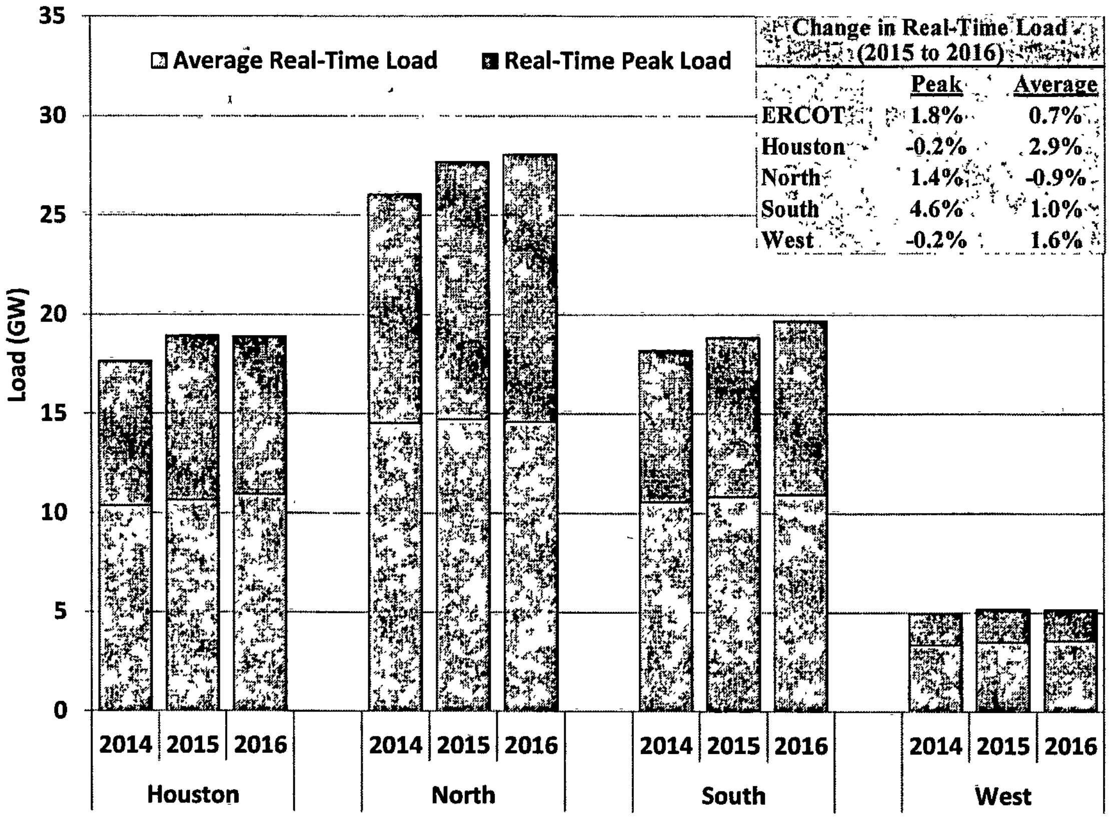
Figure 56: Annual Load Statistics by Zone

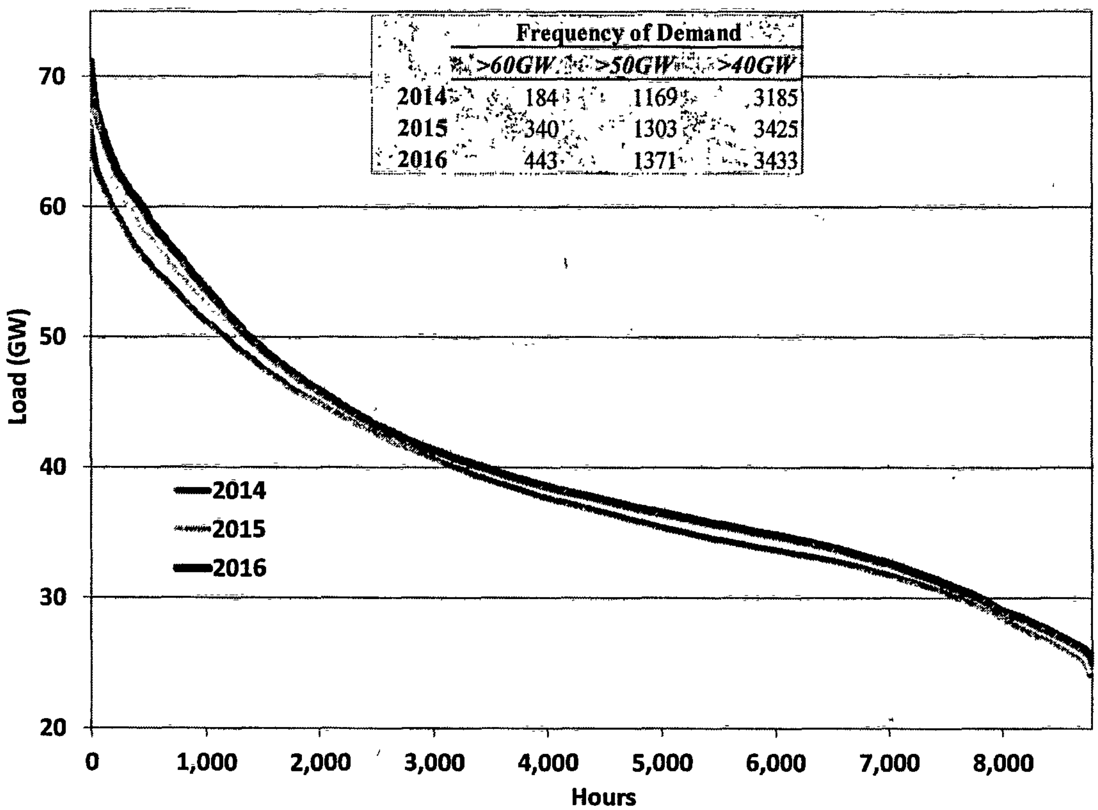


Total ERCOT load over the calendar year increased 1.1 percent (approximately 450 MW on average) to total 351.5 TWh in 2016. As 2016 was a leap year, the relative increase in the total load is higher than the increase in average load. With the exception of the North zone, all zones showed an increase in average real-time load in 2016. Houston saw the largest average load increase at 2.9 percent. Changes in average loads were largely explained by summer weather. Cooling degree days, a metric that is highly correlated with weather-related summer load, increased 4 percent on average from 2015 to 2016 in Houston and decreased 3 percent in Dallas. However, cooling degree days in 2016 were still 12 to 16 percent lower than ERCOT's hottest recent summer in 2011.

Summer conditions in 2016 also led to a new ERCOT-wide coincident peak hourly demand record of 71,110 MW on August 11, 2016. This broke the prior year's peak demand record of 69,877 MW that occurred on August 10, 2015. In fact, demand exceeded 70,000 MW five different times in 2016. The 2016 peak represents a 1.8 percent increase from the peak hourly demand of 2015. The zones experienced varying changes in peak load. Although the West zone had shown a prior trend of increasing load due to oil and gas production activity, that trend reversed in 2016 with a decrease in West zone peak load corresponding with a decline in oil and gas activity. Houston also showed a decrease in peak load. The South zone had the greatest increase in peak load at 4.6 percent.

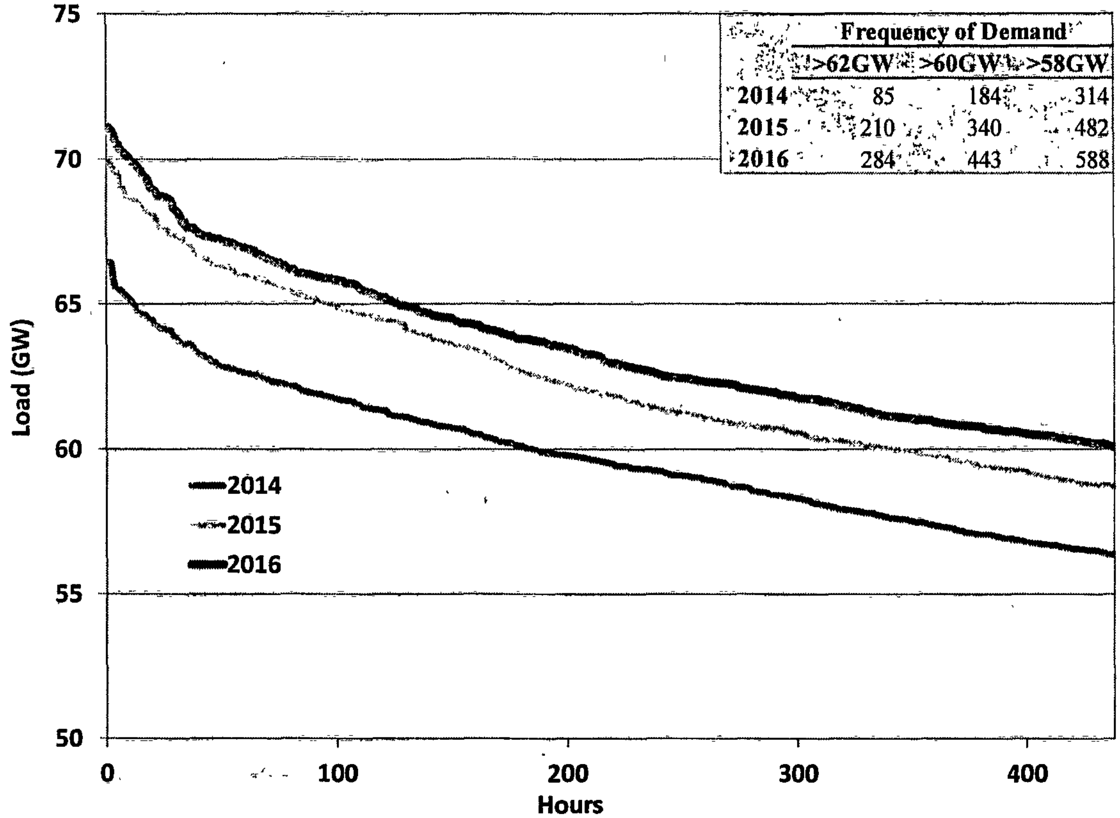
To provide a more detailed analysis of load at the hourly level, Figure 57 compares load duration curves for each year from 2014 to 2016. A load duration curve illustrates the number of hours (shown on the horizontal axis) that load exceeds a particular level (shown on the vertical axis). ERCOT has a 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 2016 is very similar to 2015, with a slight increase in the hours at the highest load levels.

Figure 57: Load Duration Curve – All Hours



To better illustrate the differences in the highest-demand periods between years, Figure 58 below shows the load duration curve for the 5 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 2016, the peak load averaged 18 percent greater than the load at the 95th percentile. These load characteristics imply that a substantial amount of capacity – more than 10 GW – is needed to supply energy in less than 5 percent of the hours.

Figure 58: Load Duration Curve – Top Five Percent of Hours



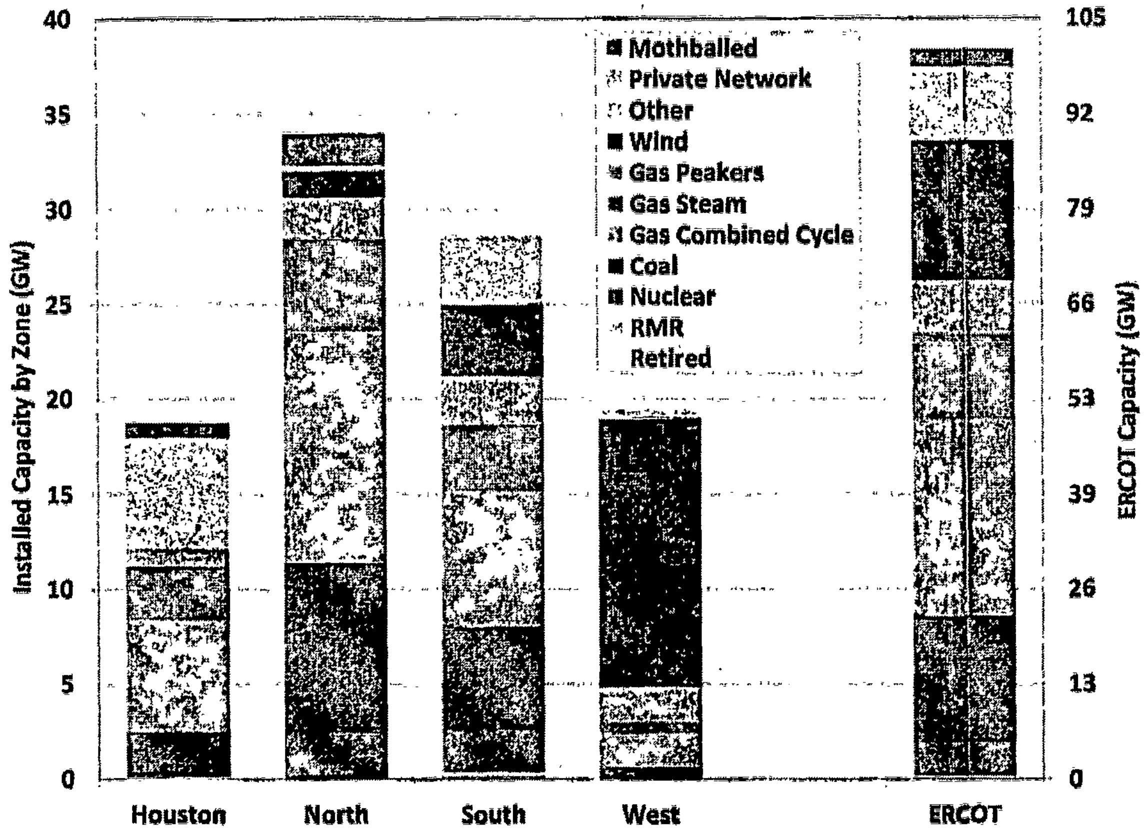
B. Generation Capacity in ERCOT

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. The North zone accounts for approximately 33 percent of capacity, the South zone 29 percent, the Houston zone 19 percent, and the West zone 19 percent. Excluding mothballed resources and including only the fraction of wind capacity available to reliably meet peak demand,[23] the North zone accounts for approximately 37 percent of capacity, the South zone 32 percent, the Houston zone 22 percent, and the West zone 9 percent. Figure 59 shows the installed generating capacity by type in each zone.

23 The percentages of installed capacity to serve peak demand assume wind availability of 14 percent for non-coastal wind and 58 percent for coastal wind.

[redacted]

Figure 59: Installed Capacity by Technology for Each Zone



Approximately 5.5 GW of new generation resources came online in 2016, but it only provided roughly 2 GW of net effective capacity. The overwhelming majority of new capacity was from wind generation. The 4.1 GW of newly installed wind capacity provides approximately 645 MW of capacity at summer peak. The remaining 1.4 GW of new capacity consisted of 370 MW of solar resources, 10 MW of storage resources, and approximately 1 GW of new natural gas combined-cycle units. Although still a small portion of the newly installed capacity, the installed solar megawatts in 2016 were more than three times the amount added in the prior year.

Considering these additions and retirements in 2016, natural gas generation decreased slightly from 48 percent of total ERCOT installed capacity in 2015 to 45 percent in 2016. The share of total installed capacity for coal generation also decreased slightly from 20 percent in 2015 to 17 percent in 2016.

Figure 60 shows the age of generation resources in ERCOT that were operational in the December 2016 Capacity, Demand, and Reserves Report.[24] The bulk of the coal fleet in ERCOT

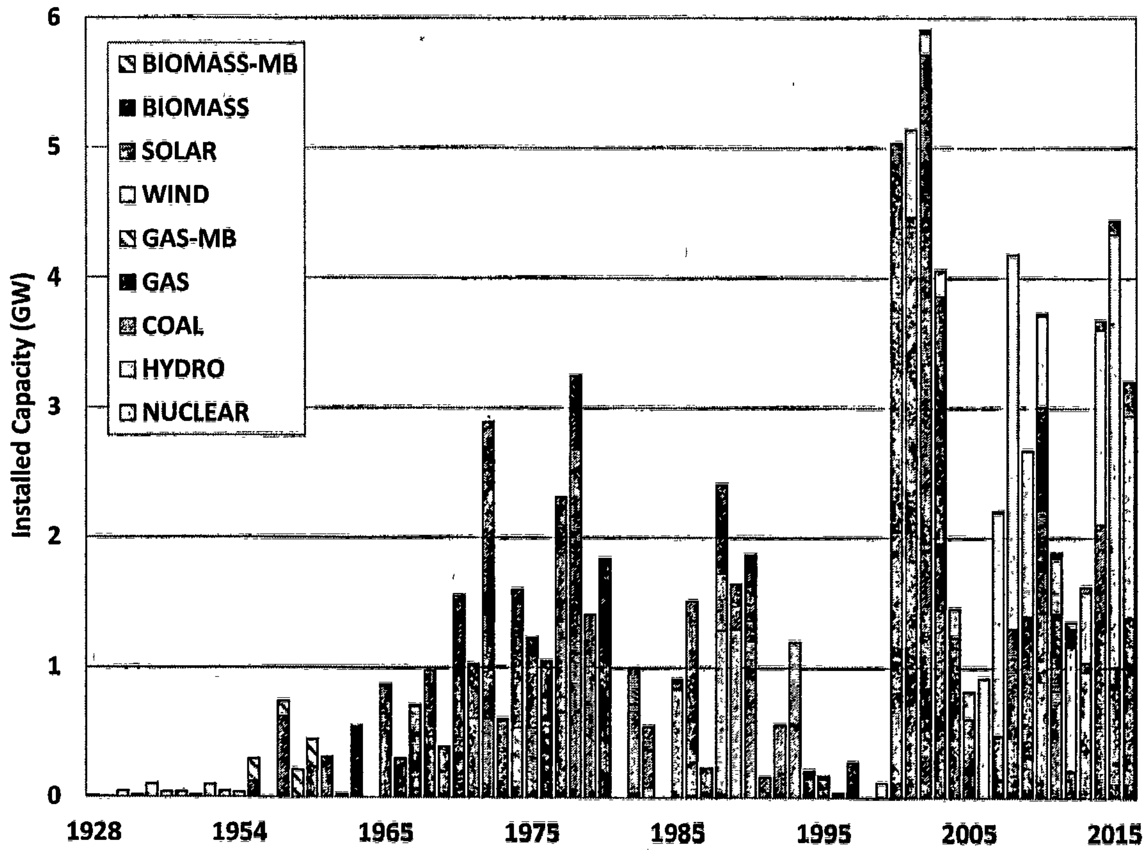
24 ERCOT Capacity, Demand, and Reserves Report (Dec. 2016), available at http://www.ercot.com/gridinfo/resource.

