# IV. DEMAND AND SUPPLY

This section reviews and analyzes the load patterns during 2015 and the existing generating capacity available to satisfy the load and operating reserve requirements. Specific analysis of the large quantity of installed wind generation is included, along with a discussion of the daily generation commitment characteristics. This section concludes with a discussion of demand response resources.

## A. ERCOT Load in 2015

The changes in overall load levels from year to year can be shown by tracking the changes in average load levels. This metric tends to capture changes in load over a large portion of the hours during the year. Separately evaluating the changes in the load during the highest-demand hours of the year is also important. Significant changes in peak demand levels play a major role in assessing the need for new resources. The level of peak demand also affects the probability and frequency of shortage conditions (i.e., conditions where firm load is served but minimum operating reserves are not maintained). The expectation of resource adequacy is based on the value of electric service to customers and the harm or inconvenience to customers that can result from interruptions to that service. Hence, both of these dimensions of load during 2015 are examined in this subsection and summarized in Figure 53.

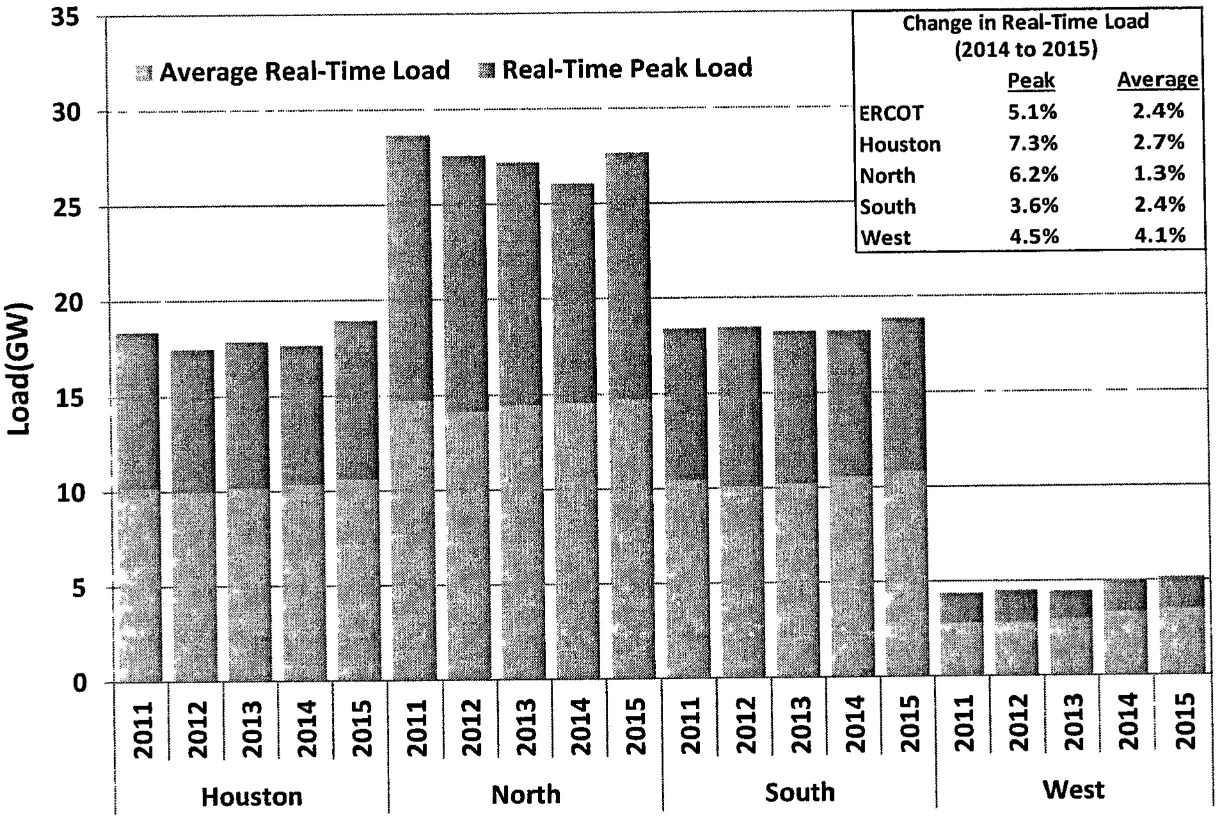
This figure shows peak load and average load in each of the ERCOT zones from 2011 to 2015.[15] In each zone, as in most electrical systems, peak demand significantly exceeds average demand. The North zone is the largest zone (with about 37 percent of the total ERCOT load); the South and Houston zones are comparable (27 percent) while the West zone is the smallest (9 percent of the total ERCOT load).

