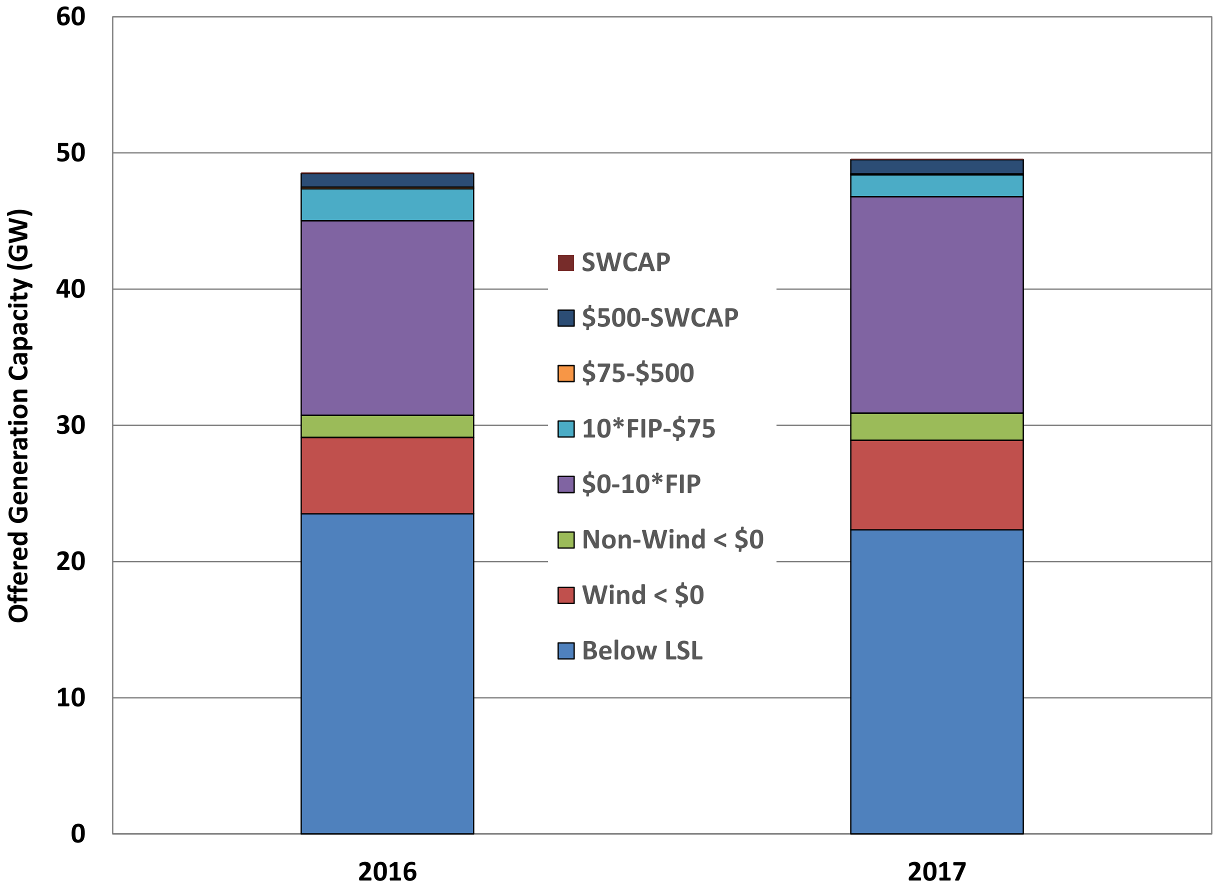


In a well-performing market, a clear positive relationship between these two variables is expected because resources with higher marginal costs are dispatched to serve higher loads. This relationship continued to exist in 2017.

C. Aggregated Offer Curves

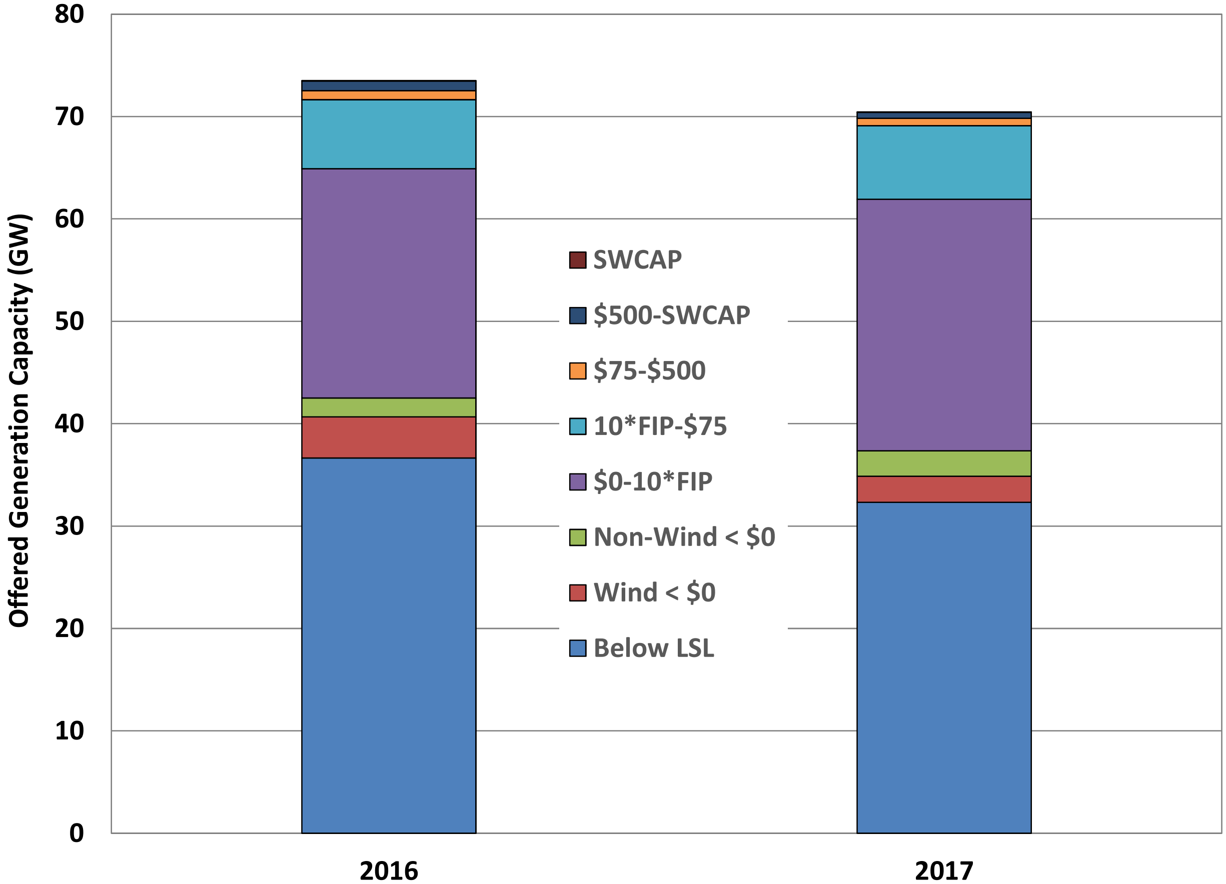
The next analysis compares the quantity and price of generation offered in 2017 to that offered in 2016. By averaging the amount of capacity offered at selected price levels, an aggregated offer stack can be assembled. Figure 14 provides the aggregated generator offer stacks for the entire year. Compared to 2016, more capacity was offered at lower prices in 2017. Specifically, continuing a trend from 2013, there was approximately 1,350 MW of additional capacity offered at prices less than zero. This increase was split between wind (70%) and non-wind (30%) generation. There was an off-setting decrease (1,200 MW) in capacity from below generators' low operating limits. At prices between zero and ten multiplied by the daily natural gas price (represented as the Fuel Index Price, or FIP), there was an increase of approximately 1,600 MW of additional capacity offered in 2017. The amount of capacity offered at prices between ten multiplied by FIP and $75 per MWh decreased by 750 MW from 2016 to 2017. With no change to the quantities of generation offered at prices above $75 per MWh, the resulting average aggregated generation offer stack was roughly 1,000 MW greater in 2017 than in 2016.

Figure 14: Aggregated Generation Offer Stack – Annual



The next analysis provides a similar comparison focused on the summer months. As shown below in Figure 15, the changes in the aggregated offer stacks between the summer of 2016 and 2017 are somewhat different than those just described. The average offer stack for the summer of 2017 was approximately 3,000 MW smaller than in the previous summer, with the biggest reduction coming from 4,300 MW less capacity from below generators' low sustained limits (LSLs). There was a further reduction of approximately 1,500 MW of capacity offered from wind units, offset by an additional 2,100 MW of capacity offered at prices between zero and ten multiplied by the daily natural gas price.

Figure 15: Aggregated Generation Offer Stack – Peak Hour



Both the annual and peak hour offer stacks display reductions in the amount of capacity below units' low dispatchable limits in 2017. 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 2017, the maximum dispatchable capacity was 37%, with 20% 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 is 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).[22] Selected as an easier to implement alternative to real-time co-optimization of energy and ancillary services, the ORDC places an economic value on the reserves being provided, with separate pricing for online and offline reserves. The ORDC curves for 2017 are shown in Figure 16 below. The curves are determined in advance for four-hour blocks that vary across seasons. This depiction shows the breadth of distribution of the ORDC values across the year. The methodology leads to some large discontinuities between the curves where for the same reserve level the adder value changes significantly between adjacent time blocks. The largest such change in 2017 occurred in the summer season between 9:59 p.m. and 10:00 p.m. where the value of the ORDC curve changed more than $800 per MWh for a 3,000 MW reserve level. Once available reserve capacity drops to 2,000 MW, prices will rise to $9,000 per MWh for all the ORDC curves.

Figure 16: Seasonal Operating Reserve Demand Curves, by Four-Hour Blocks

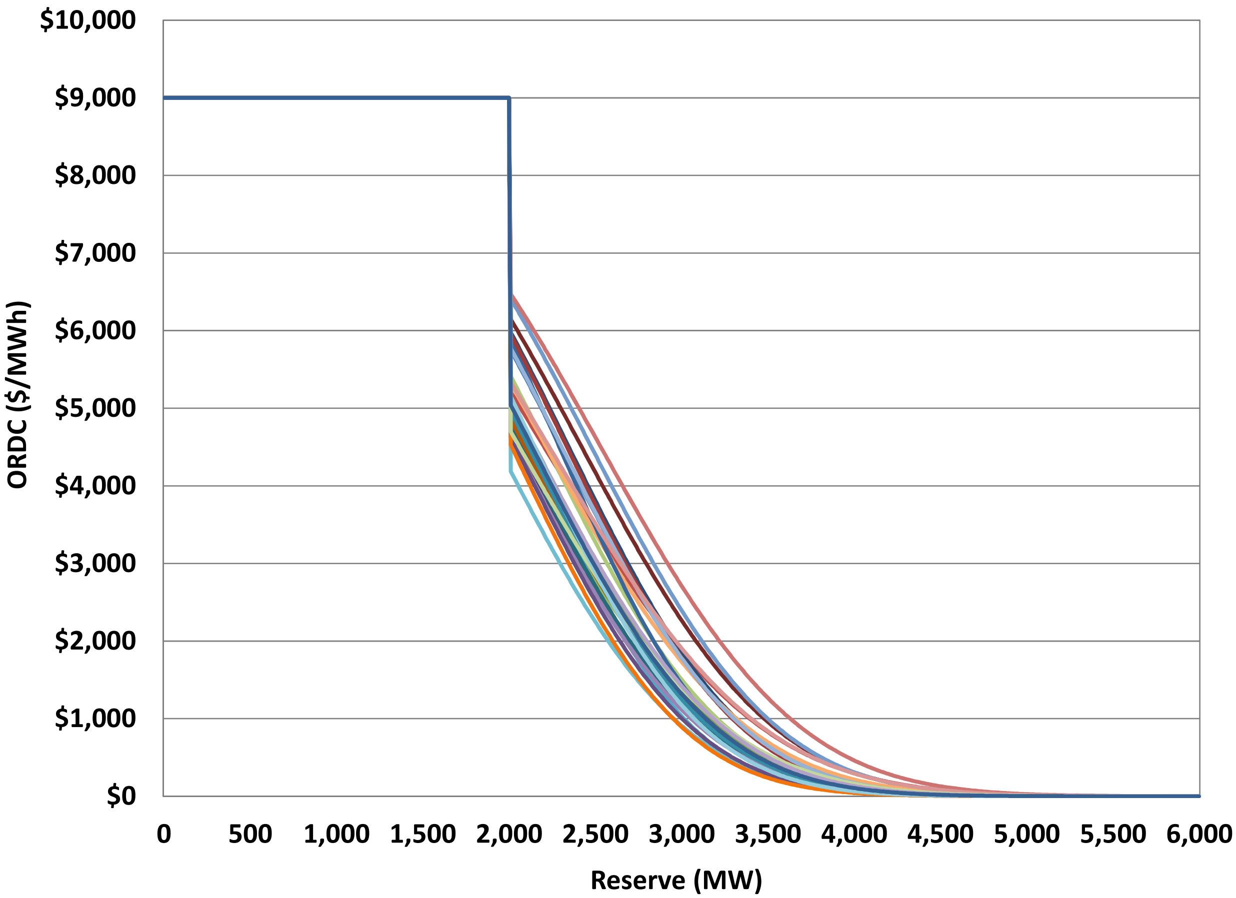
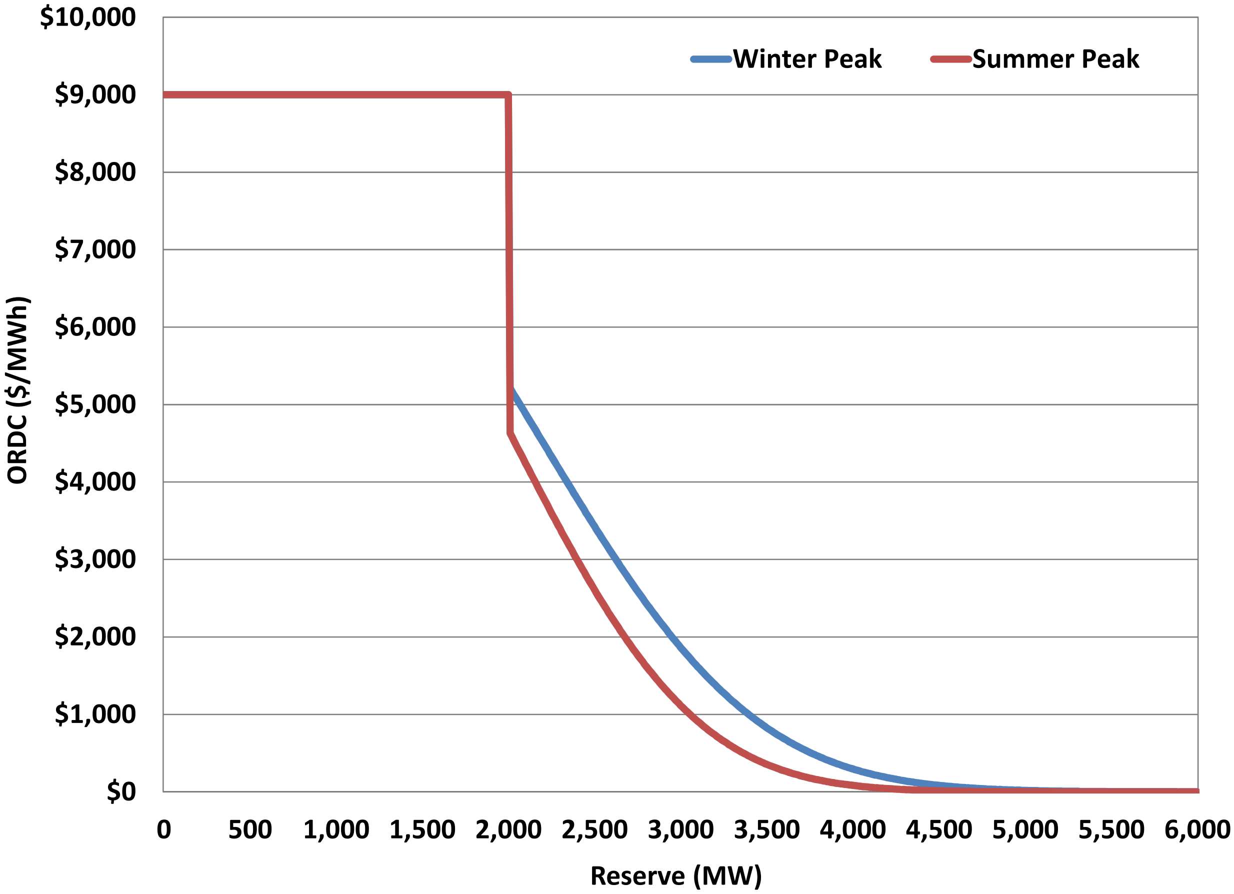


Figure 17 depicts the peak ORDCs applicable during winter and summer peak hours in 2017.

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

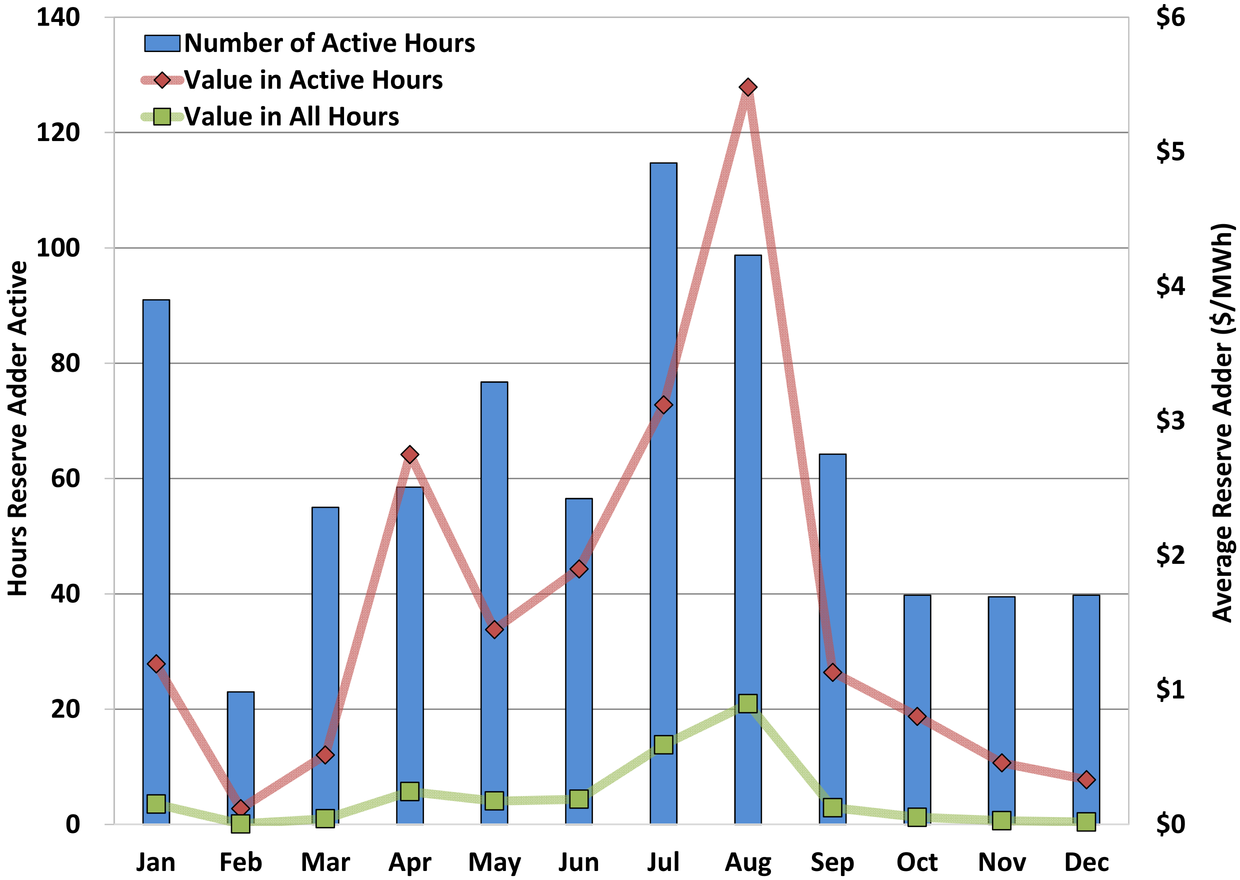
Figure 17: Winter and Summer Peak Operating Reserve Demand Curves



The following two analyses illustrate the contributions of the operating reserve adder and the reliability adder to shortage pricing. As described above in Figure 1: Average All-in Price for Electricity in ERCOT, the contributions of the energy price adders were relatively small in 2017. The first of the two adders is the operating reserve adder, is a shortage value intended to reflect the expected value of lost load (the loss of load probability, given online and offline reserve levels, multiplied by the deemed value of lost load).

Figure 18 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 2017, the operating reserve adder had the largest impacts to price during July and August. Overall, the operating reserve adder contributed $0.24 per MWh, or less than 1% to the annual average real-time energy price of $28.25 per MWh. These results do not indicate that ORDC has been ineffective or that it should be modified. The effects of the operating reserve adder are expected to vary substantially from year to year, and to have the largest effects when poor supply conditions and unusually high load conditions occur together and result in sustained shortages.

Figure 18: Average Operating Reserve Adder

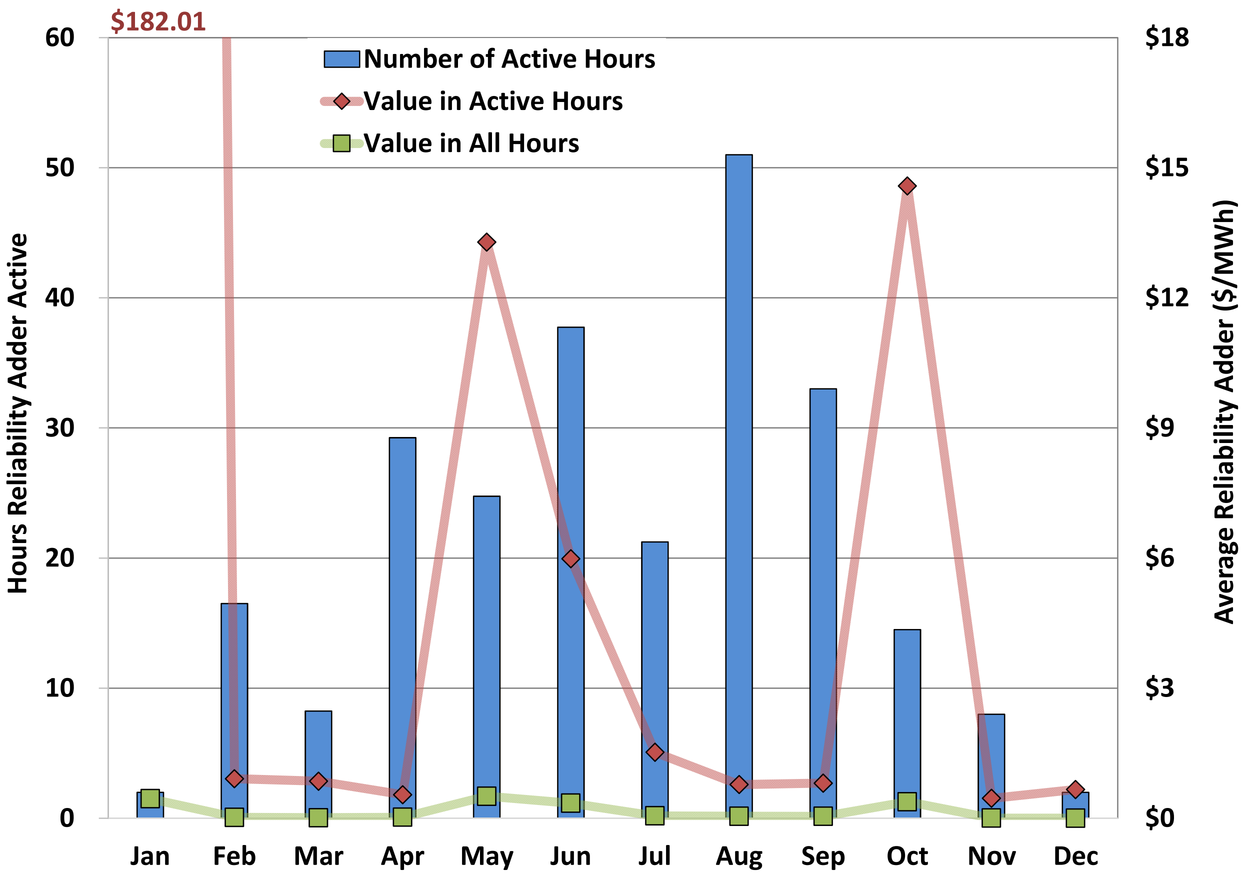


The reliability adder is intended to allow prices to reflect the costs of reliability actions taken by ERCOT, including RUCs and deployed load capacity. Absent this adder, prices will generally fall when these actions are taken.

Figure 19 below shows the impacts of the reliability adder in 2017. When averaged across only the active hours, the largest price impacts of the reliability adder occurred during two hours in January when a number of resources were issued a RUC instruction overnight between January 13 and 14. While such a RUC instruction is not common, system conditions at the time led ERCOT to call for additional capacity commitments.

The reliability adder was non-zero for fewer than 250 hours, or less than 3% of the time in 2017, most of which occurred in August. The contribution from the reliability adder to the annual average real-time energy price was $0.16 per MWh. The months with the largest impact from the reliability adder were May and January. Like the operating reserve adder, it had very little overall effect on the market outcomes in 2017 because supply conditions were rarely tight and ERCOT took fewer reliability actions.

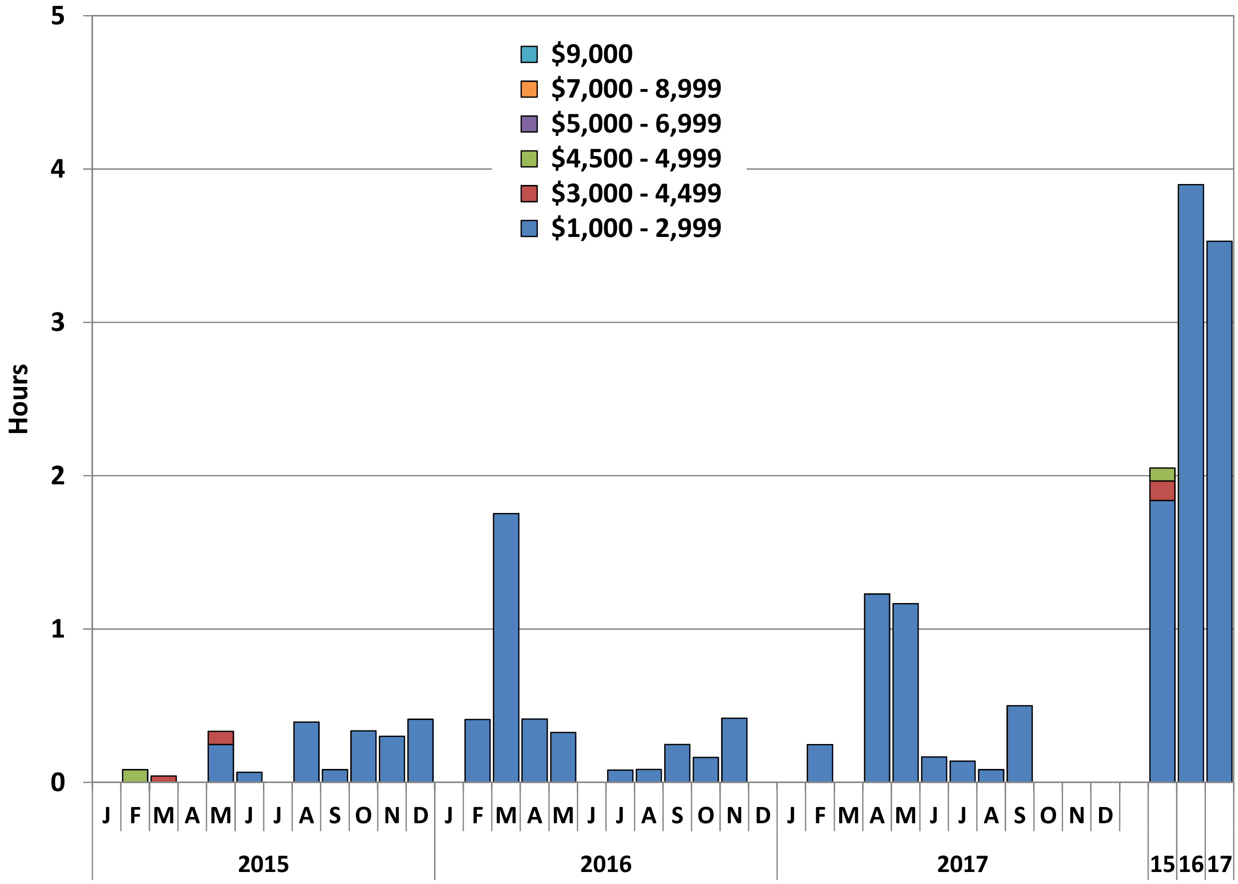
Figure 19: Average Reliability Adder



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 occurred from 2015 to 2017, Figure 20 below shows the aggregate amount of time when the real-time system-wide energy price exceeded $1,000 per MWh, by month. This figure shows that real-time pricing outcomes in 2017 were very similar to those in 2016, with the accumulation of prices greater than $1,000 per MWh occurring less than four hours over the entire year.

Figure 20: Duration of High Prices

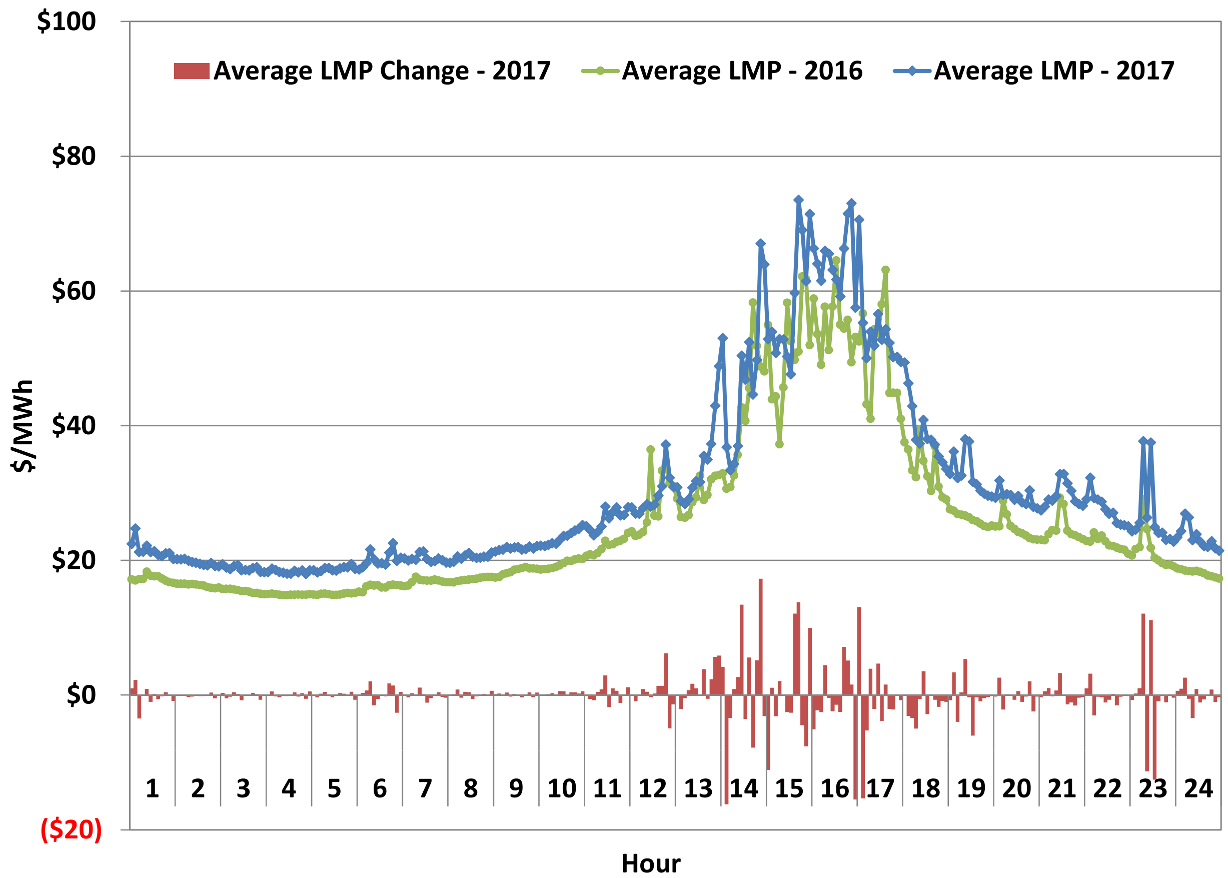


As a comparison, market prices cleared at the then in effect cap of $3,000 per MWh for 28.44 hours in 2011. Extreme cold in February 2011 and unusually hot and sustained summer temperatures led to much more frequent shortages in that year. Shortages in years with normal weather should be infrequent. As capacity margins fall, the frequency of shortages is likely to increase but will still vary substantially year-to-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 21 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 2017 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 2016 are also presented. Comparing average real-time energy prices for 2017 with those from 2016 shows very similar outcomes with greater volatility during peak hours.

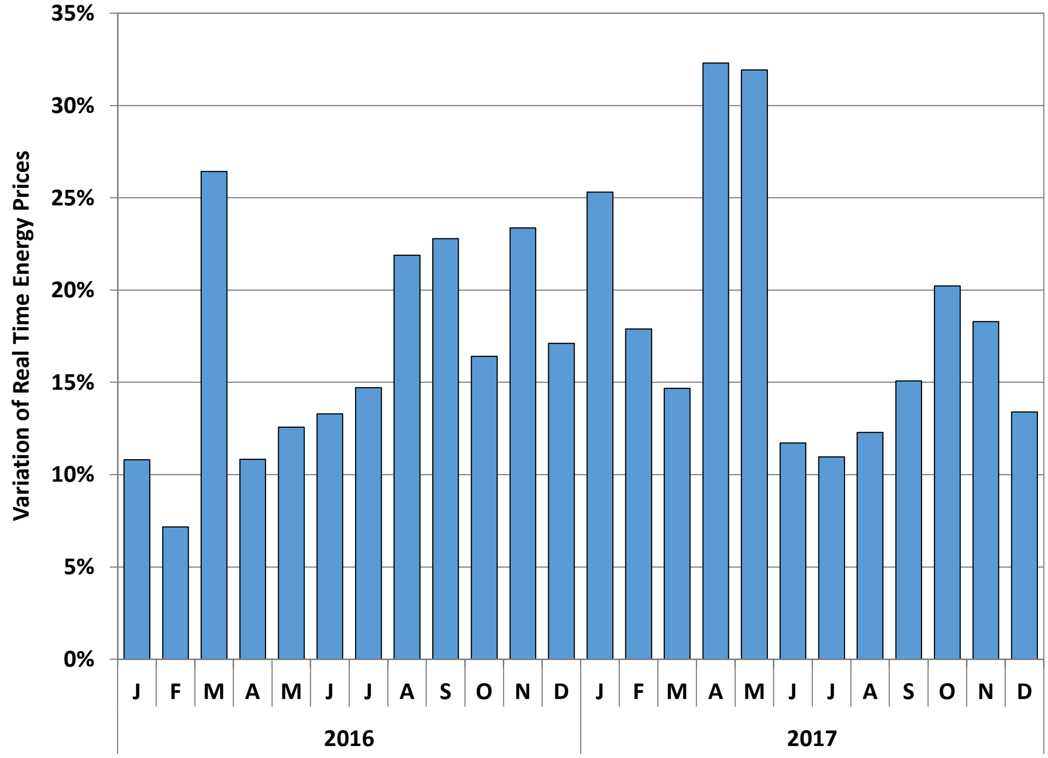
Figure 21: Real-Time Energy Price Volatility (May – August)



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 5.5% in 2017, compared to 5.4% in 2016.

Expanding the view of price volatility, Figure 22 below shows monthly average changes in five-minute real-time prices by month for 2017 and 2016. Without any prices at or close to the system-wide offer cap, the highest price variability occurred during spring and fall months when wind generation variations and load and wind generation forecast errors are the highest.

Figure 22: Monthly Price Variation



For another view of price volatility, Table 4 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 2013-2017.

Table 4: Zonal Price Variation as a Percentage of Annual Average Prices

	2013	2014	2015	2016	2017
Houston	14.8	14.7	13.4	20.8	24.9
North	15.4	15.2	14.6	19.9	26.2
South	13.7	14.1	11.9	15.5	14.8
West	17.2	15.4	12.9	16.8	17.5

These results show that price volatility is higher in 2017 for all Load Zones, except the South Load Zone. Increased percentage variation in prices can be explained by congestion pricing impacts.

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 is 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 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.

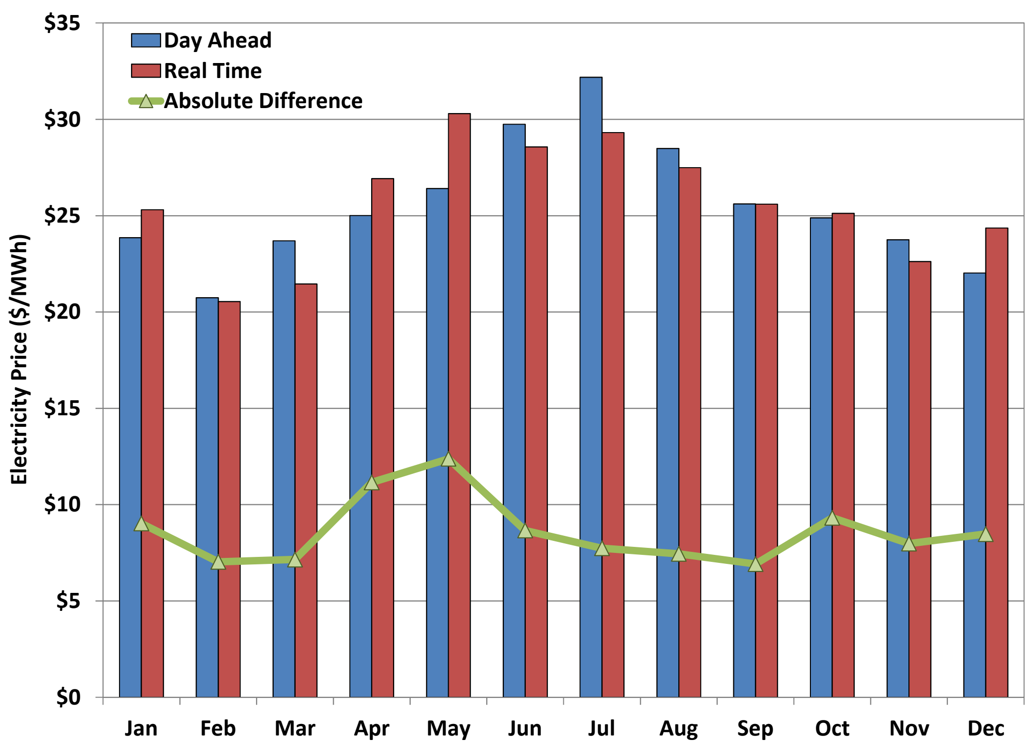
A. Day-Ahead Market Prices

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.[23]

Figure 23 summarizes the price convergence between the day-ahead and real-time markets, by month in 2017. Price convergence was very good in 2017; day-ahead and real-time prices both averaged $26 per MWh in 2017.[24] The lack of discernable day-ahead premium is likely due to the overall low energy prices and is consistent with low expectations for shortage conditions given ample installed reserves. Risk is typically 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 premium in 2017 occurred during July when the highest relative demand and highest prices occurred.

Figure 23: Convergence Between Day-Ahead and Real-Time Energy Prices



23 For instance, 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.

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

Real-time energy prices in ERCOT are allowed to rise to levels that are much higher than the shortage pricing in other organized electricity markets, which increases risk and would explain a higher day-ahead premium in ERCOT. Although most months experienced a day-ahead premium in 2017, 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 January, May, and December).

The average absolute difference between day-ahead and real-time prices was $8.60 per MWh in 2017 – a slight increase from $7.44 per MWh and $8.08 per MWh in 2016 and 2015, respectively.