[redacted]

000103

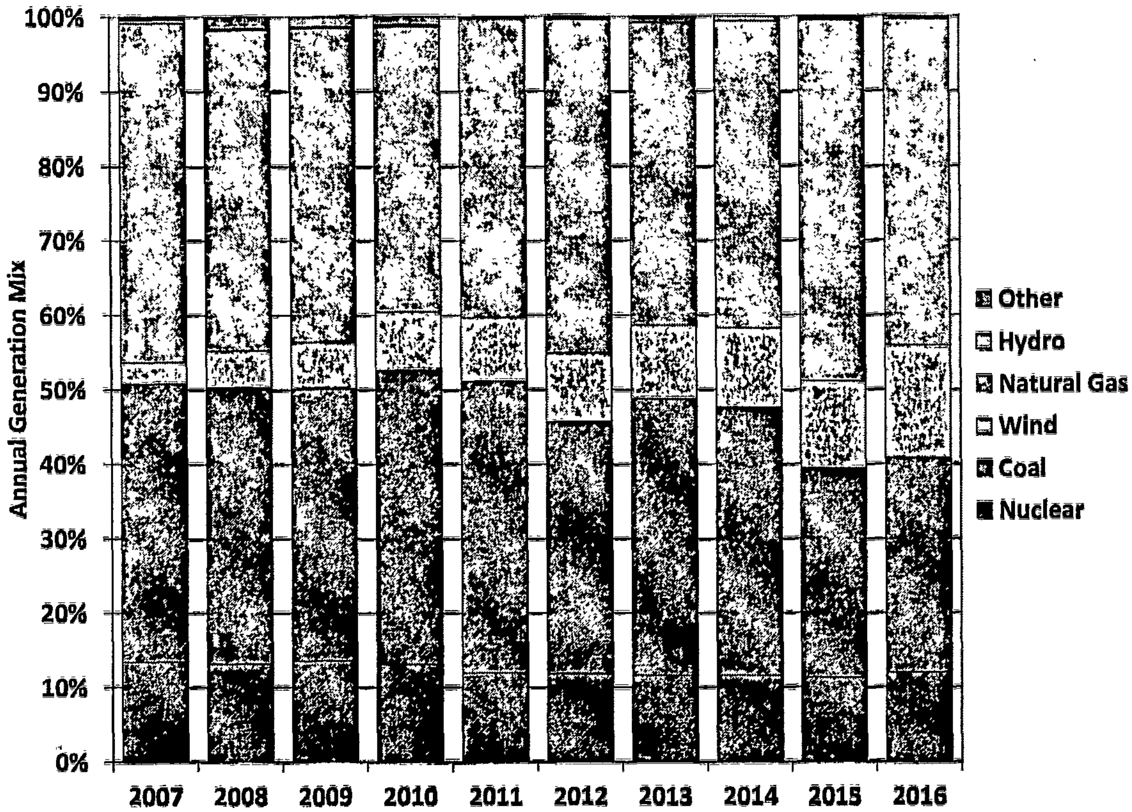
was built before 1990 and is approaching the end of useful life for this vintage of coal-fired power plants. When the ERCOT market was deregulated, there was a large increase in the construction of combined-cycle gas units. A few new coal units were added around 2010. As the figure demonstrates, wind capacity has been the dominant technology for newly installed capacity since 2006.

Figure 60: Vintage of ERCOT Installed Capacity



The shifting contribution of coal and wind generation is evident in Figure 61, which shows the percent of annual generation from each fuel type for the years 2007 through 2016.

Figure 61: Annual Generation Mix

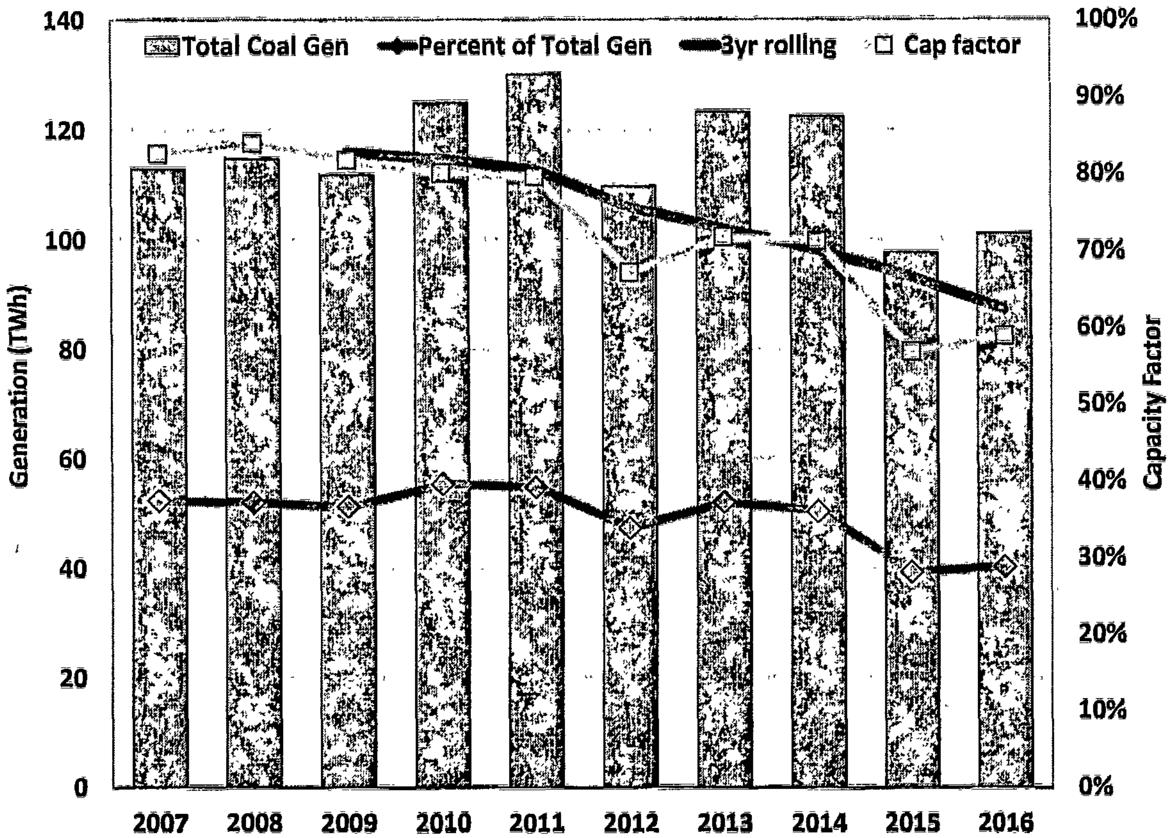


The generation share from wind has increased every year, reaching 15 percent of the annual generation requirement in 2016, up from 3 percent in 2007 and 12 percent in 2015. While the percent of generation from coal had declined significantly between 2014 and 2015, its share increased slightly to 29 percent in 2016. Natural gas declined from its high point in 2015 at 48 percent to 44 percent in 2016.

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 are approximately 24 GW of coal and nuclear generation in ERCOT. Generally, when ERCOT load is above this level, natural gas resources will be on the margin and set the real-time energy spot price.

Figure 62 shows the total coal generation, percent of total generation by coal, and the capacity factor for coal in years 2007 through 2016. The chart includes the annual capacity factor as well as the three-year rolling average capacity factor. While there was a slight increase in the coal capacity factor between 2015 and 2016, the three-year rolling average demonstrates the long-term decline in the coal capacity factor in ERCOT.

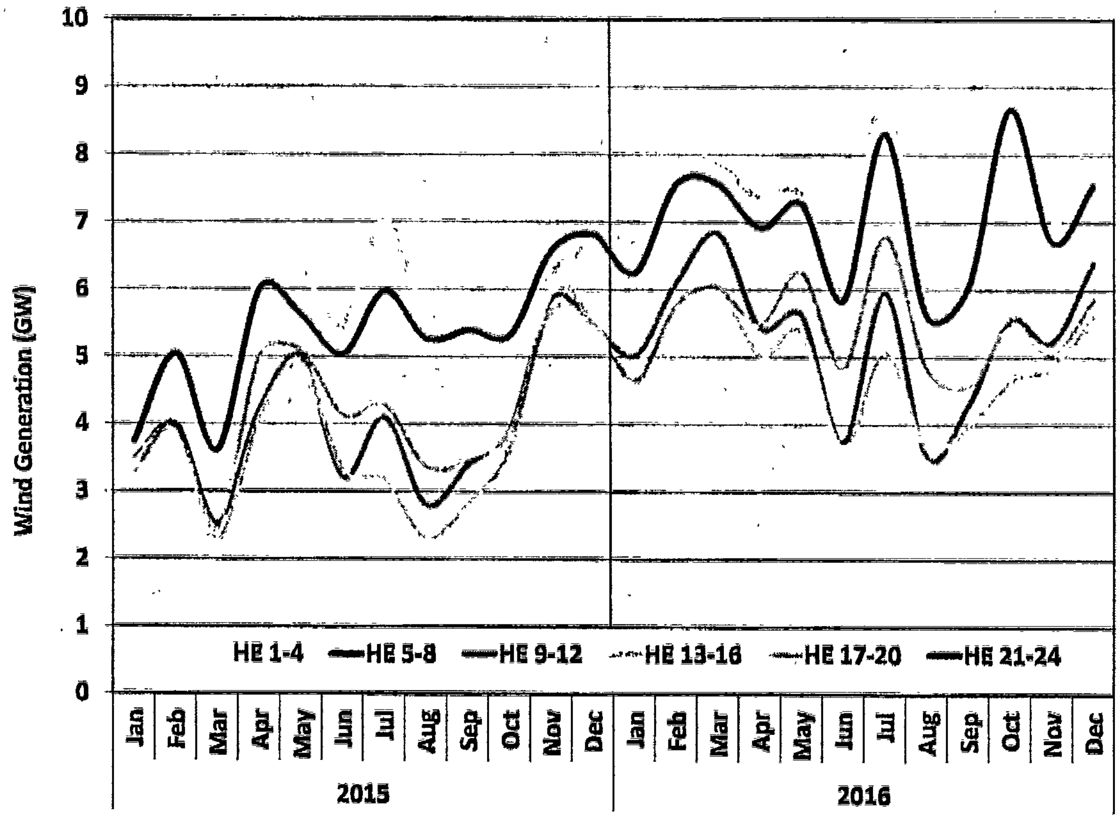
Figure 62: Historic Coal Generation and Capacity Factor



The amount of wind generation installed in ERCOT was approximately 19 GW by the end of 2016. Although the large majority of wind generation is located in the West zone, more than 3 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 2016, there were more than 50 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 63 shows average wind production for each month in 2015 and 2016, 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 4 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 63: Average Wind Production

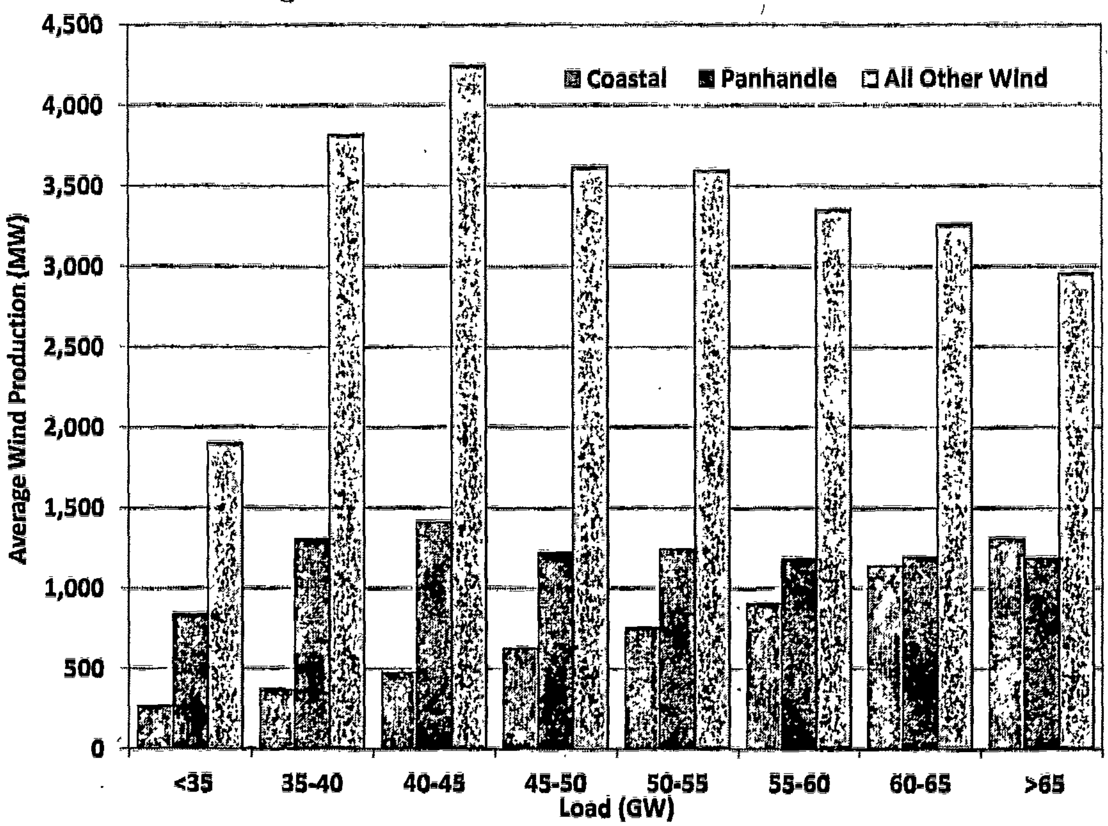


ERCOT continued to set new records for peak wind output in 2016. On December 25, wind output exceeded 16 GW, setting the record for maximum output and serving nearly 47 percent of the total load.

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 due to 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 64 below presents 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.

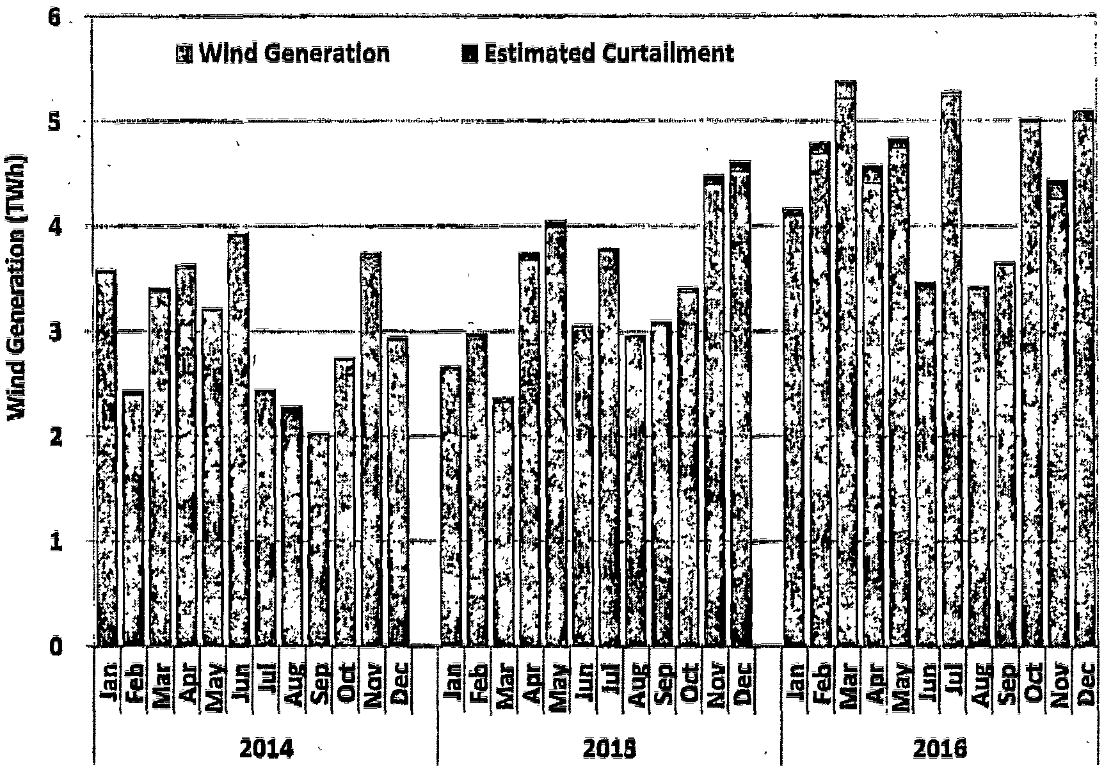
Figure 64: Summer Wind Production vs. Load



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

Figure 65 shows the wind production and estimated curtailment quantities for each month of 2013 through 2016.

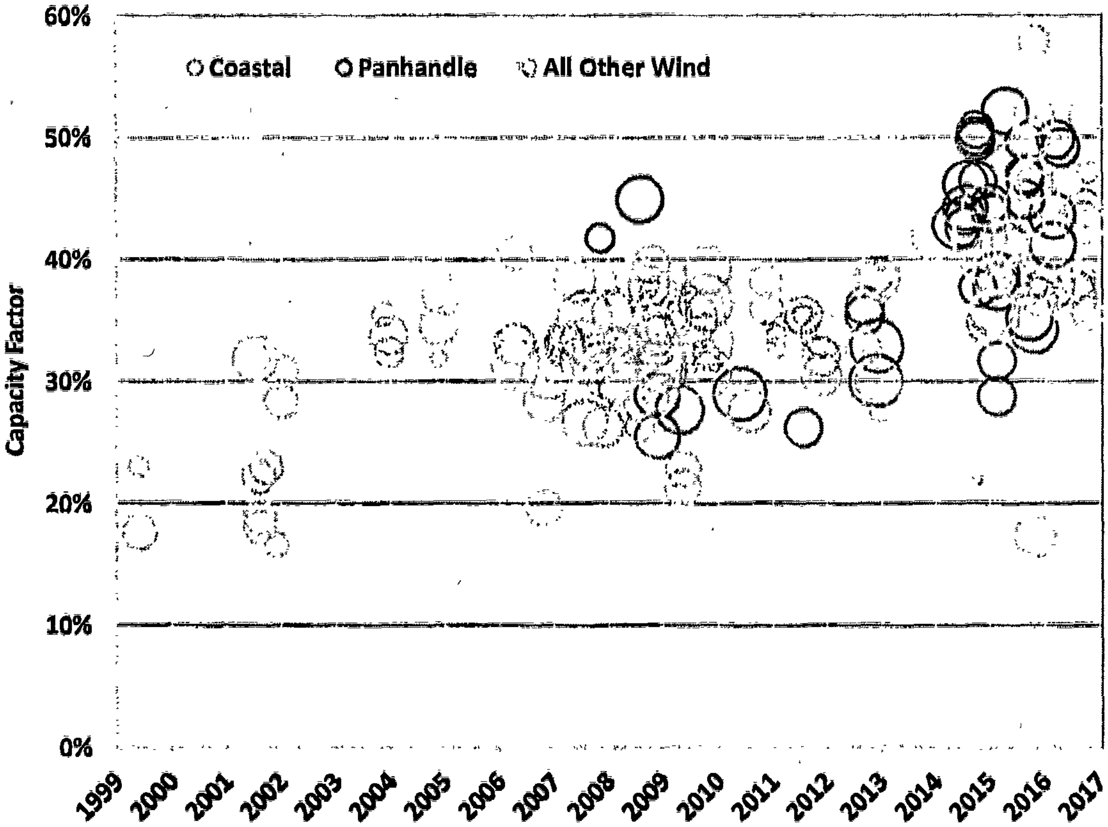
Figure 65: Wind Production and Curtailment



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 2016 was estimated at 98 percent of the total available wind, compared with 99 percent in 2015 and 99.5 percent in 2014. As a comparison, in 2009, the year with the most wind curtailment, the amount of wind delivered was only 83 percent.