Figure 53 also shows the annual non-coincident peak load for each zone. This is the highest load that occurred in a particular zone for one hour during the year; however, the peak can occur in

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

different hours for different zones. As a result, the sum of the non-coincident peaks for the zones is greater than the annual ERCOT peak load.

**Figure 53: Annual Load Statistics by Zone**




| Change in Real-Time Load (2014 to 2015) | Peak | Average |
|---|---|---|
| ERCOT | 5.1% | 2.4% |
| Houston | 7.3% | 2.7% |
| North | 6.2% | 1.3% |
| South | 3.6% | 2.4% |
| West | 4.5% | 4.1% |

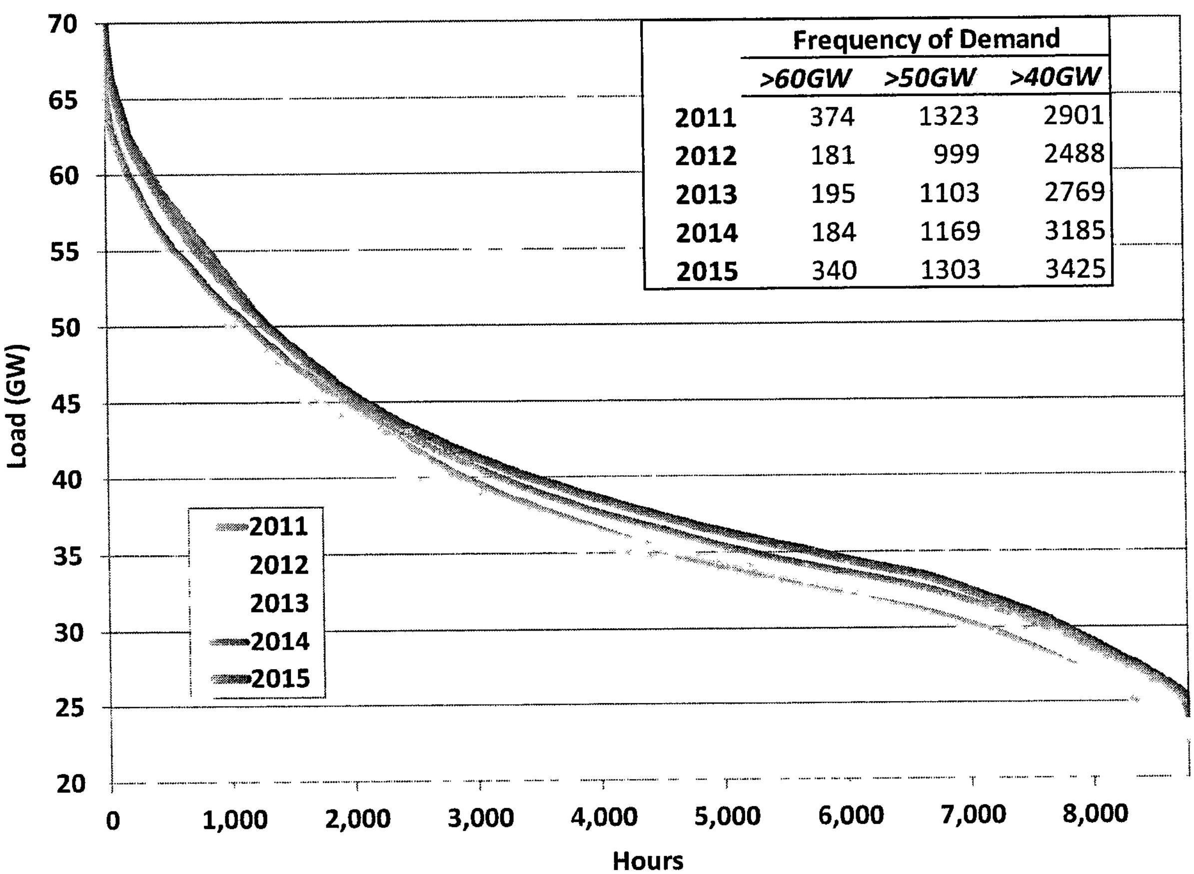
Total ERCOT load over the calendar year increased from 340 terawatt-hours (TWh) in 2014 to 348 TWh in 2015, an increase of 2.4 percent or an average of 866 MW every hour. This increase was largely driven by hotter summer temperatures in 2015. Cooling degree days, a metric that is highly correlated with weather-related summer load, increased 6 percent on average from 2014 to 2015 in Houston and Dallas. However, cooling degree days in 2015 were still 16 percent lower than ERCOT's hottest recent summer in 2011 in these locations.