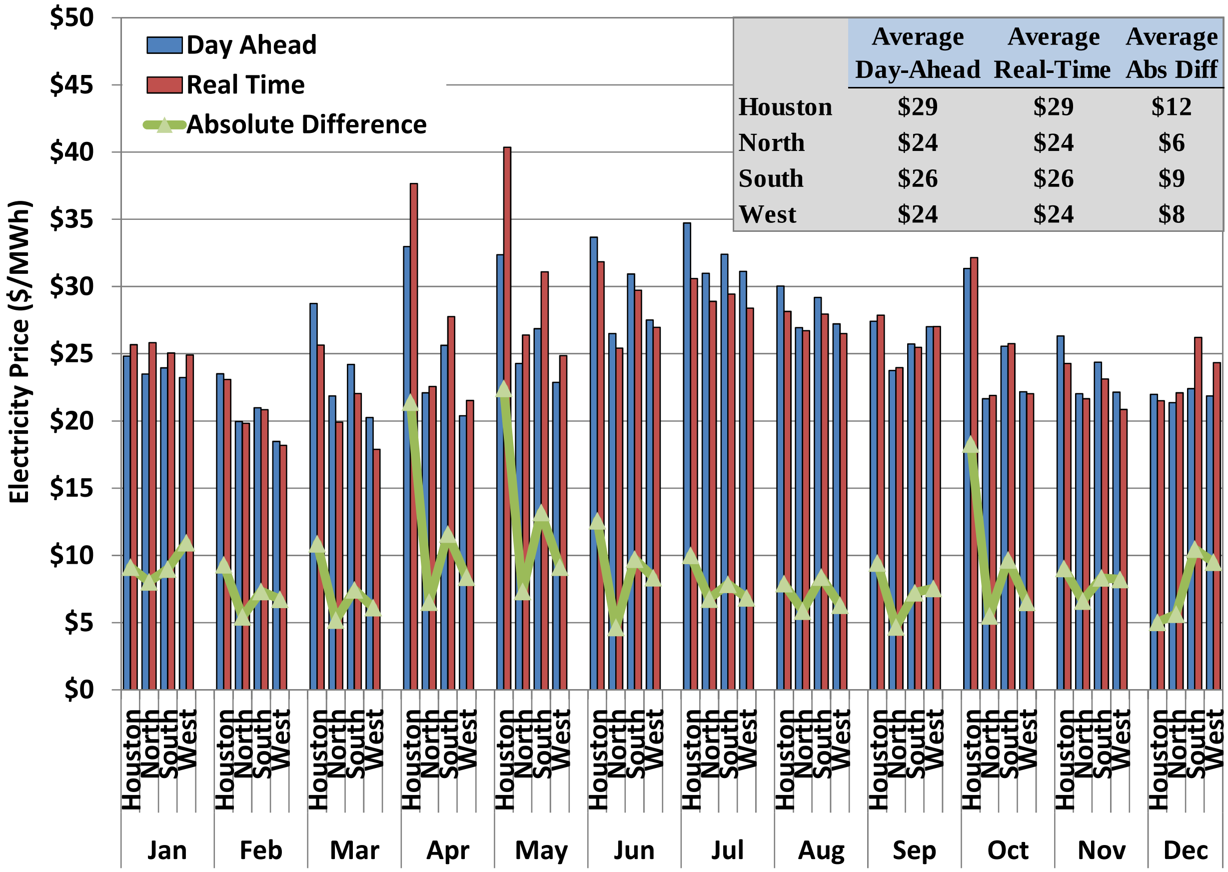
Table 5 displays the average day-ahead and real-time prices, showing the convergence for years 2011 through 2017.

Table 5: Historic Average Day-Ahead and Real-Time Prices

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

In Figure 24 below, monthly day-ahead and real-time prices are shown for each of the geographic zones. Notably, the volatility in the Houston zone increased in 2017 in contrast to the relative stability of the other zones. The larger difference between day-ahead and real-time prices observed in the Houston zone is likely associated with transmission congestion related to Houston import constraints.

Figure 24: 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 25 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 55% of real-time load in 2017, which was a slight increase compared to 53% in 2016. Although it may appear that many loads are 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.[25] When coupled with a self-

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

scheduled generating resource, the PTP obligation allows a participant to service 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.

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 25 below. Approximately 82% of QSEs' real-time load was hedged in the day-ahead market. 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 amount of real-time load hedged increased to nearly 90%.

Figure 25: Volume of Day-Ahead Market Activity by Month

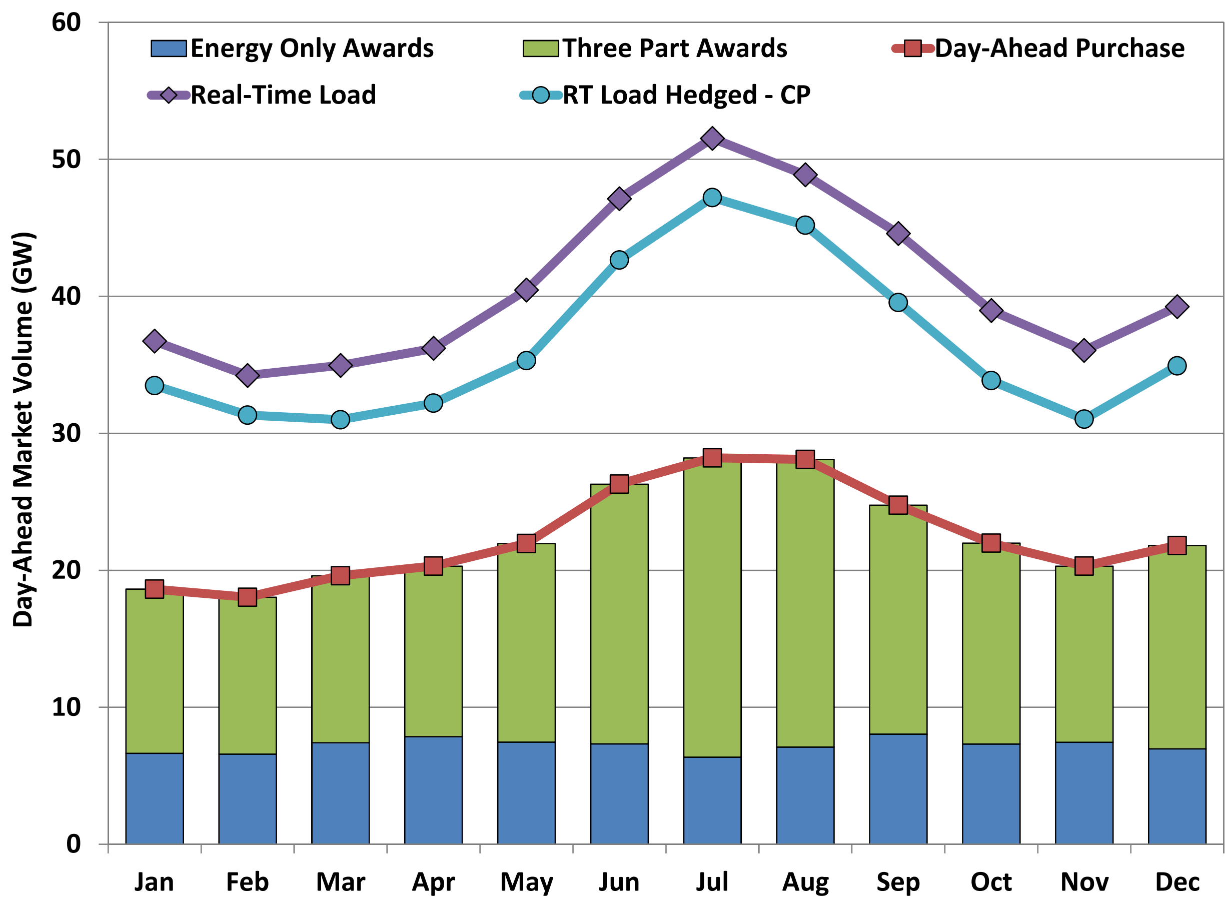
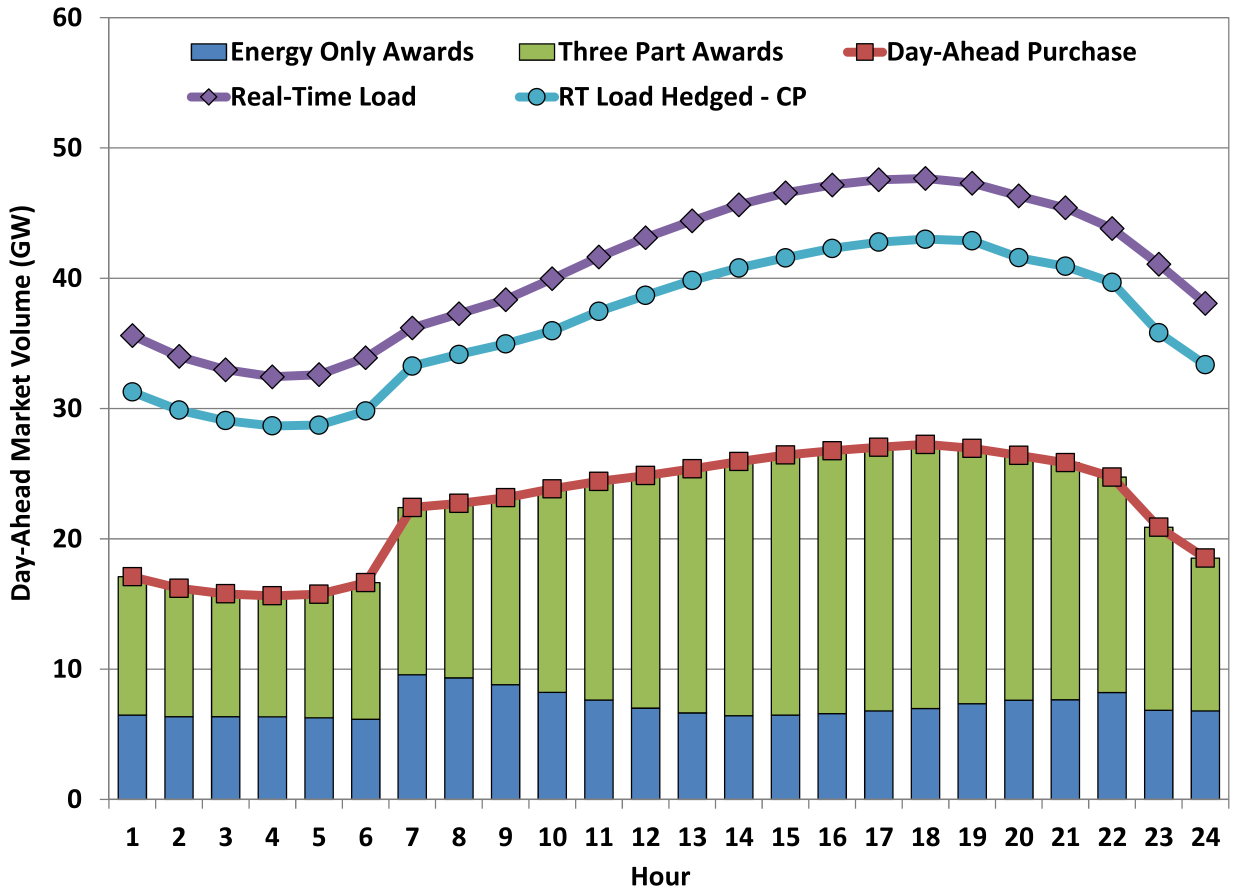


Figure 26 below, presents the same day-ahead market activity data summarized by hour of the day. In this figure the volume of day-ahead market transactions is disproportionate with load levels between the hours of 7 and 22 (hour ending). Since these times align with common bilateral and financial market transaction terms, the results in this figure are consistent with market participants using the day-ahead market to trade around those positions.

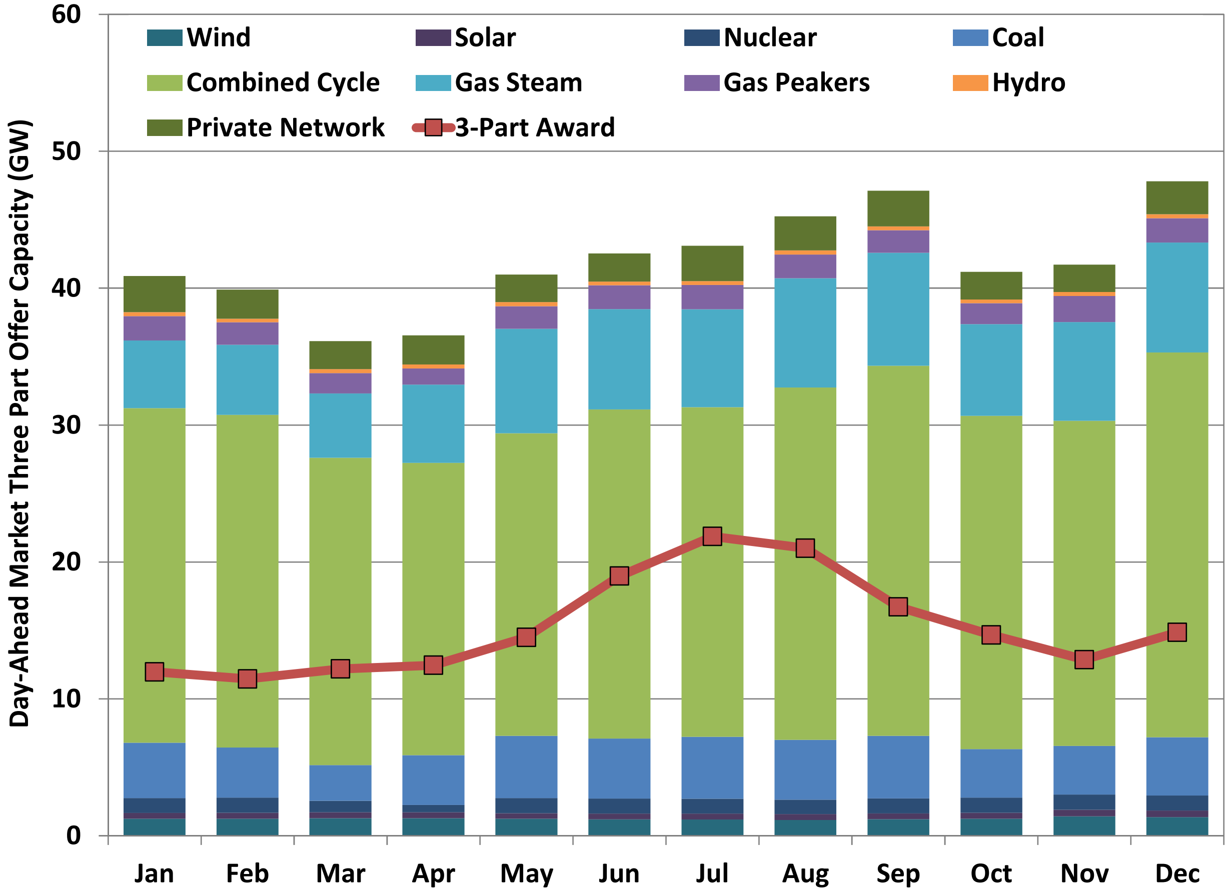
Figure 26: Volume of Day-Ahead Market Activity by Hour



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 is due to small volumes of three-part offers being submitted, the following analysis was performed.

Figure 27 below shows the total capacity from three-part offers submitted in the day-ahead market for 2017. 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 the market to determine whether the unit is economic.

Figure 27: Day-Ahead Market Three-Part Offer Capacity



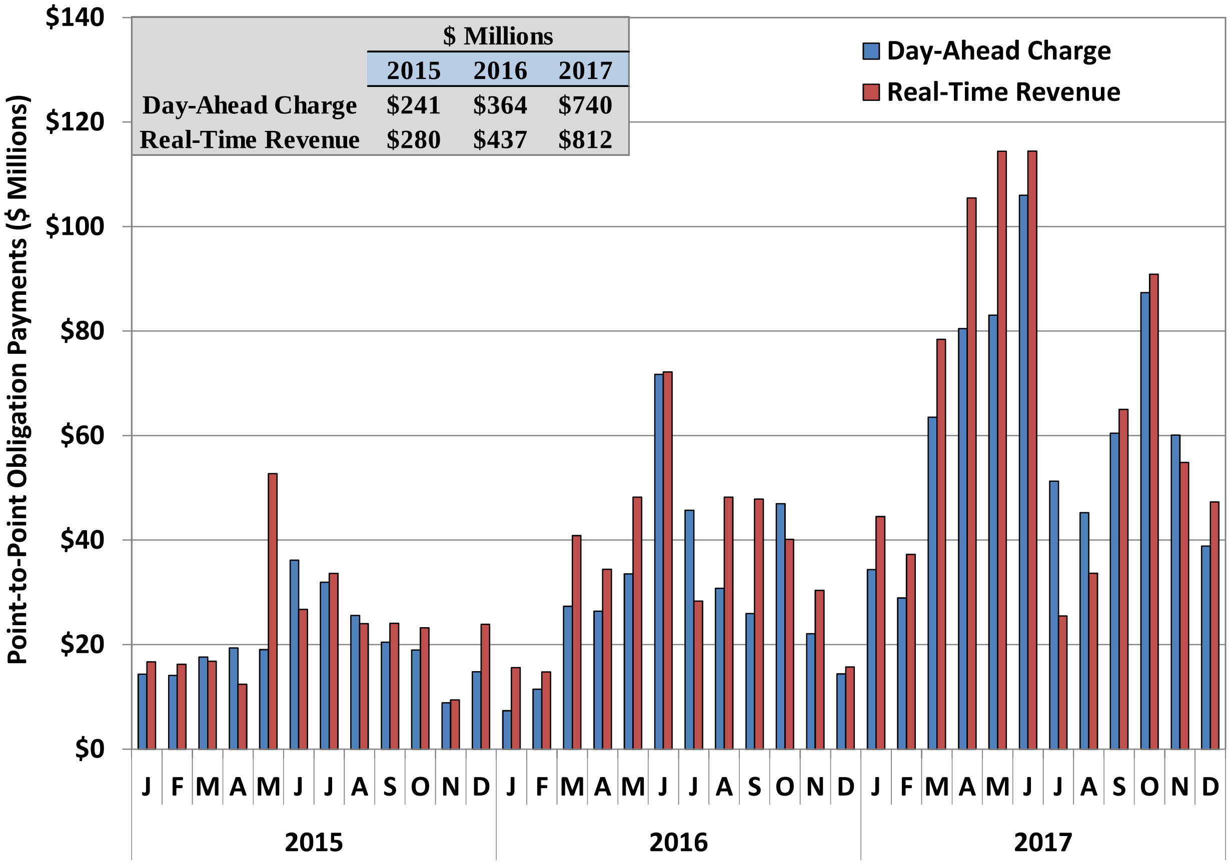
C. Point-to-Point Obligations

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. They receive 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 28, 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.

Figure 28: Point-to-Point Obligation Charges and Revenues



As in prior years, the aggregated total revenues received by PTP obligation owners in 2017 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. Across the year, and in nine of twelve months, 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 July, August 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 29: Point-to-Point Obligation Volume

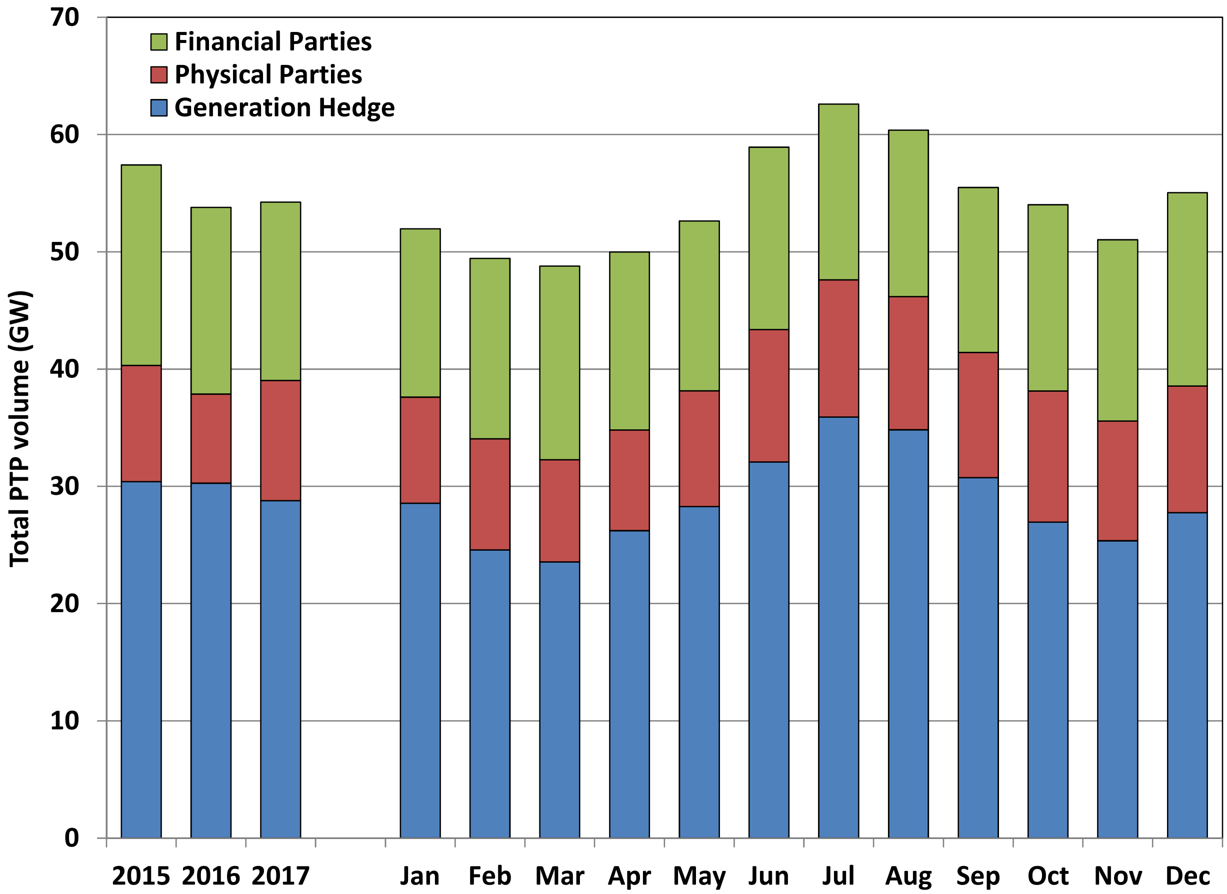
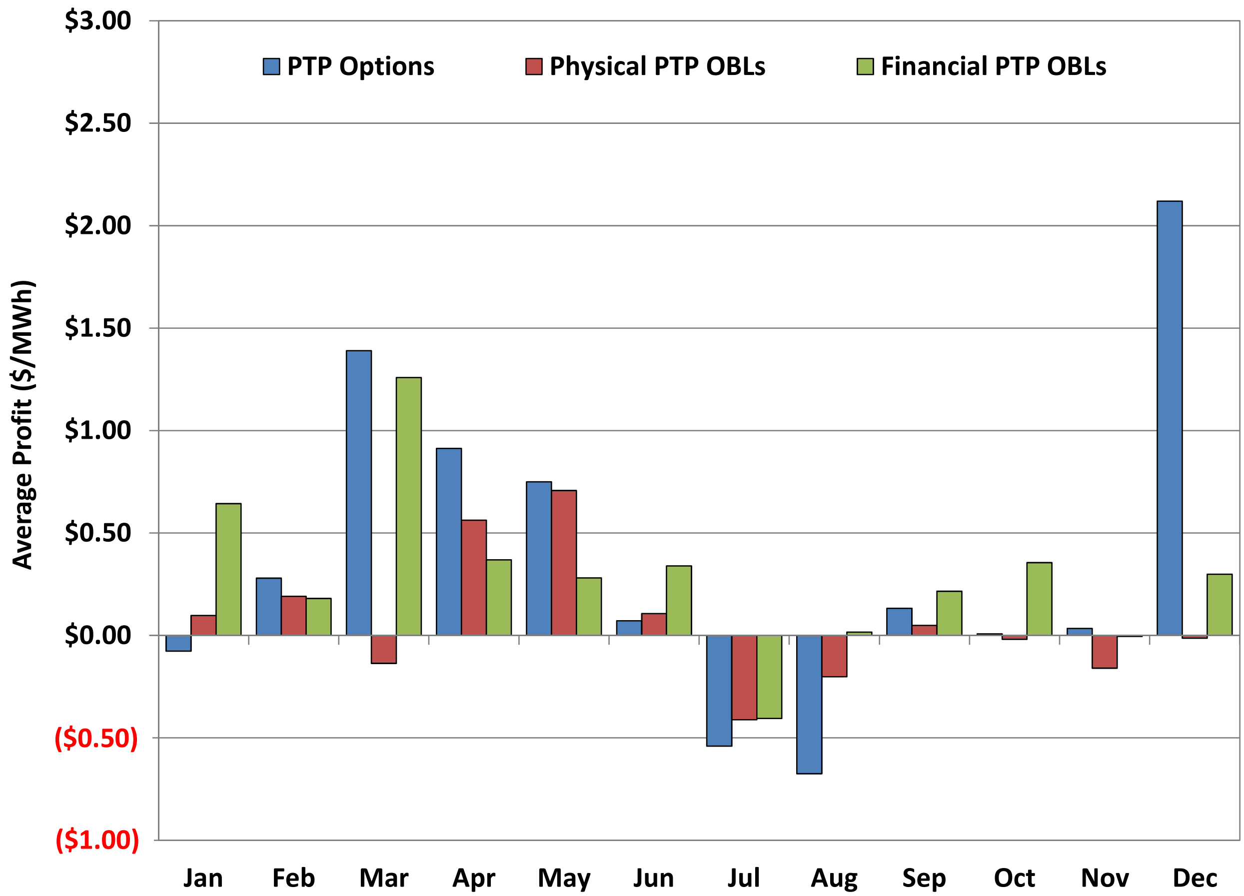


Figure 29 above presents the total volume of PTP obligation purchases divided into three categories. Different from Figure 25 and Figure 26 above, the volumes in this figure do not net out the injections and withdrawals occurring at the same location. It presents average purchase volumes on both a monthly and annual basis.

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 this 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 28% of the total volume of PTP obligations in 2017, a slight decrease from 30% in 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 30. Also shown is the profitability of "PTP obligations settled as options," which are instruments available only to Non-Opt-In Entities (NOIEs); shown below as "PTP Options".

Figure 30: Average Profitability of Point-to-Point Obligations



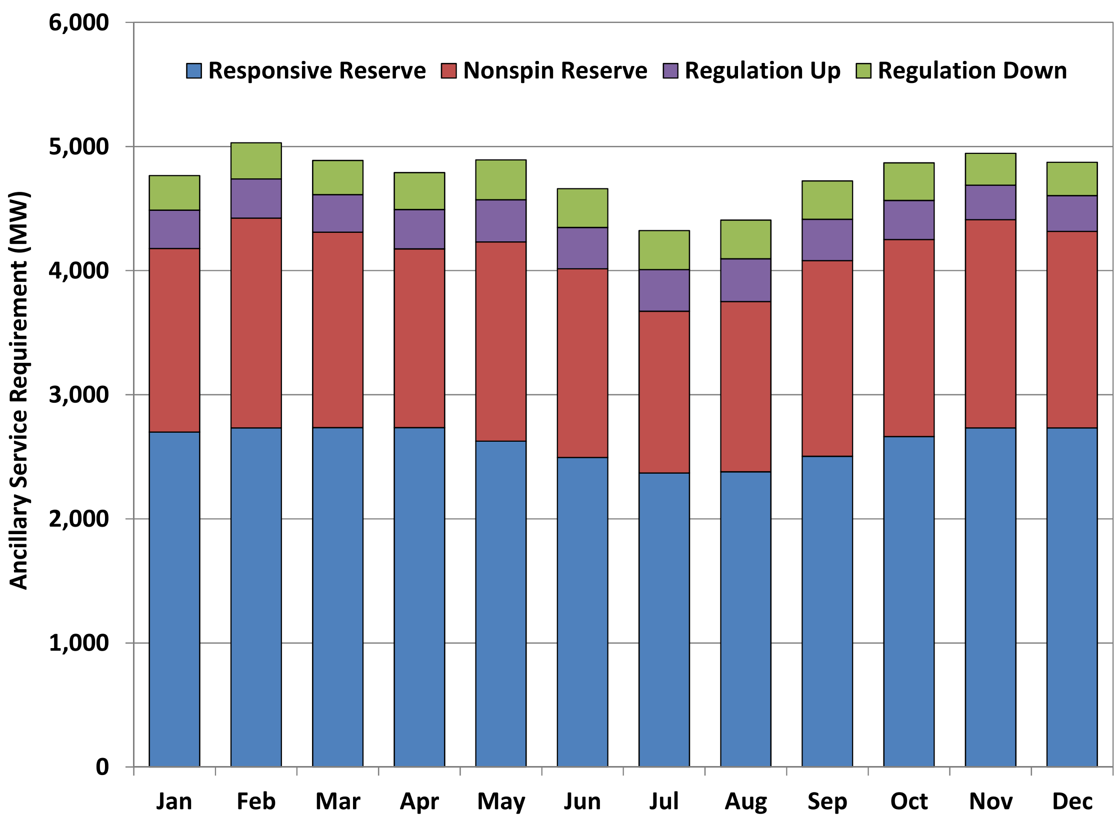
This analysis shows that in aggregate, PTP obligation transactions in 2017 were profitable for the year, yielding an average profit of $0.12 per MWh, the same average profit as in 2016. PTP obligations were profitable during 2017 for all types of parties, with average profits of $0.05 per MWh for physical parties, $0.31 per MWh for financial parties, and $0.31 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 purchase them through the ERCOT markets. 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. Total requirements for ancillary services declined again in 2017. The average total requirement in 2017 was less than 4,800 MW, a reduction from the average total requirement of approximately 4,900 MW in 2016 and 5,300 MW in 2015. The reduction in 2017 was primarily to responsive reserve. Figure 31 displays the hourly average quantities of ancillary services procured for each month in 2017.

Figure 31: Hourly Average Ancillary Service Capacity by Month



Another way to view the ancillary service requirements is by hour, averaged over the course of the year. Figure 32 presents this alternate picture of ancillary service procurement in 2017. In this view the large differences in quantities between some adjacent hours are readily apparent. For example, capacity requirements increase almost 500 MW in hour 7, decrease 260 MW in hour 8 and gradually increase for the next two hours. Hour 22 provides another example of a large increase in requirements in the hour prior to a large decrease. This pattern is 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 32: Yearly Average Ancillary Service Capacity by Hour

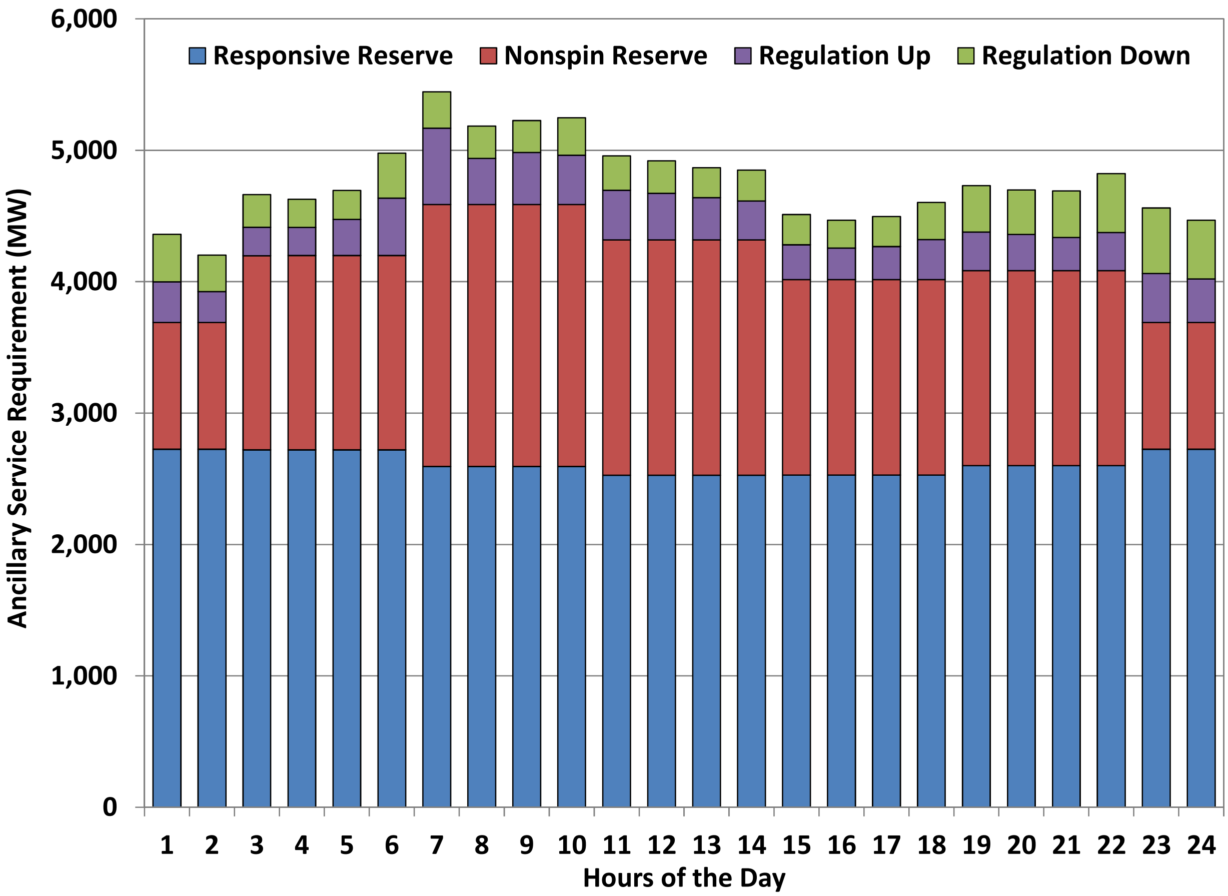
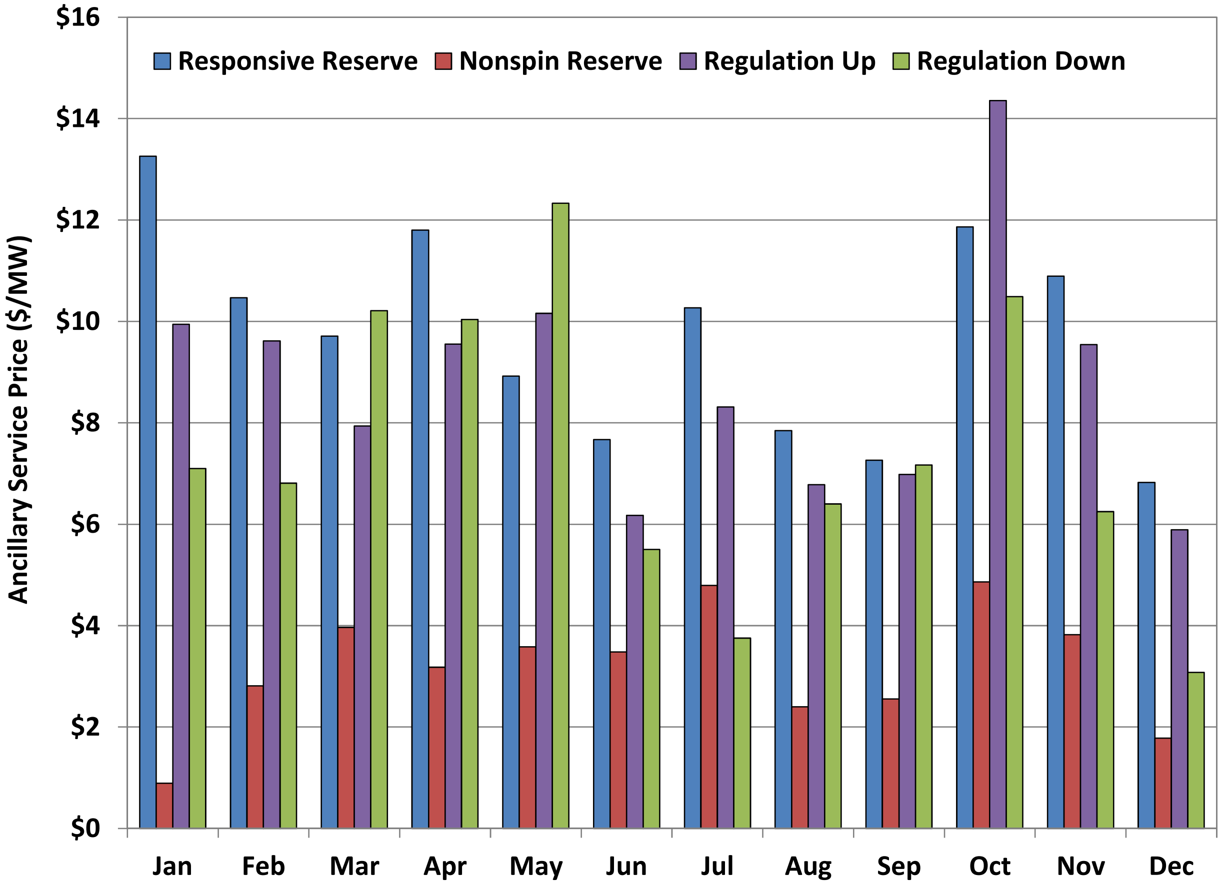


Figure 33 below presents the average clearing prices of capacity for the four ancillary services. The absence of meaningful occurrences of scarcity conditions in 2017 resulted in relatively small variation in average energy prices and correspondingly stable ancillary service prices.

Figure 33: Ancillary Service Prices



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.

During the recent period of low energy prices, this correlation has not been apparent. Monthly average prices for responsive reserve varied from $7 to $13 per MWh, with the highest price occurring in January. One possible explanation for this decoupling from day-ahead energy prices is that unit commitment patterns have changed because of high wind generation and less online capacity capable of providing reserves. This reduction in online capacity, especially in off-peak periods has led to higher prices for reserve prices in shoulder months.

Table 6 compares the average annual price for each ancillary service in 2017 with 2016. The changes in total requirements for ancillary services in 2017 led to concomitant changes in ancillary service prices. The average price for responsive reserve decreased in 2017, as did the total requirements for the service. Reductions in the average price for non-spinning reserves is consistent with the reduced requirements for this product. Average prices for regulation up and down products increased in 2017 even though requirements for the two products both decreased slightly.

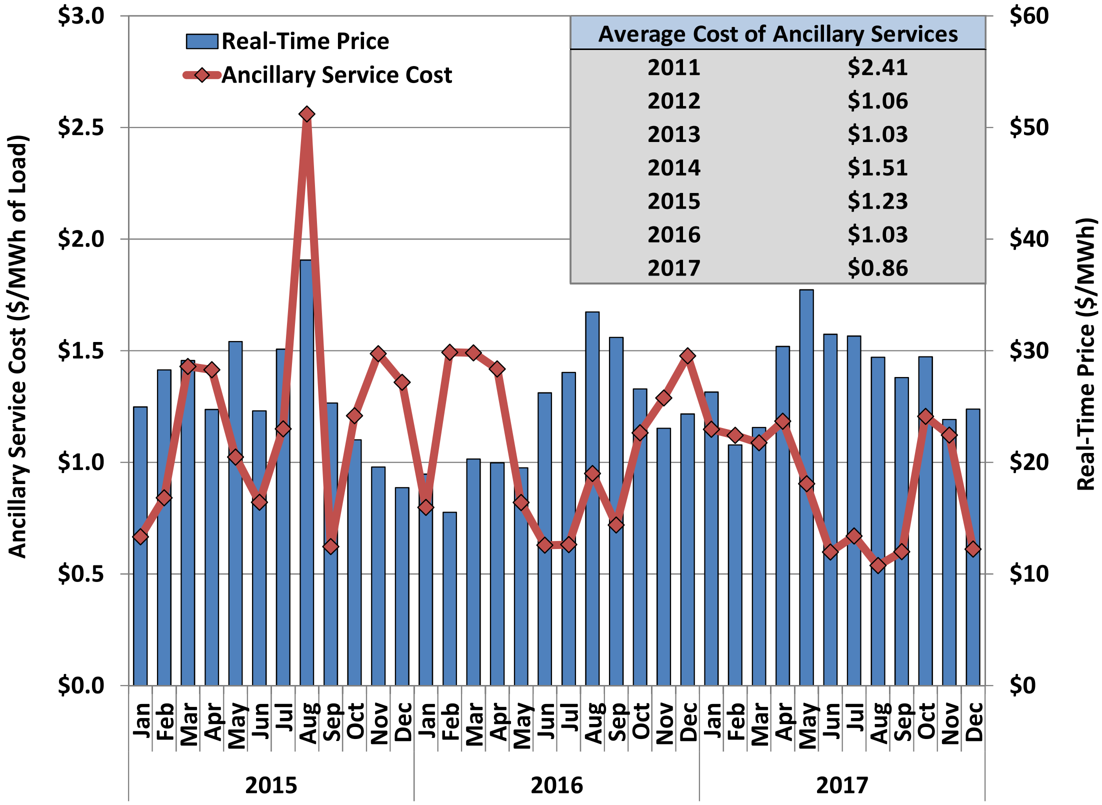
Table 6: Average Annual Ancillary Service Prices by Service

	2016 ($/MWh)	2017 ($/MWh)
Responsive Reserve	**$11.10**	**$9.77**
Nonspin Reserve	**$3.91**	**$3.18**
Regulation Up	**$8.20**	**$8.76**
Regulation Down	**$6.47**	**$7.48**

The prices for all of the ancillary service products remain modest in part because of the lack of shortages in 2017. When ERCOT experiences a shortage of operating reserves, real-time prices will rise to reflect the expected value of lost load embedded in the ORDC mechanism. The expectation of higher real-time prices will tend to drive up the day-ahead price for ancillary services. Hence, the lack of shortages contributed to the low average ancillary service prices shown in the table.

In contrast to the individual ancillary service prices, Figure 34 shows the monthly total ancillary service costs per MWh of ERCOT load and the average real-time energy price for 2015 through 2017. With no meaningful occurrences of scarcity conditions in 2017, the total cost for ancillary services was relatively low during summer months. The relatively higher costs observed during the other months may be explained by higher wind generation leading to changes in unit commitment patterns and less online capacity available to provide reserves.