Figure 66 shows the capacity factor for wind generators based on the year installed. Wind generation units located along the coast and in the panhandle are depicted with different colors because of the different wind profiles for these regions. Coastal wind generally has a lower annual capacity factor, but as previously described their 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 66: Wind Generator Capacity Factor by Year Installed

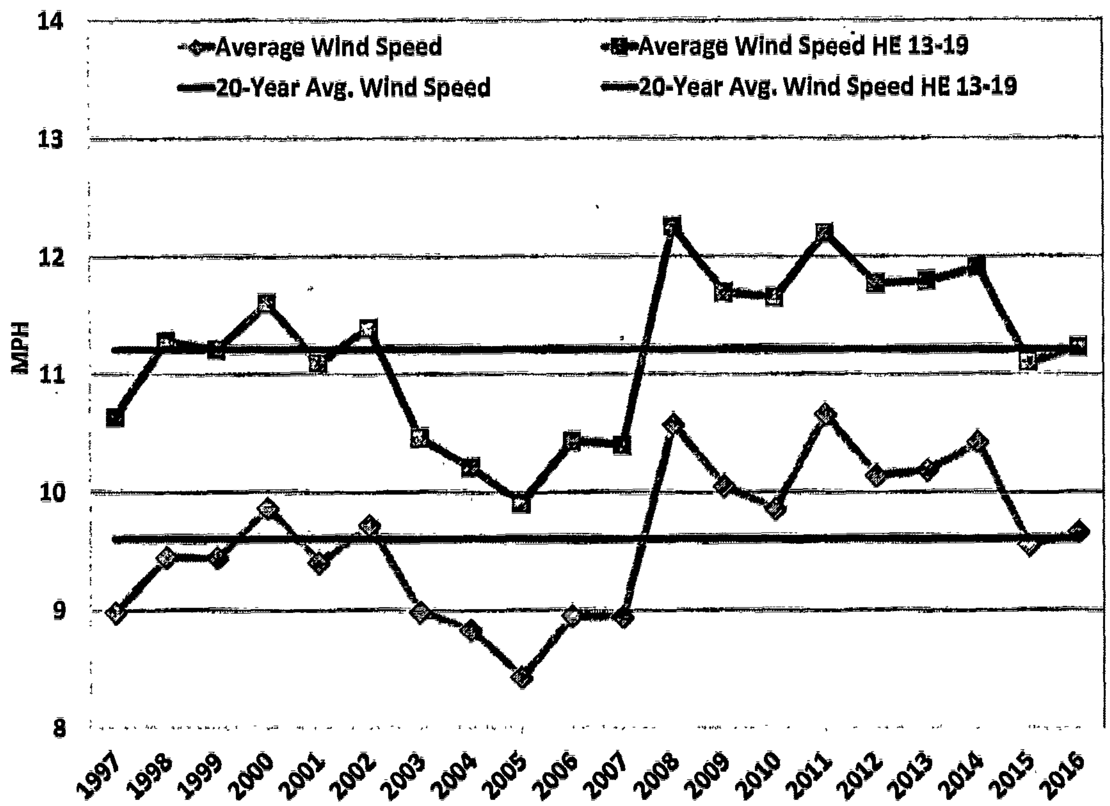


[redacted]

000110

The next figure shows average wind speeds in ERCOT, weighted by the current installed wind generation locations. Figure 67 provides a picture of the wind supply in 2016, averaged across the year and the average during peak hours, compared to the previous 20 years. The wind supply in 2016 was similar to the average over the past 20 years for all hours and for the peak hours ending 13-19. With 2016 being 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 67: 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 68 shows the net load duration curves for the years 2007, 2011, and 2016.

Figure 68 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 68: Net Load Duration Curves

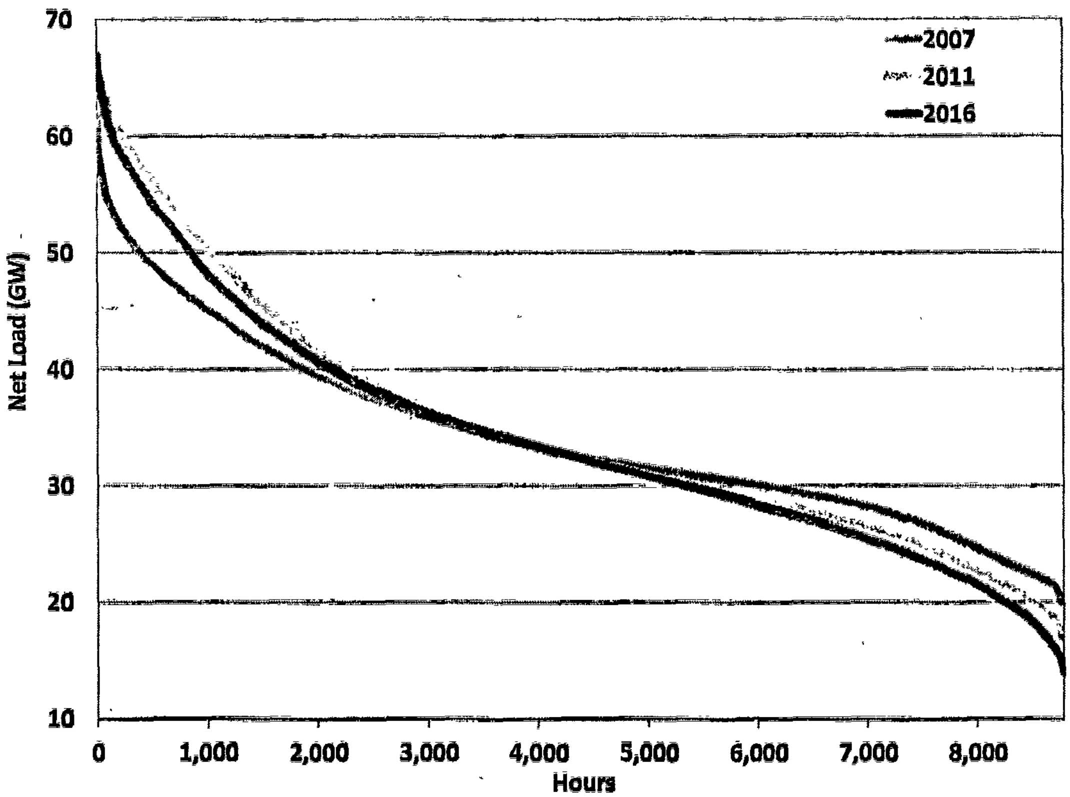
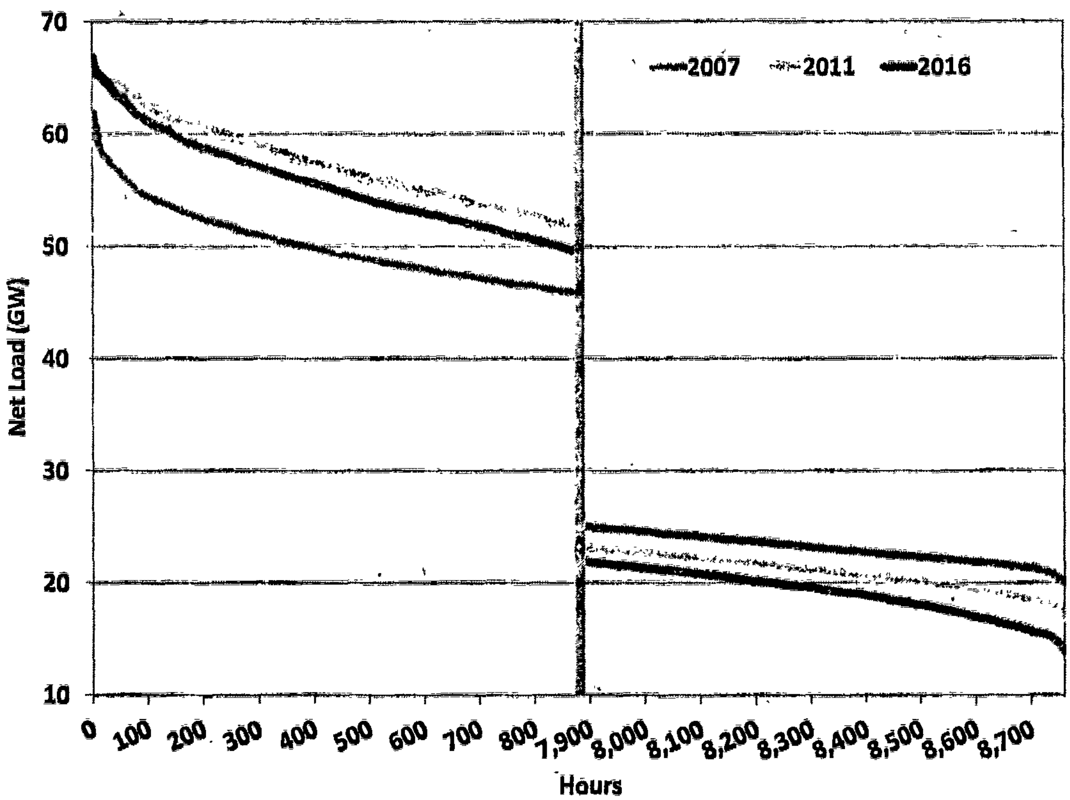


Figure 69 shows net load in the highest and lowest hours. Even with the increased development activity in the coastal area of the South zone, 73 percent of the wind resources in the ERCOT region are located in West Texas. The wind profiles in this area are such that most of the wind production occurs 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 69: Top and Bottom Ten Percent of Net Load

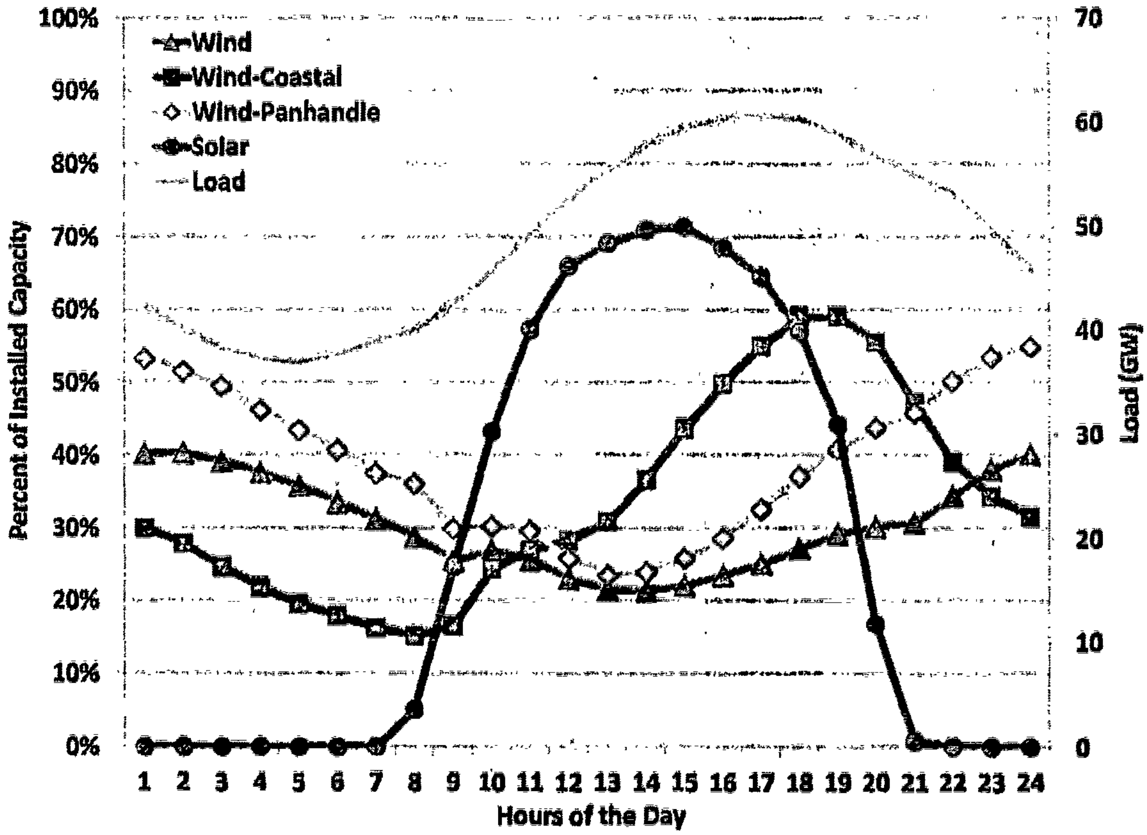


In the hours with the lowest net load (right side of the figure), the minimum net load has dropped from approximately 20 GW in 2007 to below 15.4 GW in 2016, even with the sizable growth in annual load that has occurred. This continues to put operational pressure on the 24 GW of nuclear and coal generation currently installed in ERCOT.

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 70 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 70: 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, around 70 percent, and producing more than 60 percent of its installed capacity during peak load hours.

The contrast between coastal wind and all other wind is also clearly displayed in Figure 70. Coastal wind produced over 50 percent of its installed capacity during summer peak hours. Output from Panhandle wind exceeded 30 percent, while output from all other wind (primarily West zone) was less than 30 percent during summer peak hours.

C. 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 transmission providers. Additionally, loads may self-dispatch by adjusting consumption in response to energy prices or by reducing consumption during specific hours to lower transmission charges.

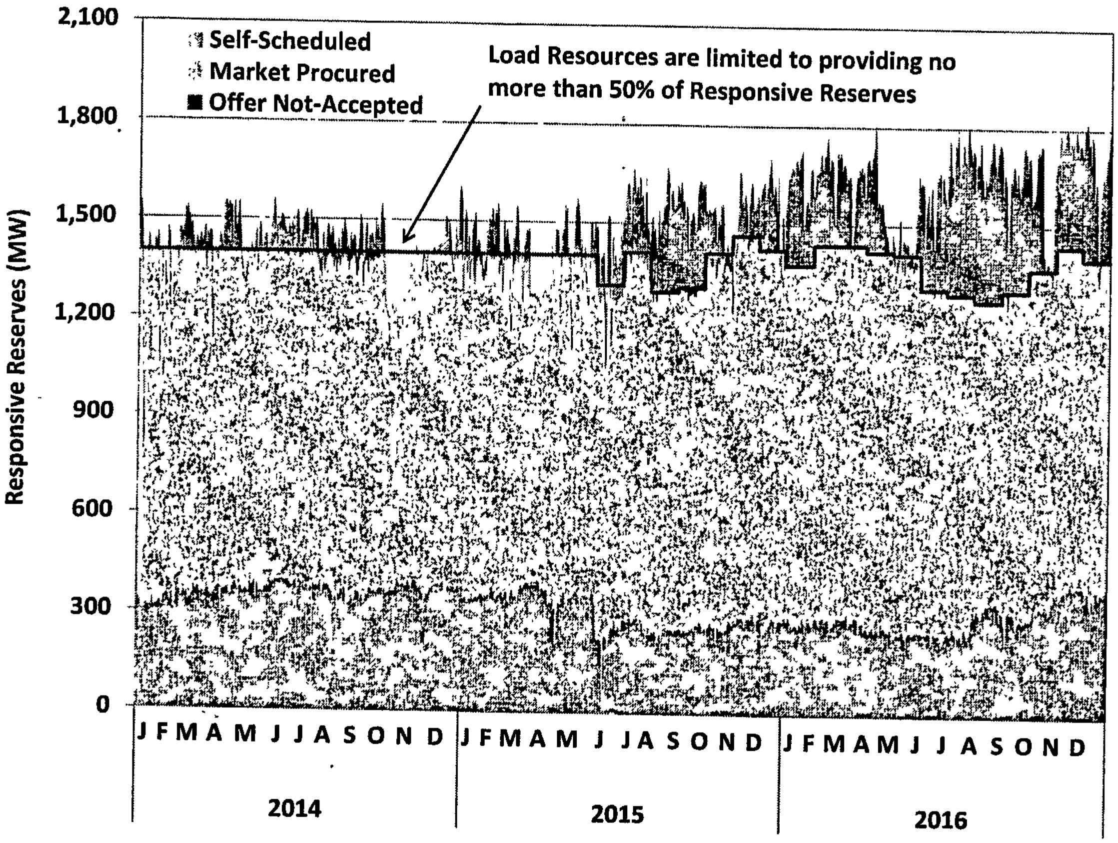
Reserve Markets

ERCOT allows qualified load resources to offer responsive reserves into the day-ahead ancillary services markets. Those providing responsive reserves have high set under-frequency relay equipment. This equipment enables the load to be automatically tripped when the frequency falls below 59.7 Hz, which will typically occur only a few times each year. As of December 2016, approximately 3,616 MW of qualified Load Resources were providing RRS, an increase of approximately 200 MW during 2016.

On June 1, 2015, ERCOT began procuring a variable amount of RRS based on season and time of day. The total amount of RRS varied between 2,300 to 3,000 MW. In 2016, the first full year with variable RRS 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 RRS. There was, however, one automatic deployment of 927 MW of frequency responsive load on May 1, 2016.