Summer conditions in 2015 also led to a new ERCOT coincident peak hourly demand record of 69,877 MW on August 10, 2015. This broke the pre-existing peak demand record of 68,311 MW that occurred during August of 2011. In fact, the 2011 demand record was broken five subsequent times during August 2015. The 2015 peak represents a 5.2 percent increase from the peak hourly demand of 66,451 MW in 2014.

The changes in load at the zonal level are not the same as the ERCOT-wide changes. The growth rate of West zone average load was once again much higher, on a percentage basis, than the other zones because of increased oil and gas production activity in this area. While all zones saw an increase in the peak demand, the increase in the Houston zone was significantly higher than others at 7.3 percent over the 2014 peak.

To provide a more detailed analysis of load at the hourly level, Figure 54 compares load duration curves for each year from 2011 to 2015. 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.

**Figure 54: Load Duration Curve – All Hours**

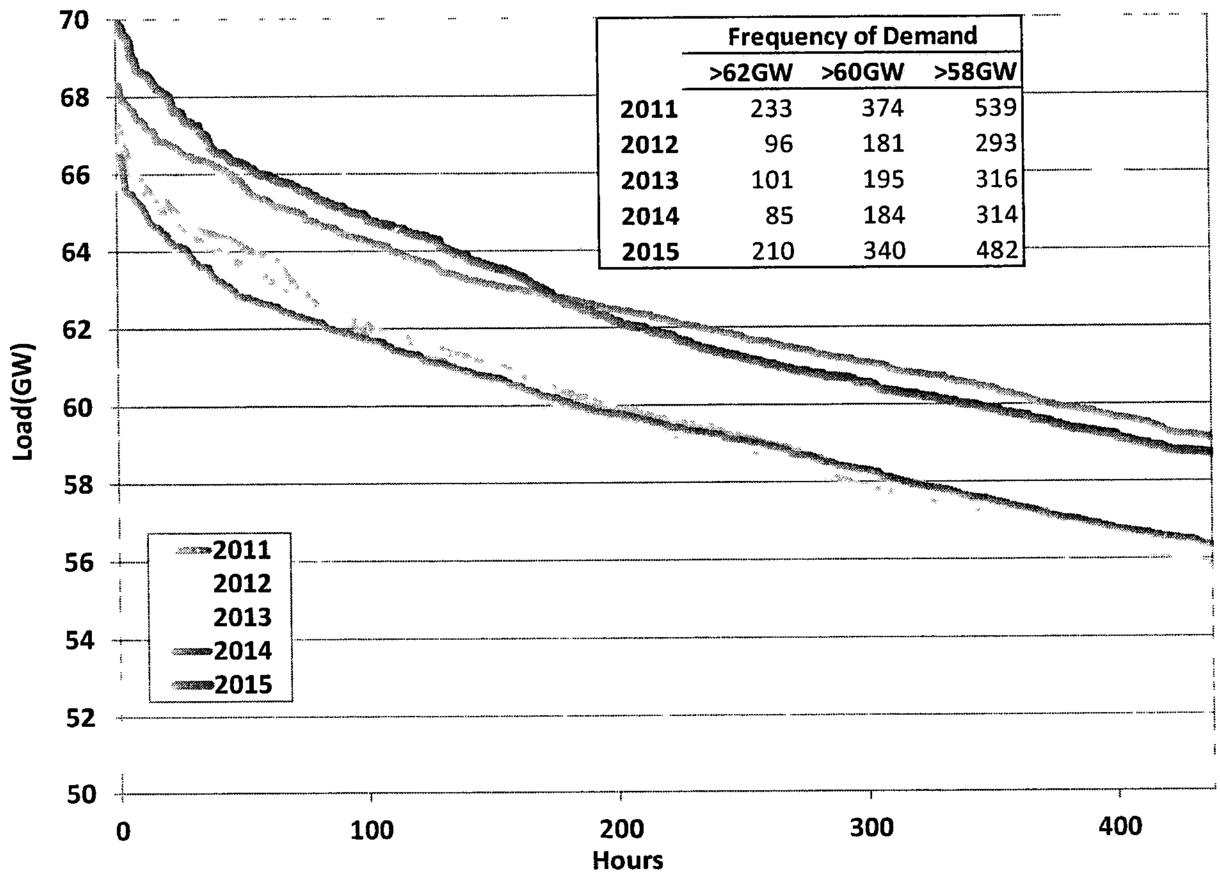



| | Frequency of Demand | | |
|---|---|---|---|
| | >60GW | >50GW | >40GW |
| **2011** | 374 | 1323 | 2901 |
| **2012** | 181 | 999 | 2488 |
| **2013** | 195 | 1103 | 2769 |
| **2014** | 184 | 1169 | 3185 |
| **2015** | 340 | 1303 | 3425 |

As shown in Figure 54, the load duration curve for 2015 is generally higher than the four previous years. However, the 2011 load duration curve remained slightly higher than 2015 for hours 500 to 1,500.

To better illustrate the differences in the highest-demand periods between years, Figure 55 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 2015, the peak load value 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.