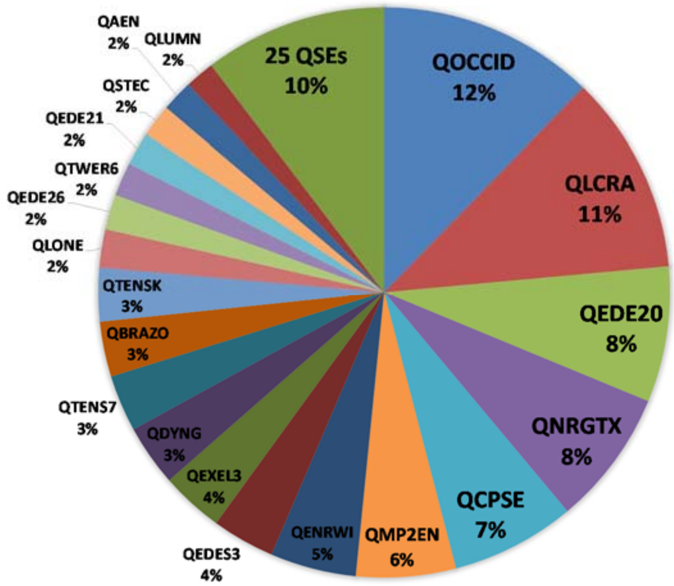
Figure 34: Ancillary Service Costs per MWh of Load



In absolute terms, the average ancillary service cost per MWh of load decreased to $0.86 per MWh in 2017 compared to $1.03 per MWh in 2016. Continued lower natural gas prices and smaller requirements for ancillary services led to further reduction in ancillary service prices in 2017. Total ancillary service costs were 3.0 % of the load-weighted average energy price in 2017, continuing the reduction seen since 2015 when they were 4.6 % and 2016 when they were 4.2 %.

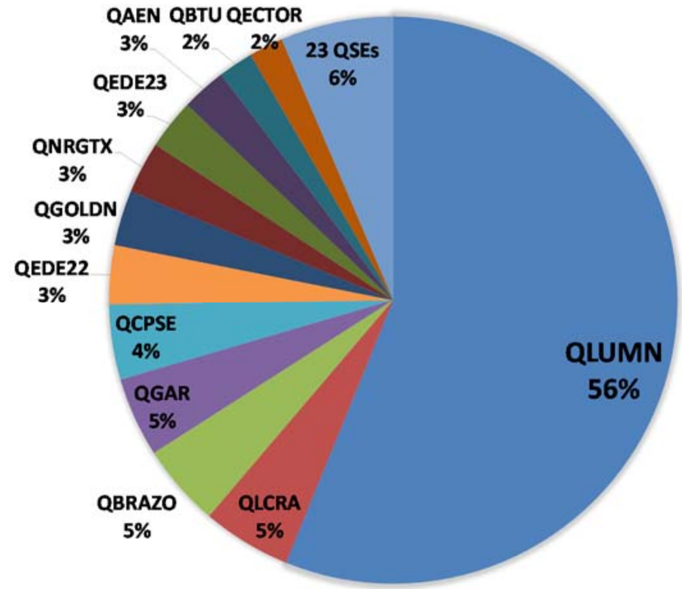
Responsive reserve service is the largest quantity purchased and typically the highest priced ancillary service product. Figure 35 below shows the share of the 2017 annual responsive reserve responsibility including both load and generation, displayed by Qualified Scheduling Entity (QSE). During 2017, 45 different QSEs self-arranged or were awarded responsive reserves as part of the day-ahead market; roughly the same as in 2016 when there were 42 separate providers and 2015 when there were 46 providers.

Figure 35: Responsive Reserve Providers



In contrast, Figure 36 below shows that the provision of non-spinning reserves is much more concentrated, with a single QSE (Luminant) bearing more than half the total responsibility.

Figure 36: 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 would allow the market to substitute its procurements between units on an interval-by-interval basis to minimize costs and set efficient prices. Additionally, it could allow higher quality reserves (e.g., responsive reserves) to be substituted for lower quality reserves (e.g., non-spinning reserves), thus reducing the reliance upon a single entity to provide this type of lower quality reserves.

Figure 37: Regulation Up Reserve Providers

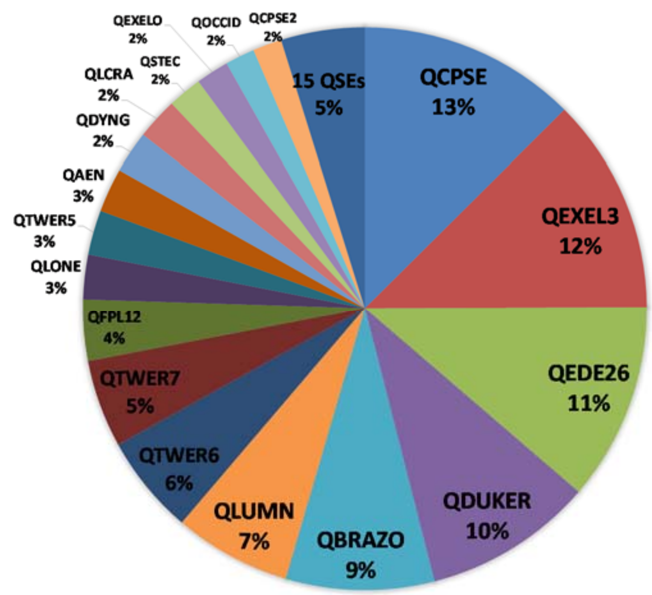
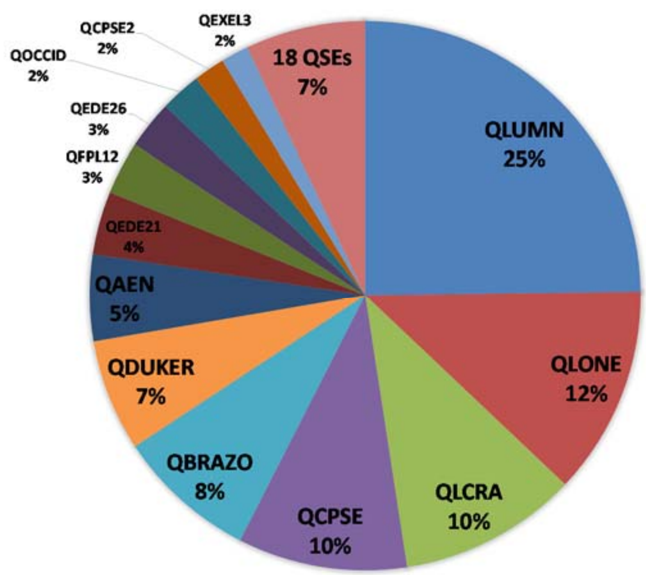


Figure 37 above shows the distribution for regulation up reserve service providers and Figure 38 shows the distribution for regulation down reserve providers. These two figures show that the provision of regulation services is somewhat more concentrated than responsive reserves, but far less so than non-spinning reserves.

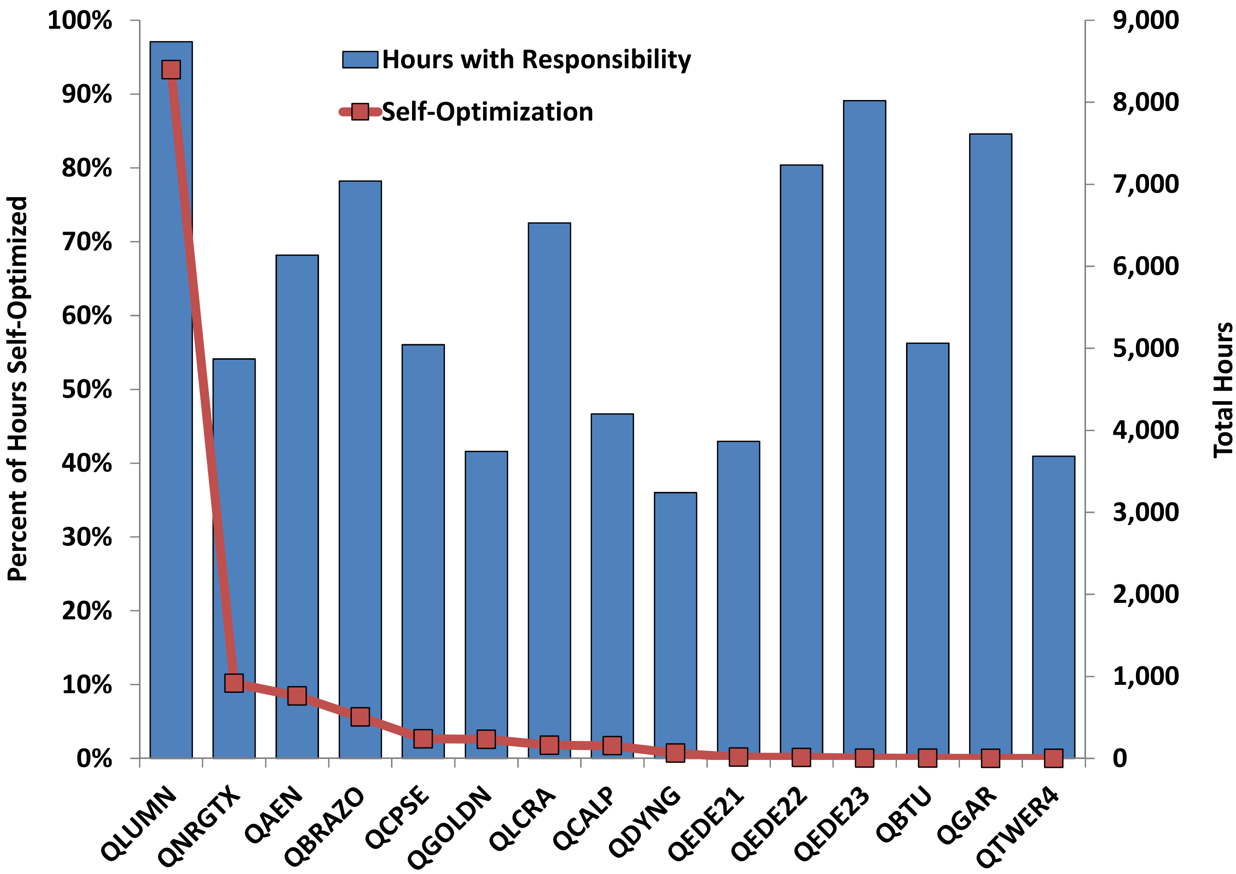
Figure 38: Regulation Down Reserve Providers



Ancillary service capacity is procured as part of the day-ahead market clearing. Between the time it 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 and as shown in the following two figures, this self-optimization is quite common.

The following two charts describe the frequency that each QSE with a unit-specific ancillary service responsibility at 16:00 day-ahead, moved any portion of its ancillary service responsibility to a different unit in its portfolio for real-time operations. Figure 39 shows the total hours each QSE has a non-spinning reserves responsibility and the percentage of time that responsibility was self-optimized.

Figure 39: Internal Management of Non-Spinning Reserve Portfolio by QSE



The QSEs are listed in descending order based on the frequency of self-optimization. This figure, taken in conjunction with Figure 36, shows that the provider with the largest share of non-spinning reserve responsibility also most frequently moved the responsibility between its units. Luminant had a responsibility to provide non-spinning reserves in almost every hour of 2017, and for nearly all of those hours they moved at least a portion of its responsibility to a unit different from the one that initially received the award.

Figure 40 below provides a similar analysis for the percentage of time when responsive reserve service was self-optimized by a QSE, that is, moving the day-ahead responsibility to a different unit before real-time.

Figure 40: Internal Management of Responsive Reserve Portfolio by QSE

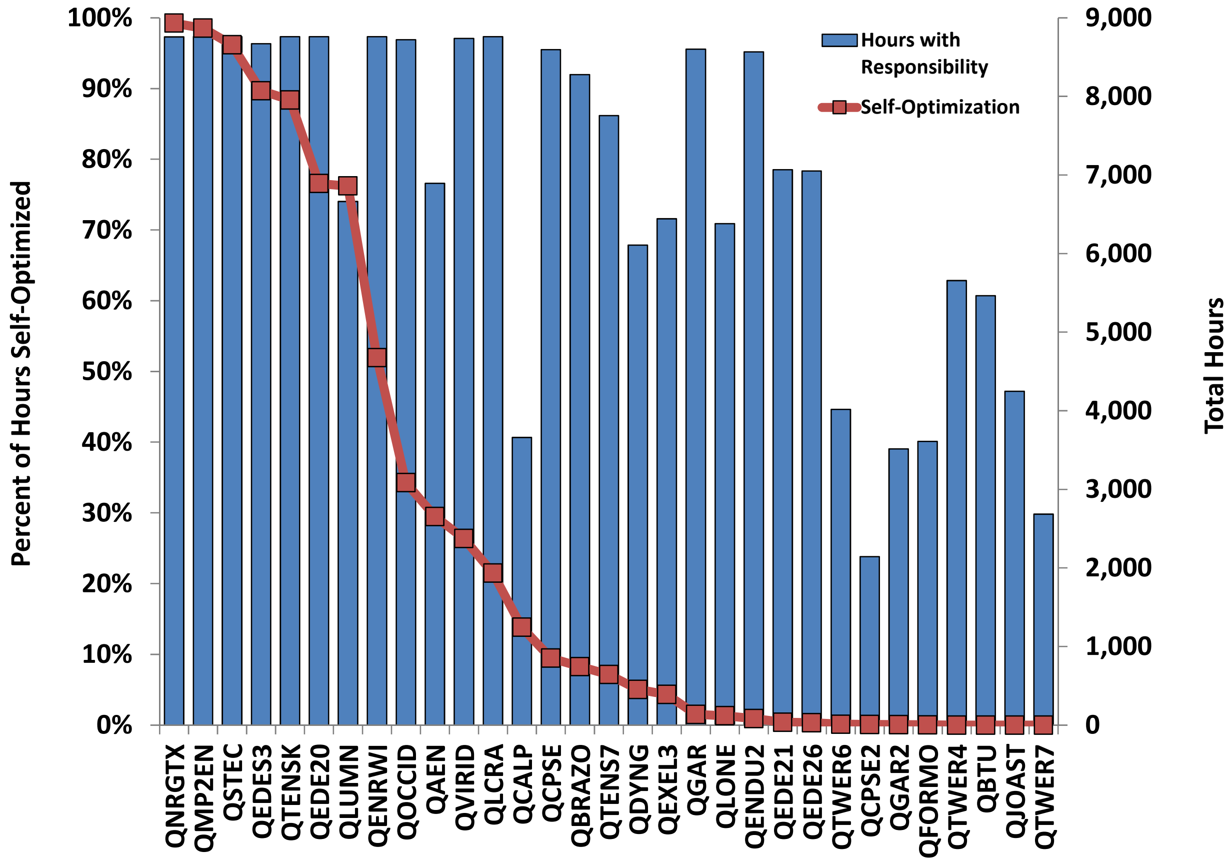


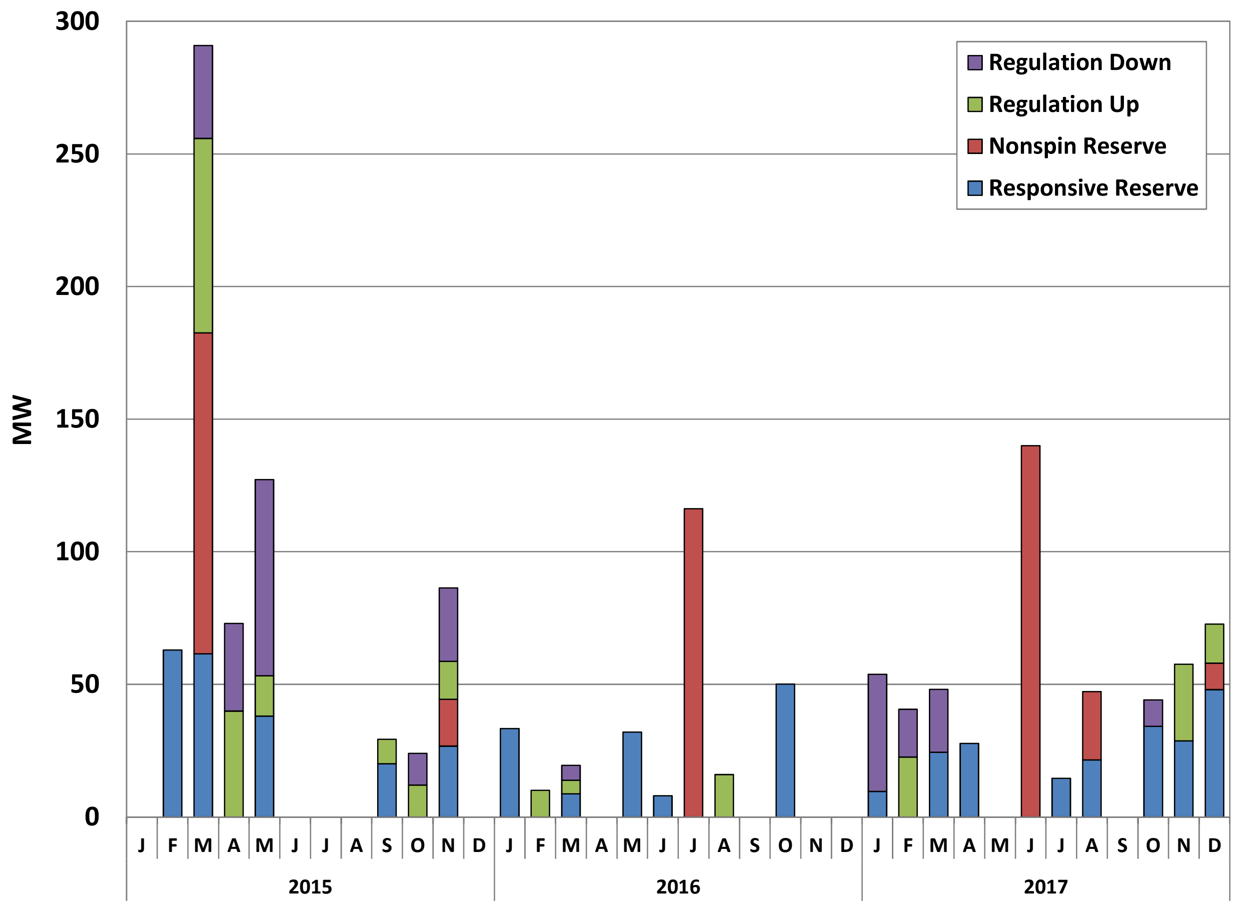
Figure 40 demonstrates that many QSEs moved responsive reserve responsibilities between units more routinely than QSEs providing non-spinning reserve service. For responsive reserve service, eight QSEs moved the responsibility more than 50% of the time; whereas only one QSE moved non-spinning reserve responsibility more than 50% of the time.

If all ancillary services could be continually reviewed and adjusted in response to changing market conditions, the efficiencies would flow to all market participants and would be greater than what can be achieved by QSEs acting individually. Since the initial consideration of ERCOT's nodal market design, the IMM has been recommending that ERCOT implement real-time co-optimization of energy and ancillary services because of this improved efficiency.

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. Without comprehensive, market-wide co-optimization, 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. They 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 they face 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 ten to thirty times greater than annual average clearing prices from the day-ahead market.

ERCOT 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 2017. A SASM was executed 17 times in 2017, providing 189 service-hours in 2017. This was more frequent than in 2016 when SASMs were executed 12 times replacing services in 80 hours. The frequency of SASMs continues to be very low, declining from a high of 9.3% of the hours in 2011, to less than 1% in 2015 and 2016, and 1.5% of the hours in 2017. The final analysis in this section, shown in Figure 41below, summarizes the average quantity of each service that was procured via SASM.

Figure 41: Ancillary Service Quantities Procured in SASM



The opportunity exists for market participants to use the SASM process as a re-configuration market. That is to move into or out of ancillary service positions awarded day ahead. SASMs were infrequent 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.

Because ancillary services are not co-optimized with energy in the SASM, potential suppliers are required to estimate opportunity costs rather than have the auction engine calculate it directly, which leads to resources that underestimate opportunity costs being 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.

Real-time co-optimization of energy and ancillary services does not require resources to estimate opportunity costs, would 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 would also reduce ERCOT's need to use RUC procedures to acquire ancillary services: its biggest benefit would 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 would provide benefits across the market.

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. The action taken to ensure operating limits are not violated is 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 as a result of necessary changes to generation output to ensure that operating limits are not violated. This increase in more expensive generation or decrease in less expensive generation, or both 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 2017, 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.

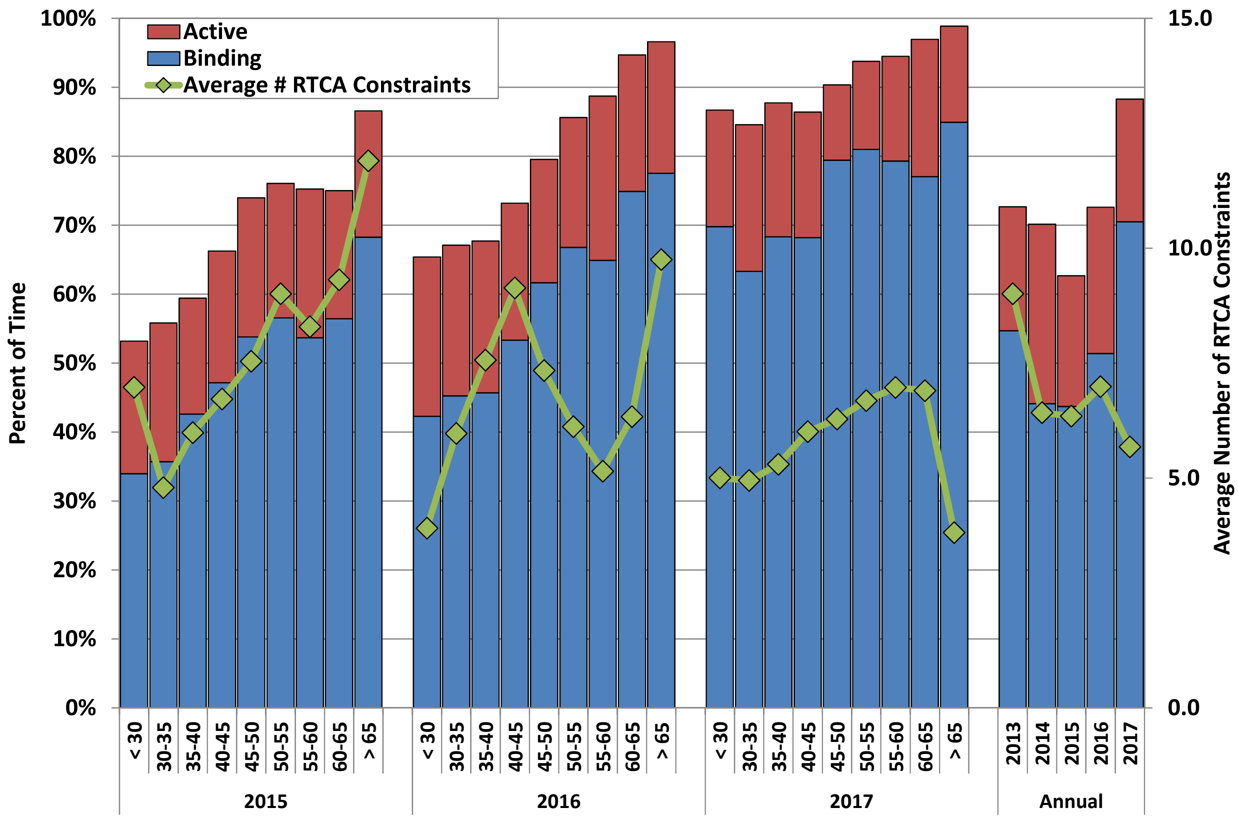
A. Summary of Congestion

The total congestion costs experienced in the ERCOT real-time market in 2017 were $967 million, almost twice the 2016 value. Three factors contributed to the substantial increase; 1) continued limitations on export capacity from the Panhandle, 2) planned outages associated with construction of the Houston Import Project[26], and 3) unusual operating conditions in the aftermath of Hurricane Harvey. Congestion was more frequent in 2017, occurring in 70% of all intervals. All zones except for the Houston zone experienced increased congestion in 2017.

26 The Houston Import Project, which consists of the installation of a Limestone-Gibbons Creek-Zenith 345 kV double circuit line to meet reliability requirements for Houston load growth. The project was approved by the ERCOT Board of Directors on April 8, 2014.

Figure 42 provides a comparison of the amount of time transmission constraints were active and binding for various load levels in 2015 through 2017. 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 in which the resulting flows on the transmission system are evaluated after systematically removing elements of the transmission system. 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 actually 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 42: Frequency of Binding and Active Constraints



Constraints were activated more frequently in 2017, occurring in 88% of all hours compared to 73% in 2016. The percentage of time with active constraints in 2017 was the highest since the implementation of the ERCOT Nodal Market in 2010 and was higher at nearly all load levels. The most notable difference between 2017 and 2016 was that, while RTCA on average showed fewer constraints in 2017, the percentage of time with an active constraint in each load level was higher in 2017 than 2016. This difference is explained by a 9% increase in the amount of time with an active Generic Transmission Constraint (GTC). A GTC was active 43% of the time 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 stability or a voltage condition. Certain GTC limits are determined in real-time using the Voltage Stability Assessment component of the Energy Management System. 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. Actions taken to resolve a GTC may also benefit other potential congestion issues, resulting in fewer thermal constraints in RTCA. This could explain the lower number of RTCA constraints overall, but also the increase in constraint activity in 2017.

Shown below in Table 7 are the GTCs that were monitored in 2017. The highlighted GTCs were either modified or terminated in 2017.

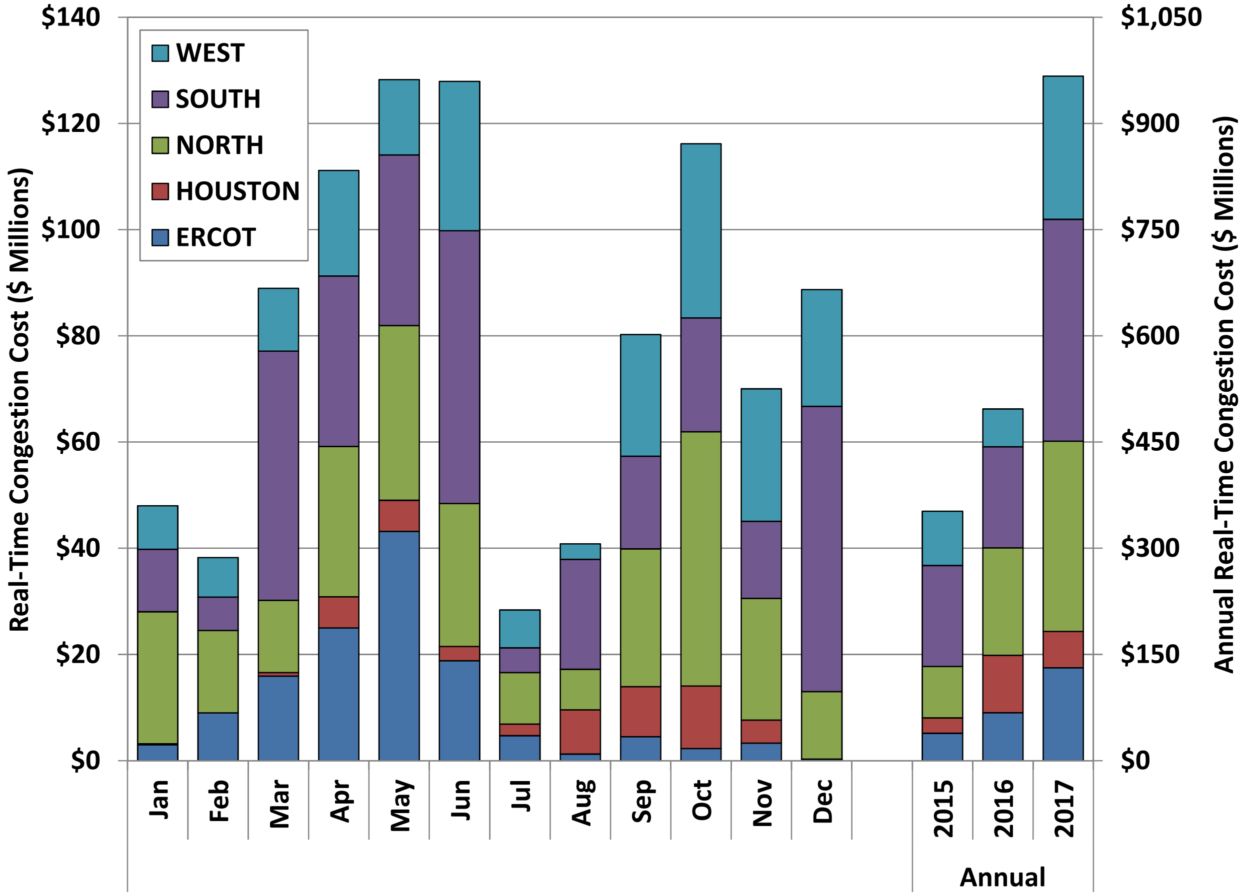
Table 7: Generic Transmission Constraints

Generic Transmission Constraint	Effective Date	Modification or Termination Date
North to Houston	December 1, 2010	
Rio Grande Valley Import	December 1, 2010	
Zorillo to Ajo	February 27, 2015	Changed to Nelson Sharpe - Rio Hondo
Panhandle	July 31, 2015	
Laredo	September 9, 2015	August 17, 2017
Liston	November 12, 2015	March 8, 2017
Pomelo Tap	October 5, 2016	Changed to North Edinburg - Lobo
Red Tap	August 29, 2016	
Bakersfield	January 25, 2017	May 4, 2017
North Edinburg - Lobo	August 24, 2017	
Nelson Sharpe - Rio Hondo	October 30, 2017	
East Texas	November 2, 2017	

Except for the North to Houston and the Rio Grande Valley Import constraints, all GTCs resulted from issues identified during the generation interconnection process.

Figure 43 displays the amount of real-time congestion costs associated with each geographic zone, with the monthly values of 2017 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 43: Real-Time Congestion Costs



The months of January, February, July, and August exhibited the least amount of congestion costs, whereas the remaining months, typically the "shoulder months," reflected much higher congestion. This trend is expected because most transmission and generation outages for maintenance and upgrades occur during the shoulder months.

Cross-zonal congestion in 2017 was the most costly since 2011 because of the increased frequency and cost associated with Houston import constraints. All zones except for the Houston zone experienced an increase in price impacts in 2017. Although the North to Houston constraint has been a significant contributor to total congestion in the past, most of the increased congestion in 2017 was attributable to conditions that materialized last year. Two of the notable new issues of 2017 were the urgent maintenance of Electric Transmission Texas (ETT) structures in the West zone and the impacts of Harvey Hurricane near Corpus Christi. North to Houston congestion was attributed to line outages to facilitate the Houston Import Project implementation. The completion of the Houston Import Project in the spring of 2018 is expected to reduce associated congestion.[27]

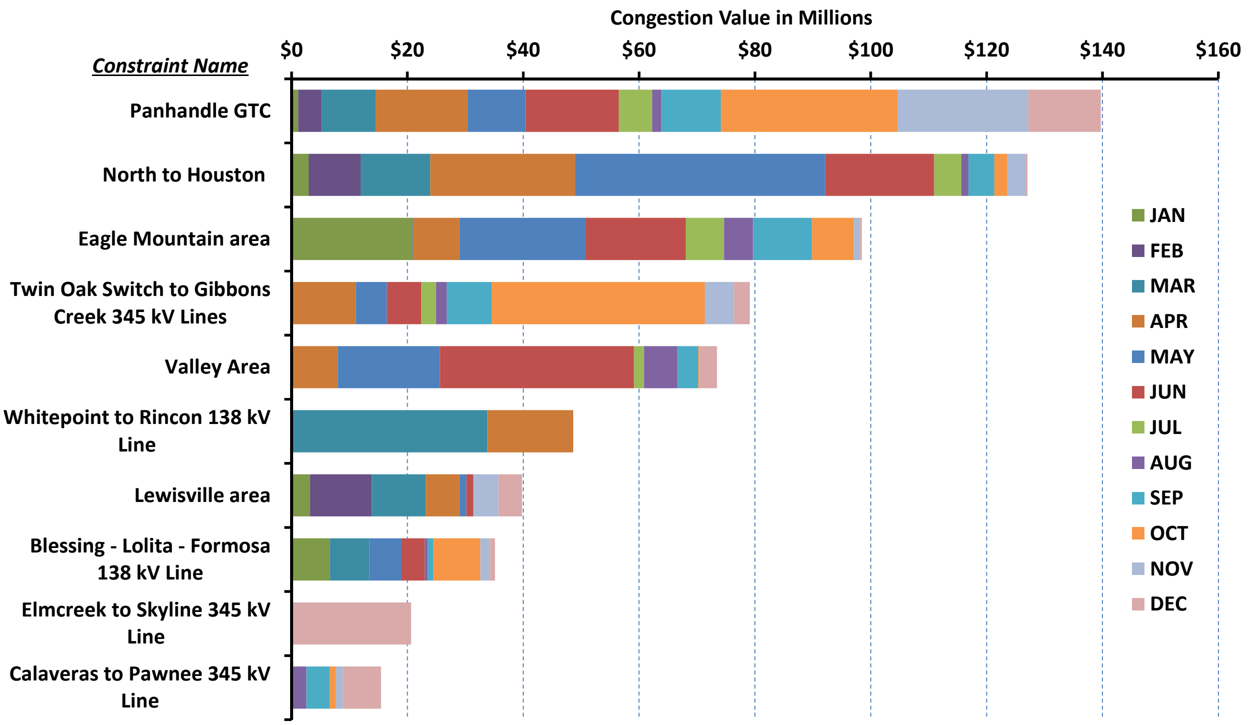
[27] The Houston Import Project was completed in April 2018, ahead of the expected completion date in June.

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 399 unique constraints that were binding at some point during 2017 with a median financial impact of approximately $235,000. In 2016, there were 320 unique constraints with a median financial impact of $150,000. The increased frequency and uniqueness of the constraints contributed to higher congestion costs in 2017.

Figure 44 displays the ten most costly real-time constraints as measured by congestion value.

Figure 44: Most Costly Real-Time Constraints



The constraint with the highest congestion value in 2017 was the Panhandle GTC at $139 million, a fivefold increase from 2016. By the end of 2017, there was almost 5 GW of generation capacity in the Panhandle area, of which 85% was wind generation. The highest GTC limit for the Panhandle was less than 4 GW, leading to frequent (16% of the intervals) and costly congestion when the wind output was high. A notable contributor to the low limit for the Panhandle GTC were outages on relatively new transmission facilities owned by ETT. These outages were required after the risk of structural damage to its transmission facilities was identified and required immediate inspection and possible repair. Outages of the facilities limited the export of the Panhandle wind generation. The average shadow price of the Panhandle GTC during binding intervals was $34 per MWh, reflecting the difference between system-wide average price and negative prices from wind generation. This, combined with the frequent need to control the Panhandle GTC, made the constraint the most costly.

The second most costly constraint in 2017 was the North to Houston constraint, comprised of a GTC and multiple thermal constraints, including the double circuit Singleton to Zenith 345 kV lines, the double circuit Jewett to Singleton 345 kV lines, and the Gibbons Creek to Singleton 345 kV lines. At $127 million, this constraint was twice as costly in 2017 as in 2016. Congestion declined sharply after June 2017 when Colorado Bend Combined Cycle Unit 3 (installed capacity of 1200 MW) came into service. Further, the considerable flooding caused by Hurricane Harvey forced load offline, also relieving congestion. Congestion in the fall months was due to outages along the North to Houston corridor, which were scheduled to facilitate the construction of the Houston Import Project.

Congestion in the Eagle Mountain area between Dallas and Fort-Worth was the next highest valued constraint. ERCOT's 2017 Regional Transmission Plan report[28] 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. Congestion in this area of the North zone is typically associated with high wind and high load conditions limiting flows from the west.

The fourth-highest congested element on this list, the double circuit Twin Oak Switch to Gibbons Creek 345 kV lines, was impacted by the North to Houston congestion. The largest impact occurred in October during construction of the north portion of the Houston Import Project, the Limestone to Gibbons Creek 345 kV lines. This constraint is noteworthy because of the dual impacts to the Gibbons Creek unit. Output from Gibbons Creek alleviates congestion on the Twin Oak Switch to Gibbons Creek 345 kV lines. However, the same generation has a negative effect on the previously described North to Houston constraints. At times, both the elements in North to Houston and a Twin Oak Switch to Gibbons Creek 345 kV line would be binding, producing opposite shift factor signals for the Gibbons Creek unit.

The Valley area constraints are located on the west side of the lower Rio Grande Valley and include the North McAllen to West McAllen 138 kV line ($51 million), the Azteca to South Edinburg 138 kV line ($14 million), and the North Edinburg 345/138 kV transformer ($8 million). These constraints were due to transmission upgrades and generation outages in the area.

[28] http://www.ercot.com/news/presentations/2017

The next five constraints were due to planned outages and/or high loads in the area. The Whitepoint to Rincon 138 kV line is located in Corpus Christi and feeds Port Aransas. This constraint was active in the spring of 2017 and was related to construction related outages. The Lewisville area, which is north of Dallas-Fort Worth, consists of the Lakepoint to Carrollton Northwest 138 kV line, the West TNP to TI TNP 138 kV line, and the Lewisville to Jones Street TNP 138 kV line. Congestion on the Blessing to Lolita to Formosa 138 kV line, located in Victoria, is mostly attributed to loads in the area and was further impacted by Hurricane Harvey damage. Congestion on the last two constraints listed above, the Elmcreek to Skyline 345 kV line and the Calaveras to Pawnee 345 kV line, was due to planned outages in San Antonio, primarily in December 2017.

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 infinity. Therefore, the shadow price is capped. Shadow price caps are based on a reviewed methodology,[29] and are intended to reflect the level of reduced reliability that occurs when a constraint is irresolvable. Currently (and throughout 2017) the shadow price caps are $5,000 per MW for base-case (non-contingency) or voltage violations, $4,500 per MW for 345 kV, $3,500 per MW for 138 kV, and $2,800 per MW for 69 kV thermal violations. GTCs are considered voltage constraints with a shadow price cap of $5,000 per MW.

When a constraint becomes irresolvable, chronically reaching the shadow price cap, ERCOT's dispatch software cannot find a dispatch combination to reduce the flows on the transmission element(s) of concern to a reliable operation level. A regional peaker net margin mechanism is used such that once local price increases accumulate to a predefined threshold because of an irresolvable constraint, the constraint's shadow price cap is re-evaluated. The shadow price is recalculated based upon the mitigated offer cap of existing resources with a defined shift factor threshold consistent with the methodology.

29 ERCOT Business Practice Manual, Setting the Shadow Price Caps and Power Balance Penalties in Security Constrained Economic Dispatch (ERCOT Board Approved 2/14/17), available at http://www.ercot.com/mktrules/obd/obdlist.

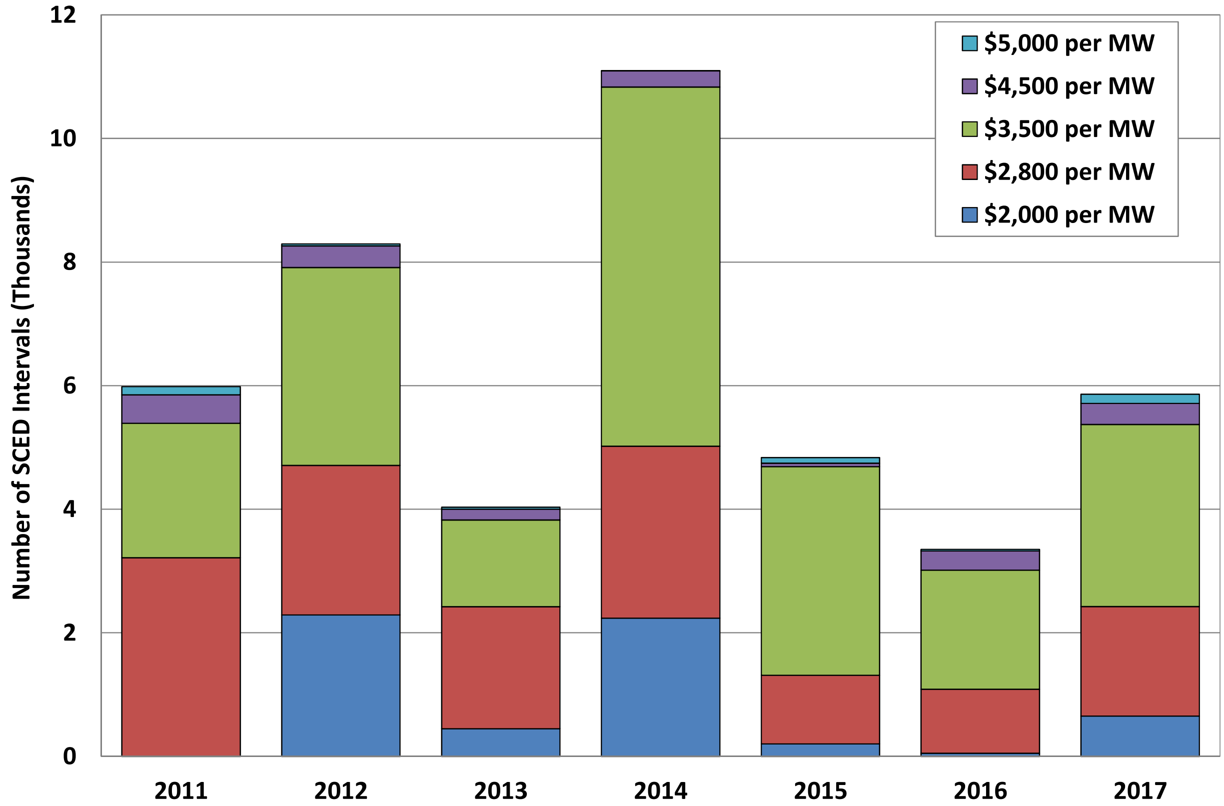
Table 8: Irresolvable Constrained Elements

Irresolvable Element	Original Max Shadow Price	2016 Adjusted Max Shadow Price	Effective Date	Termination Date	Load Zone
Valley Import	$5,000	$2,000	1/1/12	-	South
Abilene Northwest to Ely Rea Tap 69 kV Line	$2,800	$2,000	9/26/14	-	West
Harlingen to Oleander 69 kV Line	$2,800	$2,000	10/9/14	1/30/17	South
Rio Hondo to East Rio Hondo 138 kV Line	$3,500	$2,000	10/10/14	1/30/17	South
Emma to Holt Switch 69 kV Line	$2,800	$2,800	10/27/14	-	West
San Angelo College Hills 138/69 kV Autotransformer	$3,500	$2,000	7/22/15	-	West
Barilla to Fort Stockton Switch 138 kV Line	$3,500	$2,000	1/30/17	-	West

As shown above in Table 8, seven elements were deemed irresolvable in 2017 and had a shadow price cap imposed according to the irresolvable constraint methodology. The Barilla to Fort Stockton Switch constraint, located in far West Texas, was the only new irresolvable element in 2017. Two elements, the Harlingen to Oleander 69 kV line and the Rio Hondo to East Rio Hondo 138 kV line, were deemed resolvable during ERCOT's annual review and were removed from the list. All three irresolvable constraints located in the South Load Zone are located in the Valley. This list represent the smallest number of irresolvable elements since the irresolvable methodology was implemented in 2012.