Figure 71 below shows the average amount of responsive reserves provided from load resources on a daily basis for the past three years.

Figure 71: Daily Average of Responsive Reserves Provided by Load Resources



In 2016, load resources were limited to providing a maximum of 50 percent of responsive reserves. The quantity of offers submitted by load resources exceeds the 50 percent limit most of the time. The exception is when real-time prices are expected to be high. Since 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 the energy can exceed the value received from providing responsive reserves. Reduced offer quantities observed during the spring and fall months may reflect the lack of availability of load resources due to annual maintenance at some of the larger load resource facilities.

ERCOT Protocols permit load resources to provide non-spinning reserves and regulation services, but for a variety of reasons there has been minimal participation by load resources.

Reliability Programs

There are two main reliability programs in which demand can participate in ERCOT – Emergency Response Service (ERS) and transmission provider load management programs. The ERS program is defined by a PUCT Rule enacted in March 2012 setting a program budget of

$50 million.[25] 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 paid for ERS over the contract periods from February 2016 through January 2017 was $6.86 per MWh, significantly higher than the average price of $3.91 per MWh paid for non-spinning reserves in 2016. ERS was not deployed in 2016.

Beyond ERS there are slightly less than 200 MW of load participating in load management programs administered by transmission providers.[26] 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 providers 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 and/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.

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. Over the last three years, transmission costs have risen by more than 60 percent, thus significantly increasing an already substantial incentive to reduce load during probable peak

25 *See* 16 TEX. ADMIN. CODE § 25.507.

26 *See* ERCOT 2016 Annual Report of Demand Response in the ERCOT Region (Mar. 2017) at 6, available at http://www.ercot.com/services/programs/load.

intervals in the summer.[27] ERCOT estimates that 835-1,491 MW of load were actively pursuing reduction during the 4CP intervals in 2016, an increase from the estimated response in 2015.[28]

Load curtailment to avoid transmission charges may be resulting in price distortion during peak demand periods since 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.

Two recent changes in the ERCOT market have made advances in appropriately 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.

27 Transmission Cost of Service (TCOS) in 2013 was $2 billion and for 2016 it was $3.2 billion. See PUCT Docket No. 40946, Commission Staff's Application to Set 2013 Wholesale Transmission Service Charges for the Electric Reliability Council of Texas, Final Order (Mar. 28, 2013) and 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).

28 *See* ERCOT, *2016 Annual Report of Demand Response in the ERCOT Region* (Mar. 2017) at 8, 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 outcomes help to inform these decisions, but ERCOT's day-ahead market 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. There can be gaps 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 RUC commitments, and the price mitigation that occurs during RUC and local congestion. The section concludes with a discussion of the reliability must run procurement by ERCOT in 2016.

A. History of RUC-Related Protocol Changes

The RUC process has undergone several modifications since the nodal market began. 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.[29] Initially, the RUC offer floor was set at the system-wide offer cap. The RUC offer floor was subsequently adjusted to $1,000 per MWh[30] and then to the current offer floor of $1,500 per MWh.[31]

29 NPRR435, Requirements for Energy Offer Curves in the Real Time SCED for Generation Resources Committed in RUC, implemented on March 1, 2012.

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

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

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.[32] 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 commitment.[33]

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).[34] Since that date, when a resource properly telemeters a status indicating it has been RUC committed, ERCOT systems automatically set the energy offer floor at $1,500 per MWh. 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.

To provide even greater flexibility to resource owners, the RUC process will soon be modified to permit the ability to opt-out of RUC instructions given after the close of the adjustment period. NPRR744 modifies the opt-out trigger to real-time telemetry status rather than the COP submittal. This NPRR is expected to be implemented mid-year 2017.

During 2016, approximately 40 percent of RUC instructions were given after the close of the adjustment period, thereby foreclosing the opportunity for resources to self-commit the units and shoulder the market risk. The late RUC commitments, however, demonstrate ERCOT exercising restraint in waiting as long as possible for the market to respond before committing resources through RUC.

B. RUC Outcomes

ERCOT continually assesses the adequacy of market participants' resource commitment decisions using a reliability unit commitment (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 resolve a transmission constraint. The constraint may be either a thermal limit or a voltage concern.

32 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.

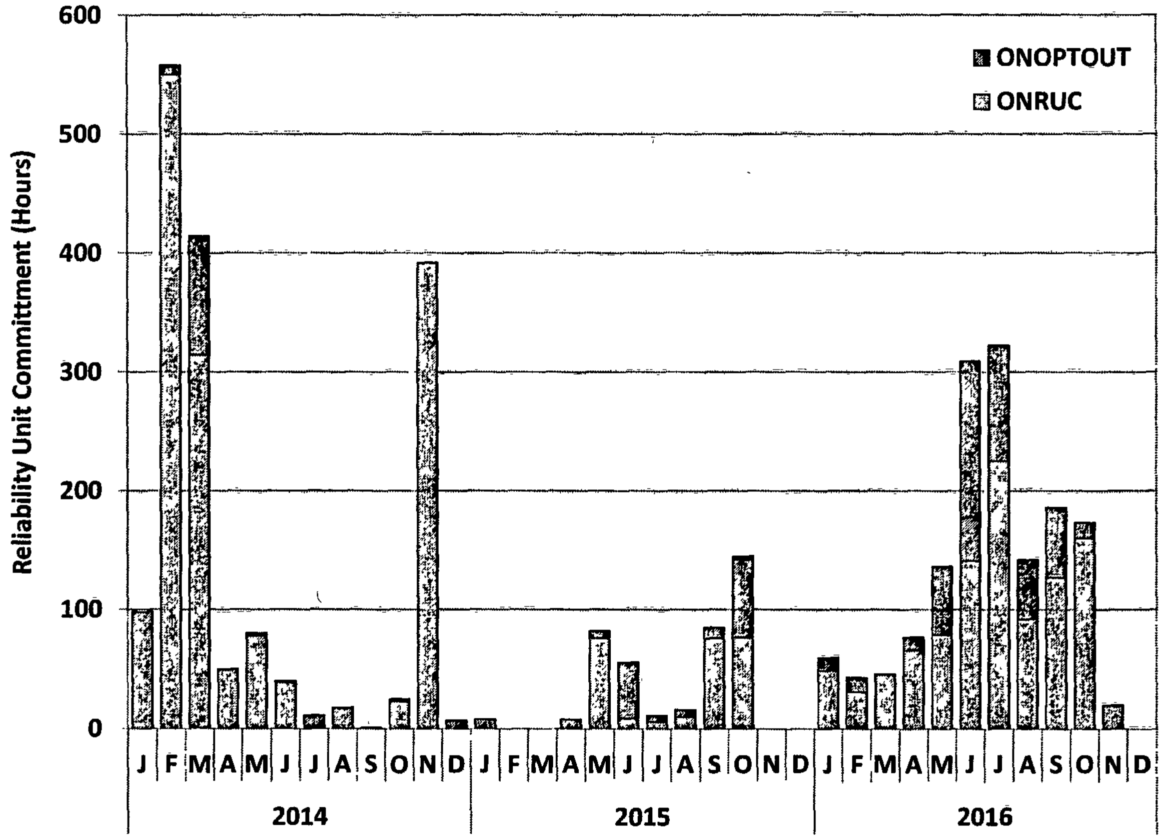
33 Note that the process for electing to opt-out of a RUC will be based on real-time telemetry when NPRR744, RUC Trigger for the Reliability Deployment Price Adder and Alignment with RUC Settlement, goes into effect in mid-2017.

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

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 that unit's costs, payment to that unit is reduced (RUC clawback charge). Beginning in January 2014, a unit receiving a RUC instructions had the choice to "opt out," meaning it would forgo all RUC make-whole payments in return for not being subject to RUC clawback charges.

Figure 72 shows how frequently these reliability unit commitments have occurred over the past three years, measured in unit-hours.

Figure 72: Frequency of Reliability Unit Commitments



RUC commitments in 2016 were more frequent than in recent years. Although the total unit-hours were similar to the unit-hours in 2014, they were much more consistent in 2016. Almost twelve percent of hours in 2016 had at least one unit receiving a reliability unit commitment instruction. The reliability commitments in 2016 were primarily made to manage transmission constraints (98 percent of unit-hours), most of which were made to manage persistent congestion in the Houston area and in the Rio Grande Valley. The RUC activity in 2014 was concentrated during cold weather events in February and March and in response to transmission outages in

March and November. In 2015, RUC commitments were most frequent in the fall due to congestion in Dallas and the Rio Grande Valley.

During 2016, QSE telemetry of RUC status served as the trigger for calculating a reliability adder. There were 740 hours in which units were settled as RUC in 2016 and less than 500 cumulative hours of pricing intervals with non-zero reliability adders that occurred coincident with a settled RUC hour.

Table 9 provides the units most frequently called upon for RUC. Also provided are the hours of RUC instruction, the number of hours in which the unit opted-out, and the average low-sustained limit (LSL) for the unit. In 2016, units receiving RUC instructions successfully opted-out of 31.5 percent of unit-hours. The units highlighted in gray on Table 9 are units that were also on the most-frequent RUC commitment list in 2015.

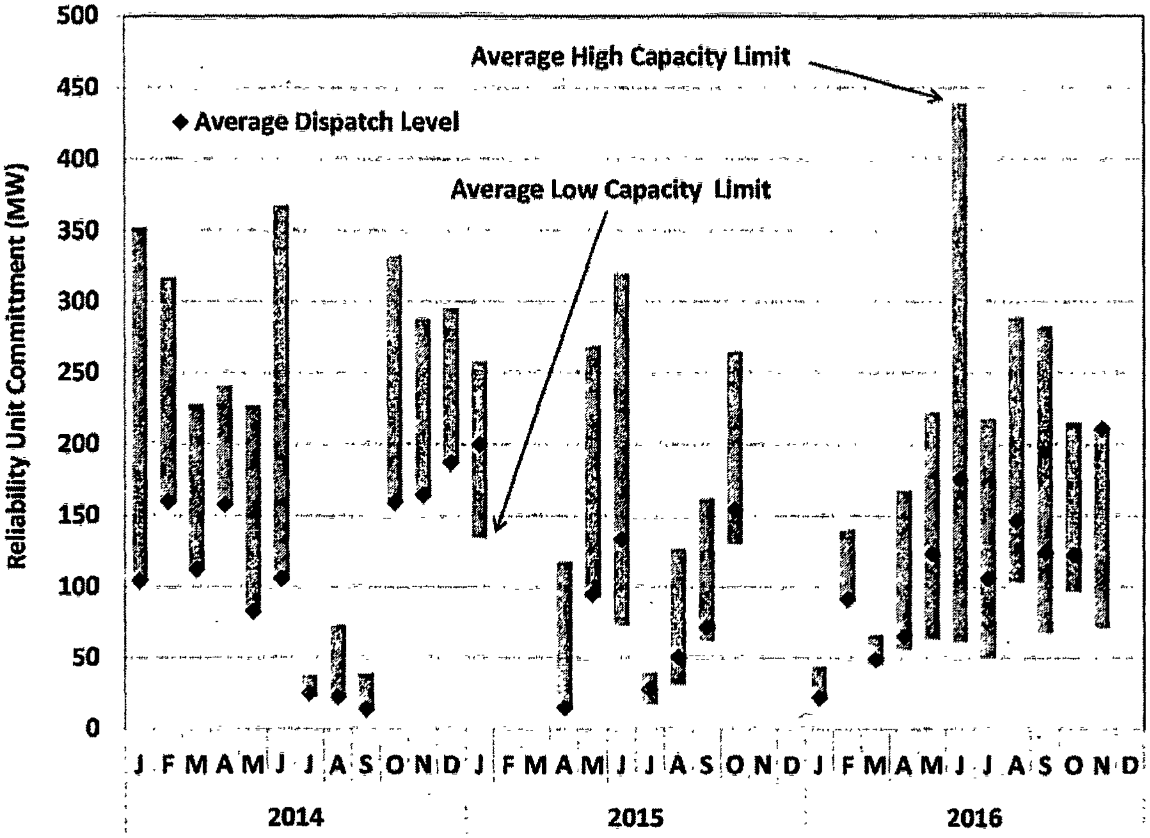
Table 9: Most Frequent Reliability Unit Commitments

Resource	Location	Unit RUC Hours	Unit OPTOUT Hours	Average LSL during RUC Hours
Silas Ray CC1	Valley	165	43	40
WA Parish G4	Houston	46	83	102
Silas Ray 10	Valley	83	28	21
WA Parish G2	Houston	53	34	29
Barney Davis G1	Corpus Christi	8	66	55
Spencer 5	Denton	54	13	17
Cedar Bayou G2	Houston	57	9	168
WA Parish G1	Houston	47	19	28
WA Parish G3	Houston	27	32	65
Cedar Bayou G1	Houston	-	51	-
Barney Davis CC1	Corpus Christi	43	3	238
North Edinburg CC1	Valley	32	8	222
Laredo G5	Laredo	35	-	35
Mountain Creek Unit 7	Dallas	33	-	15
Nueces Bay CC1	Corpus Christi	24	8	173

There were 1514 unit-hours with RUC instructions in 2016, compared with 411 unit-hours with RUC instructions during 2015. The majority of the RUC commitments were to resolve localized thermal transmission constraints (98 percent), and of those the majority were to units located in the Houston area (33 percent) and in the Rio Grande Valley (24 percent). There were 33 unit-hour commitments (2 percent) for system-wide capacity requirements. There were no commitments for voltage in 2016. Comparing 2016 to 2015 shows the same percent of RUC commitments for system-wide capacity at 2 percent; however, the total hours for system-wide capacity were significantly less in 2015 at only 8 unit-hours.

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 73 shows that the quantity of reliability unit commitment generation increased in 2016 compared to the prior two years. This figure shows that the average quantity dispatched for May through October 2016 exceeded 100 MW, and in November exceeded 200 MW.

Figure 73: Reliability Unit Commitment Capacity

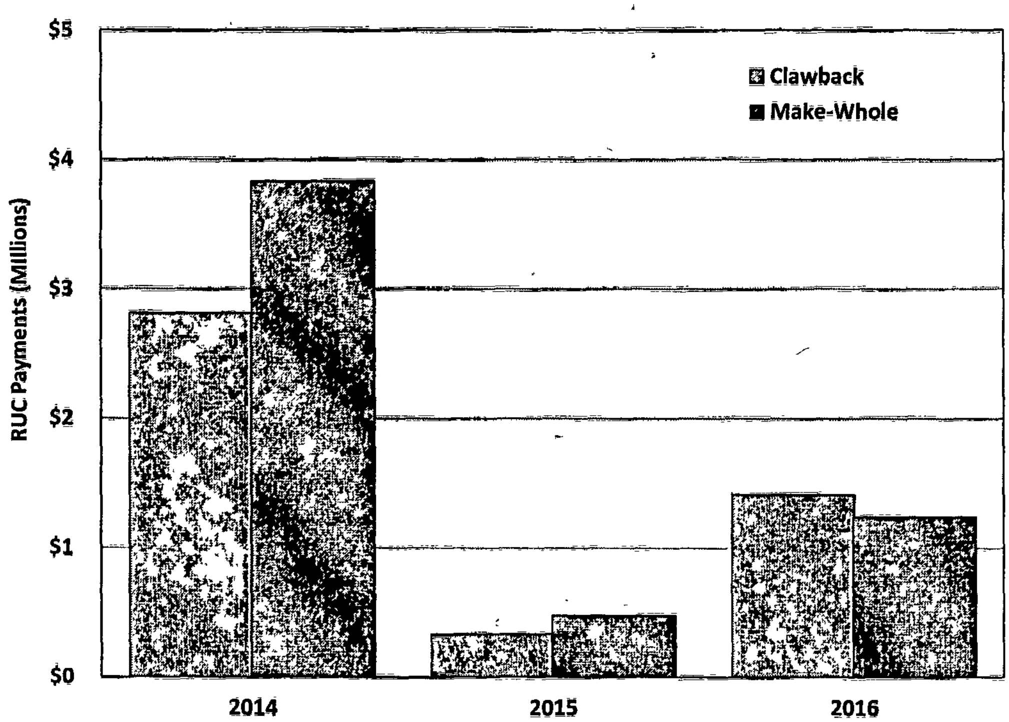


Units committed for RUC in 2016 showed a significant increase in the dispatch level compared to prior years. In twelve percent of intervals with RUC-committed resources, one or more resources were dispatched above their Low Dispatchable Limit (LDL), whereas in prior years, resources receiving RUC commitments were infrequently dispatched above LDL. Nonetheless, the higher dispatch levels in 2016 were rarely dispatched at the $1,500 per MWh offer floor because the commitments to address localized congestion were frequently mitigated.

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. Approximately 50 percent of resources in ERCOT have unit-specific verifiable costs. Of the 61 different resources that received a RUC instruction in 2016, 53 resources had approved unit-specific verifiable costs for start-up costs and minimum load costs. Those 53 resources represent 93 percent of total RUC-instructed megawatt-hours in 2016.