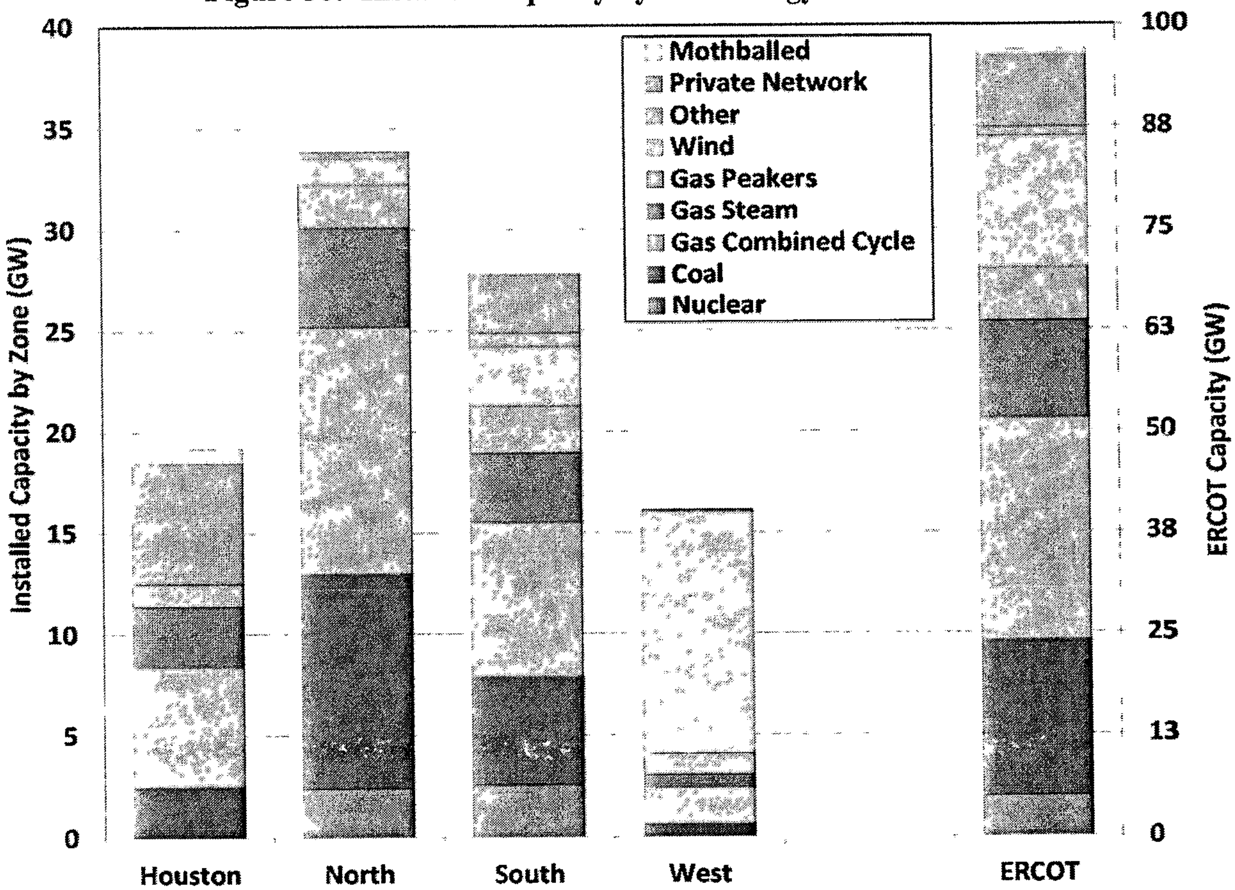
**Figure 55: Load Duration Curve – Top Five Percent of Hours**




## B. Generation Capacity in ERCOT

The generation mix in ERCOT is evaluated in this subsection. The distribution of capacity among the ERCOT zones is similar to the distribution of demand with the exception of the large amount of wind capacity in the West zone. The North zone accounts for approximately 35 percent of capacity, the South zone 29 percent, the Houston zone 20 percent, and the West zone 17 percent. Excluding mothballed resources and including only the fraction of wind capacity available to reliably meet peak demand,[16] the North zone accounts for approximately 39 percent of capacity, the South zone 32 percent, the Houston zone 22 percent, and the West zone 7 percent. Figure 56 shows the installed generating capacity by type in each zone.[17]

**Figure 56: Installed Capacity by Technology for Each Zone**




Approximately 4.8 GW of new generation resources came online in 2015, but it only provided roughly 1.7 GW of net effective capacity. The overwhelming majority of new capacity was from wind generation. The 3.7 GW of newly installed wind capacity is approximately 600 MW of

16 The percentages of installed capacity to serve peak demand assume wind availability of 12 percent for non-coastal wind and 55 percent for coastal wind.

17 For purposes of this analysis, generation located in a NOIE Load Zone has been included with the proximate geographic Load Zone.