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, SCED is not able 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 45 below shows the number of SCED intervals a constraint reached its maximum shadow price for the years 2011 to 2017.

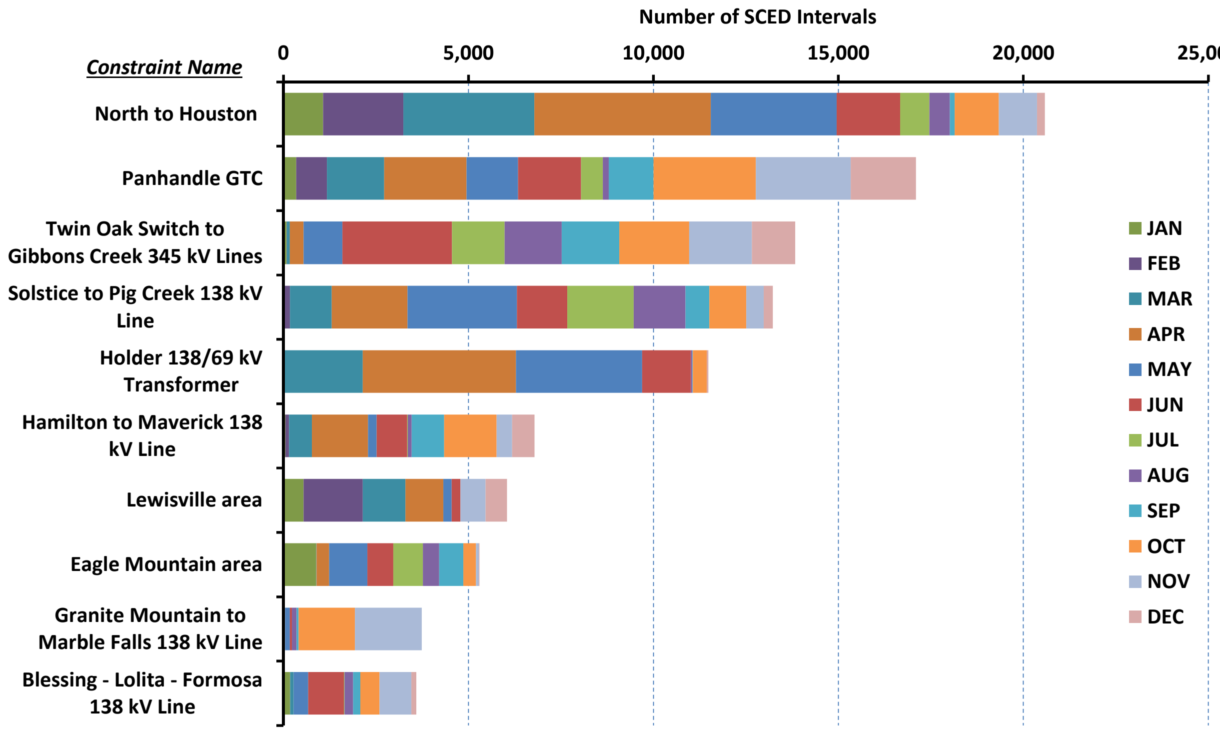
Figure 45: Frequency of Violated Constraints



Constraints were at maximum shadow prices more frequently in 2017 as compared to 2016, which was a historically low level. However, the number of constraint-intervals with violated constraints was once again a small fraction of all of the constraint-intervals. Just as in 2016, only 3% of the 2017 total constraint-intervals included violated constraints.

Figure 46 below presents a slightly different set of real-time congested areas, showing the areas that were most frequently constrained in 2017.

Figure 46: Most Frequent Real-Time Constraints



Six of the ten most frequently occurring constraints in 2017 have already been described as costly including North to Houston, Panhandle GTC, the Twin Oak Switch to Gibbons Creek 345 kV lines, Lewisville area, Eagle Mountain area, and the Blessing – Lolita – Formosa 138 kV line. Three of these constraints were also in the top ten most frequent constraints in 2016 but with a much greater frequency. From 2016 to 2017, the North to Houston constraint quadrupled in frequency, the Panhandle GTC tripled in frequency, and the Twin Oak to Gibbons Creek/Jack Creek 345 kV lines constraint doubled in frequency. The remaining constraints, although they occurred frequently, had moderate financial impacts. These high frequency constraints with minimal congestion costs occur when the generation to be re-dispatched is similarly priced.

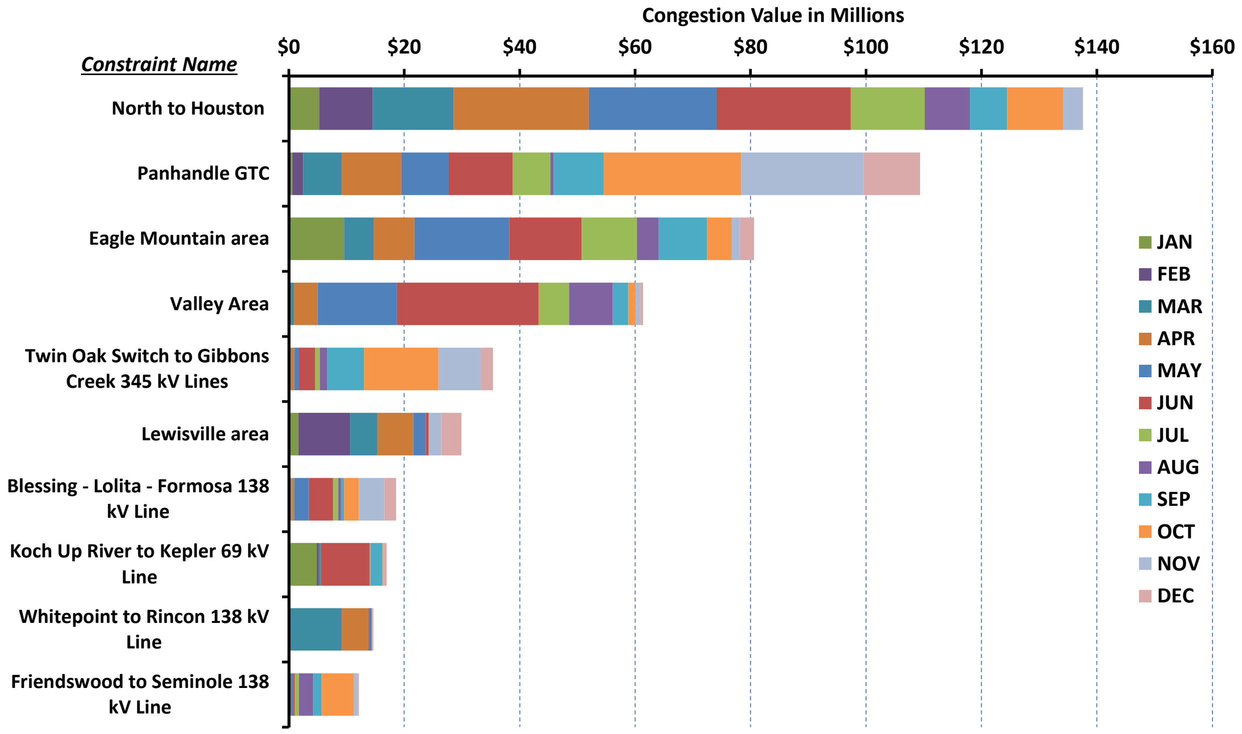
The fourth most frequent constraint in 2017 was the Solstice to Pig Creek 138 kV line located in the lower part of the Far West region where transmission is sparse. This particular area is unique because a generation resource of less than 10 MW contributes to the injection point of the constraint. Because of the nature of the modeling requirements at ERCOT, the resource is not modeled in SCED and does not receive base points. Additionally, there is not an economic incentive to alter the output to alleviate congestion as it is not calculated into the SCED dispatch. For constraints that are active, there could be an emphasis on the impact of generation outside of the SCED dispatch to be considered in their shift factor and impact on the constraint.

The next most frequent constraints in 2017 included the Holder 138/69 kV transformer located in the West zone near Comanche Peak. The congestion occurred in conjunction with planned outages in the area. The Hamilton to Maverick 138 kV line is located in the South zone and is affected by high wind output. And lastly, the Granite Mountain to Marble Falls 138 kV line is in Central Texas also tied to planned outages in the area.

C. Day-Ahead Constraints

This subsection provides a review of the transmission constraints from the day-ahead market. Figure 47 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 47: Most Costly Day-Ahead Congested Areas



Since the start of the nodal market, the day-ahead constraint list has contained many constraints that were unlikely to occur in real-time. Interestingly, 2017 was the second year that the majority of the most costly day-ahead constraints were also costly real-time constraints. A contributing factor to this convergence was that ERCOT continually hones the constraint list to monitor which constraints should be included in the day-ahead market analysis to be consistent with market activities observed in real-time.

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.

Located in Corpus Christi, the Koch Up River to Kepler 69 kV line was the eighth most costly day-ahead constraint. The Friendswood to Seminole 138 kV line is located in south Houston and was the tenth most costly day-ahead constraint.

The day-ahead market was impacted by the effects of Hurricane Harvey. The load distribution factors used by the day-ahead market to effectively spread out activity transacted at the Load Zone level to individual locations within the Load Zone are typically based on historical data. With transmission equipment damaged, historical load distribution factors were not a good representation of the system in the immediate aftermath of the hurricane. Although there were large discrepancies between the day-ahead and real-time markets immediately after the hurricane, these were rectified very quickly.

Figure 48: Day-Ahead Congestion Costs by Zone

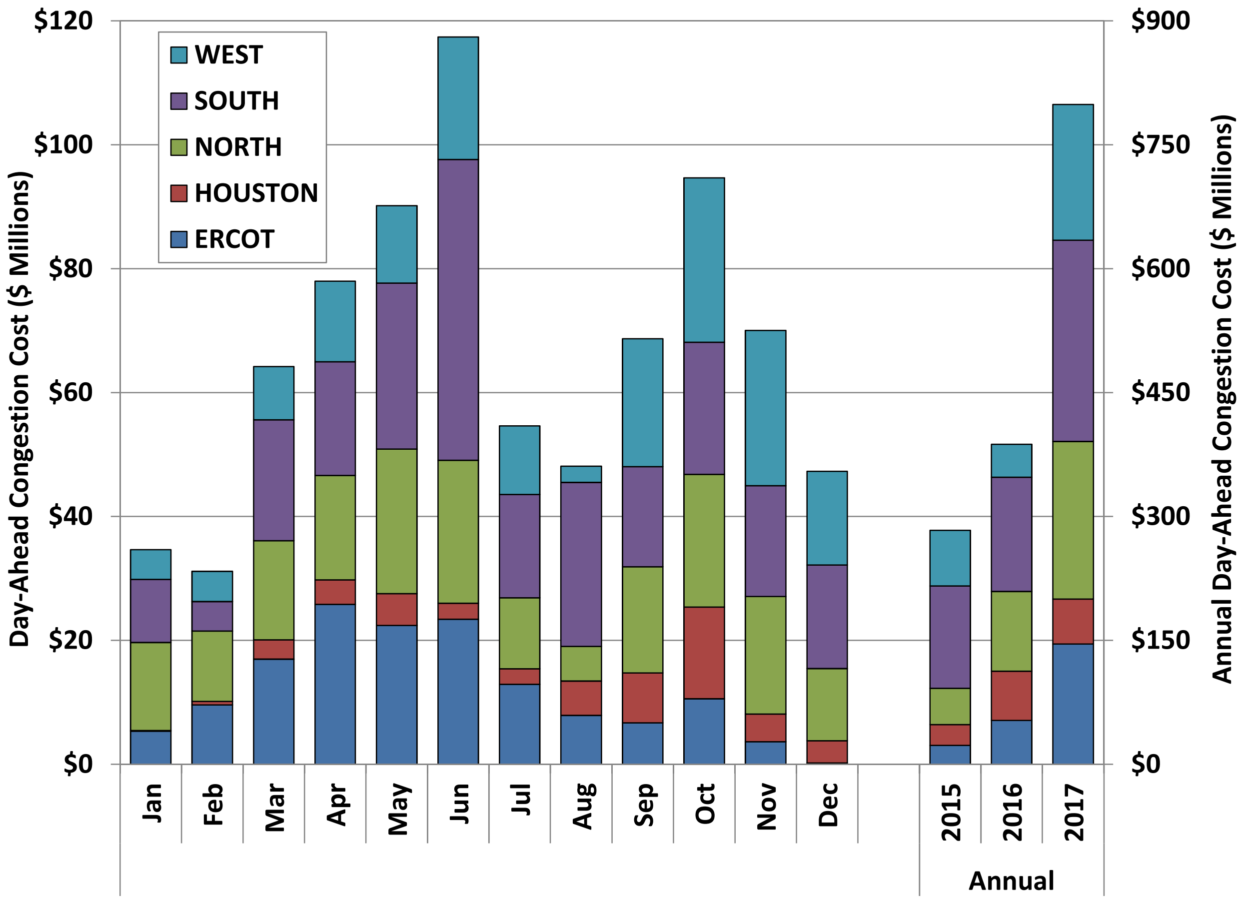


Figure 48 above presents day-ahead congestion costs by zone. Similar to real-time market outcomes, day-ahead congestion in all zones except the Houston zone was higher in 2017 than

2016. The total day-ahead congestion costs in 2017 were also almost twice as much as in 2016. The majority of the ERCOT congestion was due to North to Houston congestion caused by outages associated with the Houston Import Project. North to Houston congestion is expected to decrease in 2018 with the final implementation of the transmission upgrades. The shoulder months showed higher activity for the day-ahead congestion costs as well as in the real-time congestion values.

D. Congestion Revenue Rights Market

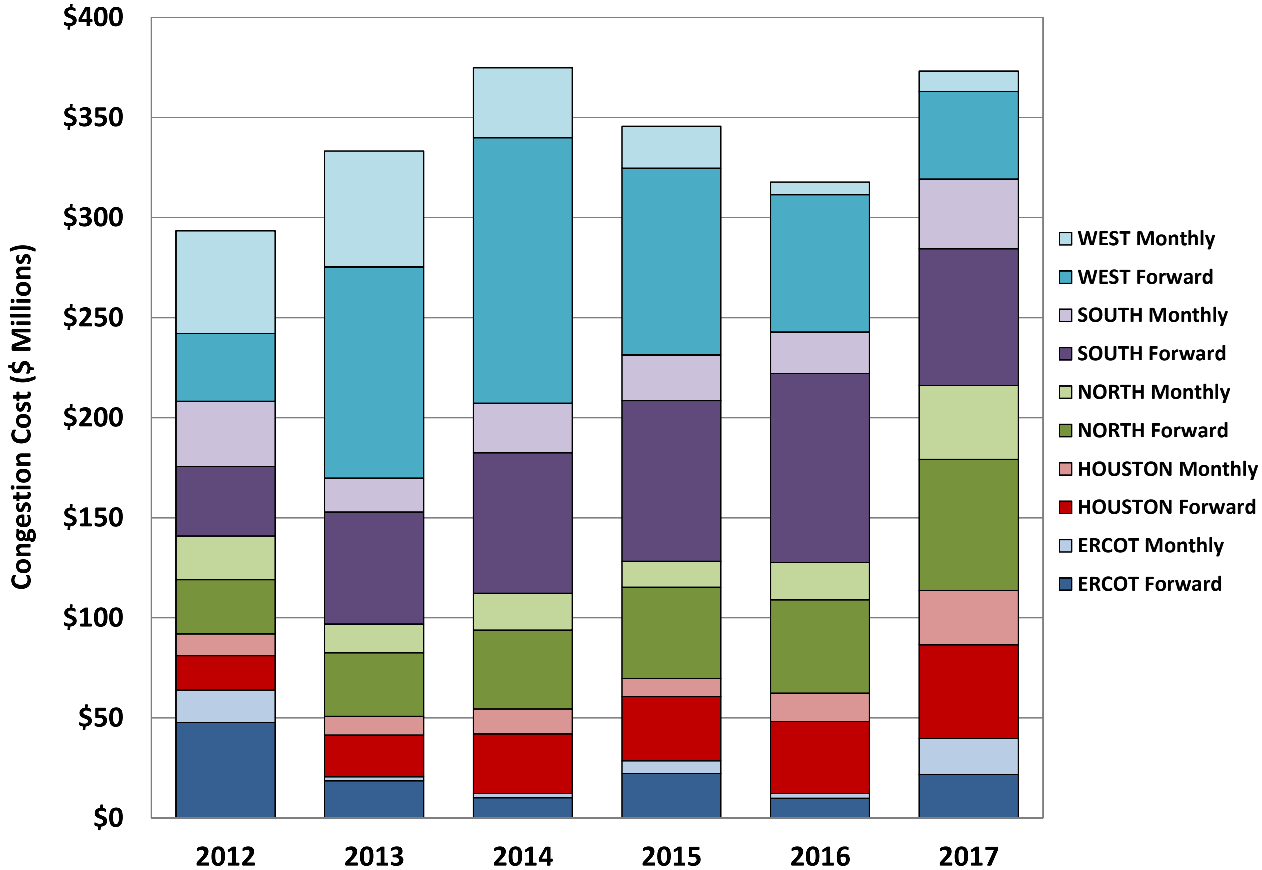
Congestion can be significant from an economic perspective, compelling the dispatch of higher-cost resources because power produced by lower-cost resources cannot be delivered because of transmission constraints. This causes different clearing prices for energy at different locations. Under the nodal market design, one means by which ERCOT market participants can hedge these price differences is by acquiring Congestion Revenue Rights (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 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 49 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).

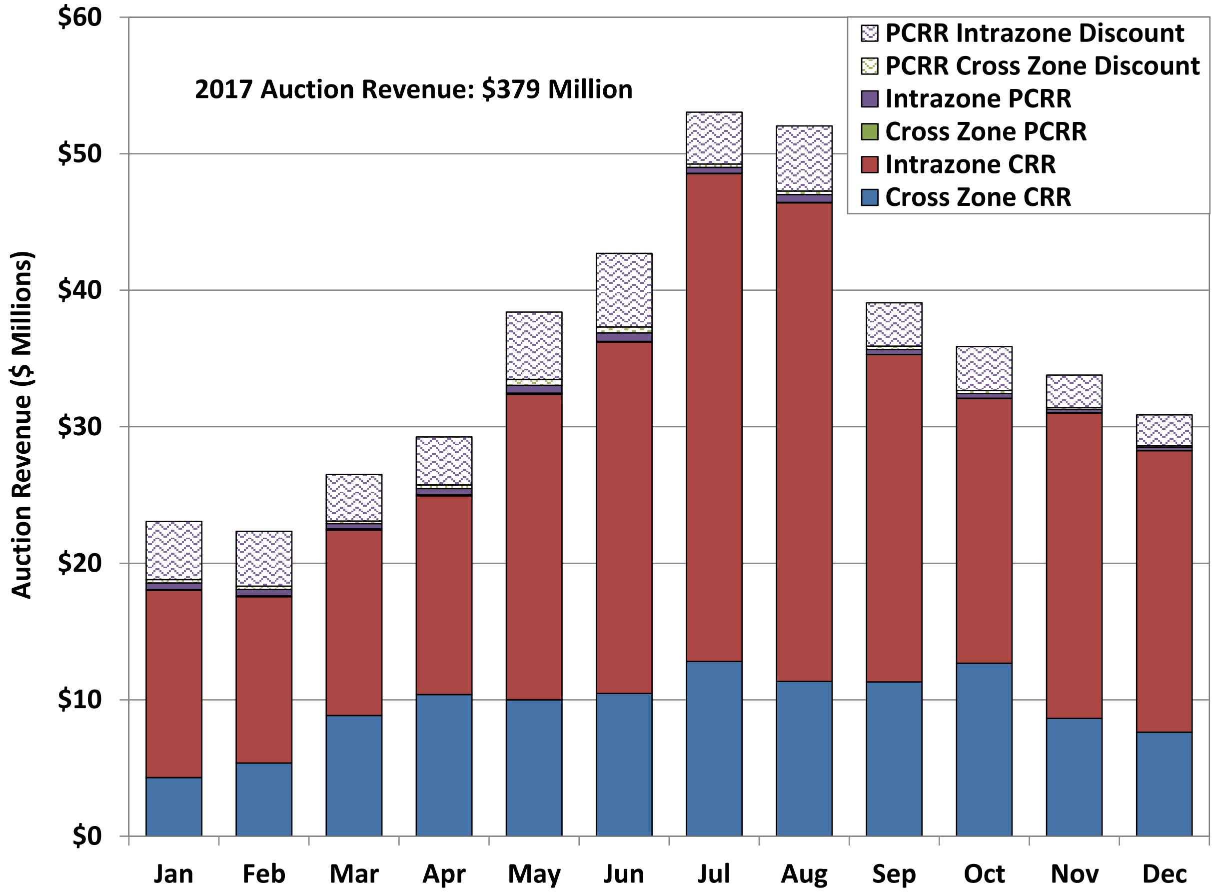
Figure 49: CRR Costs by Zone



Comparing the costs paid to acquire CRRs, shown in Figure 49 to the trends of congestion costs seen in the real-time and the day-ahead markets, indicates that the CRR market was a poor predictor of the increase of both real-time and day-ahead congestion. All zones, except South and West procured in the forward auctions, show increases in CRR congestion compared to very large increases in day-ahead and real-time congestion. CRR congestion costs in the South and West forward auctions decreased in 2017. The CRR costs for 2017 nearly equals the previous peak, seen in 2014.

Figure 50 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 50: 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, but it does so whether the congestion had raised prices in the area or lowered prices in the area. As a case in point, congestion lowered prices in the West zone to below the ERCOT average, as shown above in Figure 4: Effective Real-Time Energy Market Prices. However, because so many CRRs were purchased in the West zone to capture the value of this price lowering congestion, a higher than load-ratio share portion of the CRR revenue gets distributed to Qualified Scheduling Entities representing West zone load, thus further lowering the effective price paid by load in the West zone.

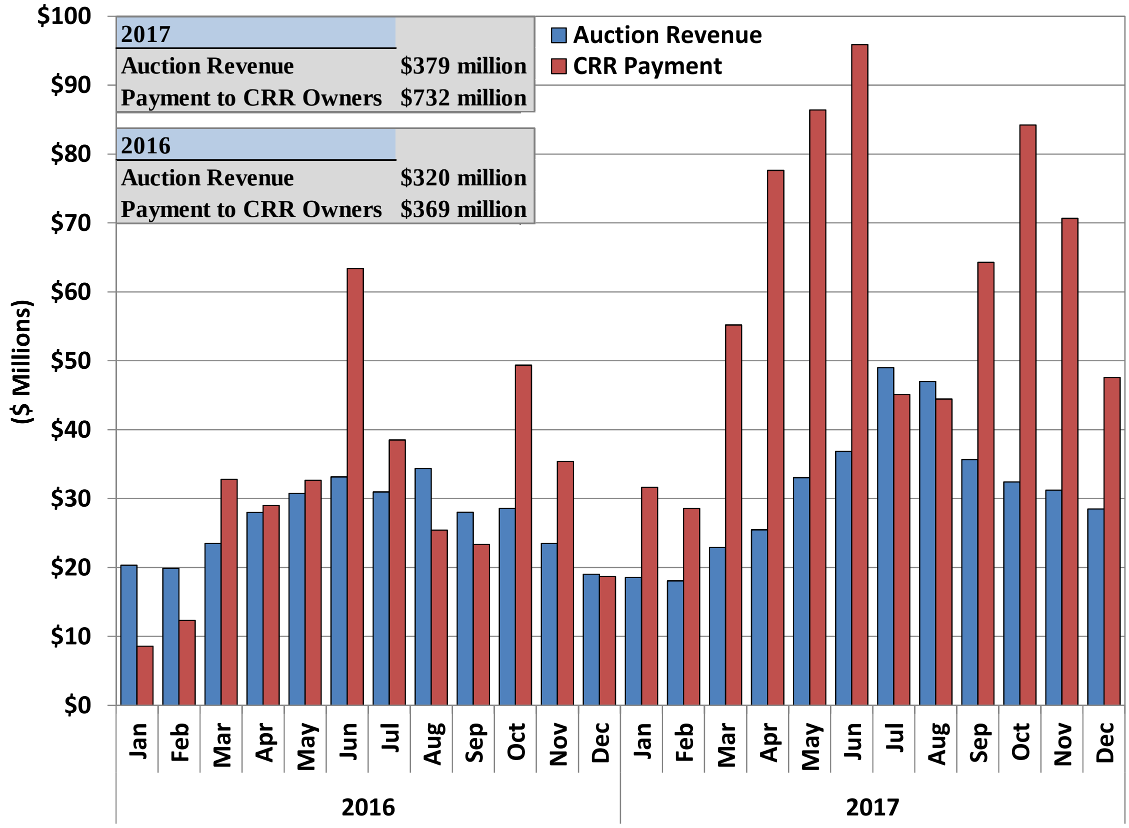
As previously mentioned in this section, the only parties eligible to receive PCRRs are Non-Opt-In Entities (NOIEs). They are charged only a fraction of the PCRR auction value. The difference between the auction value and the value charged to the purchaser is shown in Figure 50 as the PCRR Discount. Even as the total amount of CRR auction revenue increased to

$397 million in 2017 from $320 million in 2016, the total PCRR discount decreased from $70 million in 2016 to $50 million in 2017, similar to the PCRR discount in 2015.

CRR Profitability

Next, Figure 51 compares the value received by CRR owners (in aggregate) to the price paid to acquire the CRRs.

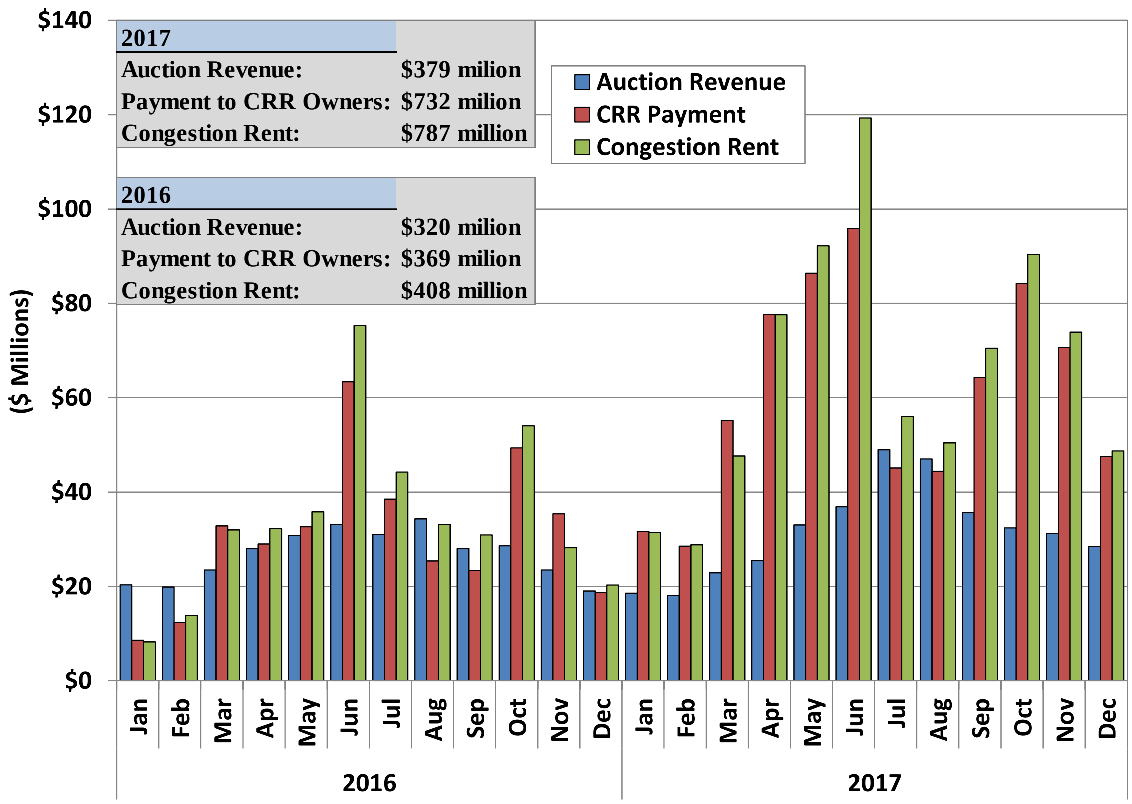
Figure 51: CRR Auction Revenue and Payment Received



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 2017 than they received in payment from the day-ahead market. For the entire year of 2017, participants spent $379 million to procure CRRs and received almost twice as much at $732 million. In general, this difference occurred because the substantial increase in congestion that occurred in 2017 was not foreseen by the market. There were two significant periods of congestion that account for this difference: March through June and September through December. In both cases, transmission outages related to construction of new facilities contributed to the substantial unforeseen increases in congestion.

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, PTP obligation bids, and PTP obligation bids linked to options in day-ahead market.[30] Day-ahead congestion rent creates the source of funds used to make payments to CRR owners. Figure 52 presents CRR auction revenues, payment to CRR owners, and congestion rent in 2016 and 2017, by month. Congestion rent for the year 2017 totaled $787 million and payment to CRR owners was $732 million.

Figure 52: CRR Auction Revenue, Payments and Congestion Rent



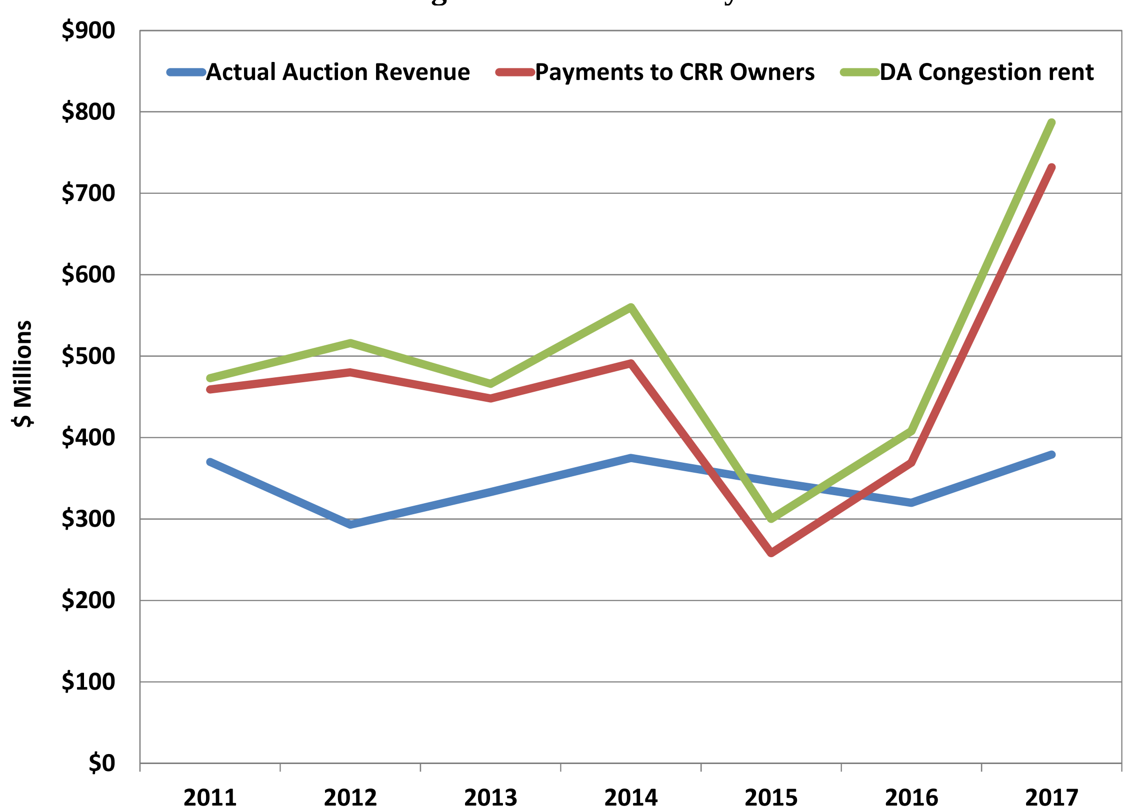
It is worth noting that because the CRR network model uses line ratings that are 90% of the expected lowest line ratings for the month, it is expected that CRRs would be somewhat undersold and that day-ahead congestion rent would be higher than the payment to CRR owners. This indeed was the case in 2017, where payments to CRR owners was 93% of day-ahead congestion rent. In 2016, this ratio was 90%.

30 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.

Figure 53 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 2017, owners of CRRs in aggregate made a substantial profit on their CRR holdings. Payments to CRR owners in 2017 were almost double the total cost paid to acquire the CRRs. As we discuss 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 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. The costs paid to acquire CRRs varied in a narrow range between $300 and $400 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 53: 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 make whole CRR account holders that received shortfall charges.

Figure 54 shows the CRR balancing fund since the beginning of 2015. Even though the amount of the fund was under $10 million in five months of 2015 and two months of 2016, it started 2017 at its capped value of $10 million and was not drawn upon during the year. While there were monthly shortfalls in day-ahead market settlement in 2015 and 2016, a surplus occurred for each month in 2017, and the total day-ahead surplus was $94.45 million. In comparison, the total annual day-ahead market surplus was only $30.85 million and $34.59 million in 2015 and 2016 respectively. Because there was enough day–ahead market surplus after paying out to the CRR owners for each month in 2017, those CRR owners who received a shortfall charge, at the total annual amount of $12.11 million, were fully refunded at the end of each month. From the perspective of the load, the monthly CRR balancing account allocation to load was always positive in 2017 and resulted in a total amount of $90.10 million at the end of the year, which almost offset the real-time revenue neutrality charge to load at the amount of $96.32 million.

Figure 54: CRR Balancing Fund

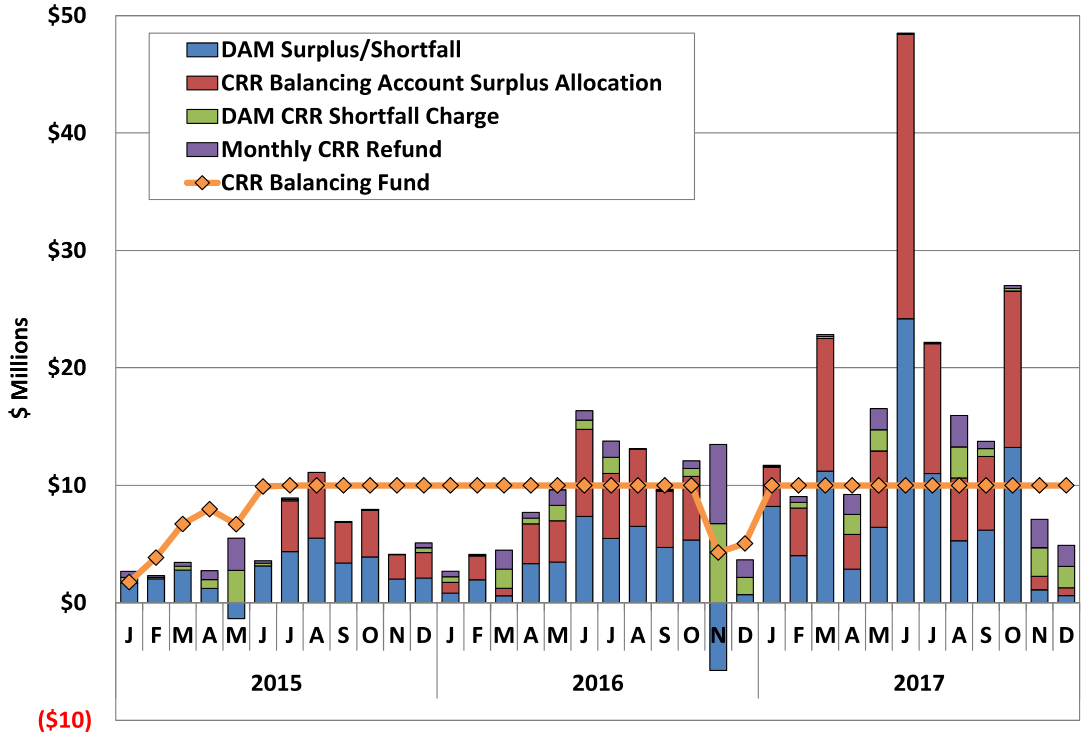
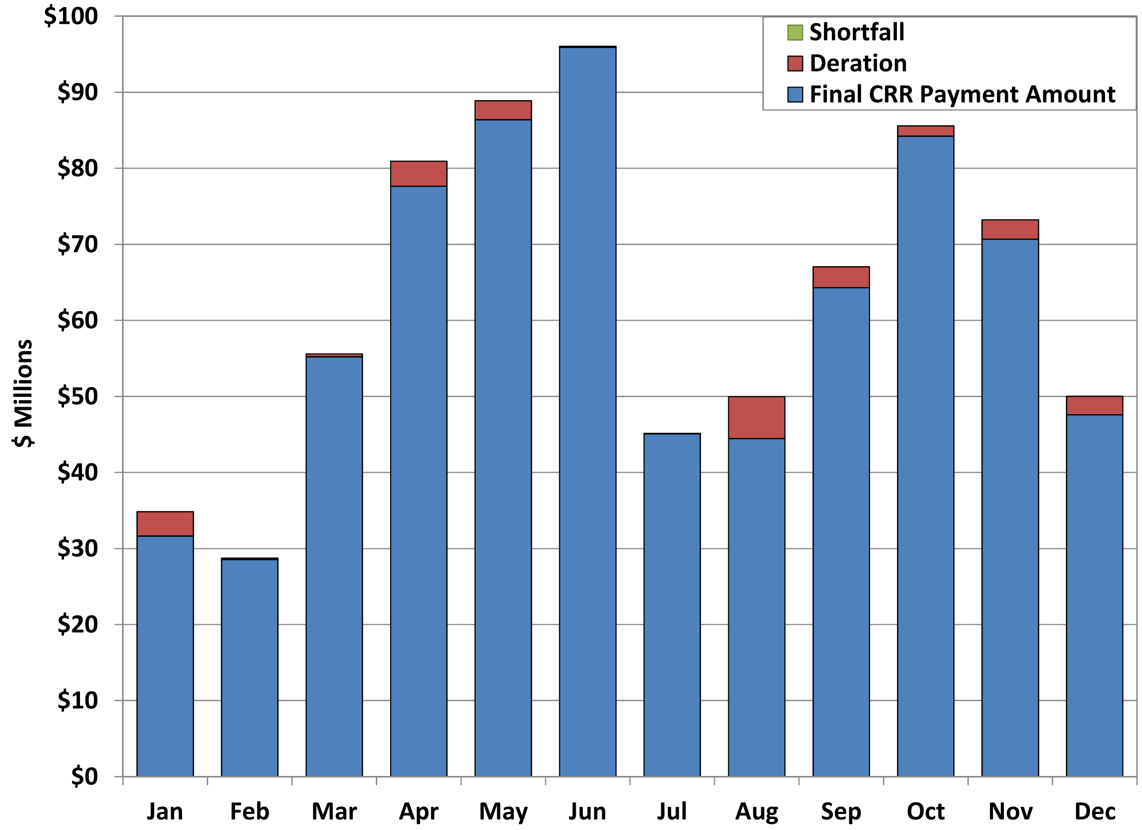


Figure 55 shows the amount of target payment, deration amount, and net shortfall charges (after make whole payments) for 2017. In 2017, the total target payment to CRRs was $756 million; however, there were $24 million of derations and no shortfall charges resulting in a final payment to CRR account holders of $732 million. This final payment amount corresponds to a CRR funding percentage of 97%.