Figure 74 displays the total amount of make-whole payments and clawback charges attributable to reliability unit commitments annually for 2014-2016. Units that are RUC committed are guaranteed to be paid start-up and minimum energy costs. To the extent that the real-time energy market does not provide sufficient revenue to cover these costs, RUC-committed resources will receive a make-whole payment. 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 74: RUC Make-Whole and Clawback

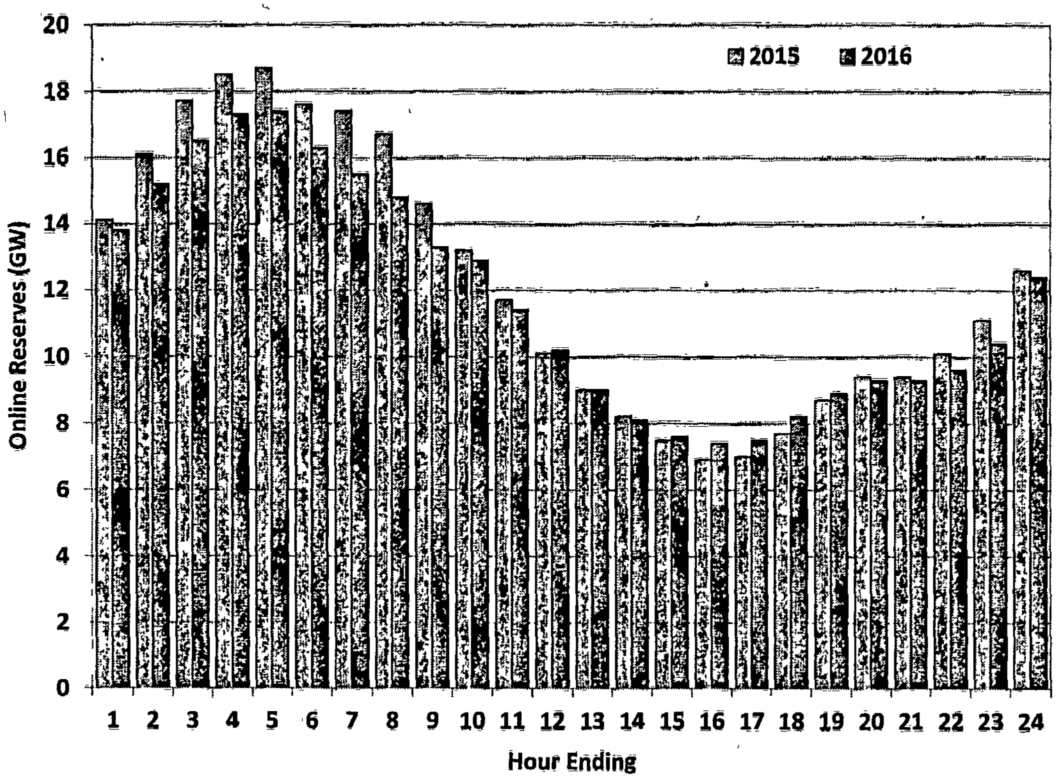


If real-time revenues received by a RUC committed resource exceed the operating costs incurred by the unit, then excess revenues are clawed-back and returned to QSEs representing load. During 2016, the make-whole and clawback amounts were nearly equal, with only slightly higher clawback charges. The source of funds for all RUC make-whole payments in 2016 were from QSEs that were capacity short. There was no general uplift to loads for RUC make-whole payments in 2016. The magnitude of both the clawback and make-whole amounts are very small in the scheme of the overall 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.

The following figure compares the amount of on-line reserves, by hour, for the summer months of June through August in 2016 and 2015. The amount of on-line reserves is equal to the amount of capacity committed in excess of expected demand. Figure 75 displays available online reserves by operating hour and shows the expected pattern of declining reserves as system load increases during peak demand hours. In 2016, the average online reserves were greater than in 2015 for hours ending 12 through 19; in all other hours, the average online reserves were less than 2015.

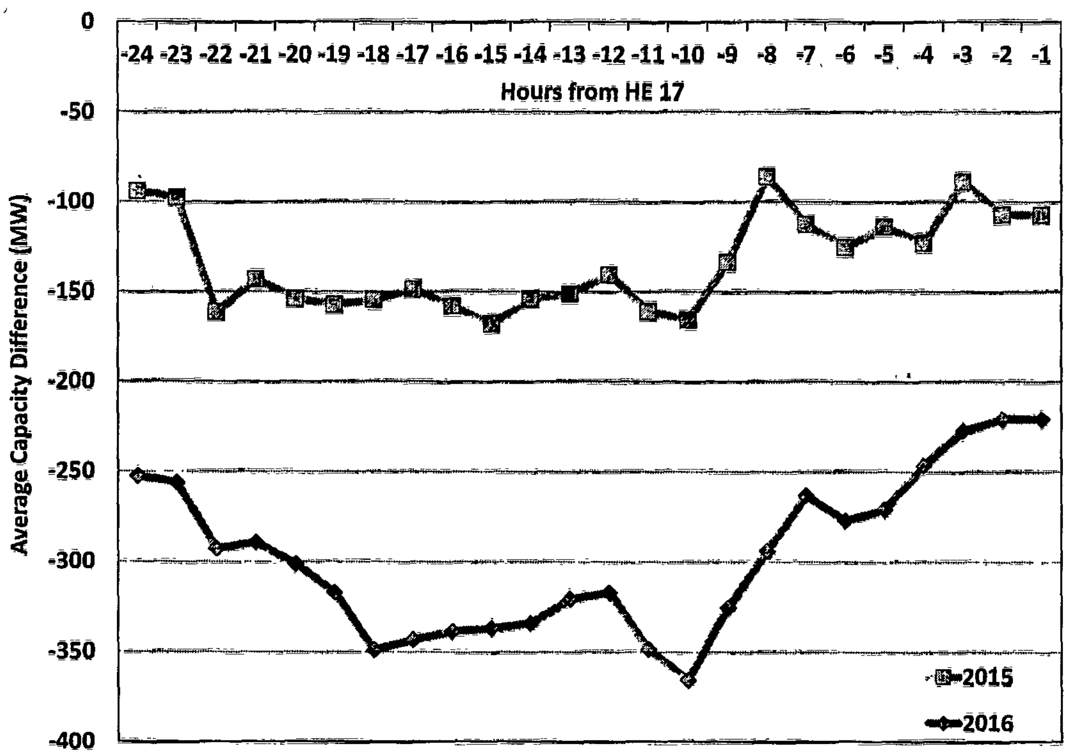
Figure 75: Average On-line Summer Reserves



The reduction in reserves during off-peak hours of the summer 2016 indicates that resource owners chose not to run units overnight. However, despite higher load levels during peak hours in 2016, average on-line summer reserves levels during peak hours were greater than in 2015. Lower energy prices are expected during periods of higher reserves.

For a different look at self-commitment during the summer of 2016, Figure 76 shows the average difference between the actual online unit capacity in the peak hour and the amount of capacity committed for the peak hour by the online units for each of the 24 hours leading up to the close of the adjustment period. This data is for hour ending 17, averaged over the months of July and August for 2015 and 2016. As can be seen from this chart, the amount of capacity committed in advance of the operating hour was less in 2016 than 2015. In 2015 about 100 MW of capacity, on average, was committed in the last hour before real time. In 2016, the amount increased to over 200 MW, with even larger deficiencies seen in the last hours leading up to real time. From an ERCOT operator perspective, the self-commitment by market participants appears deficient and may be a potential contributor to the increased RUC activity in 2016.

Figure 76: Capacity Commitment Timing – July and August Hour 17

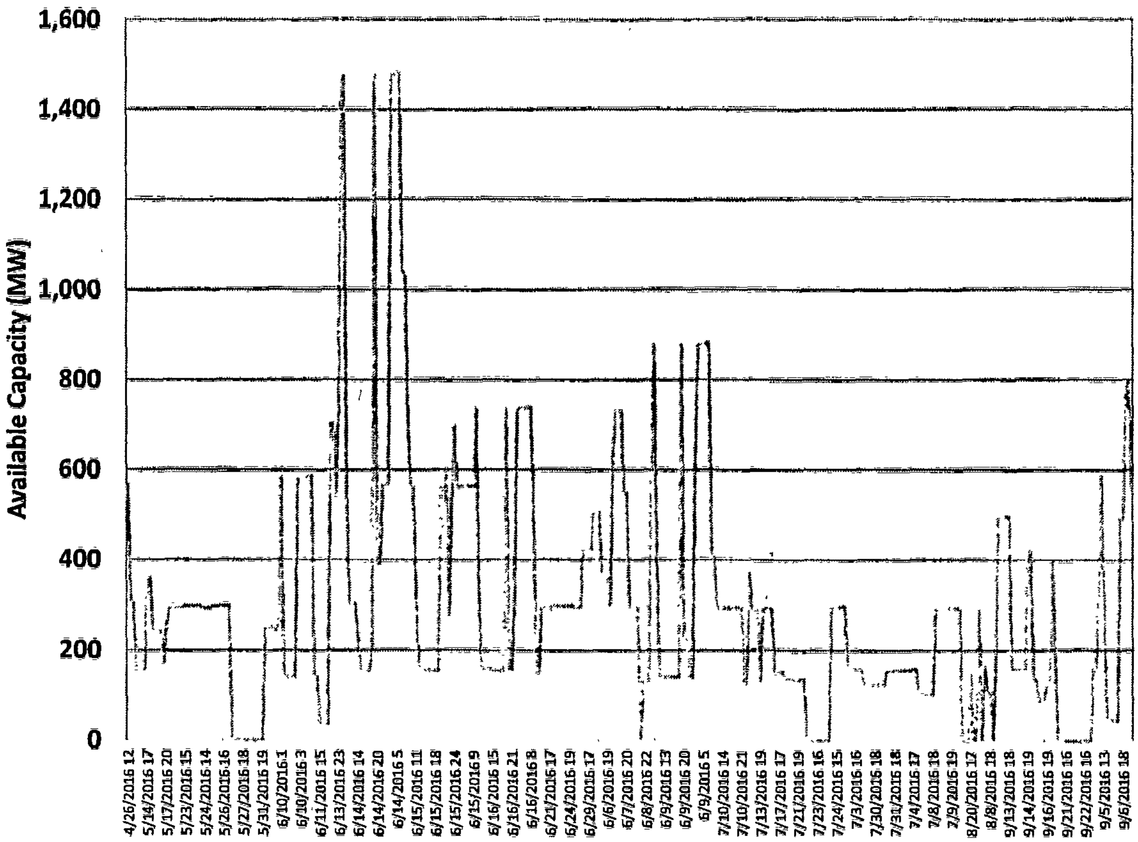


The last analysis of RUC activity in 2016 quantifies the amount of incremental combined-cycle capacity currently unavailable for RUC. Combined-cycle generators are comprised of multiple individual units, gas turbines and steam turbines that may be operated in various combinations. These different combinations, or configurations, have different operating characteristics and costs reflected in ERCOT systems. A common type of combined-cycle unit in ERCOT is

[redacted]

comprised of two gas turbines and one steam turbine. When the resource operates in a configuration with only one gas turbine and the steam turbine, ERCOT's RUC software does not recognize the additional capacity available from the second gas turbine. This inability of the RUC software to evaluate changes to combined-cycle configurations may lead to situations where other, potentially more costly units receive RUC instructions to come online. A preliminary analysis was performed to quantify the amount of additional capacity available from combined-cycle units that had self-committed in a configuration less than the unit's largest capacity configuration. Figure 77 below displays the additional combined cycle megawatts located in Houston that could have been made available to RUC during the hours that at least one unit in Houston received a RUC instruction. These values exclude any incremental capacity from private use network resources.

Figure 77: Potential for Combined Cycle Capacity Available to RUC in Houston



The changes required to the RUC process to account for larger configurations of combined-cycle resources would be complex, including changes to the RUC engine and settlement systems. In addition, market participants would be required to provide significantly more detailed information on combined-cycle configurations. Given the relatively low overall cost to the

[redacted]

market for RUC make-whole payments, implementing such a change may not be cost effective. However, the data indicates a sizable amount of incremental capacity is available.

C. Mitigation

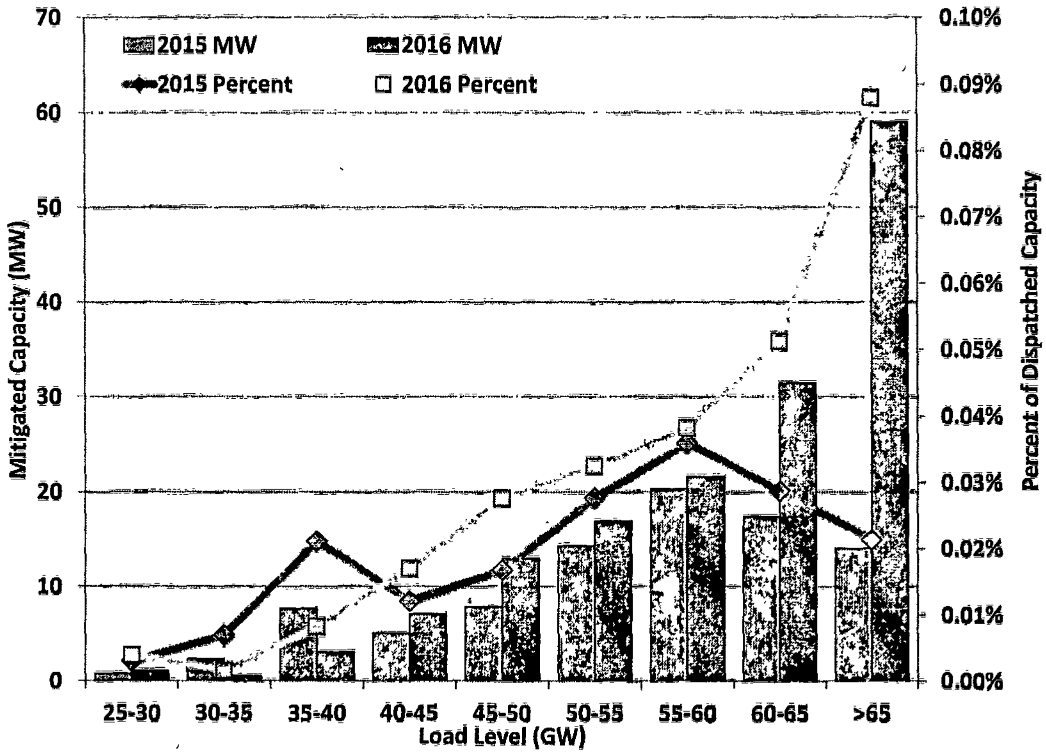
In situations where competitive forces are not sufficient, it can be necessary to mitigate prices to a level that approximates competitive outcomes. ERCOT's real-time market includes a mechanism to mitigate prices for resources that are required to resolve a transmission constraint. Mitigation applies whether the unit is self-committed or RUC committed. Units are typically RUC committed to resolve transmission constraints and as such they are typically required to resolve a transmission constraint, and therefore mitigated. As shown previously in Figure 73, it was more common for RUC-committed units to be dispatched above their low operating limits in 2016. This higher dispatch was due to the RUC-committed 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 2016 is analyzed. Although executing all the time, 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 percent of capacity, on average, that is actually mitigated during each dispatch interval. The results are provided by load level.

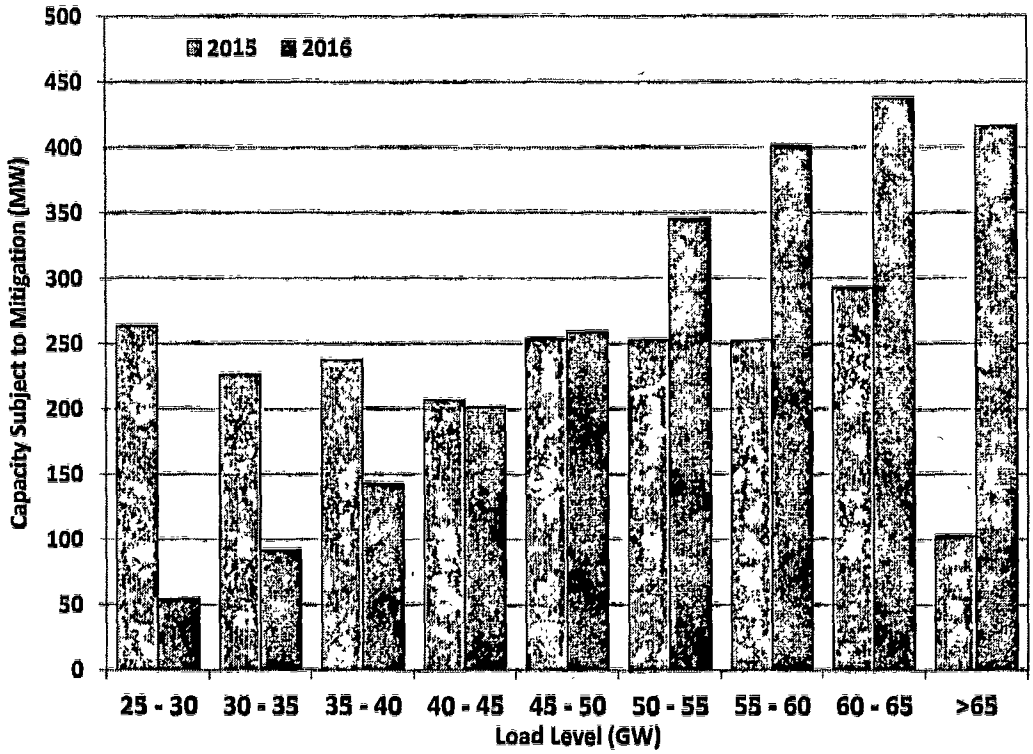
Figure 78: Mitigated Capacity by Load Level



The level of mitigation in 2016 was higher, particularly at higher load levels, than in 2015. The average amount of mitigated capacity was less than 20 MW for all load levels in 2015, but averaged almost 60 MW at loads greater than 65 GW in 2016. The greater frequency of congestion that occurred in 2016, as described in Section III: Transmission Congestion and Congestion Revenue Rights, supports the higher mitigation levels experienced in 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 2016 was higher than 2015, especially at higher 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

Five units provided notice of the intent to suspend operations with a suspension date in 2016, amounting to approximately 1,100 MW of capacity retired or mothballed during the year. For the first time since 2011 ERCOT determined that there was a reliability need that warranted putting a unit under a reliability must run (RMR) contract. Greens Bayou 5 is a 371 MW natural gas steam unit built in 1973 and located in Houston. The 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. Following changes to the RMR study parameters[35] and the earlier than expected completion of new generation in Houston, the contract with the unit was cancelled effective May 29, 2017.

Prior to Greens Bayou 5, the last time units in the ERCOT market were under RMR agreements was in 2011 – a year of extreme heat and drought. That year, ERCOT required four units that had previously been allowed to enter mothball status to return to service under RMR contracts for the peak summer demand. The protocols were changed shortly thereafter to require that any energy from RMR units be offered at the system-wide offer cap.[36] Pricing out of market energy at the system-wide offer cap ensures that energy from RMR units is dispatched last.

The Greens Bayou 5 RMR presented a different pricing issue, since it was procured to resolve a transmission constraint. The Houston import constraint is frequently a non-competitive constraint, and hence, the price of energy from the RMR unit would be mitigated. Given the unit's significant helping impact on the constraint and the relatively low mitigated price, it was likely that if the unit was committed it would be dispatched before other similarly-priced or even lower-priced units in the Houston area. NPRR784 was proposed to address mitigated offer caps for RMR units, but market participants could not reach consensus on this approach and the protocol change request was not approved. Thus, any future RMR units could still be dispatched at a mitigated price that is not reflective of the reliability value of the resource.