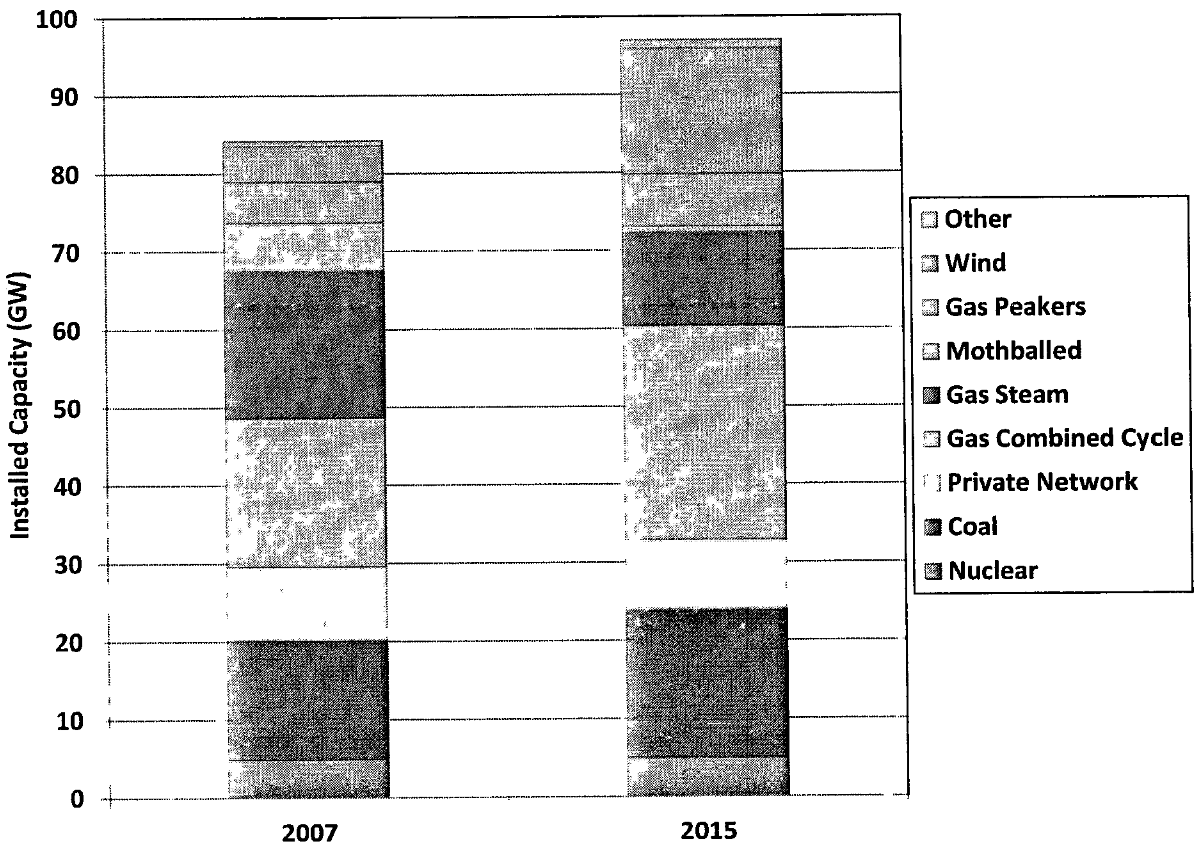
0000000105

peak capacity. The remaining 1.1 GW of new capacity consisted of 100 MW of solar resources and approximately 1 GW of new natural gas combined cycle units.

With these additions, natural gas generation continued to account for approximately 48 percent of total ERCOT installed capacity while the share of coal generation remained at 20 percent in 2015.

By comparing the current mix of installed generation capacity to that in 2007, as shown in Figure 57,[18] the effects of longer term trends can be seen.

**Figure 57: Installed Capacity by Type: 2007 Compared to 2015**




Over these eight years, additions of wind, gas combined cycle, and coal generation have been offset by retirements of older, presumably less efficient natural gas steam units. Between 2007 and 2015, twelve new combined cycle gas units were added. Four combined-cycle units totaling