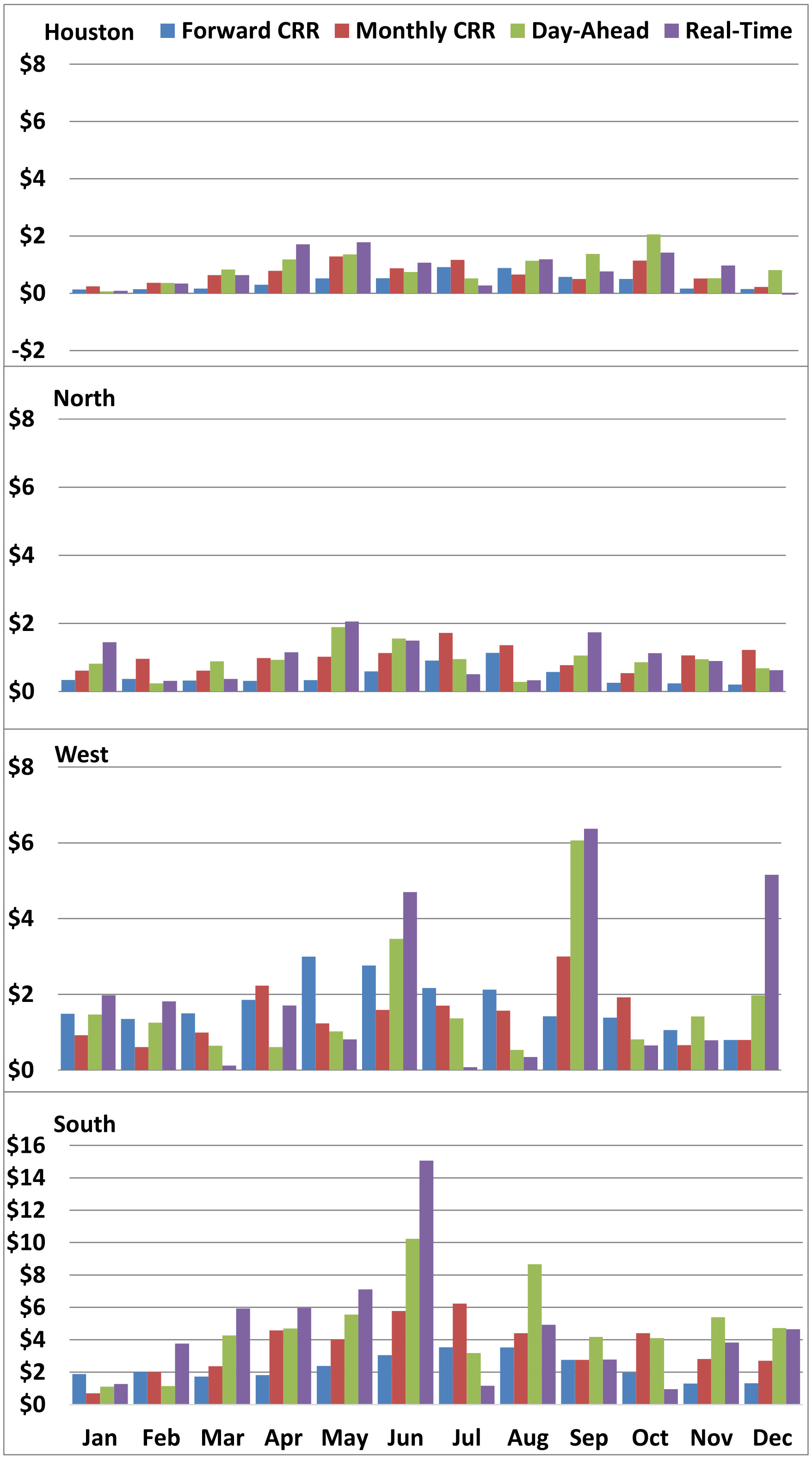
Figure 55: CRR Shortfalls and Derations



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 56 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 relatively poor convergence between the forward CRR price spreads for the West Load Zone and the actual price spreads. This may have been because of the difficulty forecasting the price impacts of variable wind output, or the added uncertainty of whether or not outages associated with ETT's structural maintenance are viable in such wind conditions. The South Load Zone still had the highest hub to zone price spread for the second year in a row, having overtaken the West Load Zone in 2016, likely because of the effects of congestion in the Valley area.

Figure 56: Hub to Load Zone Price Spreads

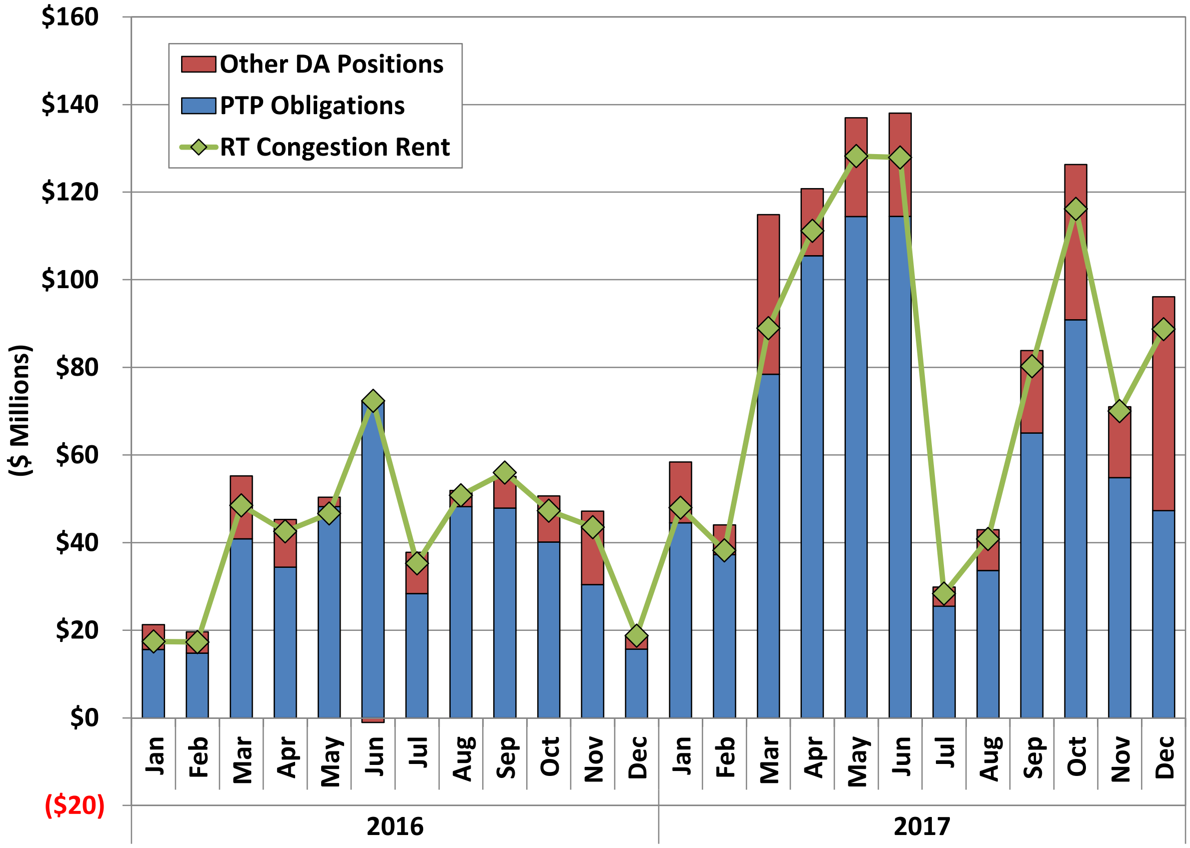


E. Revenue Sufficiency

In Figure 57, the combined payments to Point-to-Point (PTP) obligation owners and effective payments to other day-ahead positions are compared to the total real-time congestion rent. For 2017, real-time congestion rent was $967 million, payments for PTP obligations (including those with links to CRR options) were $812 million and payments for other day-ahead positions were $251 million, resulting in a shortfall of approximately $96 million for the year.

By comparison, the real-time congestion rent was $497 million in 2016. Payments for PTP obligations and real-time CRRs were $437 million and payments for other day-ahead positions were $88 million, resulting in a shortfall of approximately $28 million for the year. This shortfall is paid for by charges to load.

Figure 57: Real-Time Congestion Rent and Payments



IV. Demand and Supply

This section reviews and analyzes the load patterns during 2017 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 2017

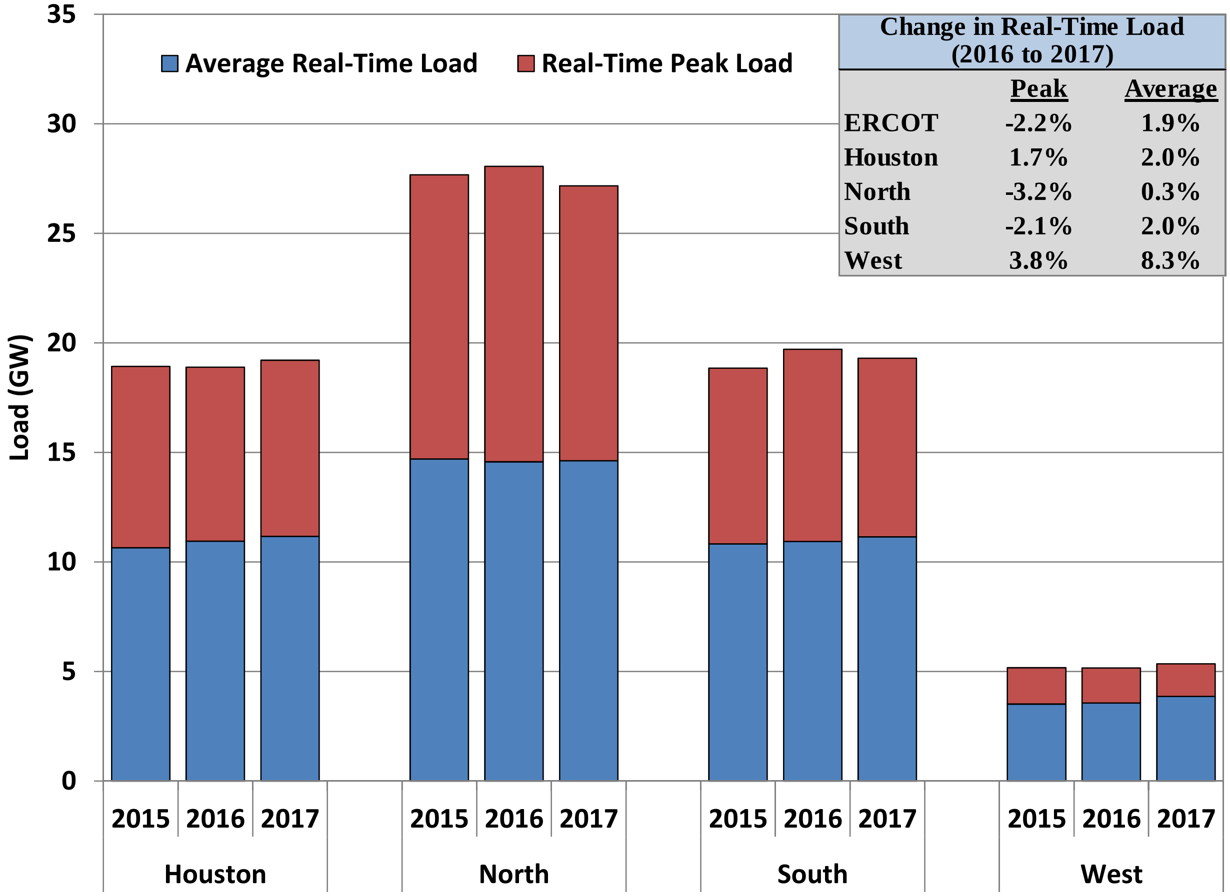
The changes in overall load levels from year to year can be shown by tracking the changes in average load levels. This metric tends to capture changes in load over a large portion of the hours during the year. Separately evaluating the changes in the load during the highest-demand hours of the year is 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 2017 are examined in this subsection and summarized in Figure 58.

This figure shows peak load and average load in each of the ERCOT geographic zones from 2015 to 2017.[31] 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 58 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.

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

Figure 58: Annual Load Statistics by Zone

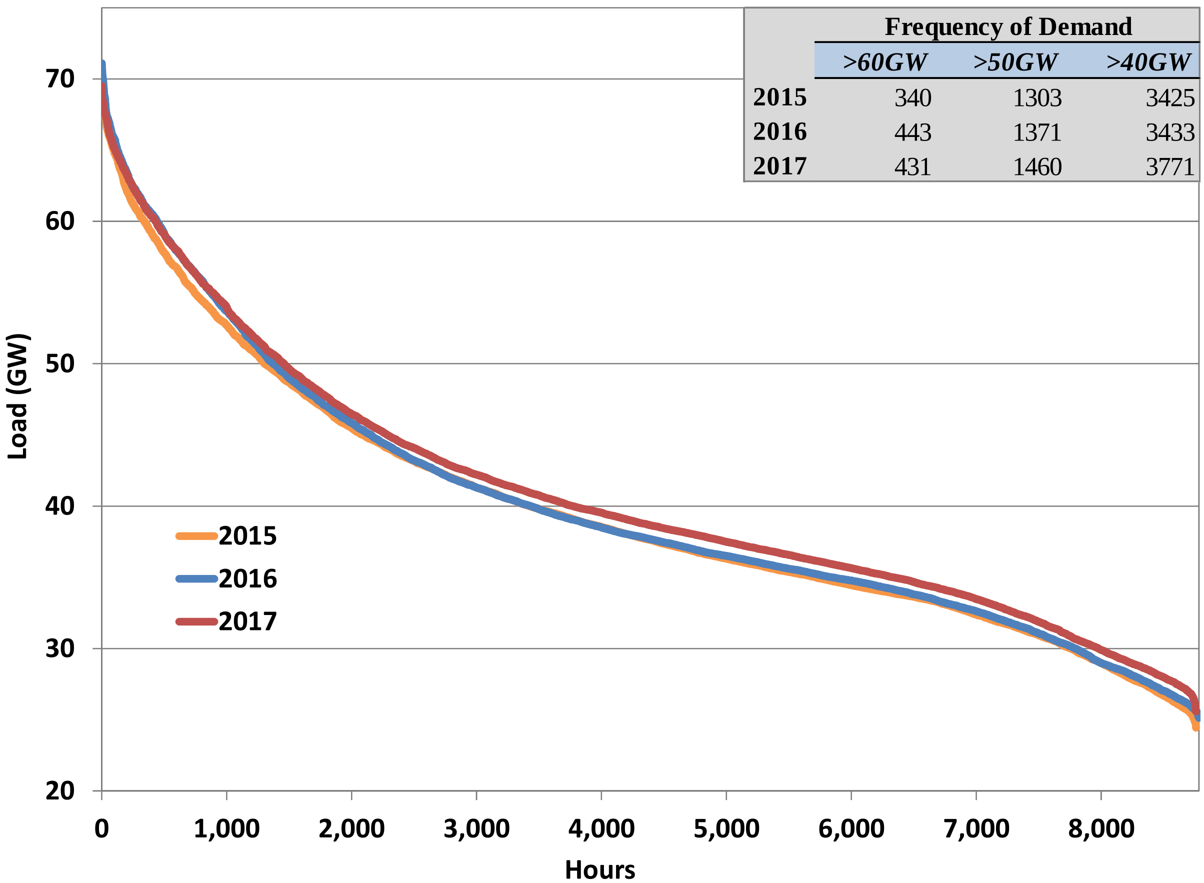


Total ERCOT load in 2017 increased 1.9% (approximately 780 MW per hour on average) to total 357.4 TWh in 2017. All zones showed an increase in average real-time load in 2017. The West zone saw the largest average load increase at 8.3%, which was likely due to continuing robust oil and natural gas production activity. Weather impacts on load in 2017 were mixed. Cooling degree days, a metric that is highly correlated with weather-related summer load, exhibited no change in Houston, decreased in Dallas and increased in Austin as compared to 2016.

Summer conditions in 2017 produced a peak load of 69,512 MW on July 28, 2017, short of the ERCOT-wide coincident peak hourly demand record of 71,110 MW set on August 11, 2016. Further, demand did not ever exceed 70,000 MW in 2017, compared to five separate hours in 2016. The zones experienced varying changes in peak load. The West zone continued to experience the highest percentage growth in peak load, which was likely driven by continuing growth in oil and natural gas production.

To provide a more detailed analysis of load at the hourly level, Figure 59 compares load duration curves for each year from 2015 to 2017. 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 2017 is very similar to 2016 and 2015.

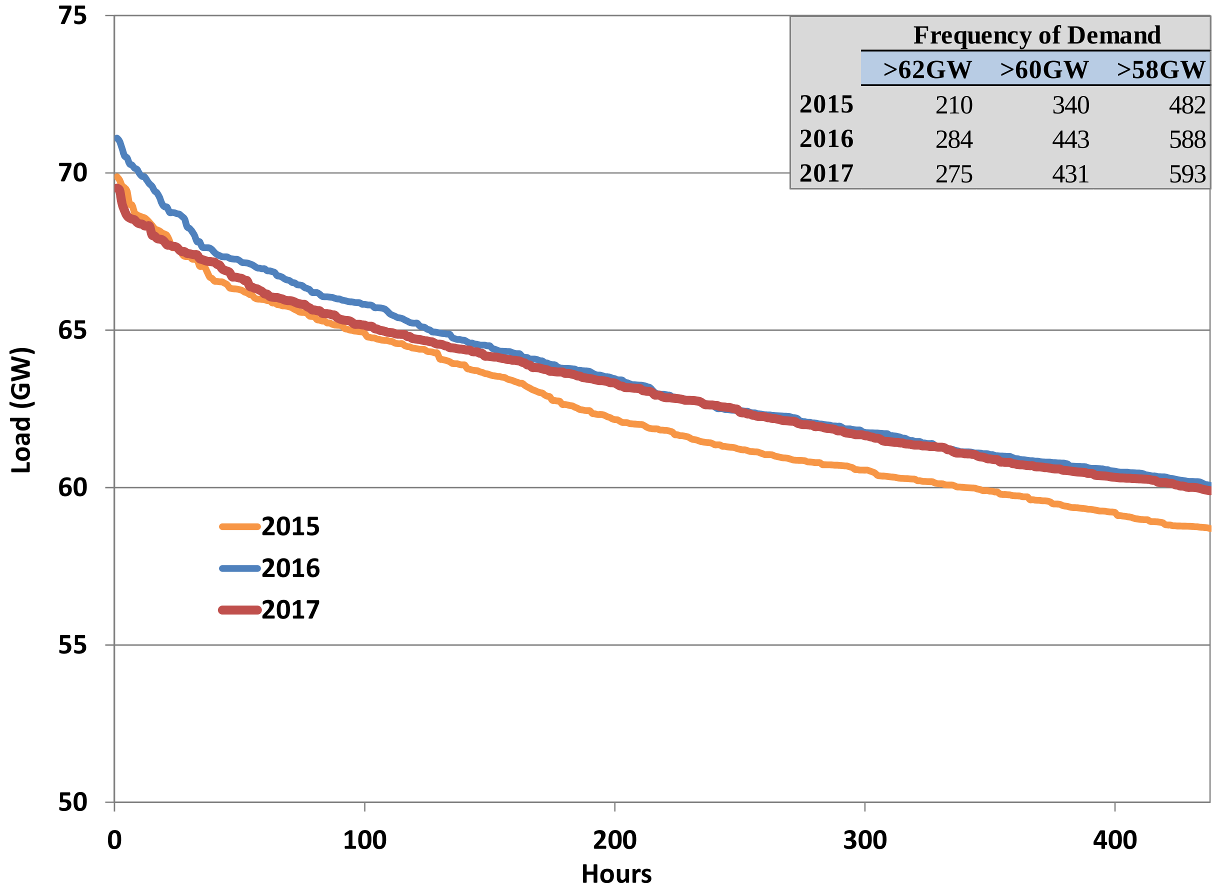
Figure 59: Load Duration Curve – All Hours



	Frequency of Demand		
	>60GW	*>50GW*	*>40GW*
2015	340	1303	3425
2016	443	1371	3433
2017	431	1460	3771

To better illustrate the differences in the highest-demand periods between years, Figure 60 below shows the load duration curve for the five percent of hours with the highest loads. This figure also shows that the peak load in each year is significantly greater than the load at the 95th percentile of hourly load. From 2011 to 2017, the peak load averaged 16% to 18% 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 five percent of the hours.

Figure 60: Load Duration Curve – Top Five Percent of Hours with Highest Load

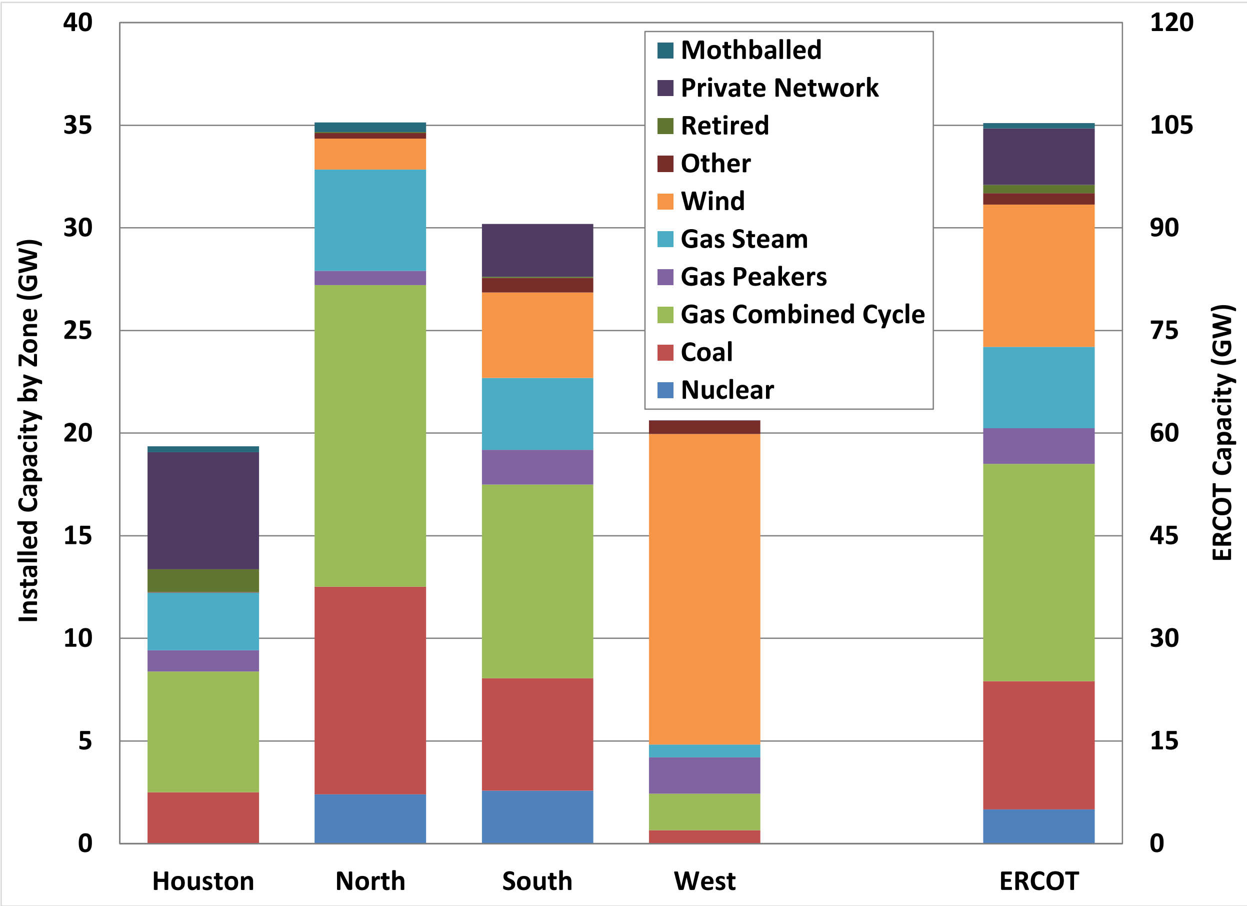


B. Generation Capacity in ERCOT

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 2017, the North zone accounted for approximately 33% of capacity, the South zone 29%, the Houston zone 18%, and the West zone 20%. Excluding mothballed resources and including only the fraction of wind capacity available to reliably meet peak demand,[32] the North zone accounted for approximately 38% of capacity, the South zone 33%, the Houston zone 20%, and the West zone 9% in 2017. Figure 61 shows the installed generating capacity by type in each zone.

32 The percentages of installed capacity to serve peak demand assume wind availability of 14% for non-coastal wind and 59% for coastal wind.

Figure 61: Installed Capacity by Technology for Each Zone



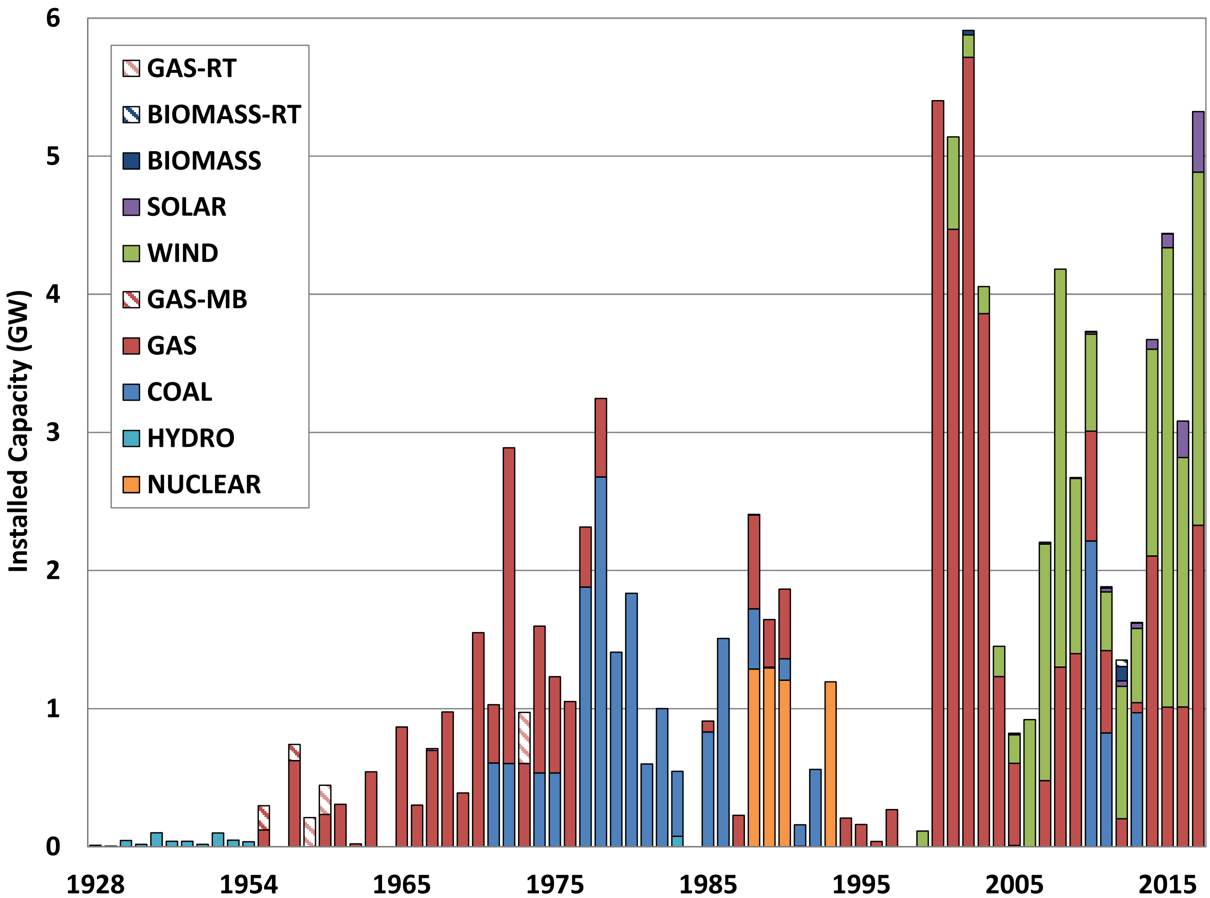
Approximately 3.6 GW of new generation resources came online in 2017; the bulk of which was two new combined cycle natural gas units with total capacity of 2.2 GW. Wind additions totaled 1.1 GW with an effective peak serving capacity of less than 300 MW. The remaining capacity additions were 180 MW of new combustion turbines and 160 MW of solar.

Fourteen generation resources totaling 1,222 MW, consisting primarily of aging natural gas generation, were retired in 2017. Five natural gas units at Calpine's Clear Lake location, totaling 280 MW, were decommissioned and retired on February 1, 2017. Aspen LLC's 45 MW LFBIO_UNIT1 biomass unit was decommissioned and retired as of February 6, 2017. South Texas Electric Cooperative, Inc.'s Pearsall Units 1, 2, and 3, totaling 61 MW of natural gas generation, were decommissioned and retired on August 1, 2017. Union Carbide Corp.'s 30 MW UCC_COGN_UCC_C1 natural gas unit was retired on September 29, 2017. NRG Energy Inc.'s previously mothballed S.R. Bertron natural gas units, totaling 435 MW, were permanently retired and decommissioned on December 31, 2017, as was the 371 MW Greens Bayou 5 natural gas unit, which had previously been deemed necessary for RMR services.

Given these additions and retirements, shares of natural gas and coal capacity did not change significantly in 2017, representing 46% and 18% of installed capacity, respectively.

Figure 62 shows the age of generation resources in ERCOT that were operational in the December 2017 Capacity, Demand, and Reserves Report.[33] The bulk of the coal fleet in ERCOT was built before 1990 and is approaching the end of useful life for this vintage of coal power plants. There was quite a large investment in combined cycle natural gas units in conjunction with deregulation of the ERCOT market. The amount of new combined cycle capacity installed in 2017 was greater than in any year since 2003. A few new coal units were added around 2010. However, wind capacity has been the dominant technology for newly installed capacity since 2006.

Figure 62: Vintage of ERCOT Installed Capacity

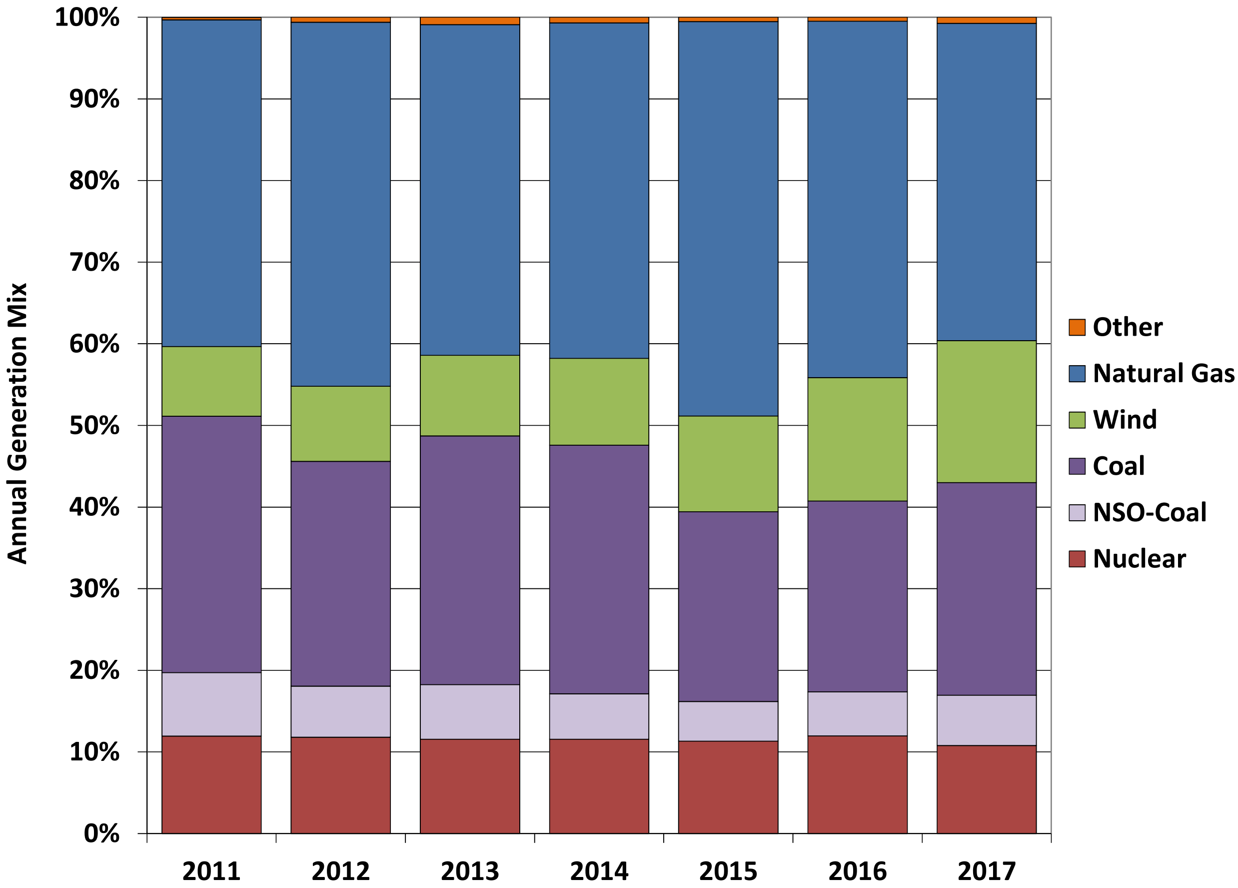


The shifting contribution of coal and wind generation is evident in Figure 63, which shows the percentage of annual generation from each fuel type for the years 2011 through 2017.

---

33 ERCOT Capacity, Demand, and Reserves Report (Dec. 2017), available at http://www.ercot.com/gridinfo/resource.

Figure 63: Annual Generation Mix

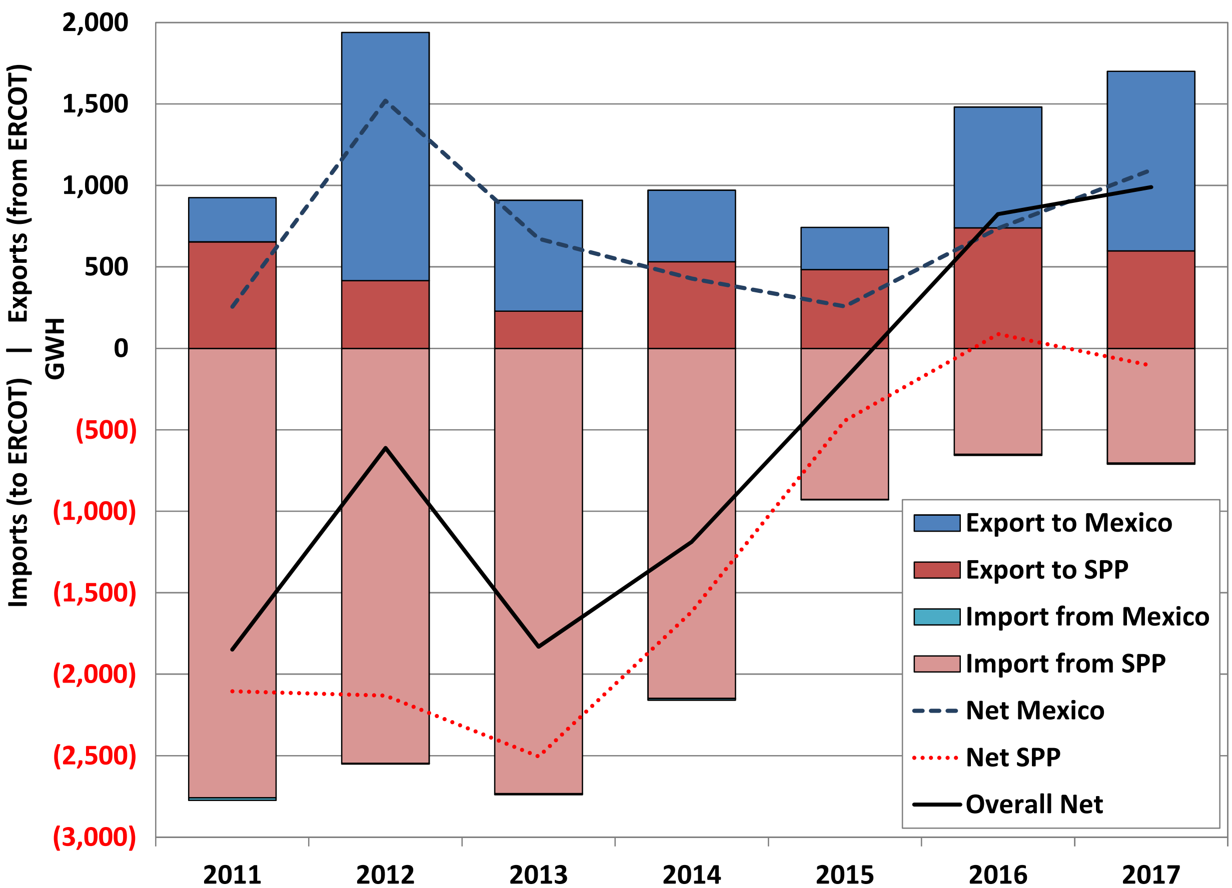


The generation share from wind has increased every year, reaching 17% of the annual generation requirement in 2017, up from 9% in 2011 and 15% in 2016. While the share of generation from coal had declined significantly between 2014 and 2015, its share has increased the last two years, up to 32% in 2017. This figure separately shows the amount of energy provided from coal units that are scheduled to be retired in 2018 (i.e., those that have submitted a Notification of Suspension of Operations or NSO). These seven units have provided an average of 6% of the total annual generation requirements over the past 7 years. Natural gas declined from its high point of 48% in 2015 down to 39% in 2017. This trend should reverse, however, once the coal resources mentioned above retire.

While coal/lignite and nuclear plants operate primarily as base load units in ERCOT, it is the reliance on natural gas resources that drives the high correlation between real-time energy prices and the price of natural gas fuel. There were approximately 24 GW of coal and nuclear generation in ERCOT in 2017. 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 64 below shows the total energy transacted across the ties for each of the past several years.

Figure 64: Energy Transacted Across DC Ties in August



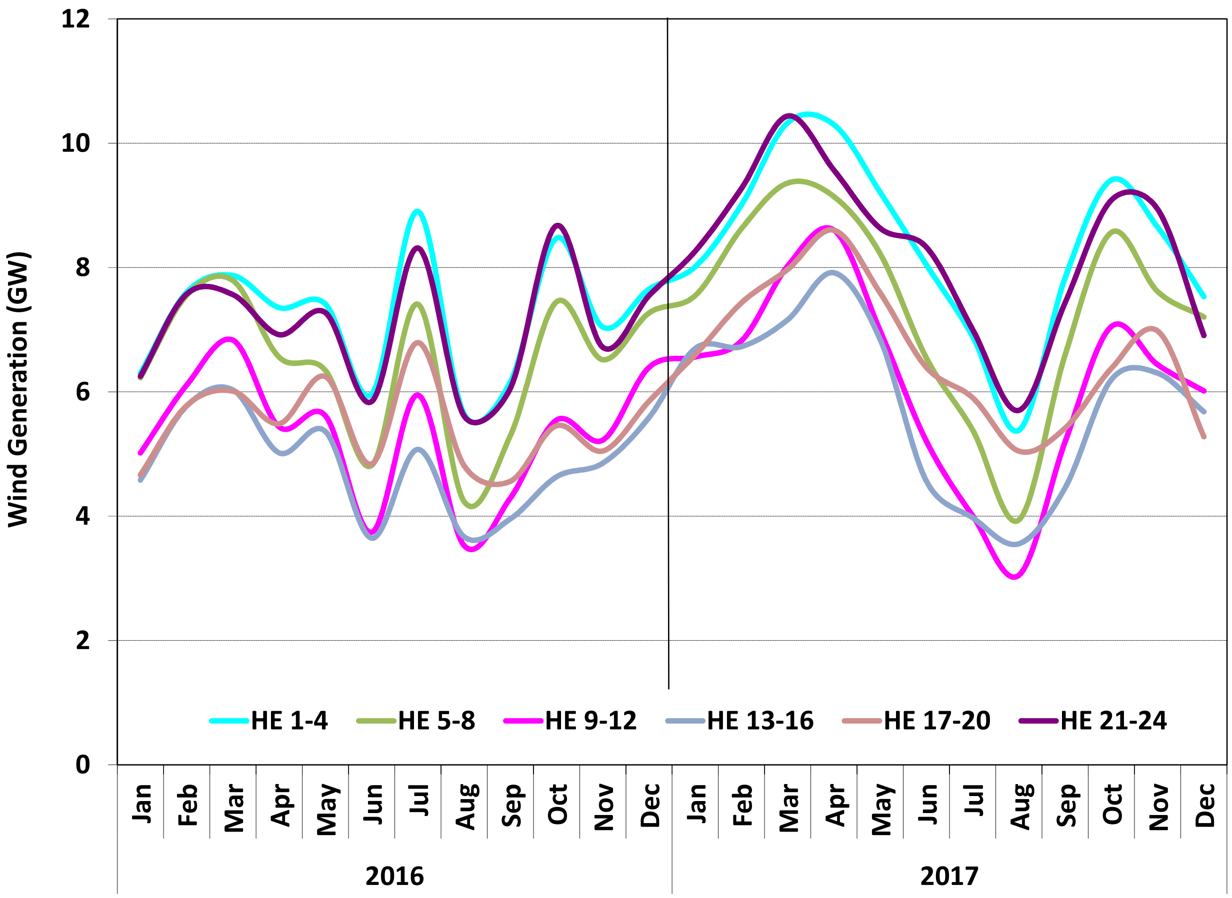
Between 2011 and 2014, ERCOT imported far more energy into its market than it exported into Mexico and SPP combined. In 2011, ERCOT was a net importer by 1,848 GWh, largely because of the high loads and tight conditions in ERCOT. Increased exports to Mexico led to decreased net imports in 2012, but return to previous levels in 2013. Since then there has 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 potential for higher prices in 2018, it is likely that this trend will reverse.