The Greens Bayou 5 RMR drew significant scrutiny from market participants on the RMR process. In addition to NPRR784, there were other Protocol changes put in place as a result of the RMR contract. The ERCOT evaluation criteria for potential RMR units was adjusted to require that RMR units have a material impact on the expected transmission overload in order to be procured under an RMR contract.[37] A material impact was defined to mean more than a two percent helping shift factor and more than a five percent unloading factor on the transmission facility that is overloaded. This Protocol change facilitated ERCOT's re-evaluation of the RMR contract for Greens Bayou 5 and ultimately resulted in early termination of the contract. Other protocol changes clarified the ERCOT commitment process for RMR units,[38] updated the contracting and reimbursement process for RMR units,[39] and created a mechanism for clawback of capital contributions from an RMR unit if the unit returns to the market.[40]

35 *See* NPRR788, RMR Study Modifications.

36 *See* NPRR442, Energy Offer Curve Requirement for Generation Resources Providing Reliability Must-Run Service.

37 NPRR788, RMR Study Modifications.

38 NPRR793, Clarification to RMR RUC Commitment and Other RMR Cleanups.

39 *Id.*

40 NPRR795, Provisions for Refunds of Capital Contributions Made in Connection with an RMR Agreement.

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 system demands and reliability 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

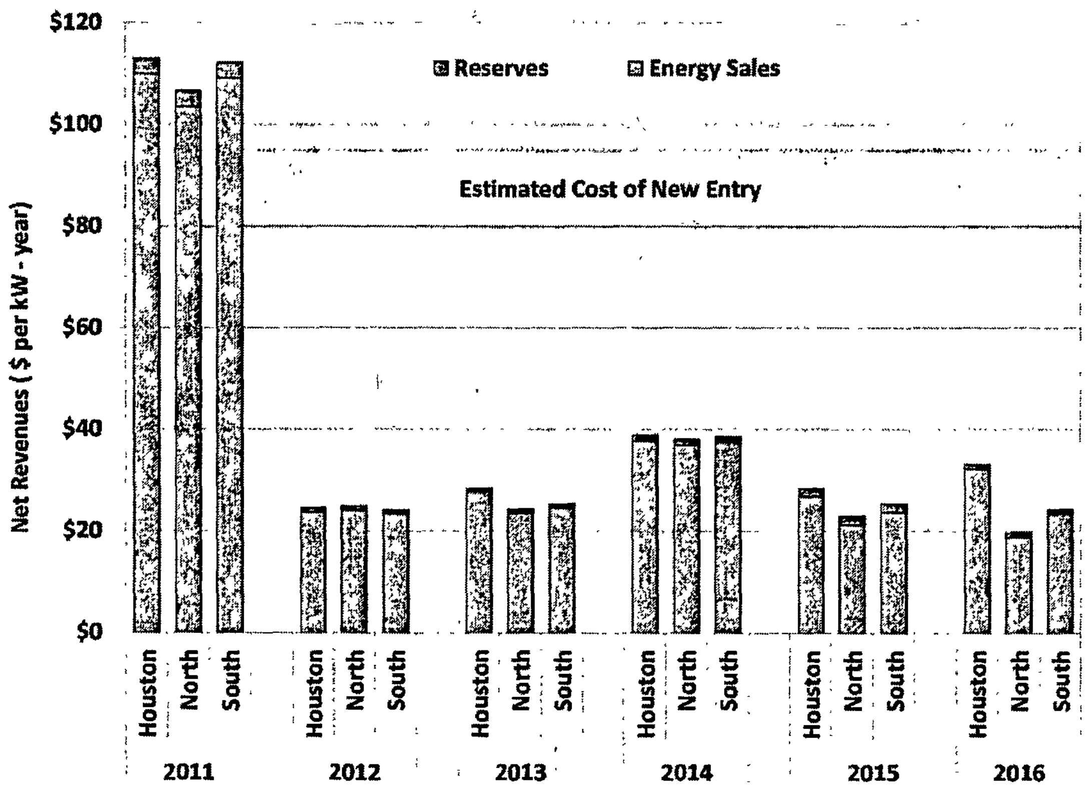
Net revenue is calculated by determining the total revenue that could have been earned by a generating unit less its variable production costs. Put another way, it is the revenue in excess of short-run operating costs that is available to recover a unit's fixed and capital costs, including a return on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

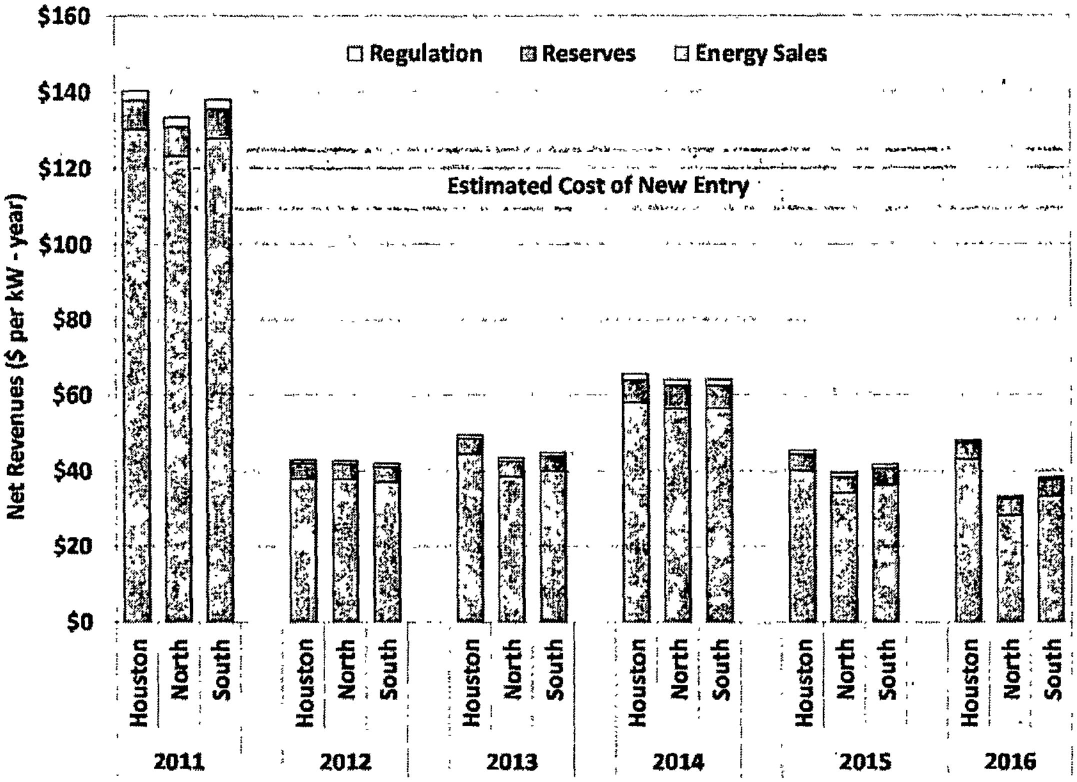
The next two figures provide an historical perspective of the net revenues available to support investment in a new natural gas combustion turbine (Figure 80) and combined cycle generation (Figure 81), selected to represent the marginal new supply that may enter when new resources are needed. The figure 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 gas turbine unit ranges from $80 to $95 per kW-year. The net revenue in 2016 for a new gas turbine was calculated to be approximately $20 to 33 per kW-year, depending on the zone, which are well below the estimated cost of new gas turbine generation.

Figure 81: Combined Cycle Net Revenues



For a new combined cycle gas unit, the estimate of net revenue requirement is approximately $110 to $125 per kW-year. The net revenue in 2016 for a new combined cycle unit was calculated to be approximately $33 to 48 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 the current surplus capacity, which contributed to infrequent shortages in 2015 and 2016. In an energy only market, shortages play a key role in delivering the net revenues an investor would need to recover its investment. Such shortages will tend to be clustered in years with unusually high load and/or poor generator availability. Hence, these results alone do not raise substantial concern regarding design or operation of ERCOTs ORDC mechanism for pricing shortages.

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

Table 10: Settlement Point Price by Fuel Type

Generation Type	Output-Weighted Price
Combined-cycle > 90 MW	$24.59
Combined-cycle <= 90 MW	$27.74
Coal and lignite	$23.98
Diesel	$45.60
Gas steam non-reheat	$53.53
Gas steam reheat boiler	$44.60
Gas steam supercritical boiler	$35.12
Hydro	$22.04
Nuclear	$21.46
PhotoVoltaic Generation Resources	$31.95
Power Storage	$22.75
Renew	$28.21
Simple-cycle > 90 MW	$23.91
Simple-cycle <= 90 MW	$39.68
Wind	$16.18

Given the very low energy prices during 2016 in non-shortage hours, 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 2016 was $24.62 per MWh. The generation-weighted average price for the four nuclear units in ERCOT - approximately 5 GW of capacity - was only $21.46 per MWh in 2016, down from $24.56 per MWh in 2015. According to the Nuclear Energy Institute (NEI), total operating costs for all nuclear units across the U.S. averaged $27.17 per MWh in 2016, a slight decrease from the reported costs for 2015.[41] Assuming that operating costs in ERCOT are similar to the U.S. average, it is likely that these units were not profitable in 2016 based on the fuel and operating and maintenance costs alone. To the extent nuclear units in ERCOT had any associated capital costs, it is likely those costs were not recovered. Compared to other regions with larger 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

41 NEI Whitepaper, "Nuclear Costs in Context," April 2017, available at https://www.nei.org/www.nei.org/files/fe/fed92b11-8ea6-40df-bb0e-29018864a668.pdf.

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 be monitored.

The generation-weighted price of all coal and lignite units in ERCOT during 2016 was $23.98 per MWh. Although specific unit costs may vary, index prices for Powder River Basin coal delivered to ERCOT were approximately $2.50 per MMBtu in 2016, a decrease from approximately $2.60 per MMBtu in 2015. For the past two years delivered coal costs in ERCOT have been about $0.03 to $0.05 per MMBtu higher than natural gas prices at the Houston Ship Channel. Given that the coal units generally have higher heat rates and more expensive non-fuel operations and maintenance costs, it follows that they have been losing market share to natural gas. As with nuclear units, it appears that coal units were likely not profitable in ERCOT during 2016. With the bulk of the coal fleet in ERCOT being more than 30 years old, the retirement or suspended operation of some of these units could cause ERCOT's capacity margin to fall to unreliable levels more quickly than anticipated. While both nuclear and coal are feeling the pressure of an increased reliance on lower-priced natural gas units, coal units appear to be at greater risk of retirement than the nuclear units in ERCOT. This may be due to their relative age and inefficiency.

These results indicate that during 2016 the ERCOT markets would not have provided sufficient revenues to support profitable investment in any of the types of generation technology evaluated, which may seem inconsistent with the fact that new generation continues to be added in the ERCOT market. This 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 2016. Given the level to which prices will rise under shortage conditions, buyers may enter bilateral contracts to hedge against high shortage pricing.

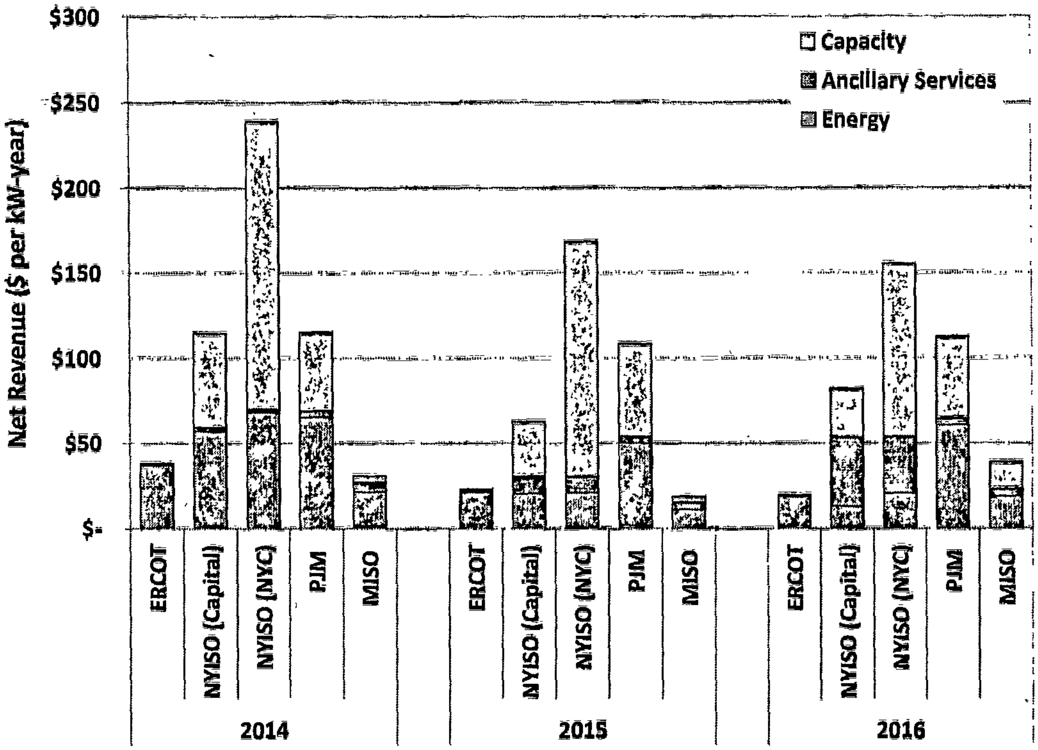
Third, net revenues in any one year may be higher or lower than an investor would require over the long term. In 2016, shortages were much less frequent than would be expected over the long term. Shortage revenues play a pivotal role in motivating investment in an energy-only market like ERCOT. Hence, 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 with centrally-cleared capacity markets. The technologies are differentiated by assumed heat rate; 7,000 MMBtu per MWh for combined cycle and 10,500 MMBtu per MWh for simple-cycle combustion turbine.

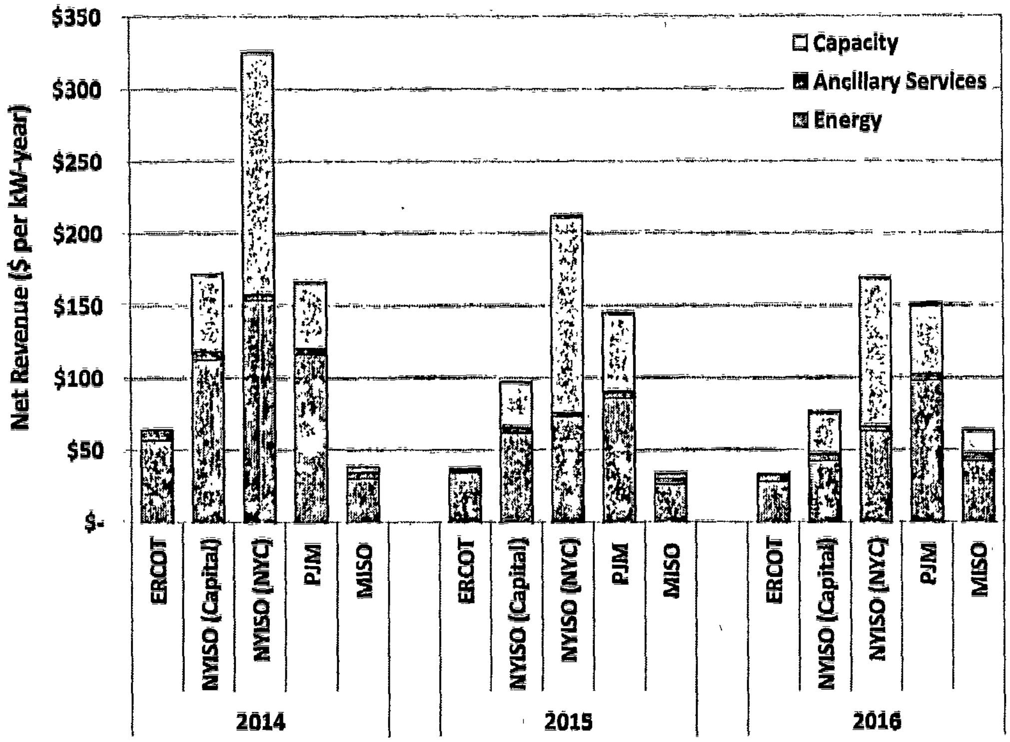
The next two figures compare estimates of net revenue for these two types of natural gas generators for the ERCOT North zone, PJM, two locations within the New York ISO, and the Midcontinent ISO. Figure 82 provides a comparison of net revenues for a combustion turbine and Figure 83 provides the same comparison for a combined cycle unit.

Figure 82: Combustion Turbine Net Revenue Comparison Between Markets



The figures include 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. Most of the locations shown are central locations, but there are load pockets within each market where net revenue and the cost of new entry may be higher. The NYC zone of the New York ISO is an example of much higher value in a load pocket. Thus, even if new investment is not generally profitable in a market, it may be economic in certain areas.

Figure 83: Combined Cycle Net Revenue Comparison Between Markets



Both figures indicate that across all markets, with the exception of New York ISO (Capital) for combustion turbine, net revenues decreased substantially in 2016 because of low natural gas prices across the country and sufficient installed reserves, typically a result of flat or no load growth. With the exception of MISO, capacity revenues 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 Public Utility Commission of Texas (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,[42] this subsection assesses the Scarcity Pricing Mechanism (SPM) in 2016 under ERCOT's energy-only market structure.

Revisions to 16 TEX. ADMIN. CODE § 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 on page 20, there have been very brief periods when energy prices rose to the cap since the system-wide offer

42 *See* 16 TEX. ADMIN. CODE § 25.505(g)(6)(D).

cap was increased to greater than $3,000 per MWh. There have been no instances of prices rising above $5,000 per MWh.

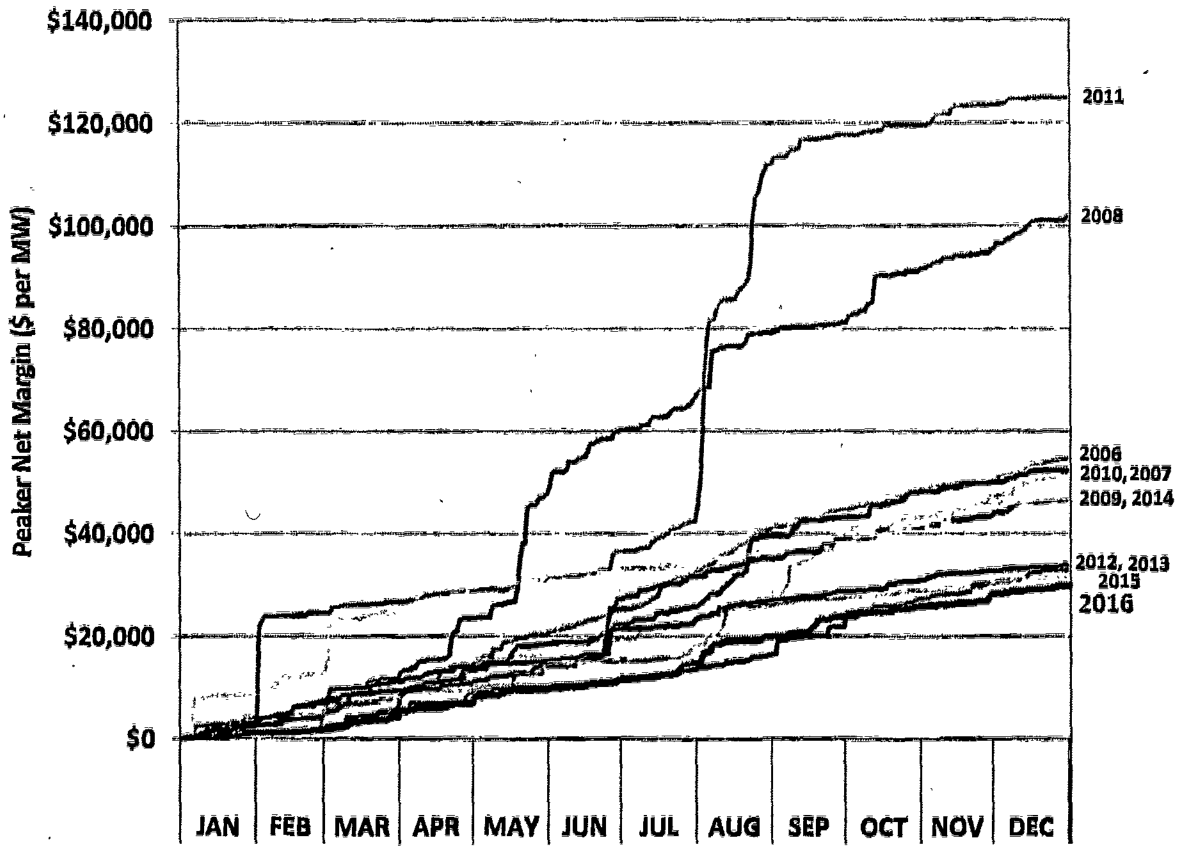
The SPM includes a provision termed the Peaker Net Margin (PNM) that is designed to provide a fail-safe pricing measure, which if exceeded would cause the system-wide offer cap to be reduced. If the PNM for a year reaches a cumulative total of $315,000 per MW, the system-wide offer cap is then reduced to the higher of $2,000 per MWh or 50 times the daily natural gas price index.[43] PNM also serves as a simplified measure of the annual net revenue of a hypothetical peaking unit.[44]

Figure 84 shows the cumulative PNM results for each year from 2006 through 2016 and shows that PNM in 2016 was the lowest it has been since it became effective in 2006. 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.

43 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 analysis prepared by Brattle dated June 1, 2012, and will remain in place until there is a change identified in the cost of new entry of new generation plants.

44 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.[45] 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 improve resource adequacy signals, by directing ERCOT to implement the Operating Reserve Demand Curve (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.

45 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."[46] 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 and/or generation availability is poor; neither of which was the case in 2015 or 2016. The PUCT has taken comment from stakeholders, but to date the PUCT has not directed modification of the reserve adder component of ORDC.[47]

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.

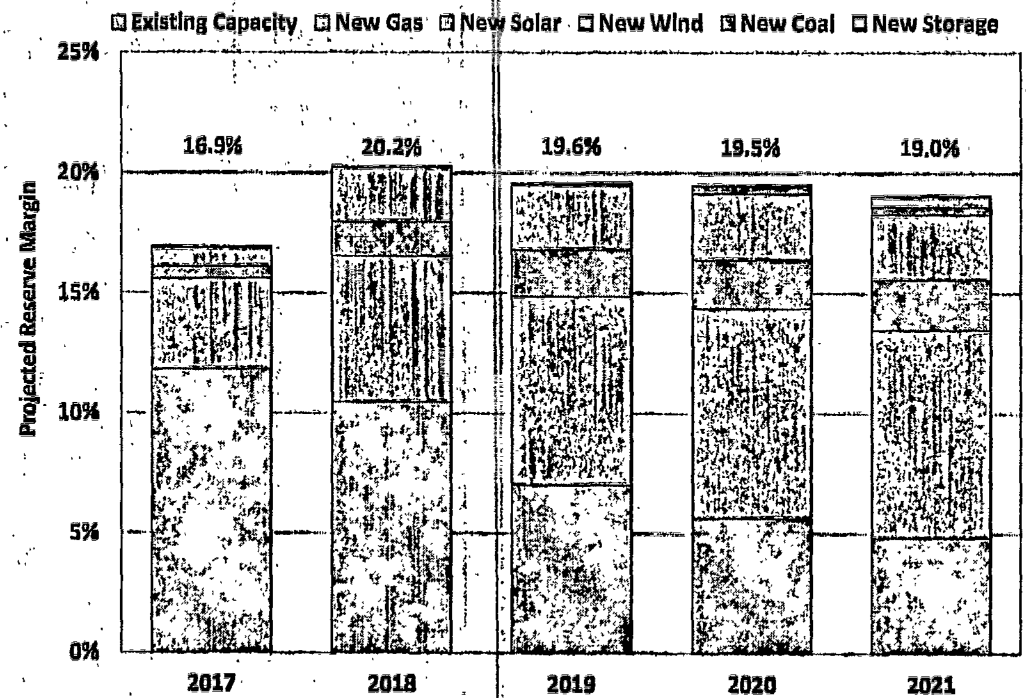
C. Planning Reserve Margin

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

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

47 See PUCT Docket No. 45572, Review of the Parameters of the Operating Reserve Demand Curve.

Figure 85: Projected Planning Reserve Margins



Source: ERCOT Capacity, Demand and Reserves Report - December 2016

Figure 85 above indicates that the region will have a 16.9 percent reserve margin heading into the summer of 2017. While these projections are slightly lower than those developed last year, the current outlook is very different than it was in 2013, when planning reserve margins were expected to be below the then-existing target level of 13.75 percent for the foreseeable future.[48]

This current projection of planning reserve margins combined with relatively infrequent shortage pricing may raise doubts regarding the likelihood of announced generation coming on line as planned. Given the projections of continued low prices, investors of some of the new generation included in the Report on the Capacity, Demand and Reserves in the ERCOT Region (CDR) may choose to delay or even cancel their project. Additionally, the profitability analysis of existing base load resources casts doubt on the assumption embedded in the CDR that all existing generation will continue to operate. Hence, it is likely that the planning reserve margins will be lower than forecasted in the figure above.

[48] The target planning reserve margin of 13.75 percent was approved by the ERCOT Board of Directors in November 2010, based on a 1 in 10 loss of load expectation (LOLE). The PUCT recently 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).

With expectations for future natural gas prices to remain relatively low, the pressure on the ability of coal units in the ERCOT market to economically operate is not expected to subside any time soon. These challenging fuel market economics 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 in any way be viewed as failure to provide resource adequacy. Having the right pricing signals to encourage sufficient and efficient generation signals is the goal. Most of the coal units facing the greatest price and environmental pressure have been operating for more than thirty-five years. 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.

D. Ensuring Resource Adequacy

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

Generators earn revenues from three sources: energy prices during non-shortage, energy prices during shortage and capacity payments. The capacity payments generators receive in ERCOT are related to the provision of ancillary services. Ancillary service payments are a small contributor, 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" 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 for 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 2016. 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 2016.

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.[49] 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. The RDI also does not indicate whether it would have been profitable for a pivotal supplier to exercise market power. However, it does identify conditions under which a supplier could raise prices significantly by withholding resources.

Figure 86 shows the ramp-constrained RDI relative to load for all hours in 2016. The trend line indicates a strong positive relationship between load and the RDI. The analysis shown below is done at the QSE level because the largest suppliers that determine the RDI values own a large majority of the resources they are offering. To the extent that the resources scheduled by the largest QSEs are not controlled by or provide revenue to the QSE, the RDIs will tend to be slightly overstated.

49 For the purpose of this analysis, "quick-start" includes off-line simple cycle gas 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

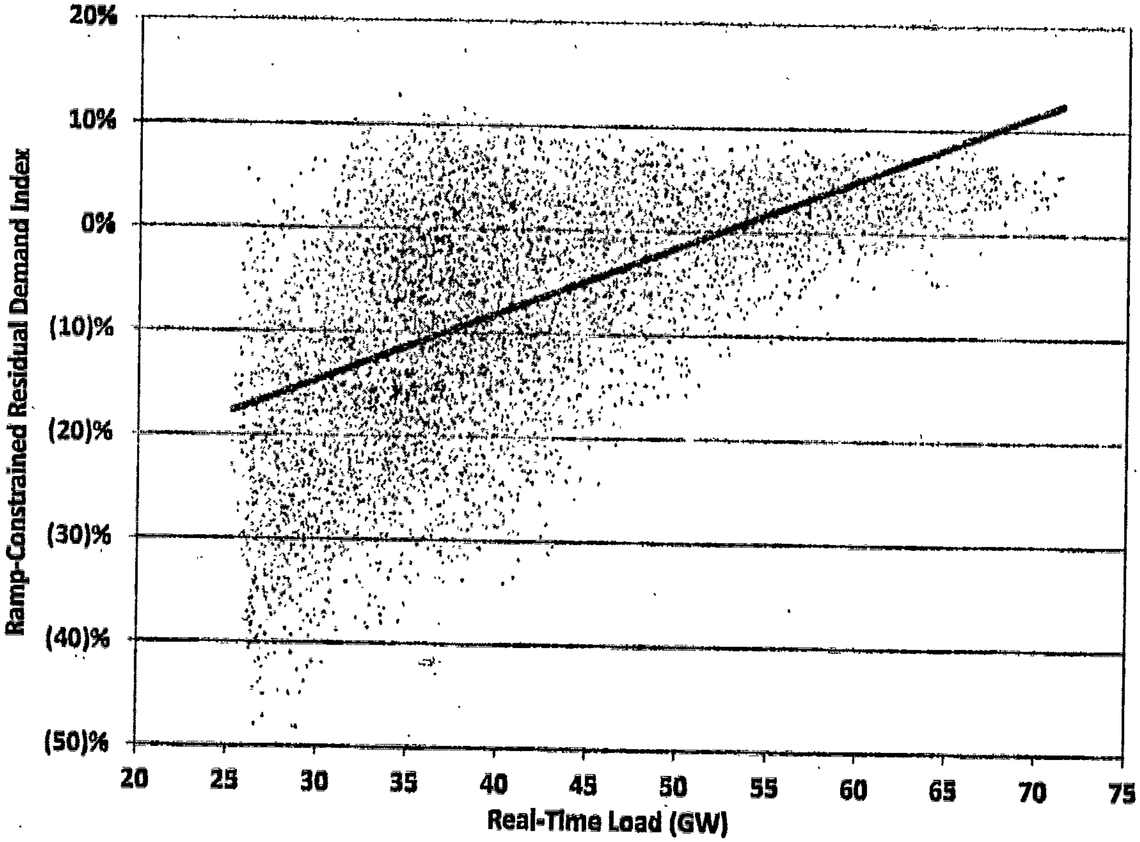
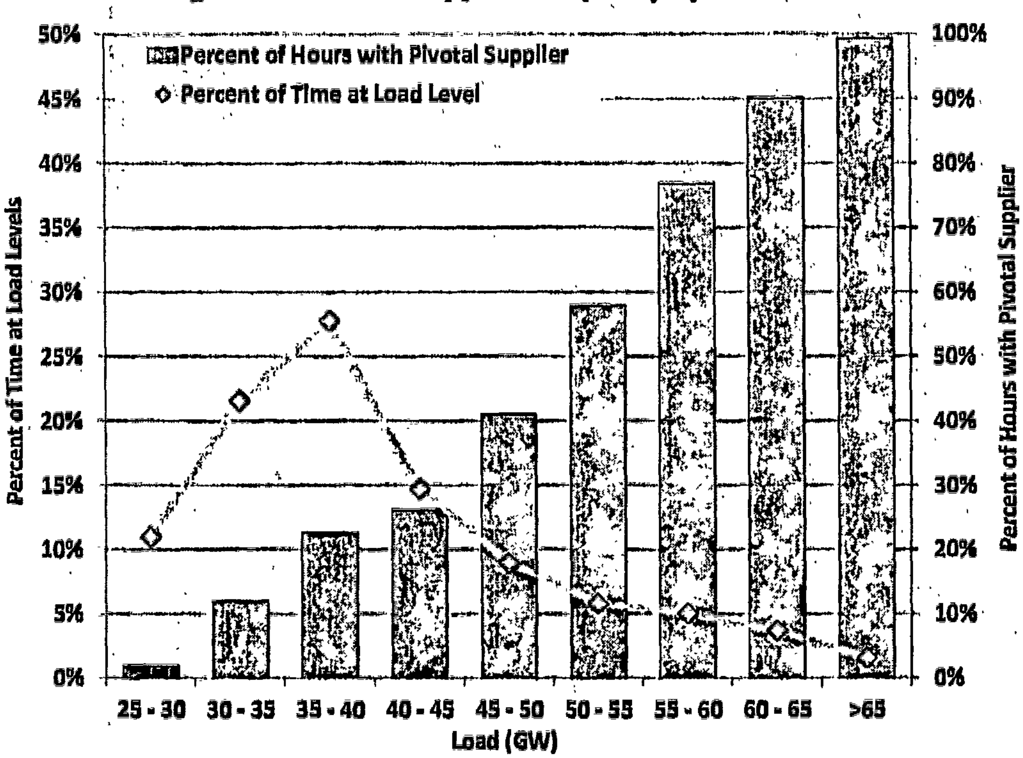


Figure 87 below summarizes the results of the RDI analysis by displaying the percent of time at each load level there was a pivotal supplier. The figure also displays the percent of time each load level occurs.

Figure 87: Pivotal Supplier Frequency by Load Level



At loads greater than 65 GW there was a pivotal supplier 99 percent of the time. This is expected because at high load levels, larger suppliers are more likely to be pivotal because other suppliers' resources are more fully utilized serving the load. The frequency of relatively high loads increased in 2016. This led to an increase in the pivotal supplier frequency to 28.5 percent of all hours in 2016, up from 26 and 23 percent of all hours in 2015 and 2014, respectively. This indicates that 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 three market participants in 2016. 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 TEX. ADMIN. CODE § 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.

In 2016, there were three market participants with approved VMPs – NRG, Calpine, and Luminant. NRG's plan, initially approved in June 2012 and modified in May 2014,[50] allows the company to offer some of its capacity at prices up to the system-wide offer cap. Specifically, up to 12 percent of the difference between the high sustained limit and the low sustained limit – the dispatchable capacity – for each natural gas unit (5 percent for each coal/lignite unit) may be offered no higher than the higher of $500 per MWh or 50 times the natural gas price. Additionally, up to 3 percent 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.

Calpine's VMP was approved in March of 2013.[51] Because its generation fleet consists entirely of natural-gas fueled combined cycle units, the details of the Calpine plan are somewhat different than NRG. Calpine may offer up to 10 percent of the dispatchable capacity of its portfolio at prices up to $500 per MWh. Additionally, Calpine may offer up to 5 percent of the dispatchable capacity of its portfolio at prices no higher than the system-wide offer cap. When approved, the

---

50 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).

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

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.

Luminant received approval from the PUCT for a VMP in May 2015.[52] The Luminant plan is similar in many respects to the NRG plan. Under the VMP, Luminant is permitted to offer a maximum of 12 percent of the dispatchable capacity for its natural gas units (5 percent for coal/lignite units) at prices up to $500 per MWh and offer a maximum of 3 percent of the dispatchable capacity for natural gas units up to the system-wide offer cap. The amount of capacity covered by these provisions is slightly more than 500 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.

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 three 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 element in the VMPs is 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."[53] 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.

52 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).

53 PURA § 39.157(a).

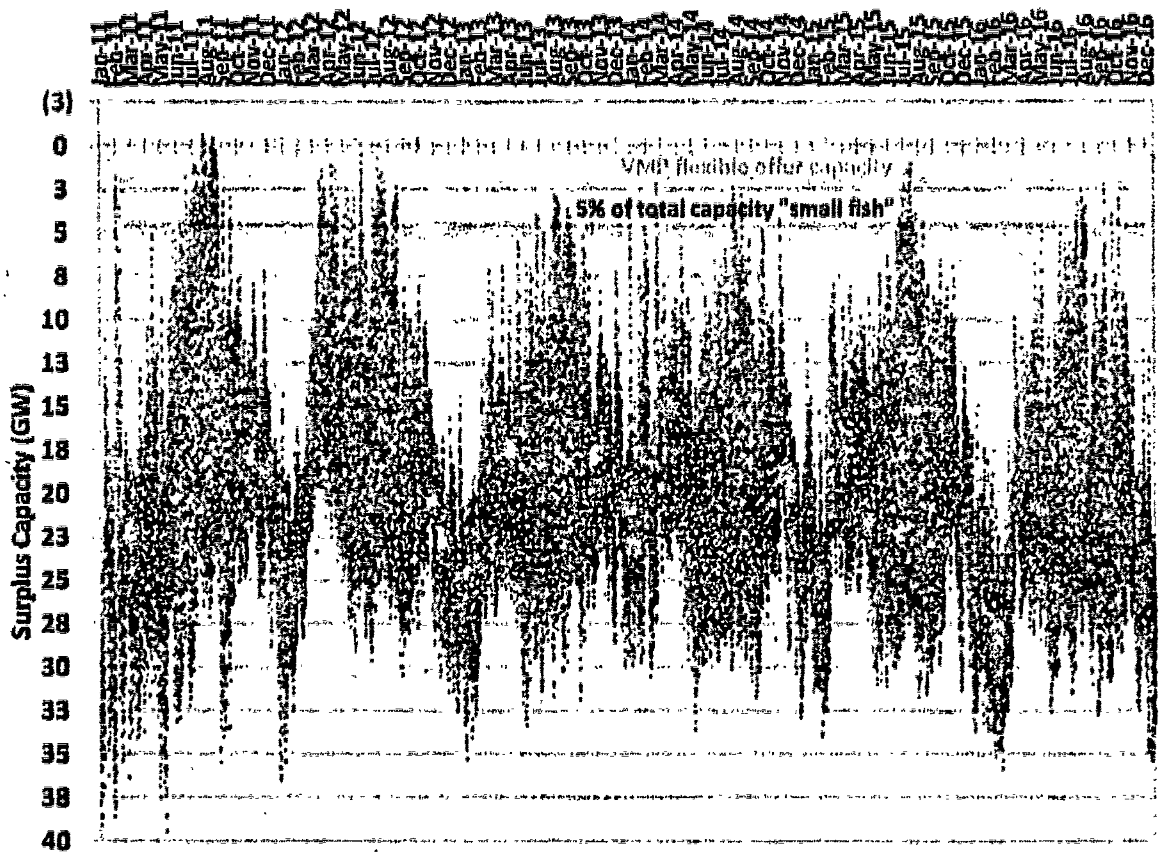
The amount of offer flexibility afforded by the VMPs is small when compared to the offer flexibility that small participants – those with less than 5 percent of total ERCOT capacity – are granted under 16 TEX. ADMIN. CODE § 25.504(c). Although 5 percent of total ERCOT capacity may seem like a small amount,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

The figure below shows the amount of surplus capacity available in each hour of every day from 2011 to 2016. For this analysis, surplus capacity is defined as online generation plus any offline capacity that was available day ahead, plus DC Tie imports (minus exports), minus responsive reserves provided by generation, regulation up capacity, and load. Every hour of the past four years has had surplus capacity. Only during 2011 (12 hours) and for one hour in 2012 was ERCOT was unable to meet load and maintain all operating reserve obligations.