C. Wind Output in ERCOT

The amount of wind generation installed in ERCOT was approximately 21.5 GW by the end of 2017. 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 2017, there were 55 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 65 shows average wind production for each month in 2016 and 2017, 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 in excess of 5 GW. This may be a small fraction of the total installed capacity but is now a non-trivial portion of generation supply, even at its lowest outputs.

Figure 65: Average Wind Production

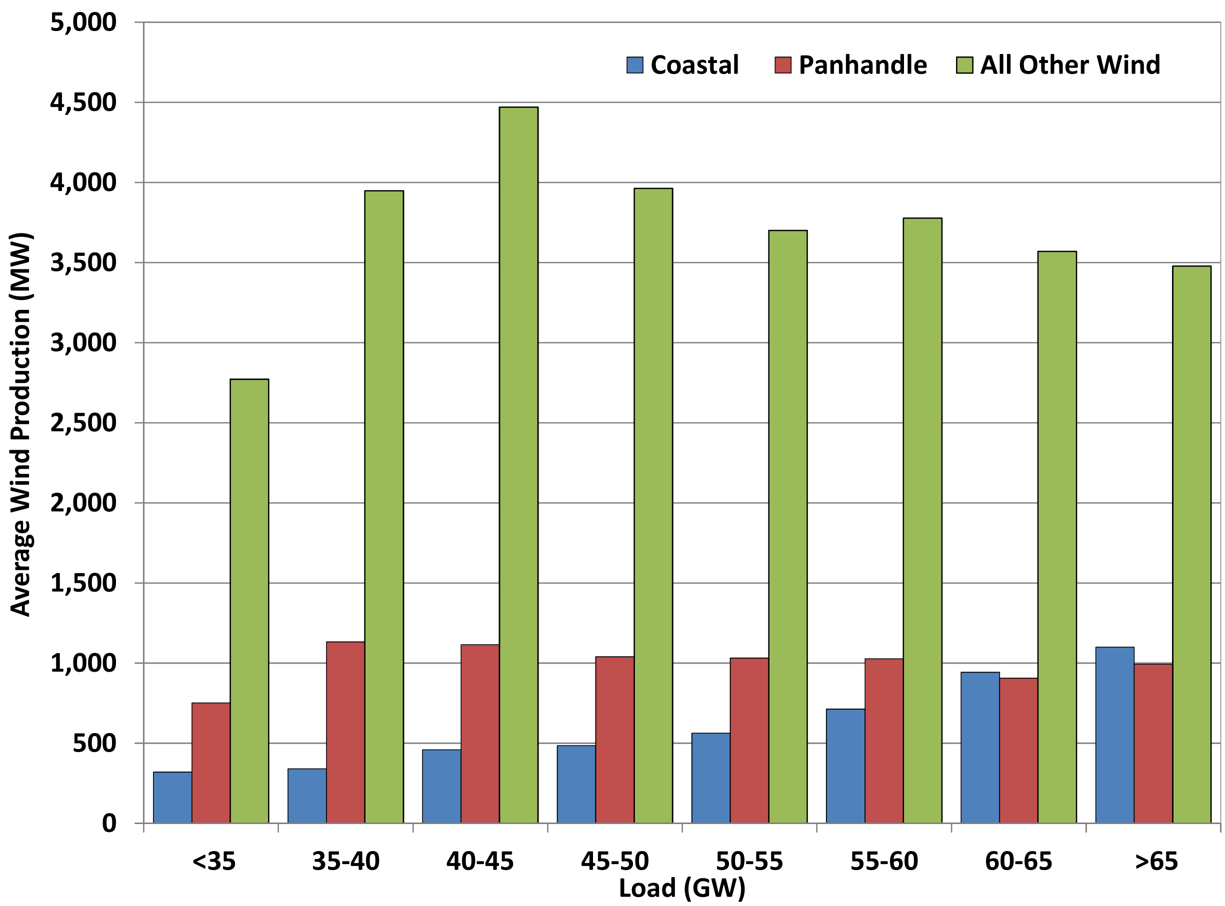


ERCOT continued to set new records for peak wind output in 2017. On November 17, 2017, wind output exceeded 16 GW, setting the record for maximum output and providing nearly 42% of the total load.[34]

Examining wind generation in total masks the different wind profiles that exist for locations across ERCOT. The attraction to sites along the Gulf Coast of Texas is due to the higher correlation of the wind resource in that location with electricity demand. More recently, the Texas Panhandle has attracted wind developer interest because of its abundant wind resources. The differences in output for wind units located in the coastal area of the South zone and those located elsewhere in ERCOT are compared below.

Figure 66 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 66: Summer Wind Production vs. Load

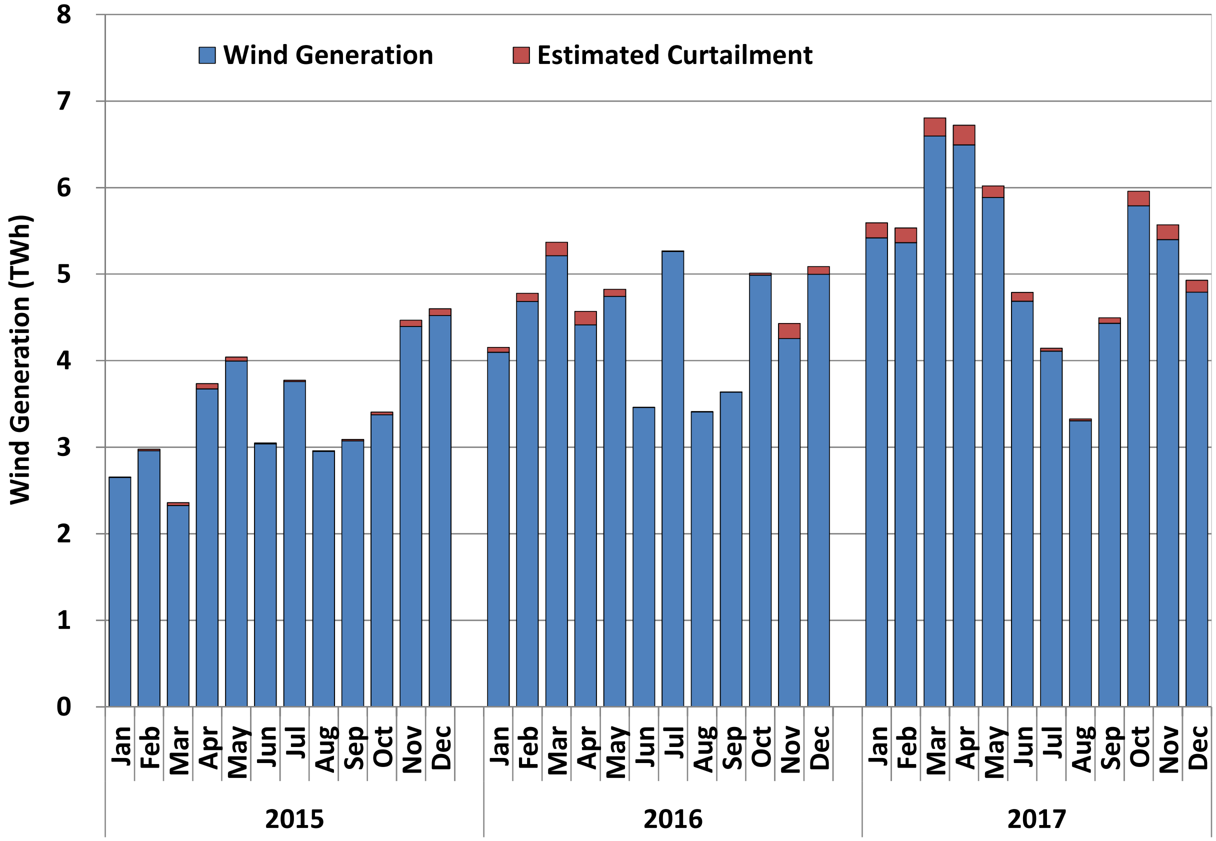


[34] Peak hourly wind generation was 16,035 MW on November 17, 2017 at 10:00 p.m.

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 67 shows the wind production and estimated curtailment quantities for each month of 2015 through 2017.

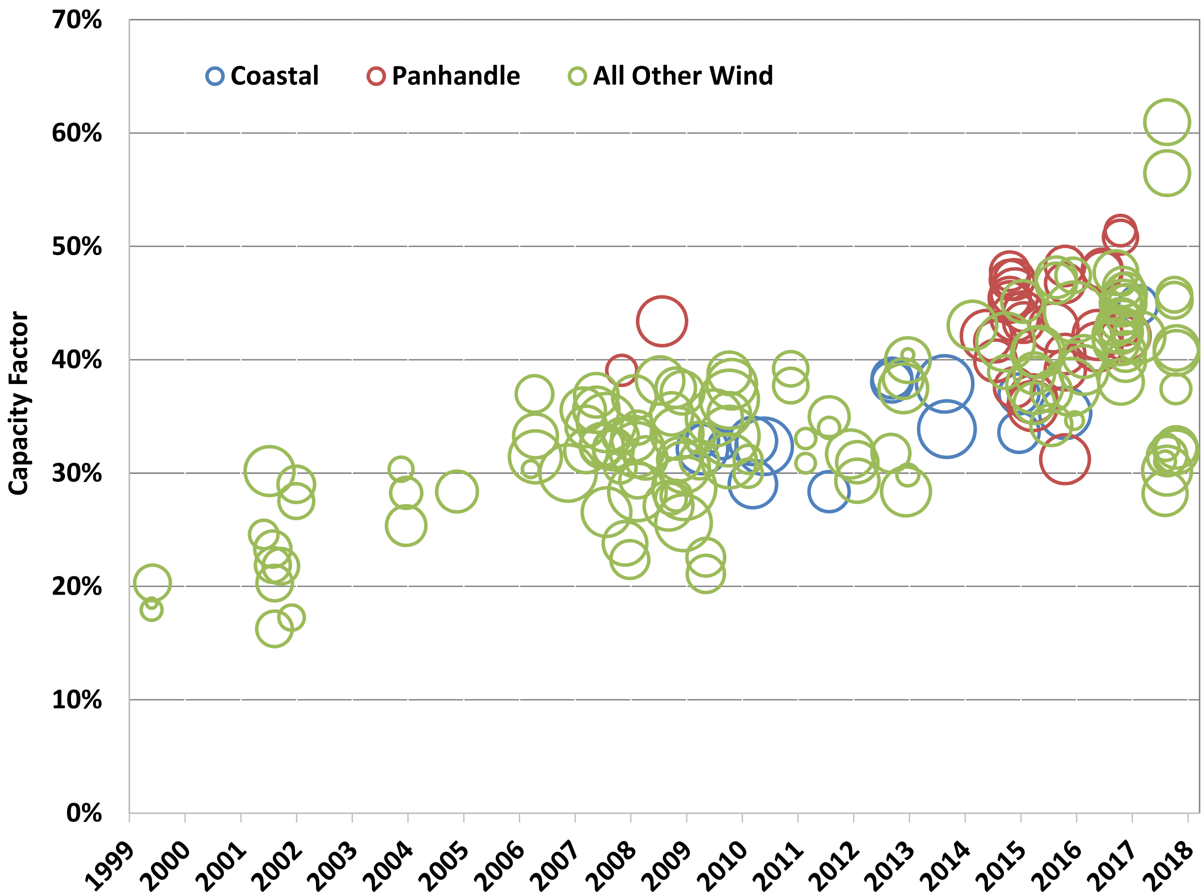
Figure 67: Wind Production and Curtailment



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 2017 was estimated at 98% 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%.

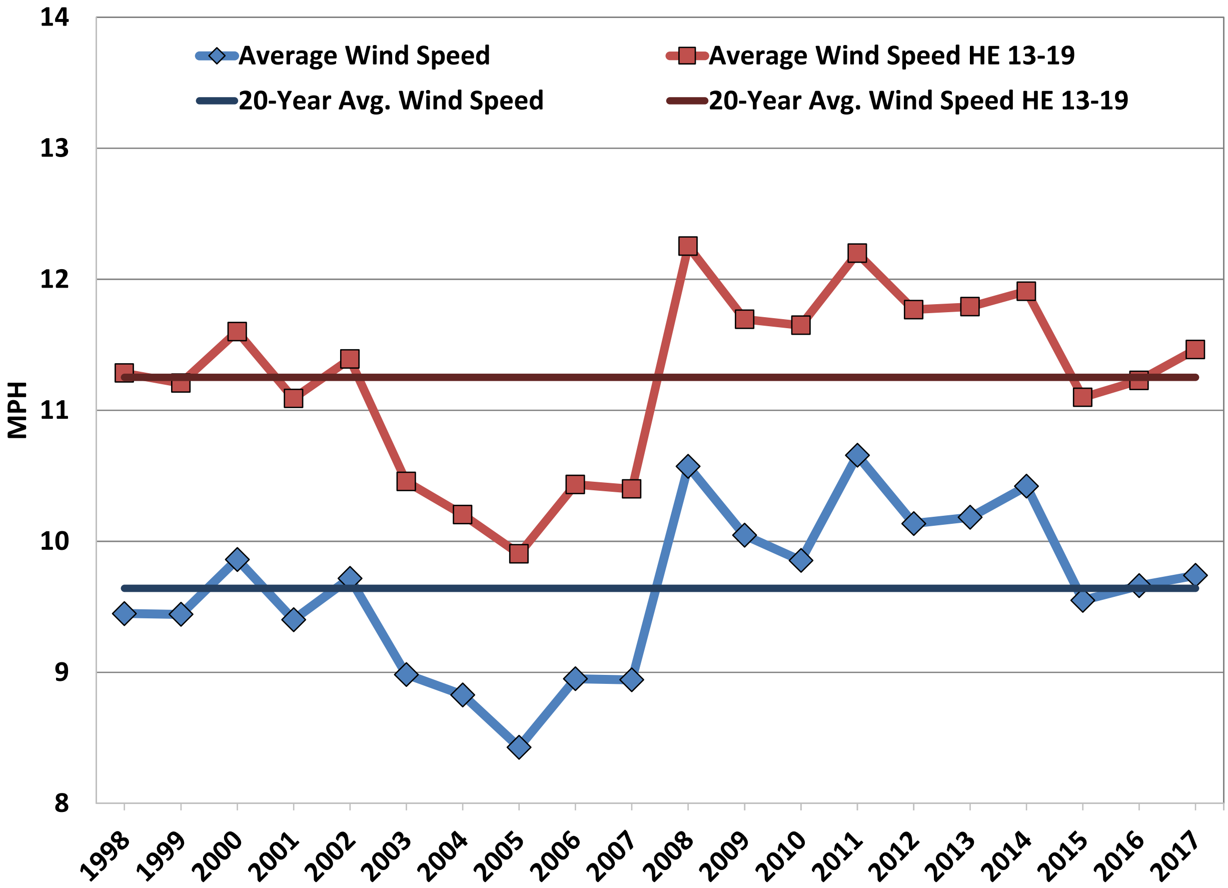
Figure 68 shows the capacity factor and relative size for wind generators by year installed. The chart also distinguishes wind generation units by location, with coastal units in blue and Panhandle resources in red, because of the different wind profiles for these regions. Coastal wind generally has a lower annual capacity factor, but as previously described its output is generally more coincident with summer peak loads. Completion of CREZ transmission lines has enabled more wind units to locate in the windier Panhandle area. The figure also shows a trend toward greater capacity factors for newer units.

Figure 68: Wind Generator Capacity Factor by Year Installed



The next figure shows average wind speeds in ERCOT, weighted by the current installed wind generation locations. Figure 69 provides a picture of the wind supply in 2017, averaged across the year and the average during peak hours, compared to the previous 19 years. The wind supply in 2017 was similar to the average over the past 20 years for all hours and for the peak hours of 13-19. With 2017 being close to an average wind supply year, if the existing fleet of wind generation had existed in prior years, total wind production could have been much greater. Notably, one of the years with higher than average wind speeds was 2011.

Figure 69: Historic Average Wind Speed



Increasing wind output also has important implications for the net load served by non-wind resources. Net load is the system load minus wind production. Figure 70 shows the net load duration curves for the years 2007, 2015, and 2017.

Figure 70 shows the reduction of remaining energy available for non-wind units to serve during most hours of the year, even after factoring in several years of load growth. The impact of wind on the highest net load values is much smaller.

Figure 70: Net Load Duration Curves

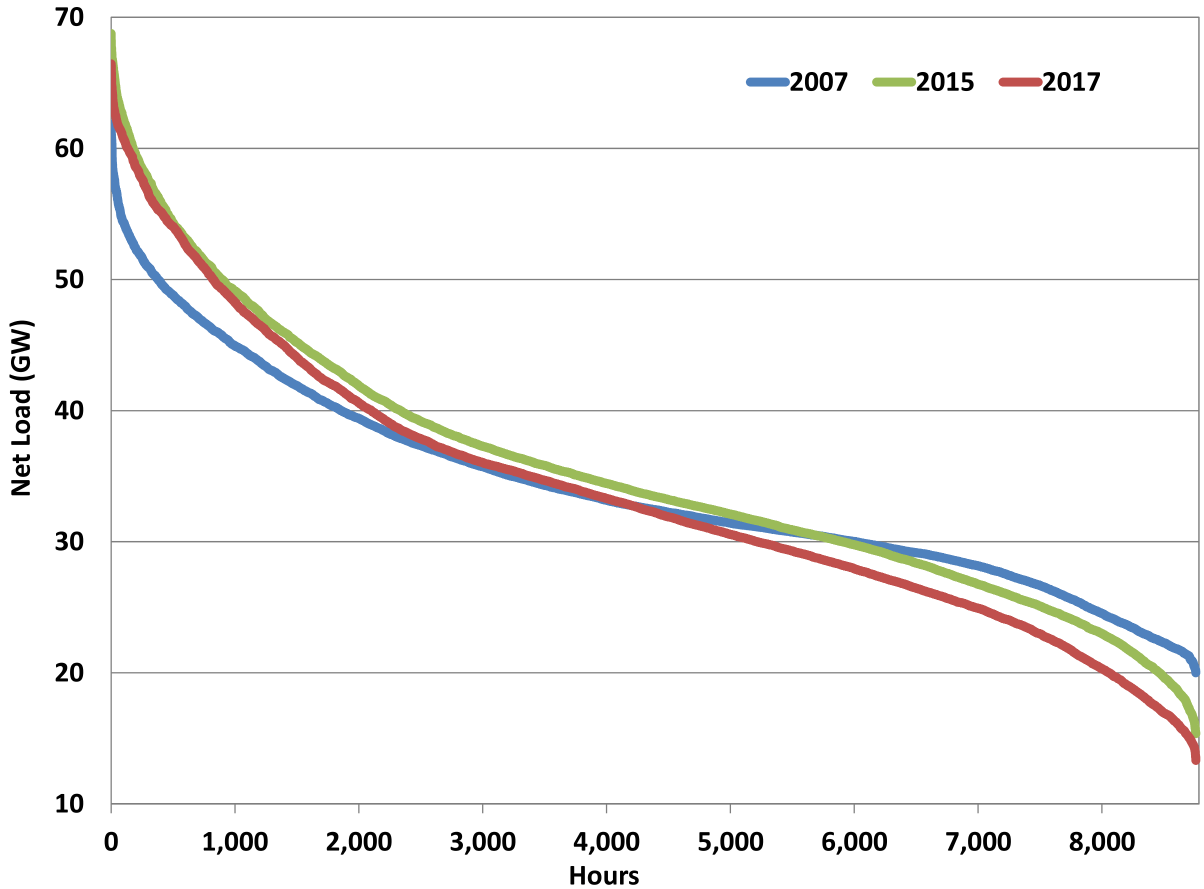
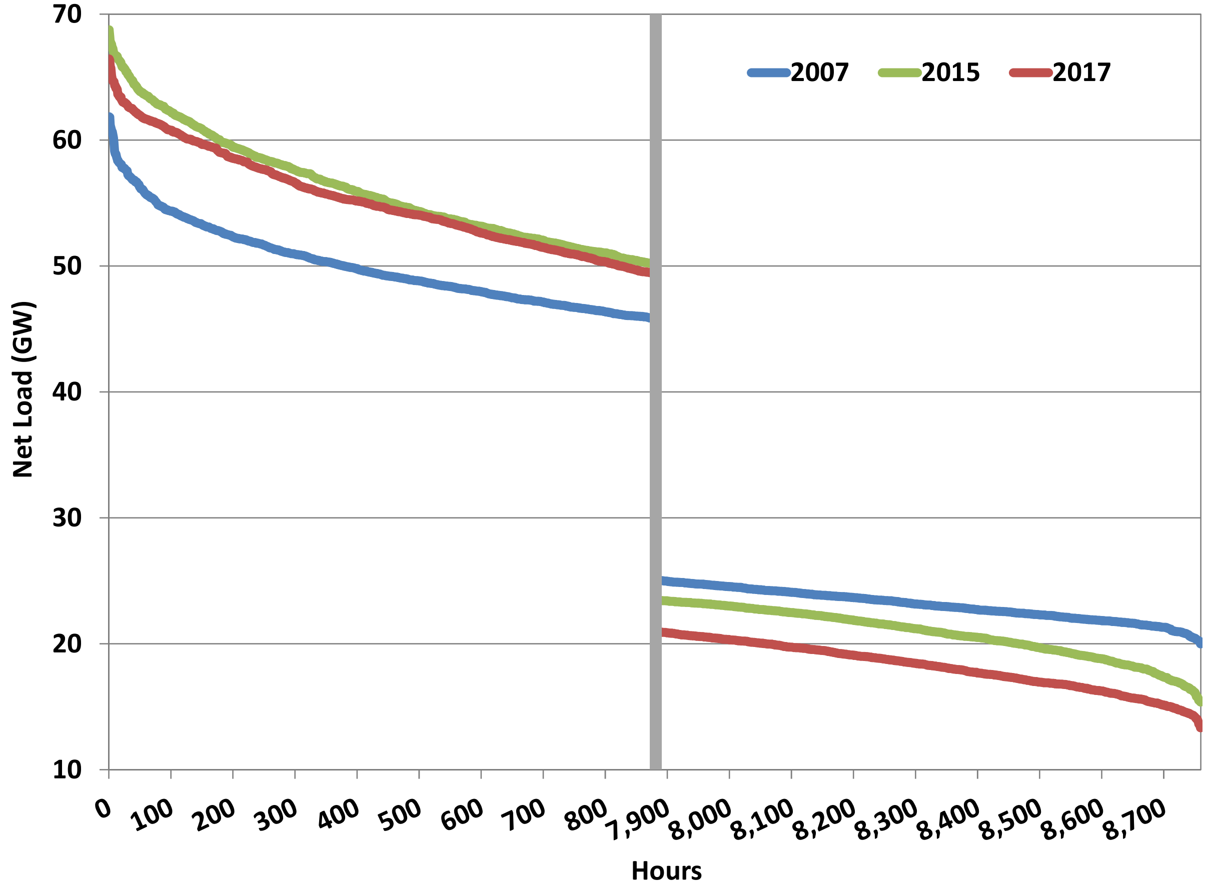


Figure 71 shows net load in the highest and lowest hours. Even with the increased development activity in the coastal area of the South zone, 73% of the wind resources in the ERCOT region are located in West Texas. The wind profiles in this area are such that most of the wind production occurs during off-peak hours or other times of low system demand. This profile 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 above), the difference between peak net load and the 95th percentile of net load has averaged 12.3 GW the past three years. This means that 12.3 GW of non-wind capacity is needed to serve load less than 440 hours per year.

Figure 71: Top and Bottom Deciles (Hours) of Net Load

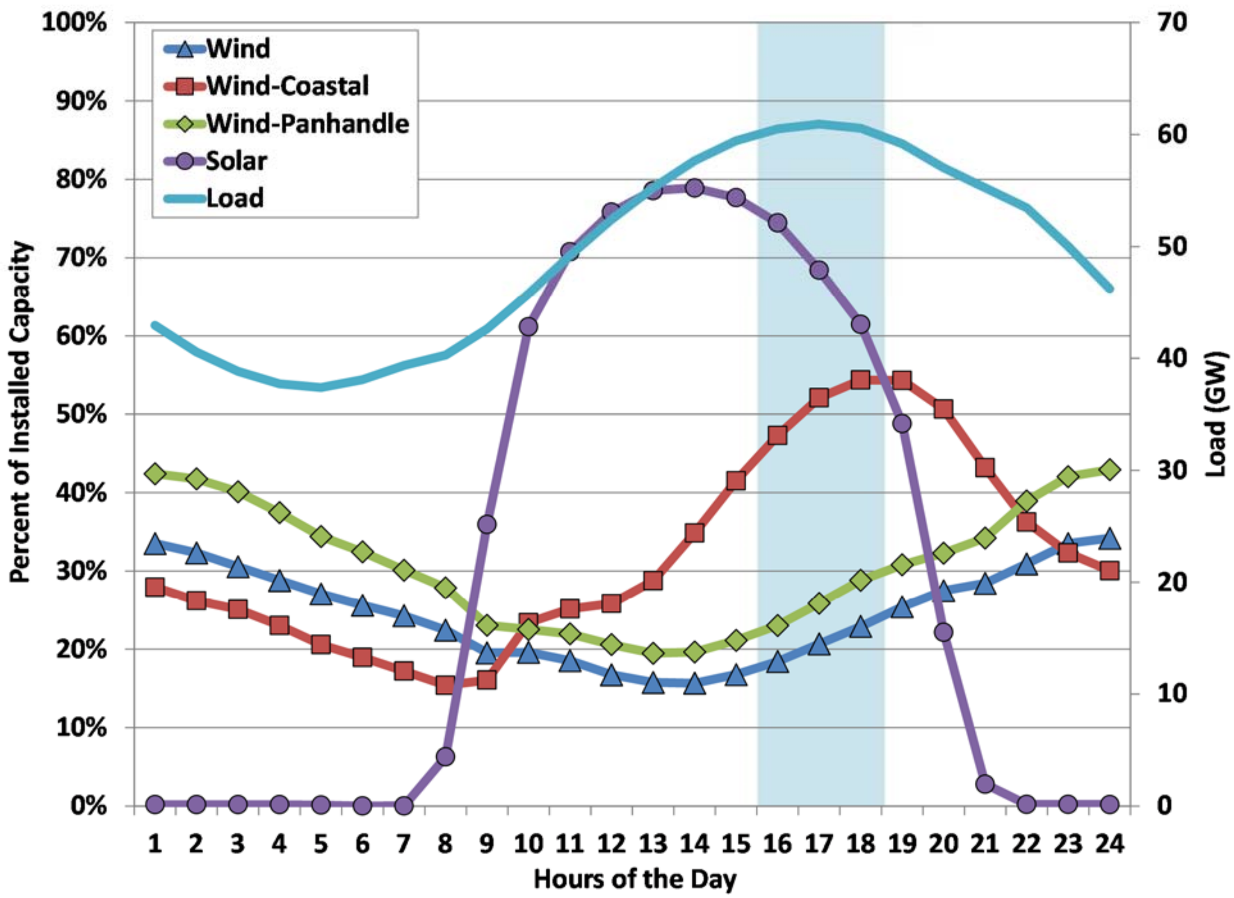


In the hours with the lowest net load (right side of the figure), the minimum net load has dropped from approximately 20 GW in 2007 to below 13.3 GW in 2017, even with the sizable growth in annual load that has occurred. This trend has put operational pressure on the almost 25 GW of nuclear and coal generation that were in-service in 2017. This operational pressure was certainly one of the contributors to the recent retirement of more than 4 GW of coal.

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 72 compares average summertime (June through August) hourly loads with observed output from solar and wind resources. Generation output is expressed as a ratio of actual output divided by installed capacity.

Figure 72: Summer Renewable Production



This figure shows that while the total installed capacity of solar generation is much smaller than that of wind generation, its production as a percentage of installed capacity is the highest in the early afternoon, approaching 70%, and producing almost 70% of its installed capacity during peak load hours.

The contrast between coastal wind and all other wind is also clearly displayed in Figure 72. Coastal wind produced over 50% of its installed capacity during summer peak hours. Output from Panhandle wind and all other wind (primarily West zone) was less than 30% 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 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

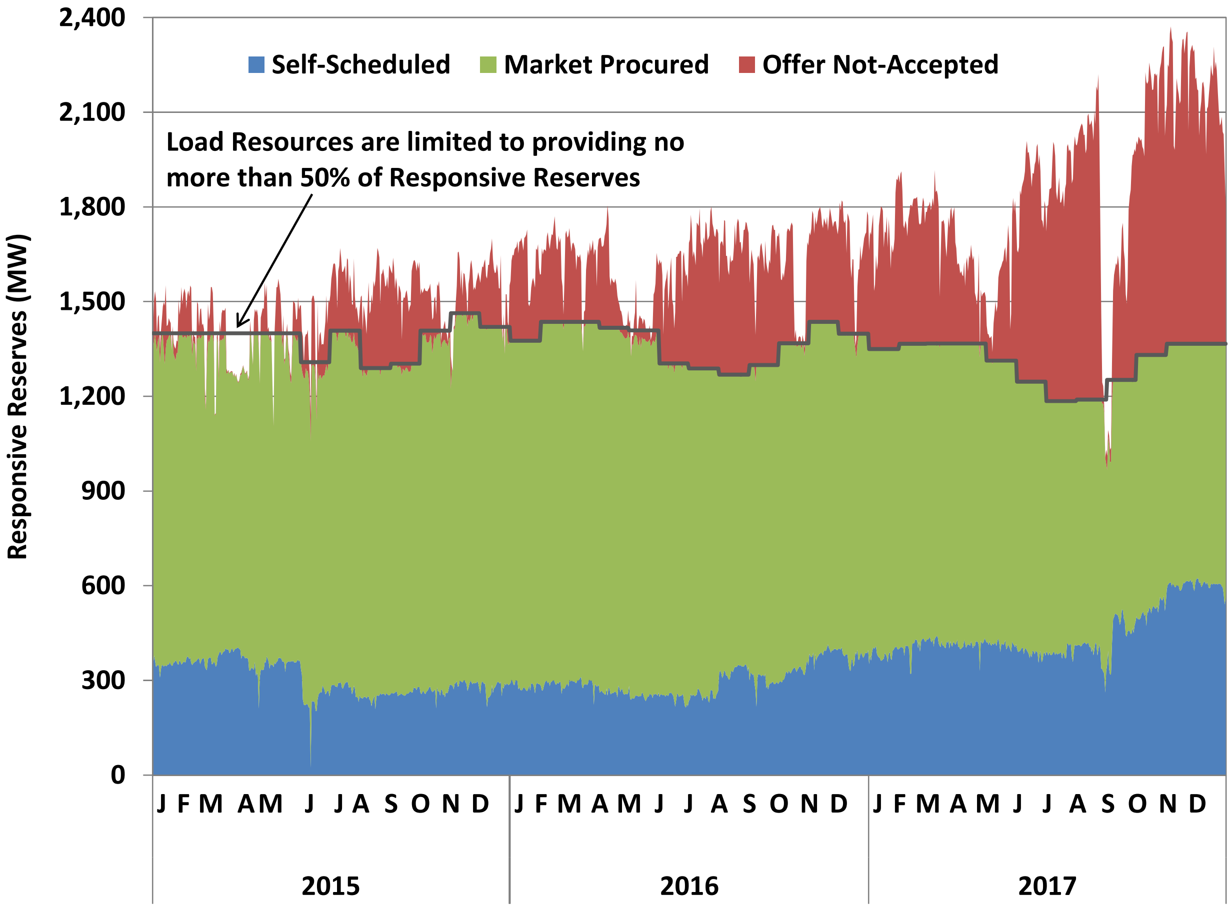
ERCOT allows qualified load resources to offer responsive reserves into the day-ahead ancillary services markets. Tripping load has the effect of increasing system frequency and can be a very effective mechanism for maintaining system frequency at 60Hz. Load resources providing responsive reserves have 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 2017, approximately 4,715 MW of qualified Load resources were capable of providing responsive reserve service, an increase of approximately 890 MW during 2017.

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 2017, the total amount of responsive reserves procured by ERCOT varied between 2,300 MW and 2,808 MW per hour. During 2017, there were no system-wide manual or automatic deployments of load resources providing responsive reserve service.

Figure 73 below shows the average amount of responsive reserves provided from load resources on a daily basis for the past three years.

Figure 73: Daily Average of Responsive Reserves Provided by Load Resources



Load resources are limited to providing a maximum of 50% of responsive reserves and the quantity of offers submitted by load resources exceeded the limit most of the time in 2017. 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 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. The significant reduction in offers from load resources observed in late August and early September is caused by the effects of Hurricane Harvey interrupting industrial processes along the Gulf Coast.

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 – Emergency Response Service (ERS) and load management programs offered by the transmission and distribution utilities. The ERS program is defined by a PUCT rule enacted in March 2012 setting a program budget of $50 million.[35] 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 2017 through January 2018 was $6.86 per MWh, exactly the same outcome as the previous program year. This price is significantly higher than the average price of $3.18 and $3.91 per MWh paid for non-spinning reserves in 2016 and 2017. ERS was not deployed in either year.

On March 30, 2017, the Public Utility Commission of Texas adopted an amendment to 16 TAC §25.507, permitting ERS resources to participate in Must Run Alternative (MRA) arrangements to replace the need for Reliability Must Run (RMR) generation resources.[36]

Beyond ERS there were slightly more than 200 MW of load participating in load management programs administered by transmission and distribution utilities in 2017.[37] Energy efficiency and peak load reduction programs are required under state law and PUCT rule and most commonly take the form of load management, where participants allow electricity to selected appliances (typically air conditioners) to be curtailed. 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 the allocation of transmission costs.

---

35 *See* 16 TAC § 25.507.

36 *See* Project No. 45927, *Rulemaking Regarding Emergency Response Service*.

37 *See* ERCOT 2017 Annual Report of Demand Response in the ERCOT Region (Mar. 2018) at 6, available at http://www.ercot.com/services/programs/load.

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.[38] ERCOT estimates that as much as 1500 MW of load were actively pursuing reduction during the 4CP intervals in 2016 and 2017.[39]

Load curtailment 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 curtailments 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 load curtailments on peak load days in June, August and September when real-time prices were less than $100 per MWh.

Two recent changes 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 first step, there are currently no loads qualified to participate in SCED. Second, the reliability adder, discussed in more detail in Section I: Review of Real-Time Market Outcomes, performs a second pricing run of SCED to account for the amount of load deployed, including ERS.

---

38 *See* 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) and 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).

39 *See* ERCOT, 2017 Annual Report of Demand Response in the ERCOT Region (Mar. 2018) at 7, available at http://www.ercot.com/services/programs/load.

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), the outcomes of RUCs, and the price mitigation that occurs during RUC and local congestion. The section concludes with a discussion of the Reliability Must Run (RMR) process revisions in ERCOT in 2017.

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.[40] Initially, the RUC offer floor was set at the system-wide offer cap. The RUC offer floor was subsequently

40 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[41] and then to the current offer floor of $1,500 per MWh.[42] 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.[43] This buyback or "opt-out" mechanism for RUC requires a resource to update its Current Operating Plan (COP) before the close of the adjustment period for the first hour of a RUC.[44]

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).[45] 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 Qualified Scheduling Entity (QSE) to accurately telemeter an ONRUC status.[46] To provide even greater flexibility, resources now have the ability to opt-out of RUC instructions given after the close of the adjustment period.

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 SCED-dispatchable interval within the first RUC-hour of the commitment block instruction. During 2017, approximately 28% of RUC instructions were given after the close of the

---

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

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

43 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.

44 Note that the process for electing to opt-out of a RUC will be based on real-time telemetry when NPRR744, RUC Trigger for the Reliability Deployment Price Adder and Alignment with RUC Settlement, goes into effect in mid-2017.

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

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

adjustment period. By comparison, 40% of RUC instructions were issued after the close of the adjustment period in 2016.

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.

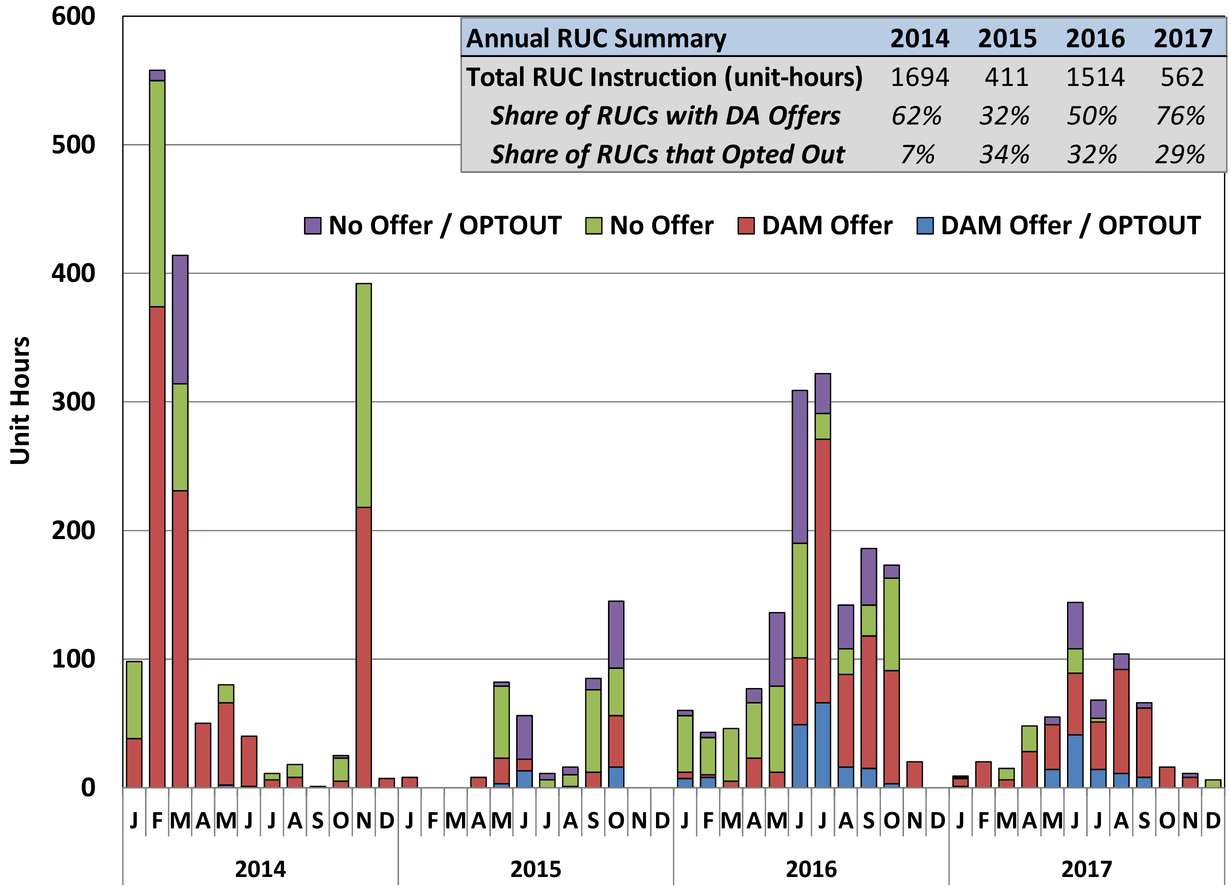
The number of RUC instructions in 2017 dropped considerably from 2016. The 562 unit-hours of RUC instructions in 2017 represent a 63% decrease from 1514 unit-hours in 2016. These 2017 RUC instructions were geographically diverse as well, with 41% to generators in the South zone in a variety of locations: San Antonio, Corpus Christi and the Rio Grande Valley (the Valley), 33% were to generators in the Houston zone, 24% were to generators in the North zone, and the remaining 2% were to generators in the West zone.

Like 2016, most reliability commitments in 2017 were made primarily to manage transmission constraints in 2017 (84% of unit-hours), including 7% to manage congestion in the aftermath of Hurricane Harvey. Only 13% of RUC instructions were made to ensure sufficient system-wide capacity and 2% for voltage support. The RUC activity in previous years was driven by a variety of other factors; in 2014, RUC activity was concentrated during cold weather events in February and March and in response to transmission outages in March and November. In 2015, RUCs were most frequent in the fall because of congestion in Dallas and the Valley. The high amount of RUC activity in 2016 was primarily for localized transmission congestion mainly to units located in Houston and the Valley.

Although the total volume of RUC instructions was much lower in 2017 compared to 2016, the amount of RUC instructions for system-wide capacity was greater in 2017. There were 73 unit-hours of RUC instructions to ensure system-wide adequacy, which represents 13% of the total in 2017. In 2016, there were 33 unit-hours, representing 2% of the total.

Figure 74 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.