Currently, the 5 percent "small fish" threshold is roughly 4,000 MW, as indicated by the red line in Figure 88. There were 572 hours over the past six years with less than 4,000 MW of surplus capacity.[54] During these times a large "small fish" would have been pivotal and able to increase the market clearing price through its offer, potentially as high as the system-wide offer cap. In contrast, the combined amount of capacity afforded offer flexibility under the VMPs granted to NRG, Calpine, and Luminant totals less than 1,800 MW of capacity. This amount of capacity would have been pivotal for a total of 120 hours across the past six years, with none occurring in 2016.

54 Surplus capacity was less than 4,000 MW for 296 hours in 2011, 154 hours in 2012, 15 hours in 2013, 26 hours in 2014, 56 hours in 2015, and 25 hours in 2016.

Figure 88: Surplus Capacity



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.

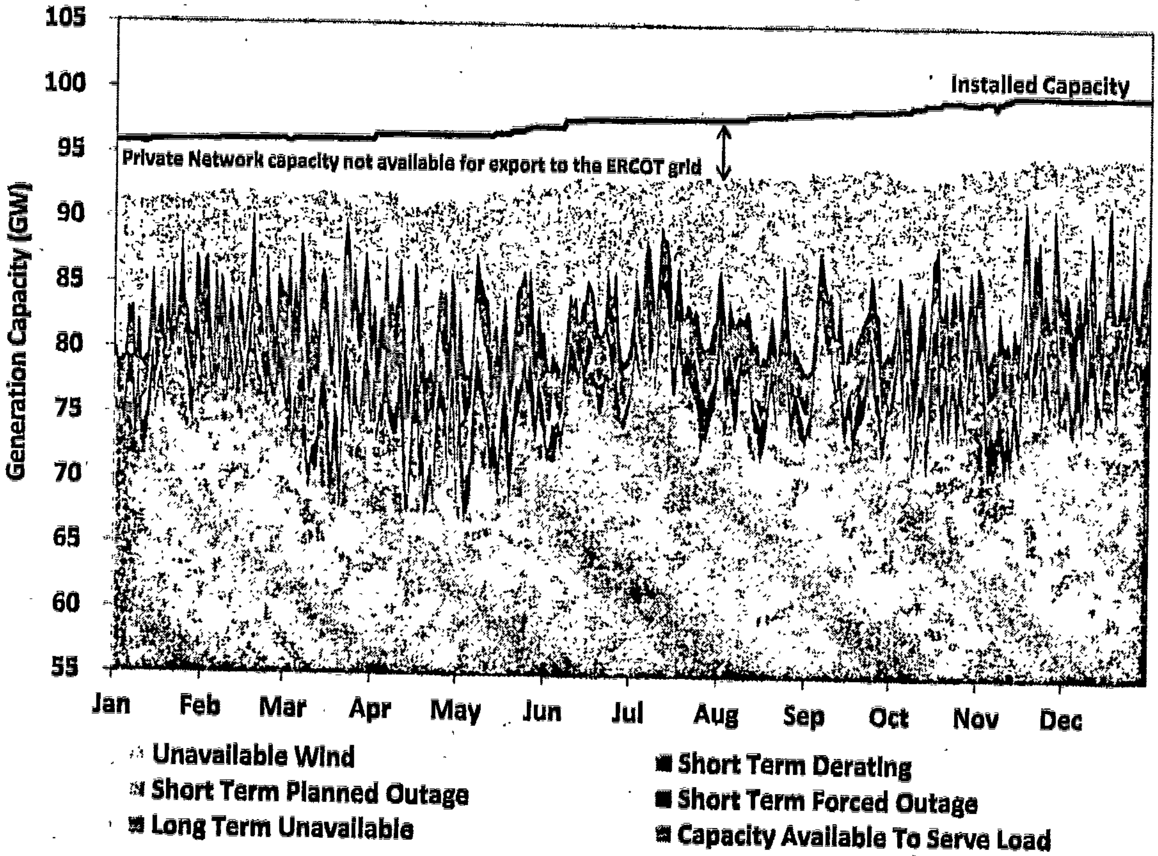
In a single-price auction like the real-time energy market, suppliers may attempt to exercise market power by withholding resources. The purpose of withholding is to cause more expensive resources to set higher market clearing prices, allowing the supplier to profit on its other sales in the real-time 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 due to 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. Due to 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 outages. 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 (e.g., by 5 to 10 percent) because the resource cannot achieve its installed capacity level due to technical or environmental factors (e.g., component equipment failures or ambient temperature conditions). Wind generators rarely produce at the installed capacity rating due to 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 89 shows a breakdown of total installed capacity for ERCOT on a daily basis during 2016. 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 due to 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 89: Reductions in Installed Capacity



Outages and deratings of non-wind generators fluctuated between 4 and 19 GW, as shown in Figure 89, while wind unavailability varied between 1 and 15 GW. Short-term planned outages were largest between March and April and smallest during the summer months, which is 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 7 GW, reduced to less than 1 GW during the summer months, and increased to 5 GW in November. This pattern reflects the 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 90 shows the average magnitude of the outages and deratings lasting less than 30 days for the year and for each month during 2016.

Figure 90: Short-Term Outages and Deratings

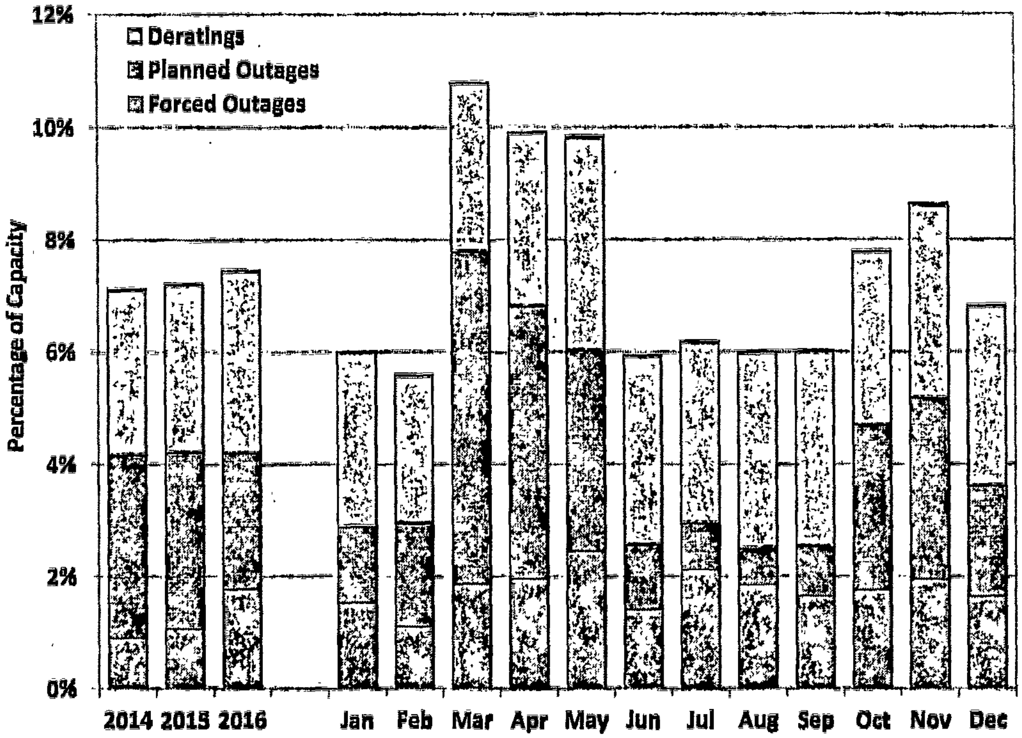


Figure 90 shows that total short-term deratings and outages were as large as 10.8 percent of installed capacity in March, and averaged around 6 percent during the summer. Most of this fluctuation was due to anticipated planned outages. The amount of capacity unavailable during 2016 averaged 7.5 percent of installed capacity. This is a slight increase from 7.2 percent experienced in 2015 and 7.1 percent experienced in 2014. Overall, the fact that 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

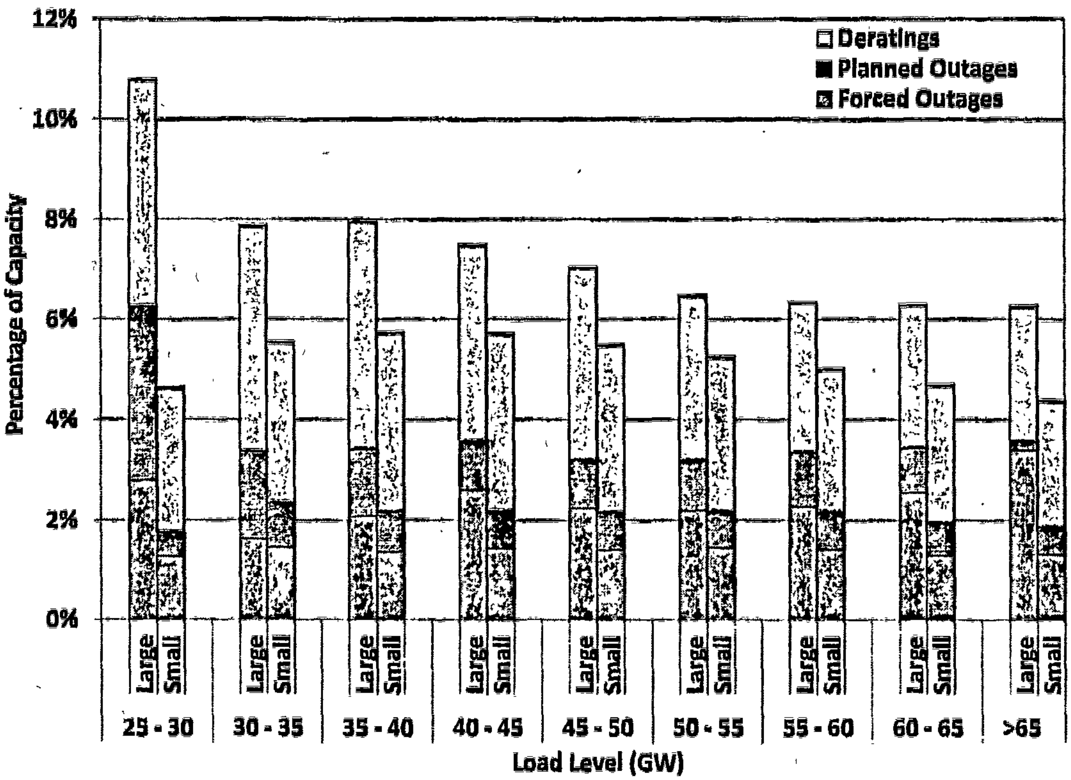
Physical withholding occurs when a participant makes resources unavailable for dispatch that are otherwise physically capable of providing energy and that are economic at prevailing market prices. This can be done either by derating a unit or declaring it as forced out of service. Because generator deratings 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 since their output is generally most profitable in peak periods.

Figure 91 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 91: Outages and Deratings by Load Level and Participant Size, June-August



000157

Figure 91 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. Since 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. For large suppliers, the percent of derated capacity declined at higher load levels, whereas for small providers the percent of derated capacity was fairly constant across all load levels. Although large providers had slightly higher forced outage rates than small providers, their level – 2.4 percent – does not raise potential competitive concerns.

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

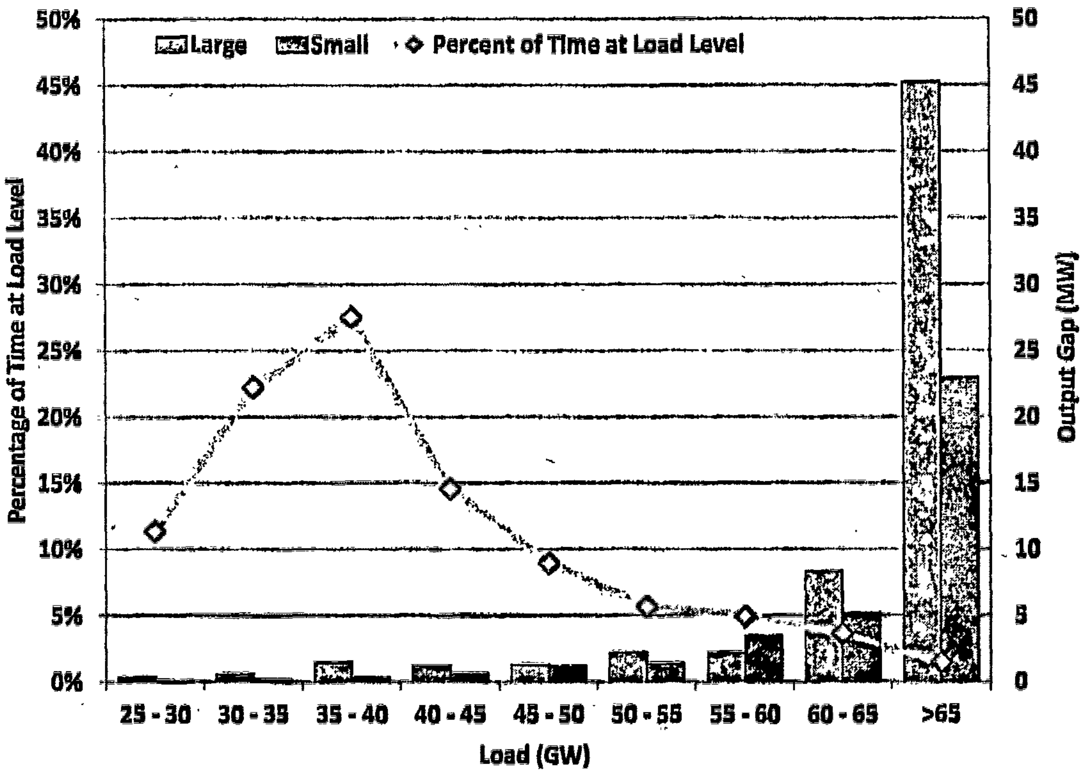
Before presenting the results of the output gap analysis, a description of ERCOT's two-step dispatch software is required. In the first step, the dispatch software calculates output levels (base points) and associated locational marginal prices using the participants' offer curves and only considering 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

55 Given the low energy prices during 2016, the output gap margin was reduced to $30 for purposes of this analysis. Prior to 2015, the State of the Market report used $50 for the output gap margin.

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.

Figure 92: Incremental Output Gap by Load Level and Participant Size – Step 1

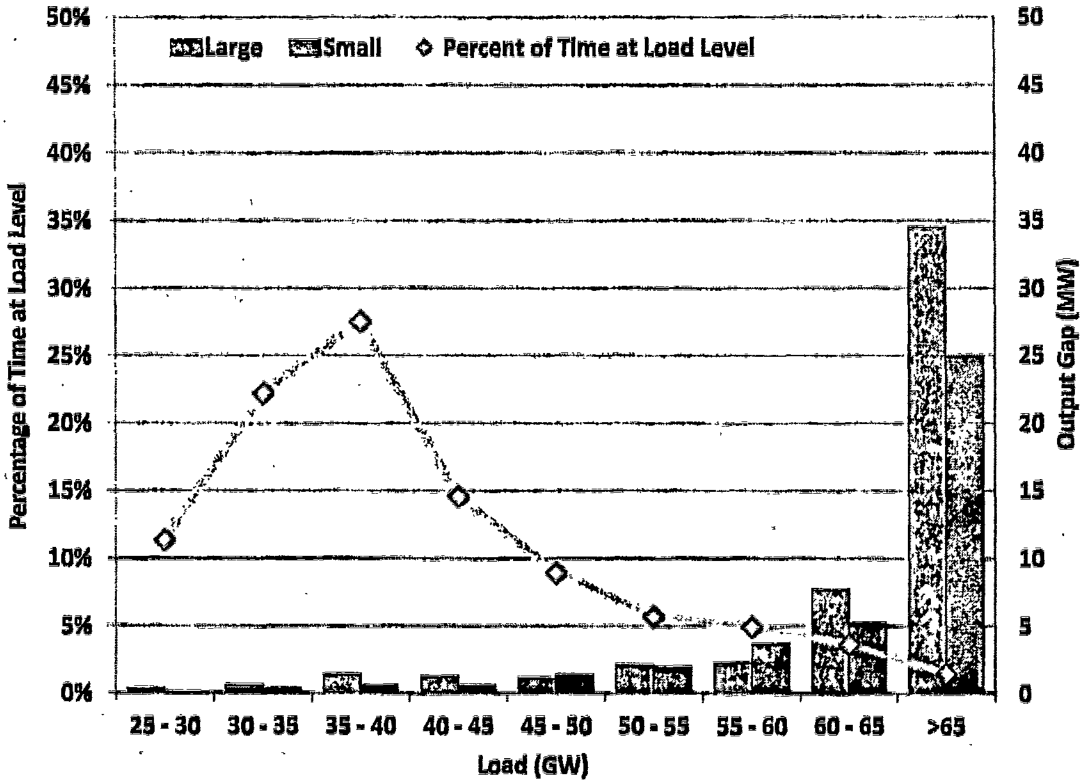


The results of the analysis shown in Figure 92 indicate that only very small amounts of capacity would be considered part of the first step output gap.

Figure 93 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 93 also shows very small quantities of capacity that would be considered part of this output gap.

Figure 93: 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 2016. 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 2016.