Figure 74: Day-Ahead Market Activity of Generators Receiving a RUC



A unit that receives 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 into the day-ahead market, it is somewhat surprising that all units do not submit day-ahead offers. In 2017, only 76% of the generators receiving RUC instructions had day-ahead offers, a relatively low percentage considering the incentive to provide day-ahead offers inherent in the RUC claw-back rules. This low percentage was still an increase from 2016 when the ratio was 50%. This may indicate that some reduction in the RUC activity in 2017 was due to a larger share of the units needed for reliability being committed through the day-ahead market.

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 2017 that chose to opt-out was 29%, similar to the 32% of generators that chose to opt-out in 2016.

During the first half of 2017, QSE telemetry of a generator's RUC status served as the trigger for calculating a reliability adder. There were 397 hours in which units were settled as RUC in 2017 and 201.6 hours of pricing intervals with non-zero reliability adders that occurred coincident with a settled RUC hour.

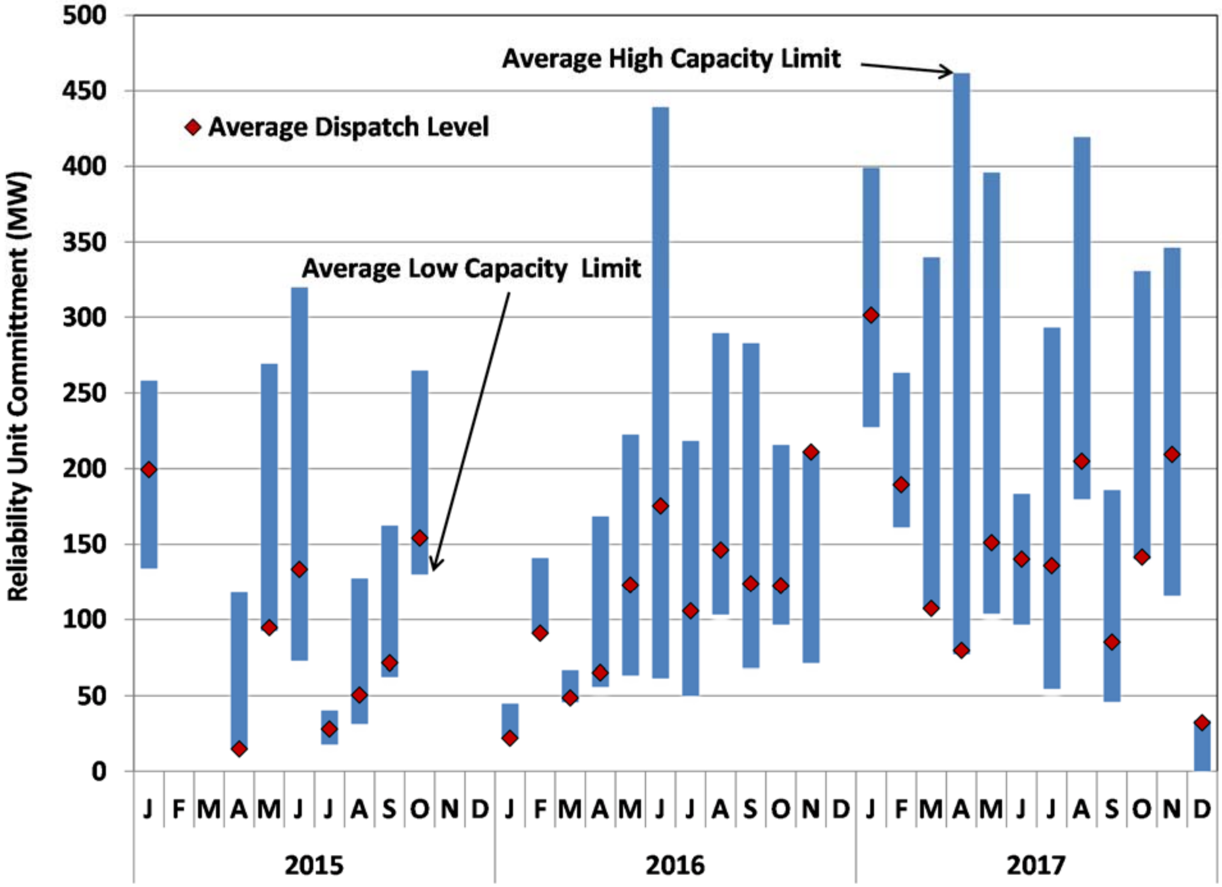
Table 9 lists the generators receiving the most RUC instructions in 2017. Also provided in the table are the total hours of RUC instruction, the number of hours in which the unit opted-out, and the average LSL for the unit. The units highlighted in gray in Table 9 are generators that most frequently received RUC instructions in 2016.

Table 9: Most Frequent Reliability Unit Commitments

Resource	Location	Unit RUC Hours	Unit OPTOUT Hours	Average LSL during RUC Hours
WA Parish G4	Houston	40	24	138
Duke CC1	Valley	31	21	177
Mountain Creek Unit 6	DFW	32	8	15
Silas Ray 10	Valley	2	36	24
WA Parish G3	Houston	12	24	90
Silas Ray CC1	Valley	21	12	47
WA Parish G2	Houston	19	8	27
Handley Unit 5	DFW	26	-	120
Coleto G1	Victoria	24	-	300
Handley Unit 4	DFW	21	1	120
Barney Davis G1	Corpus Christi	21	-	58
Cedar Bayou G2	Houston	16	-	94
Ennis Tractebel CC1	DFW	16	-	140
Barney Davis CC1	Corpus Christi	13	-	244
WA Parish G1	Houston	5	6	25

The next analysis compares the average dispatched output of the reliability-committed units, including those that opted-out, with the operational limits of the units. Figure 75 shows that the monthly average magnitude of RUC generation increased in 2017 compared to the prior two years. This figure shows that the average quantity dispatched during most months of 2017 exceeded 100 MW. In January, the average dispatch level was 300 MW because of a number of large generators receiving RUC instructions for a brief period.

Figure 75: Reliability Unit Commitment Capacity

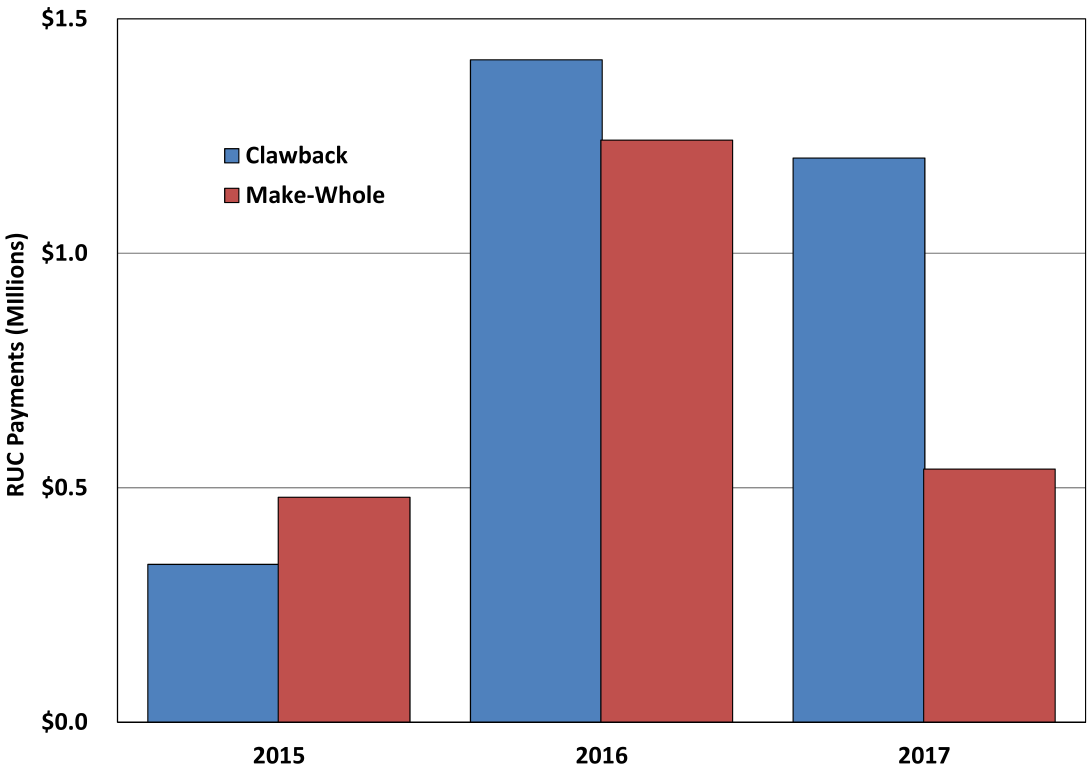


Units committed for RUC in 2017 showed a significant increase in the dispatch level compared to prior years. In 21% of intervals with RUC resources, one or more resources were dispatched above their Low Dispatchable Limit (LDL), whereas in prior years, resources receiving a RUC were infrequently dispatched 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 receive payment based on their mitigated offer caps. It is rare for a generator receiving a RUC instruction to be dispatched above LDL with their offer above the $1,500 per MWh offer floor and it did not occur during 2017.

When a unit is committed for RUC, the unit will receive a make-whole payment if the real-time revenues 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 43 different resources that received a RUC instruction in 2017, 34 resources had approved unit-specific verifiable costs for start-up costs and minimum load costs. Those 34 resources represent 80% of total RUC-instructed megawatt-hours in 2017.

Figure 76 displays the total annual amount of make-whole payments and clawback charges attributable to RUCs for 2015-2017. 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.

Figure 76: RUC Make-Whole and Clawback



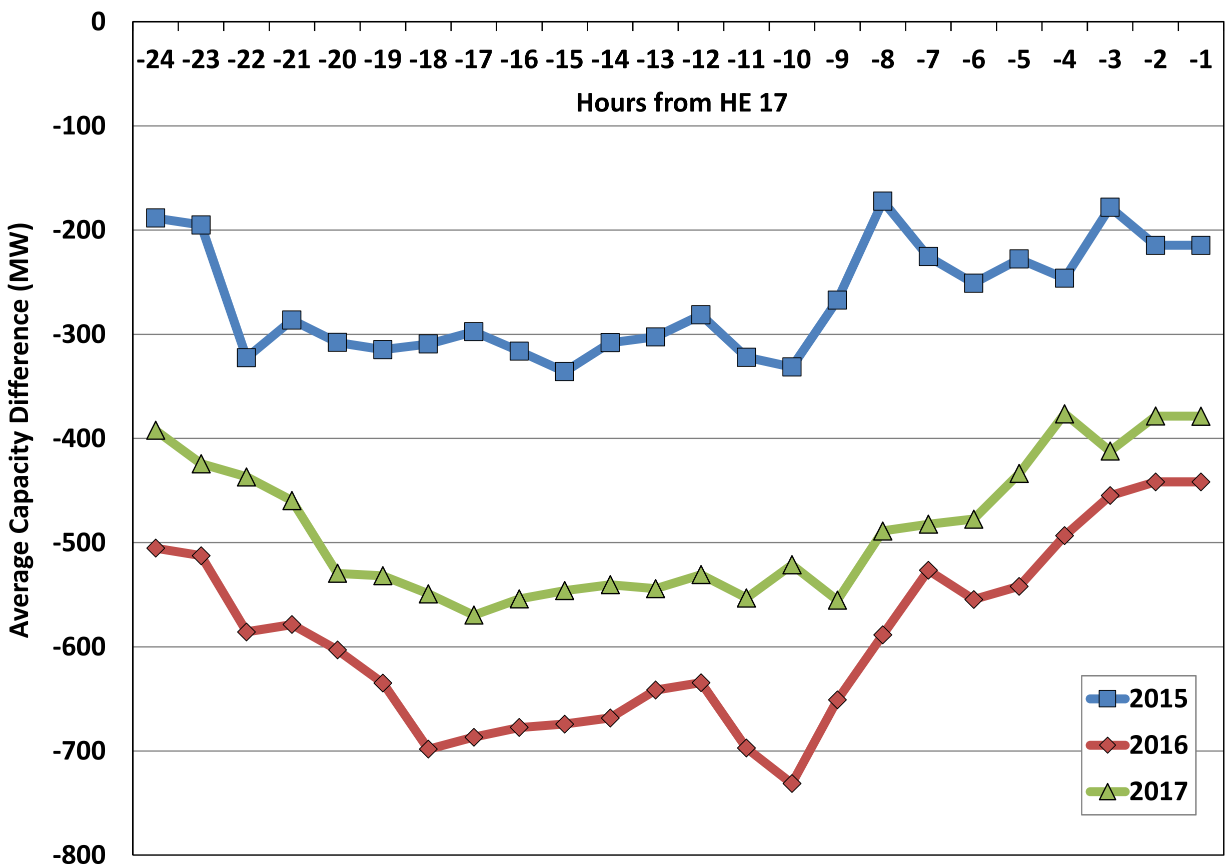
As stated above, if real-time revenues received by a RUC resource exceed the operating costs incurred by the unit, then excess revenues are clawed-back and returned to QSEs representing load. During 2017, $1.2 million was clawed back from RUC units while only $0.5 million in make-whole payments were made to RUC units. All RUC make-whole payments in 2017 were collected from QSEs that were capacity short. The magnitude of both the clawback and make-whole amounts are very small compared to the size of the ERCOT real-time energy market.

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.

Figure 77 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 for each of the 24 hours leading up to the close of the adjustment period. This data is derived from current operating plan submissions and averaged for hour ending 17 in the months of July and August, for each year 2015 through 2017. As shown in the figure below, the amount of capacity committed in advance of the operating hour for 2017 was greater than in 2016, but much less than in 2015. In 2015, on average, about 200 MW of capacity was committed in the last hour before real time. In 2016, the amount increased to over 420 MW, with even larger deficiencies seen in the last hours leading up to real time. The increase in self-committed capacity seen for summer 2017 may have been a reaction to the increased RUC activity observed in 2016.

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 make their own commitment decisions with regard to the nearly 400 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 77: Capacity Commitment Timing – July and August Hour 17



C. Mitigation

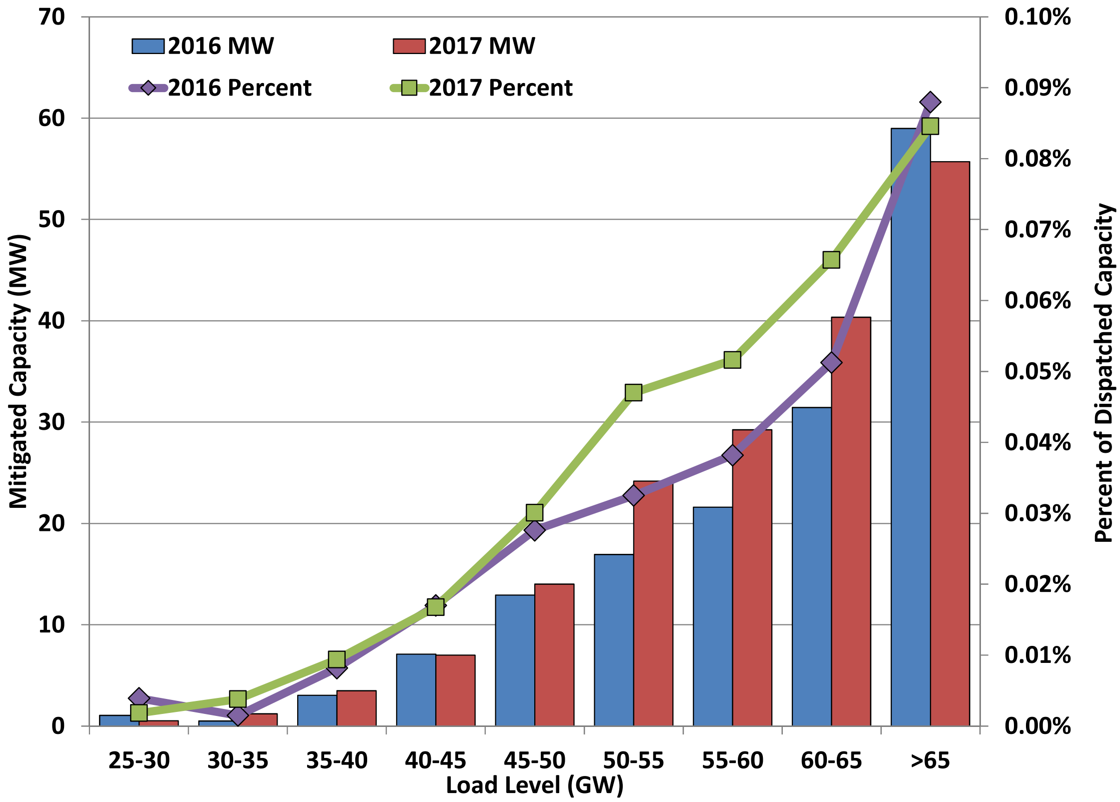
In situations where competitive forces are not sufficient, it can be necessary to mitigate prices to a level that approximates competitive outcomes. ERCOT’s real-time market includes a mechanism to mitigate prices for resources that are required to resolve a transmission constraint. Mitigation applies whether the unit is self-committed or receives a RUC instruction. Units typically received a RUC instruction to resolve transmission constraints and as such they are typically required to resolve a transmission constraint, and therefore mitigated. As shown previously in Figure 75, units that received a RUC instruction were frequently dispatched above their low operating limits in 2017. This higher dispatch was due to the RUC units being dispatched based on their mitigated price, not 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 2017 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. With 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 now subject to mitigation. This change has significantly reduced the amount of capacity subject to mitigation.

The analysis shown in Figure 78 computes the percentage of capacity, on average, that is actually mitigated during each dispatch interval. The results are provided by load level.

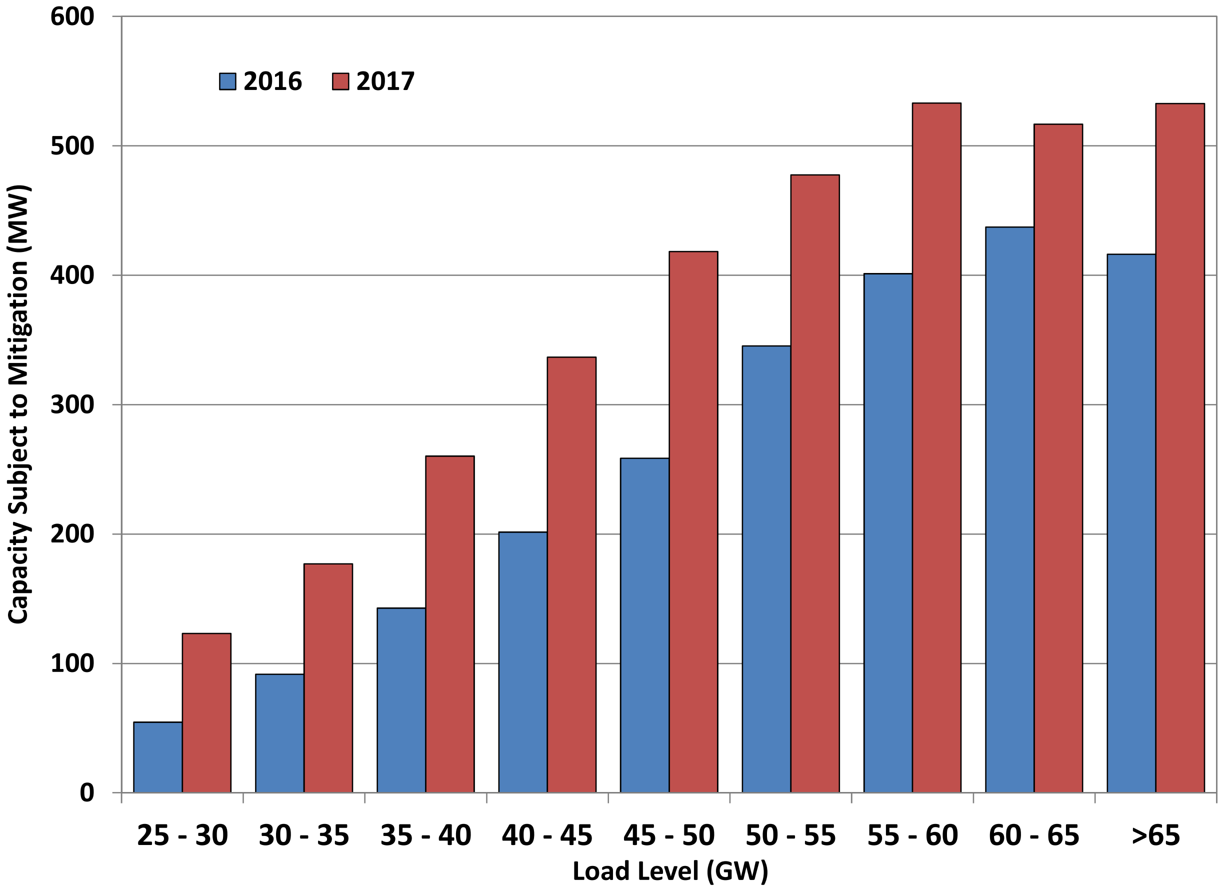
Figure 78: Mitigated Capacity by Load Level



The level of mitigation in 2017 was very similar to 2016. The average amount of mitigated capacity averaged almost 60 MW at loads greater than 65 GW in both 2017 and 2016.

In the previous figure, only the amount of capacity that could be dispatched within one interval was counted as mitigated. 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 79.

Figure 79: Capacity Subject to Mitigation



The amount of capacity subject to mitigation in 2017 was higher than 2016 in all load levels. In 2015 and 2014, the largest amount of capacity subject to mitigation did not exceeded 300 MW. 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.

D. Reliability Must Run

A total of eight generation resources provided Notifications of Suspension of Operations (NSOs) with suspension dates in 2017, accounting for approximately 2,000 MW of the capacity being retired or mothballed during the year.[47] ERCOT determined that the units were not necessary to support ERCOT transmission system reliability, and as a result no new reliability must run (RMR) contracts were awarded in 2017. However, review of the RMR process remained active

[47] Calpine Corp (RE), Aspen LLC, Pearsall Units 1, 2, and 3, Union Carbide Corp (RE), Gibbons Creek and Barney Davis.

throughout the year, including continued scrutiny of the RMR contract for Greens Bayou 5 executed in 2016.

Greens Bayou 5 is 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[48] 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.

As a result of the ongoing review of the RMR process, several protocols changes were implemented in 2017. Effective May 1, 2017, NPRR810 removed the applicability of the RMR Incentive Factor to reservation and transportation costs associated with firm fuel supplies, which will now be considered fuel costs.[49] The protocols were also changed to separate costs in the RMR Standby Payment equation based on Incentive Factor applicability.[50]

In addition to the protocol revisions contemplated in the stakeholder process, the Commission-directed rulemaking proceeding to evaluate certain aspects of RMR service in ERCOT concluded in 2017.[51] The amendments to 16 TAC §25.502 adopted by the Commission[52] adjust the notice requirements and complaint timeline applicable to suspending a resource's operation. They also gives ERCOT the discretion to decline to enter into an RMR agreement based on the economic value of lost load, requires ERCOT approval of RMR and MRA agreements and requires refunds

48 *See* NPRR788, RMR Study Modifications.

49 NPRR810, Applicability of RMR Incentive Factor on Reservation and Transportation Costs Associated with Firm Fuel Supplies.

50 *Id.*

51 *See* Project No. 46369, *Rulemaking Relating to Reliability Must-Run Service*.

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

in some instances for capital expenditures related to those agreements. An NPRR to incorporate these rule changes into the ERCOT Protocols is currently in progress.[53]

Further, several new proposed Protocol revisions were initiated in 2017, including reevaluation of the process for determining the Mitigated Offer Cap for RMR resources, previously contemplated in NPRR784.[54] The proposal would allow the RMR resource to be dispatched but be priced above other resources that solve the same constraint. Another 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.[55] And finally, a proposal was submitted that would allow 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.[56] This protocol change would include 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.

53 *See* NPRR862, Updates to Address Revisions under PUCT Project No. 46369.

54 NPRR826, Mitigated Offer Caps for RMR Resources.

55 NPRR838, Updated O&M Cost for RMR Resources.

56 NPRR845, RMR Process and Agreement Revisions.

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 and other regions are presented, followed by a description of the factors necessary to ensure resource adequacy in an energy-only market design.

A. Net Revenue Analysis

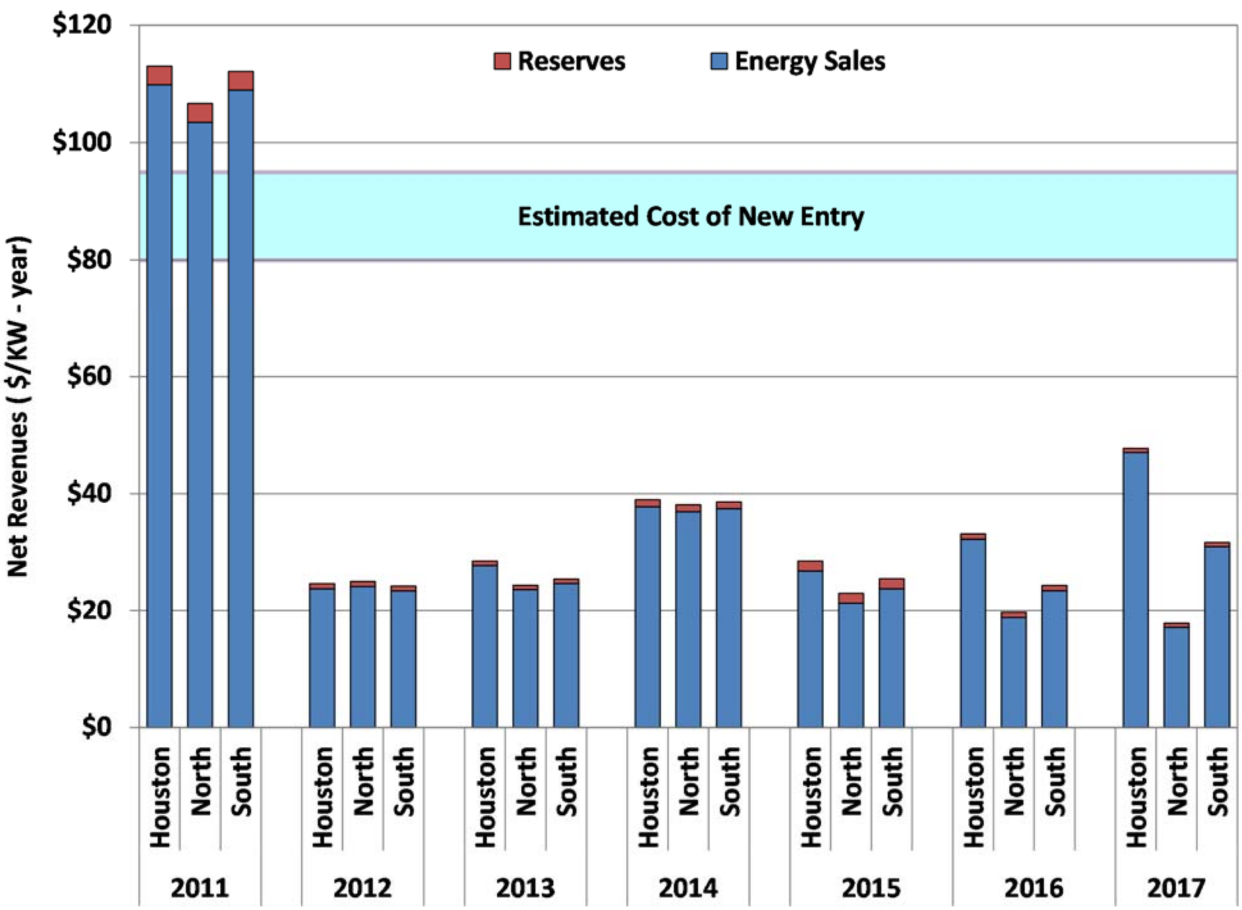
Net revenue is calculated by determining the total revenue that could have been earned by a generating unit less its variable production costs. Put another way, it is the revenue in excess of short-run operating costs that is available to recover a unit's fixed and capital costs, including a return on 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 necessitates reliance 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.

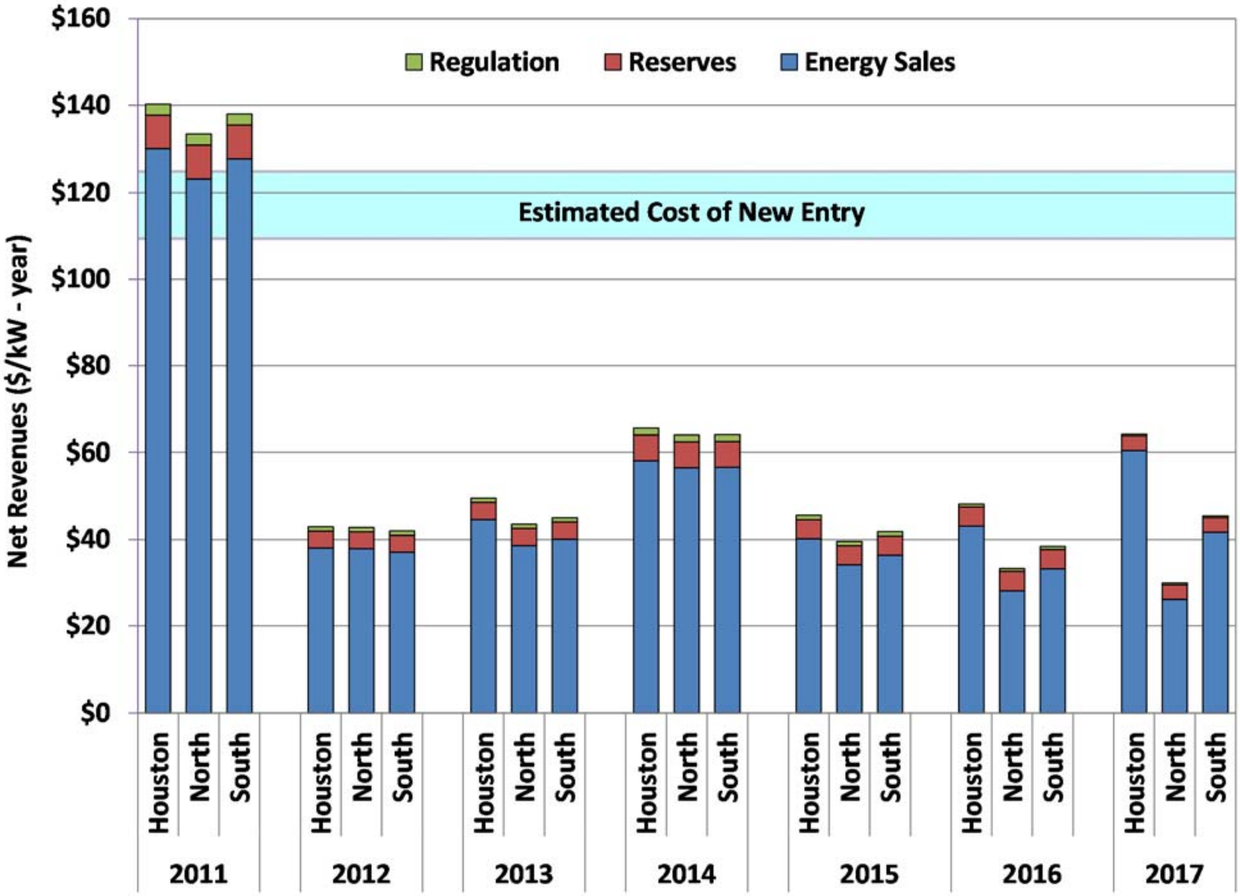
The next two figures provide an historical perspective of the net revenues available to support investment in a new natural gas combustion turbine (Figure 80) and combined cycle generation (Figure 81), selected to represent the marginal new supply that may enter when new resources are needed. Values 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 80: 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. The ERCOT market continued to provide net revenues well below the level needed to support new investment, ranging from below $20 per kW-year in the North Zone to almost $48 per kW-year in Houston.

Figure 81: 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 revenue in 2017 for a new combined cycle unit was calculated to be approximately $30 to $64 per kW-year, depending on the zone. These values are well below the estimated cost of new combined cycle generation.

These results are consistent with continued surplus of capacity, which contributed to infrequent shortages over the past three 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. Hence, these results alone do not 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 may well be a year with significantly more occurrences of shortage pricing.

Given the low natural gas and resulting energy prices in 2017, 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 2017 was $28.25 per MWh. The generation-weighted average price for the four nuclear units in ERCOT (approximately 5 GW of capacity) was lower at $24.73 per MWh. This is similar to nuclear prices in 2016 and 2015, which were also lower than the ERCOT-wide prices in those years. Nuclear prices were $21.46 per MWh in 2016, down from $24.56 per MWh in 2015.

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	$26.32
Combined Cycle	$28.45
Gas Peakers	$50.22
Gas Steam	$43.34
Hydro	$27.48
Nuclear	$24.73
Power Storage	$47.66
Private Network	$30.07
Renewable	$23.91
Solar	$24.34
Wind	$16.57

Assuming that operating costs in ERCOT are similar to the U.S. average, it is likely that these units were not profitable in 2017 based on the fuel and operating and maintenance costs alone. Hence, it is unlikely that these nuclear units covered any capital costs that may have been incurred. However, unlike other regions with large amounts of nuclear generation, the four nuclear units in ERCOT are relatively new and owned by four entities with sizable load obligations. Although 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. Nonetheless, the economic pressure on these units does potentially raise a resource adequacy issue that will need to continue to be monitored.

The generation-weighted price of all coal and lignite units in ERCOT during 2017 was $26.32 per MWh, an increase from $23.98 per MWh in 2016. Although specific unit costs may vary, index prices for Powder River Basin coal delivered to ERCOT were approximately $2.59 per MMBtu in 2017; returning to 2015 levels after decreasing to $2.51 per MMBtu in 2016. 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, the spread between coal and natural gas increased to nearly $0.40 per MMBtu. However, given coal units generally have higher heat rates and more expensive non-fuel operations and maintenance costs, economic pressure remain. During 2017 one coal unit was seasonally mothballed and Luminant declared its intention to retire seven other coal units in early 2018. The IMM reviewed each of these actions and found them to be supported by the unit specific financials.

These results indicate that during 2017 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 62, 2017 saw the highest level of non-renewable capacity additions since 2010, which may seem inconsistent with the low levels of scarcity pricing present in the ERCOT market in recent years. However, the fact that new generation continues to be added in the ERCOT market can be explained by a number of 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.

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 2017. 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 2017, shortages were again 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 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 82 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, MISO, NYISO, and PJM. 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.

Figure 82: Combustion Turbine Net Revenue Comparison Between Markets

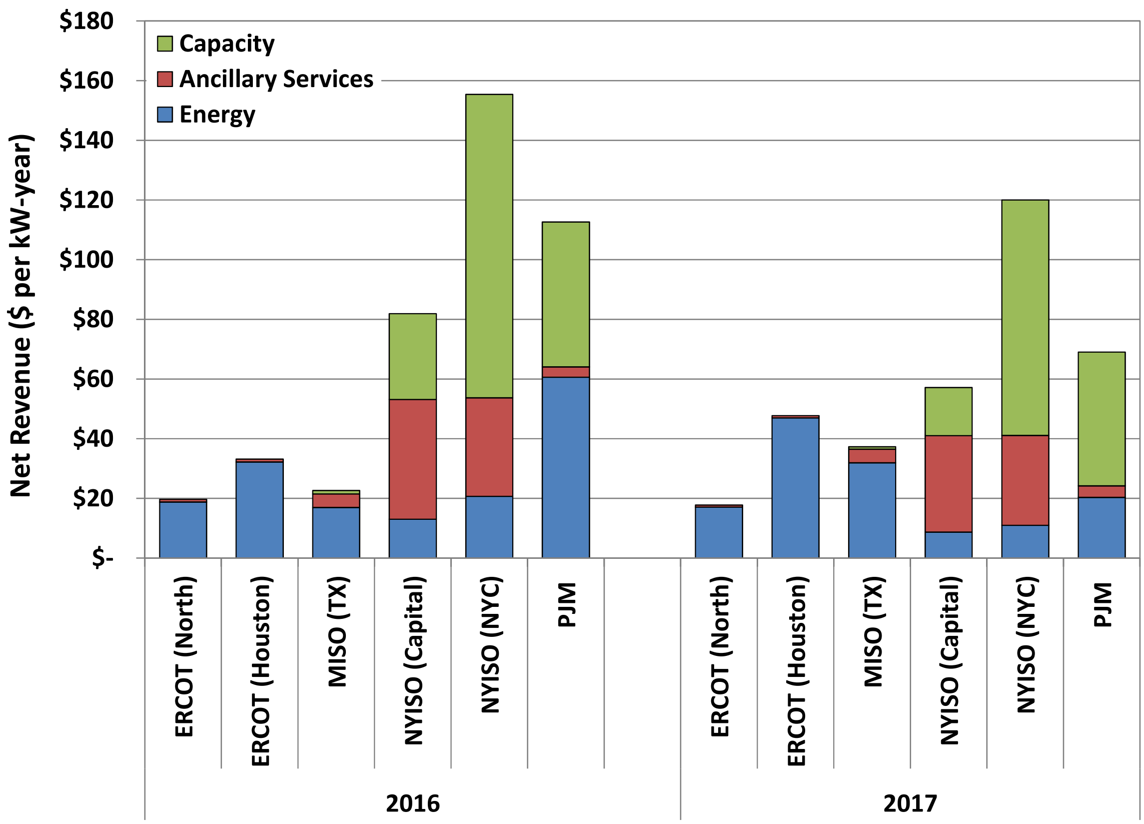
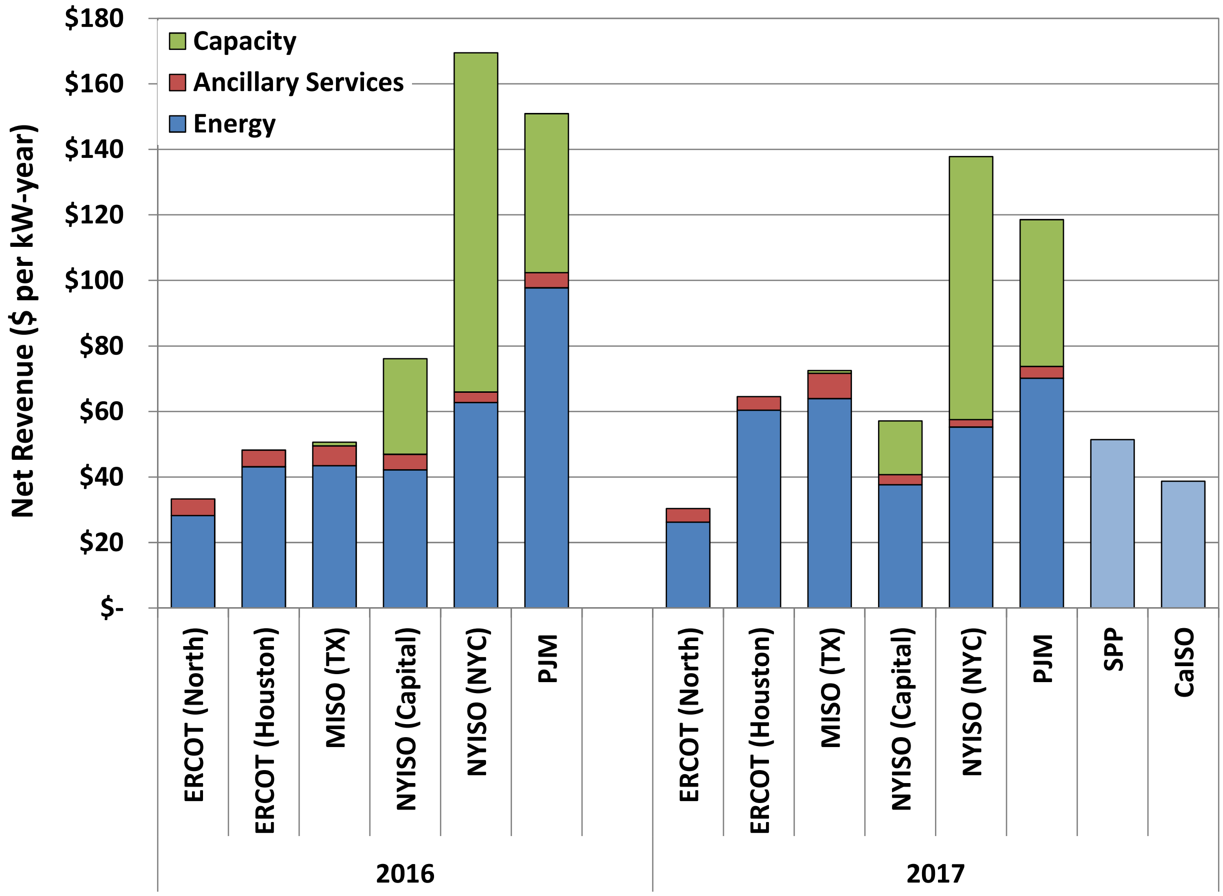


Figure 83 provides the net revenues for a hypothetical combined cycle unit with an assumed heat rate of 7,000 MMBtu per MWh installed in ERCOT, MISO, NYISO, and PJM. Both figures display estimates of net revenue from energy, reserves and regulation, and capacity. ERCOT does not have a capacity market, and thus, does not have any net revenue from capacity sales. Additionally, Figure 83 includes estimated total net revenues for a combined cycle generator located in SPP and CaISO, shown without the component values.

Figure 83: Combined Cycle Net Revenue Comparison Between Markets



Both figures indicate a general decline in net revenues across all markets. The exceptions to this trend were ERCOT's Houston zone and MISO's TX zone. Most other markets also have sufficient installed reserves, typically a result of flat or no load growth. The increase in Houston was related to transmission congestion limiting imports to the area. 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.

B. Effectiveness of the Scarcity Pricing Mechanism

The PUCT 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,[57] this subsection assesses the Scarcity Pricing Mechanism (SPM) in 2017 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 20: Duration of High Prices on page 23, there have been very brief periods when energy prices rose to the cap since the system-wide offer cap was increased to greater than $3,000 per MWh, and none since 2015.

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.[58] PNM also serves as a simplified measure of the annual net revenue of a hypothetical peaking unit.[59]

Figure 84 shows the cumulative PNM results for each year from 2006 through 2017 and shows that PNM in 2017 increased slightly from 2015 and 2016 levels. Considering the purpose for which the PNM was initially defined, that is to provide a "circuit breaker" trigger for lowering the system-wide offer cap, it has not approached levels that would dictate a needed reduction in the system wide offer cap.

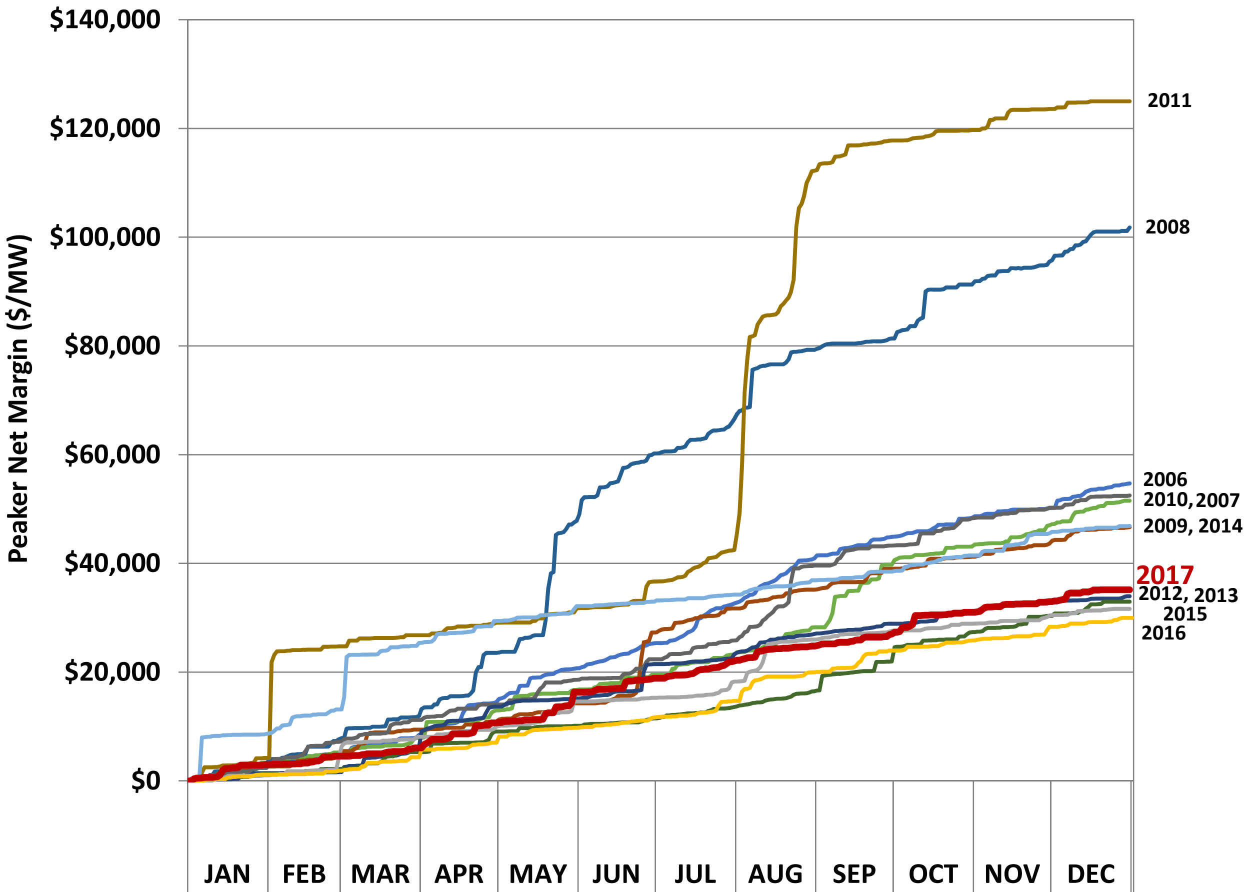
---

57 *See* 16 TAC § 25.505(g)(6)(D).

58 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."

59 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 84: 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.[60] 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 PUCT 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 small impact on real-time prices.

60 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.

In October 2015, the PUCT signaled its interest in reviewing ORDC "in order to examine how it has functioned and whether there is a need for minor adjustments to improve its efficiency."[61] Given the short time period with ORDC in effect, it is difficult to evaluate whether adjustments are warranted. As previously discussed, shortages are generally clustered in periods when weather-dependent load is unusually high or generation availability is poor; neither of which has occurred since the ORDC was implemented.

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, regardless of the ORDC parameters that are selected. Jointly optimizing all products would 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. Hence, the IMM continues to recommend that ERCOT make the investment necessary to achieve the full benefits of real-time co-optimization across all resources.

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 85 below shows ERCOT's current projection of planning reserve margins.

61 PUCT Project No. 40000, *Commission Proceeding to Ensure Resource Adequacy in Texas*, Memorandum from Commissioner Kenneth W. Anderson, Jr. (Oct. 7, 2015).

Figure 85: Projected Planning Reserve Margins

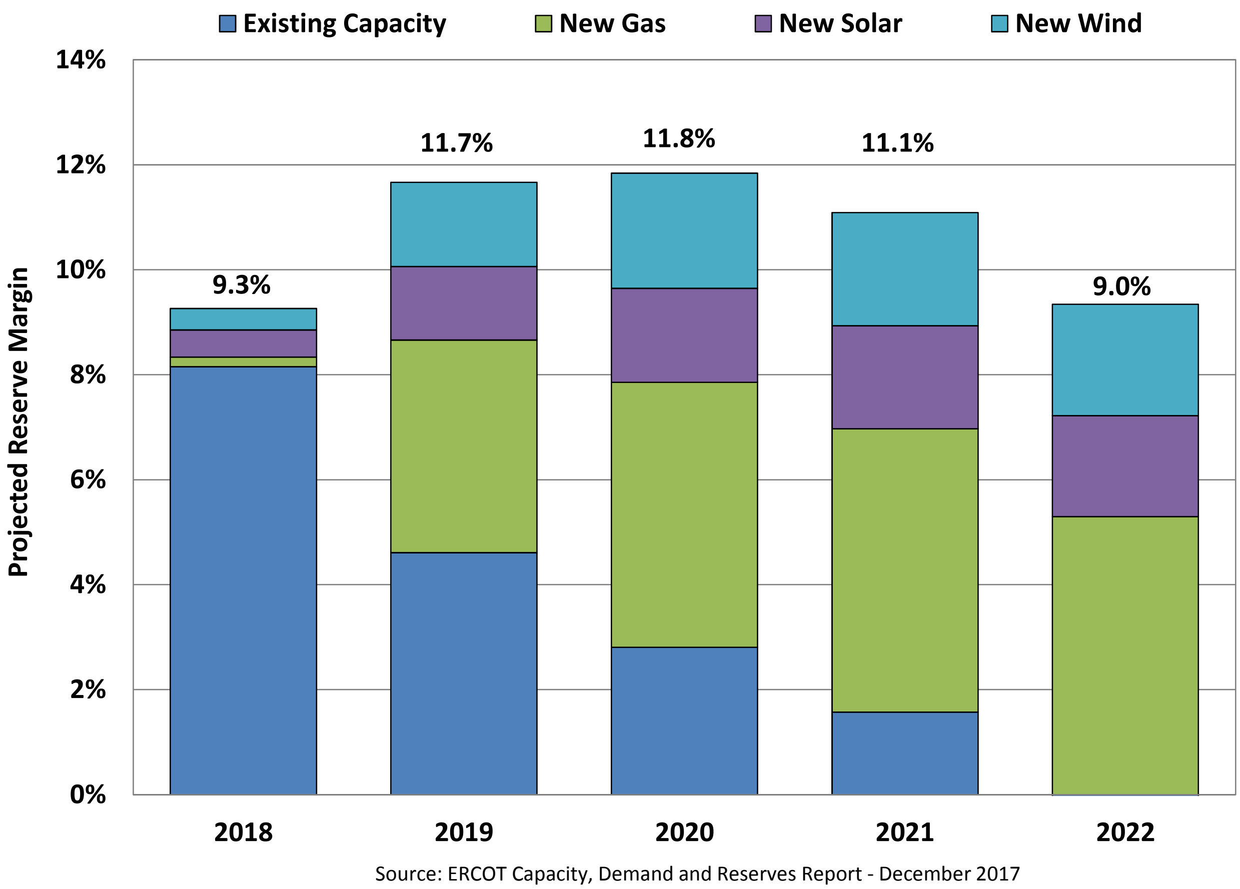


Figure 85 indicates that the region will have a 9.3% reserve margin heading into the summer of 2018. These projections are noticeably lower than those developed since May of last year,[62] which is due in large part to the approximately 5 GW of capacity taken offline by early 2018, with an expectation that the reserve margin will continue to be below the existing target level of 13.75% for the foreseeable future.[63]

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

62 See Report on the Capacity, Demand and Reserves in the ERCOT Region (May 2, 2017); http://www.ercot.com/content/wcm/lists/114798/CapacityDemandandReserveReport-May2017.pdf

63 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 PUCT 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.

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 scheduled to retire or mothball since the May 2017 CDR, eight of those units (totaling approximately 4,500 MW) were coal units.[64] 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. 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. This economic pressure will exist regardless of the future of environmental regulations that could require additional capital investment for existing coal units.

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. The market will achieve both objectives by establishing good economic price signals.

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 2017 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 generation and reduced amounts of natural gas generation.

Because the surplus has now disappeared and shortages are likely to be more frequent in 2018, 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 Figure 85 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

64 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 seasonally mothballed in October 2017.

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, approximately $5 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.

Expectations for energy pricing under non-shortage conditions are 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, in 2012 and 2013 the PUCT devoted considerable effort deliberating issues related to resource adequacy. In September 2013, the PUCT Commissioners directed ERCOT to move forward with implementing ORDC, a mechanism designed to ensure effective shortage pricing when operating reserve levels decrease. 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 2017. 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 2017.

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.[65] 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, although it is important to recognize its limitations. As a structural indicator, it 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, it does identify conditions under which a supplier could raise prices significantly by withholding resources.

Figure 86 shows the ramp-constrained RDI, calculated at the QSE level, relative to load for all hours in 2017. The trend line indicates a strong positive relationship between load and the RDI.

65 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 86: Residual Demand Index

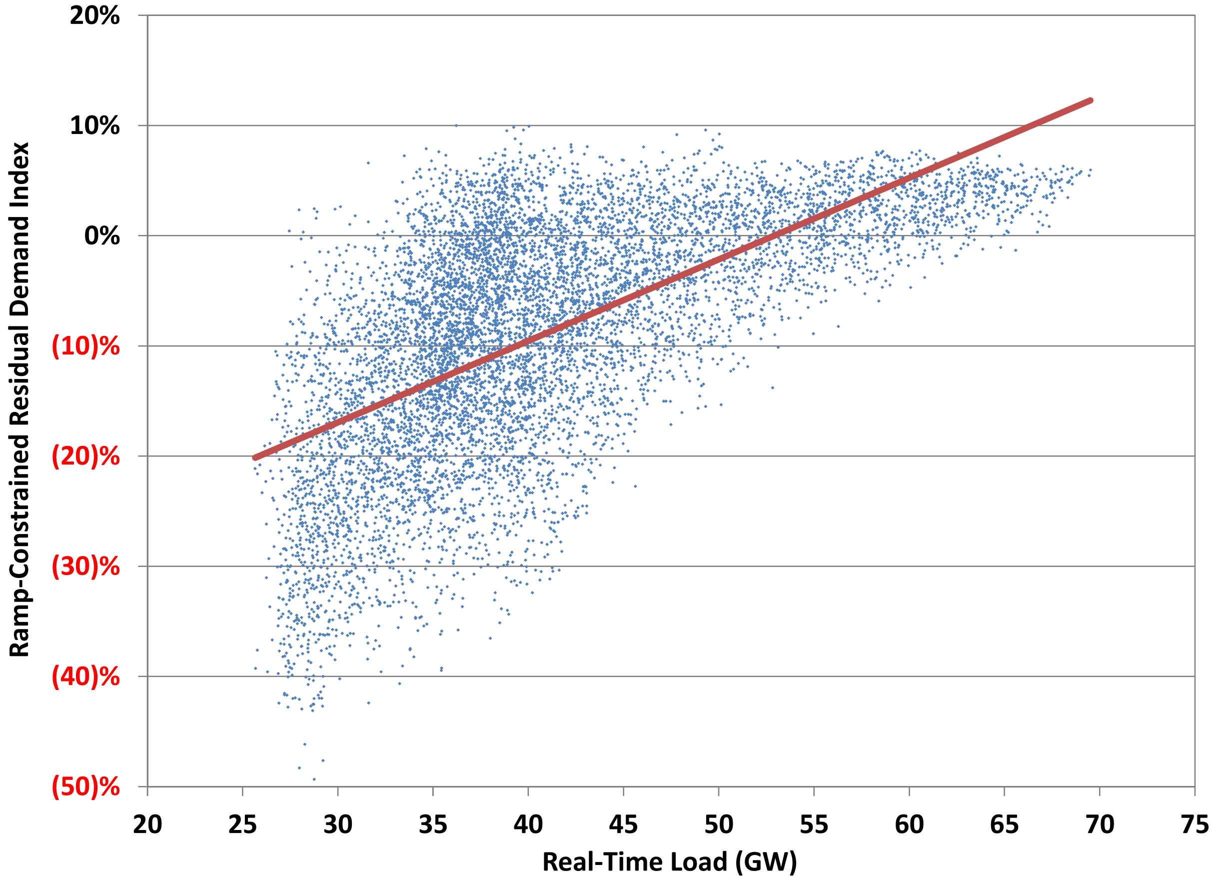
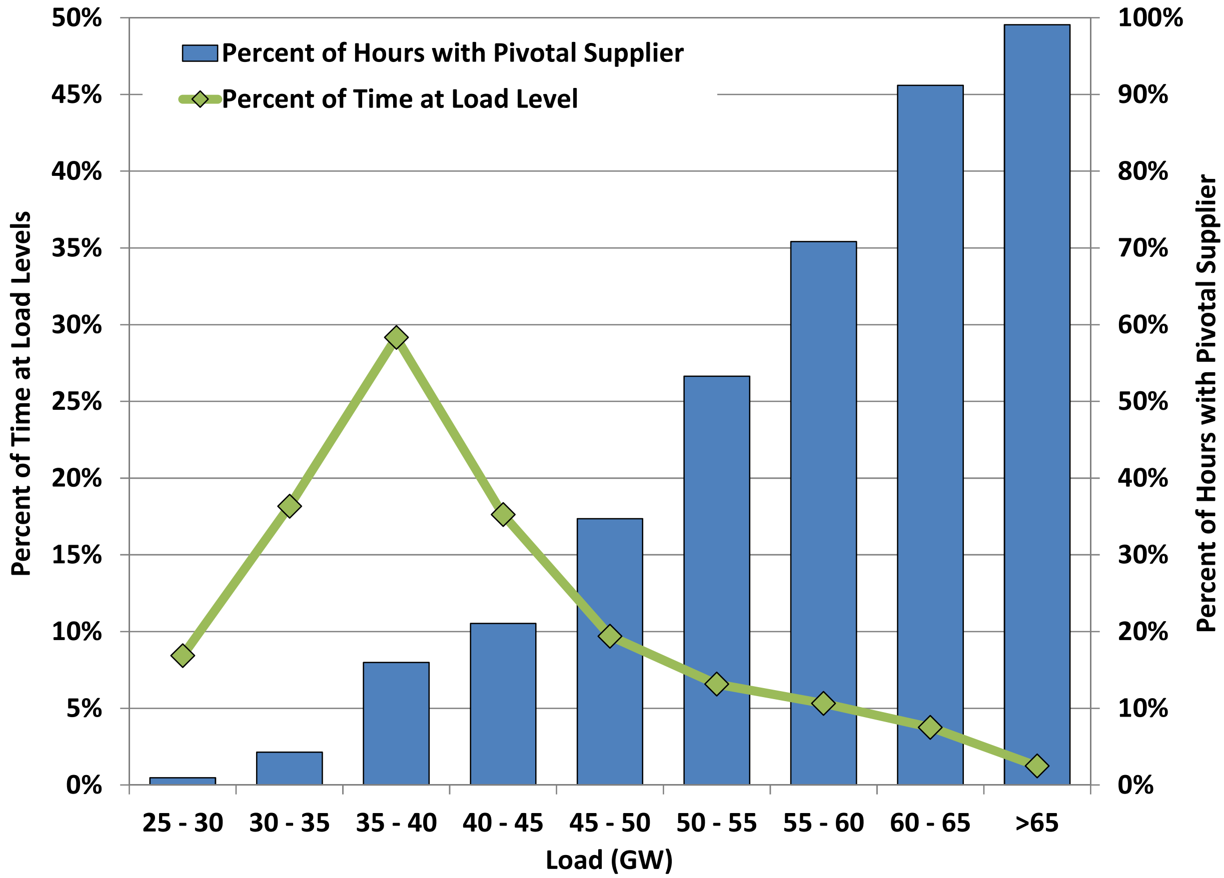


Figure 87 below summarizes the results of the RDI analysis by displaying the percentage of time at each load level there was a pivotal supplier. The figure also displays the percentage of time each load level occurs.

Figure 87: Pivotal Supplier Frequency by Load Level



At loads greater than 65 GW there was a pivotal supplier 99% 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. There was a noticeable decrease in the percentage of time with a pivotal supplier at loads below 50GW in 2017. This led to a decrease in the pivotal supplier frequency to 24.5% of the time in 2017, down from 28.5% and 26% of all hours in 2016 and 2015, respectively. Even with the slight decrease, market power continues to be a potential concern in ERCOT and underscores the need for 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 do 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 in 2017. Generation owners are motivated to enter into VMPs because adherence to a plan approved by the PUCT 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 the 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 2017, the four market participants with approved VMPs were Calpine, NRG, Luminant and Exelon. Calpine’s VMP was approved in March of 2013.[66] 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 recent additions to Calpine’s generation fleet 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.

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

NRG's plan, initially approved in June 2012 and modified in May 2014,[67] 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 PUCT for a VMP in May 2015.[68] The Luminant plan is similar in many respects to the NRG plan. Under the VMP, Luminant is 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 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 contains 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's VMP was in effect for all of 2017, with a termination clause requiring that it would stay in effect until terminated by the Executive Director of the Commission or by Luminant.[69]

Approved on August 31, 2017,[70] Exelon's VMP provides 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 may be offered at prices up to and including the high system-wide offer cap (HCAP). The amount of capacity covered by these provisions is slightly less than 600 MW. Exelon's VMP shall remain in effect from the

67 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).

68 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).

69 Luminant terminated its VMP on April 9, 2018, upon closing of the proposed transaction approved by the Commission in the Order in PUCT Docket No. 47801.

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

date it is approved by the Commission until terminated by the Executive Director of the Commission or Exelon, or terminated automatically upon the earlier of: (a) three years from the date of the Commission's August 31, 2017 Order, or (b) the day Exelon's Installed Generation Capacity drops below 5% of the total ERCOT Installed Generation Capacity.

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, all four 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 PUCT 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."[71] The exercise of market power may not rise to the level of an abuse of market power if it does 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.

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 roughly 4,000 MW.[72] The combined amount of capacity afforded offer flexibility under the VMPs granted to Calpine, NRG, Luminant and Exelon totals less than 2,800 MW of capacity.

---

71 PURA § 39.157(a).

72 For purposes of the 5% exemption, the estimated total installed generation capacity is currently 80,423 MW; see Project No. 39870, *Estimate of Installed Generation Capacity in ERCOT*, PUC Competitive Markets' Estimate of Installed Generation Capacity in ERCOT at 1 (May 25, 2018).

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.

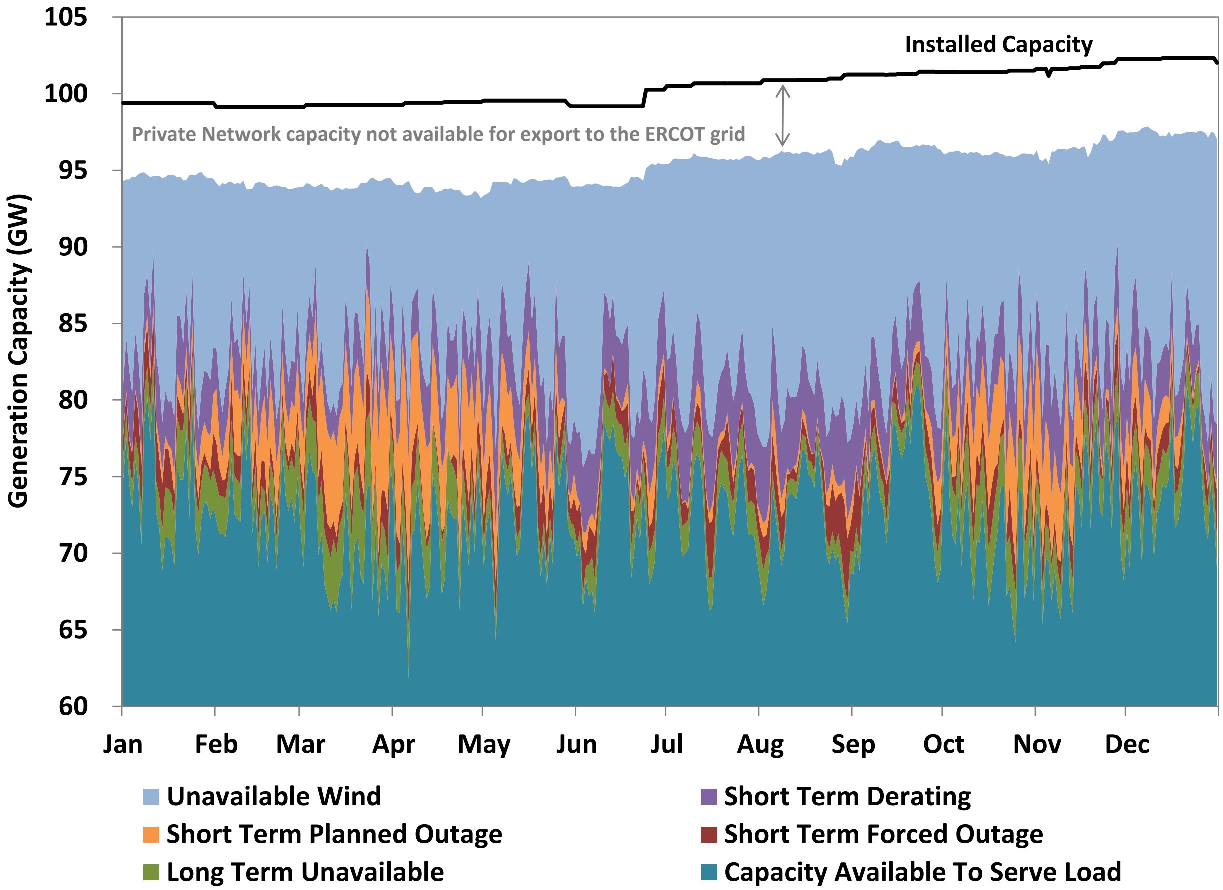
In a single-price auction like the real-time energy market, suppliers may attempt to exercise market power by withholding resources. The purpose of withholding is to cause more expensive resources to set higher market clearing prices, allowing the supplier to profit on its other sales in the real-time 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.

Generation Outages and Deratings

Some portion of installed capacity is commonly 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 88 shows a breakdown of total installed capacity for ERCOT on a daily basis during 2017. 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 88: Reductions in Installed Capacity



Outages and deratings of non-wind generators fluctuated between 5 and 17 GW, as shown in Figure 88, while wind unavailability varied between 4 and 20 GW. Short-term planned outages were largest in the shoulder months of March, 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.

The quantity of long-term (greater than 30 days) unavailable capacity, peaked in March at 4.7 GW and dropped to below 1 GW in late May. In early June, one of the Comanche Peak nuclear units experienced a long term forced outage lasting until early August, driving the long-term unavailable capacity to just over 2 GW. Unavailable capacity reduced to 1 GW with the return to service of Comanche Peak unit 2 before increasing to 3.3 GW in October. With the exception of the impacts of the Comanche Peak outage, this pattern reflects the continued choice by generation owners to schedule long duration outages during the spring and fall so as to ensure the units are available during the high load summer season when the units have a higher likelihood of operating.

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 89 shows the average magnitude of the outages and deratings lasting less than 30 days for the year and for each month during 2017.

Figure 89: Short-Term Outages and Deratings

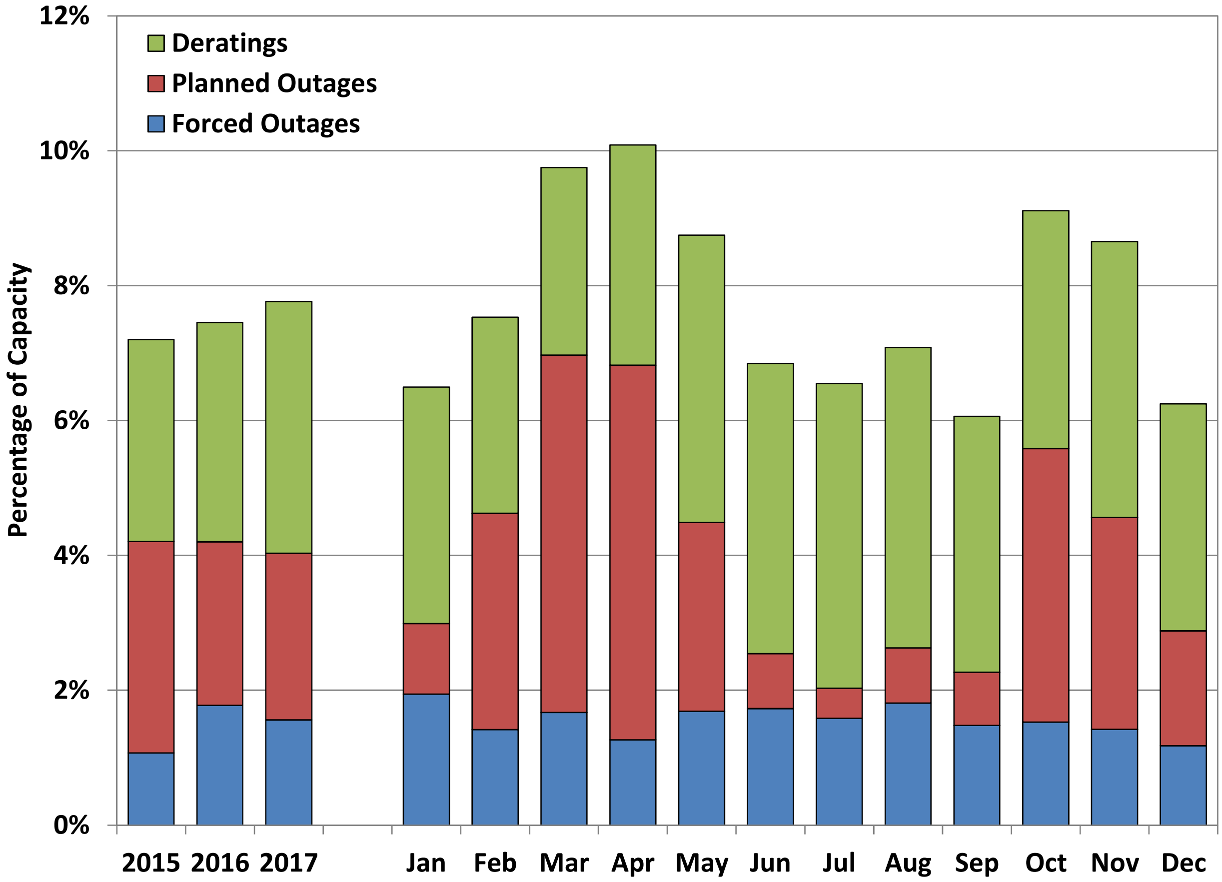


Figure 89 shows that total short-term deratings and outages were as large as 10% of installed capacity in April, and averaged around 6.5% during the summer. Most of this fluctuation was due to anticipated planned outages. The amount of capacity unavailable during 2017 averaged 7.7% of installed capacity. This is a slight increase from 7.5% experienced in 2016 and 7.2% experienced in 2015. Excluded from this analysis was a lengthy forced outage of Comanche Peak unit 2, which occurred from early June to mid-August. Including the effects of this long-term forced outage of this large unit increases the monthly forced outage rates in June through August to almost 3%, and raises the annual forced outage rate from 1.6% to approximately 1.8%. Even with including the Comanche Peak outage, outages and deratings are lowest during the summer when load is expected to be highest is consistent with expectations in a competitive market.

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. This can be done either by derating a unit or declaring it 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 86 and Figure 87 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 90 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 90: Outages and Deratings by Load Level and Participant Size, June-August

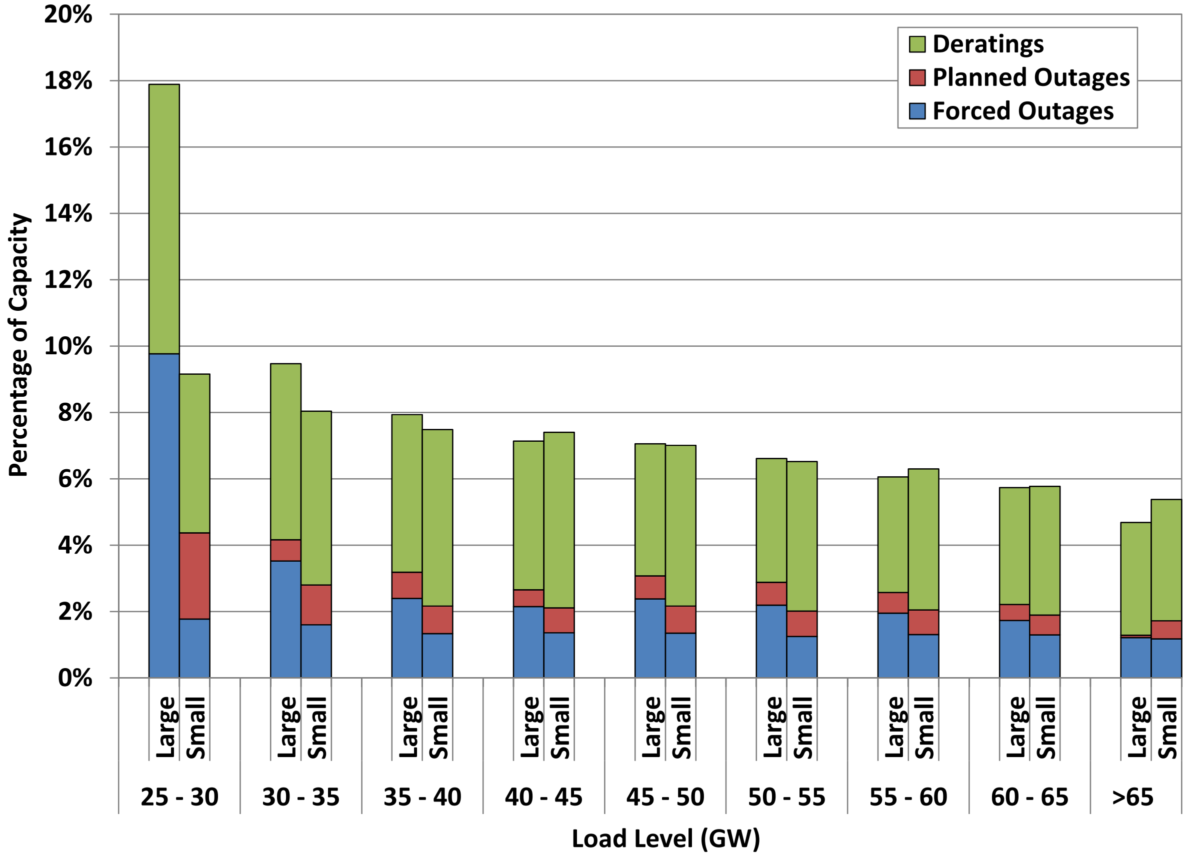


Figure 90 suggests that as demand for electricity increases, all market participants tend to make more capacity available to the market by scheduling planned outages during low load periods. Because small participants have less incentive to physically withhold capacity, the outage rates for small suppliers serves as a good benchmark for competitive behavior expected from the larger suppliers.

As in the previous analyses, the lengthy forced outage of Luminant's Comanche Peak nuclear unit is excluded from the analysis shown in Figure 90. If included, the effects of that outage would have approximately doubled the forced outage rates for large parties during the higher load periods. The higher forced outage rate for large parties at the lowest load levels reflects the impacts of Hurricane Harvey. Setting these two issues aside because they raise no competitive concerns, outage and deration rates for large suppliers were less than those of the smaller suppliers in 2017.

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.[73] The mitigated offer cap serves as a proxy for the marginal production cost of energy from that resource.

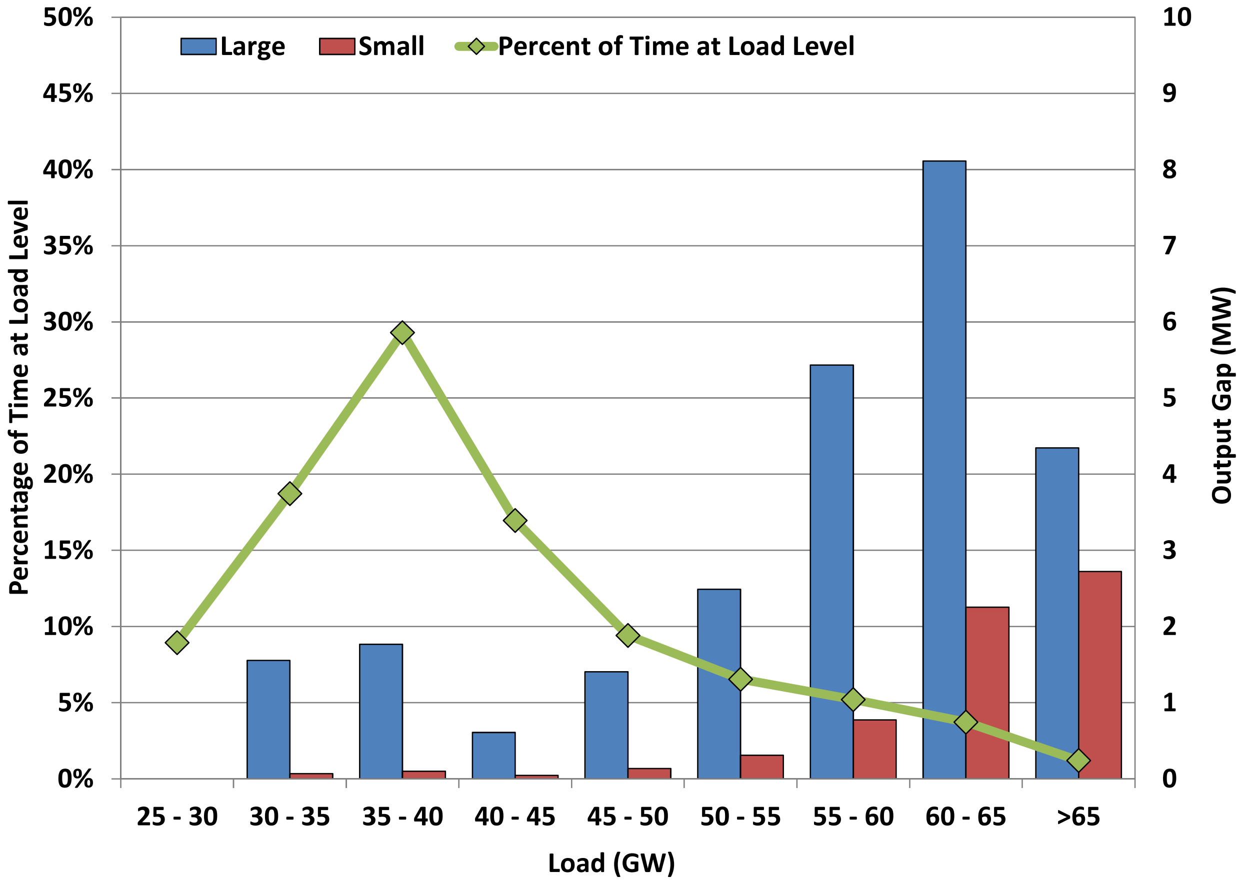
Before presenting the results of the output gap analysis, a description of ERCOT's two-step dispatch software is required. In the first step, the dispatch software calculates output levels (base points) and associated locational marginal prices using the participants' offer curves and only considering 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. It is only used to screen whether a market participant is withholding in a manner that may influence the reference price.

---

73 Given the low energy prices since 2016, the output gap margin has been reduced to $30 for purposes of this analysis. Prior to 2015, the State of the Market report used $50 for the output gap margin.

Figure 91: Incremental Output Gap by Load Level and Participant Size – Step 1

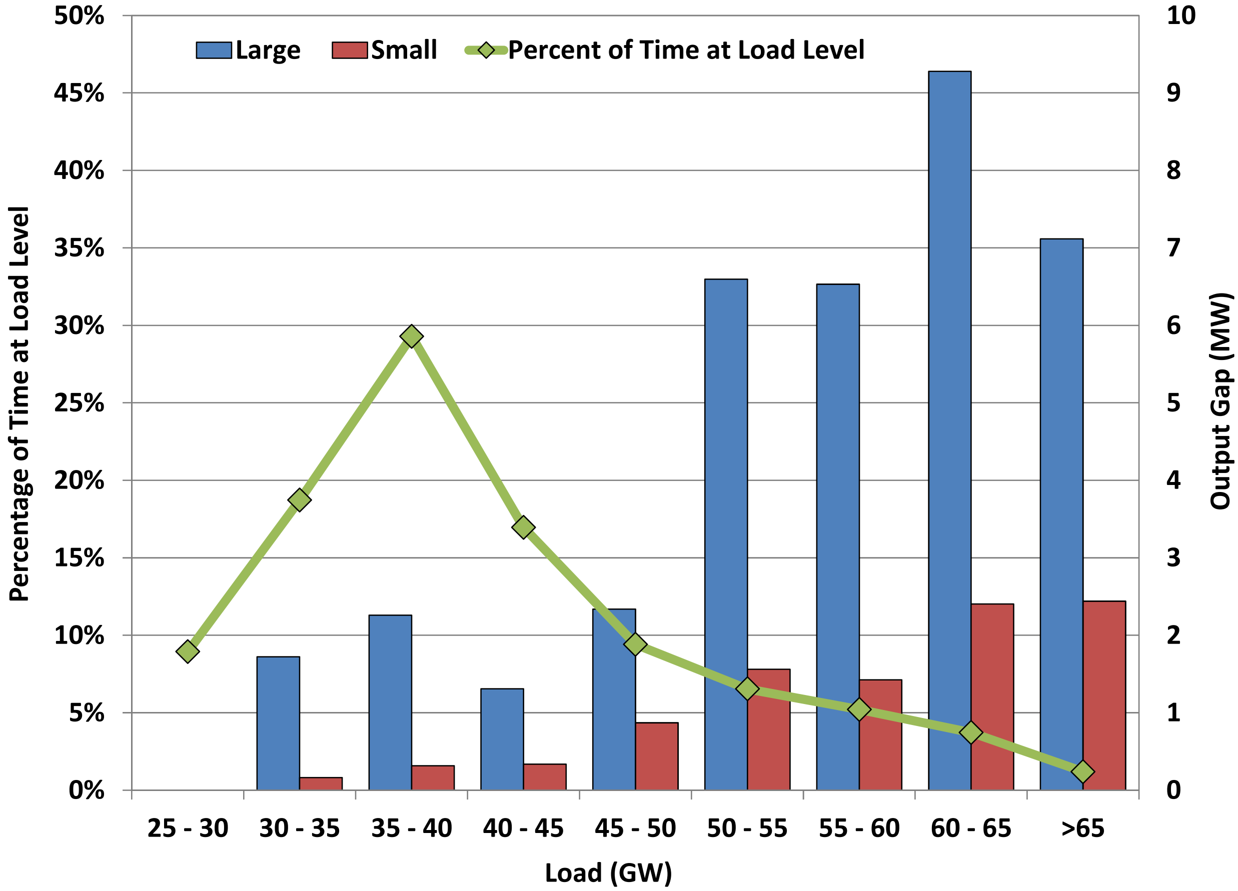


The results of the analysis shown in Figure 91 indicate that only very small amounts of capacity would be considered part of the first step output gap.

Figure 92 below 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 as a result of a high first-step reference price being influenced by a market participant raising prices.

Similar to the previous analysis, Figure 92 also shows very small quantities of capacity that would be considered part of this output gap.

Figure 92: Incremental Output Gap by Load Level and Participant Size – Step 2



These results show that potential economic withholding levels were extremely low for the largest suppliers and small suppliers alike in 2017. 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 2017.