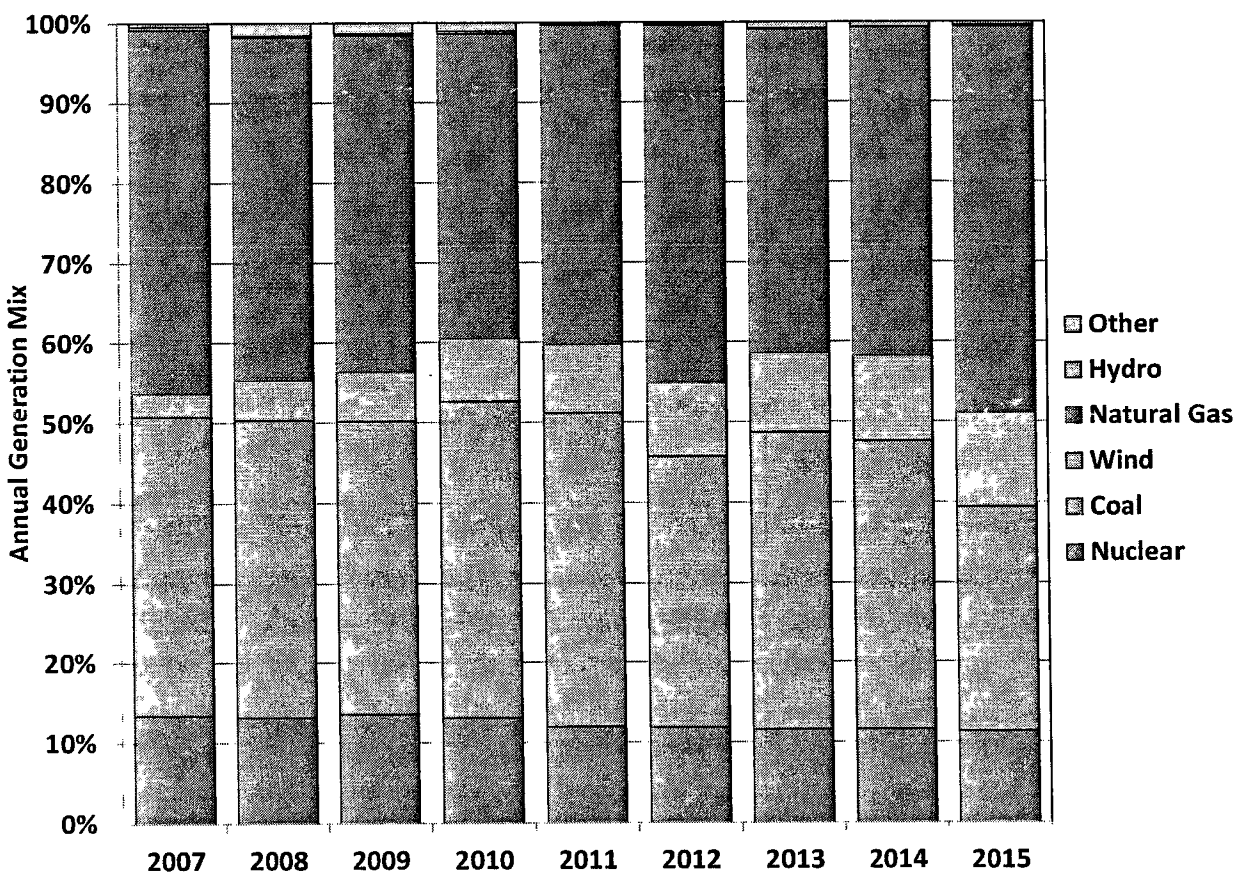
[18] Wind and Private Network capacity is shown at its full installed capacity in this chart.

2.6 GW of capacity have been added in just the past two years. The amount of new wind generation installed since 2007 is 11.5 GW. The effective peak load serving capability of this new wind generation is calculated to be 2.1 GW.

These new additions and the return from mothball of 5.4 GW of resources, less the 6.9 GW of natural gas steam unit retirements has resulted in the installed capacity in 2015 growing by 12.8 GW compared to 2007. However, the increase in peak load serving capability of all net changes to installed capacity is 3.4 GW from 2007 to 2015, whole peak load was 7.7 GW higher in 2015 than in 2007. Hence, although ERCOT's generation base is growing, installed reserve margins have decreased.

The shifting contribution of coal and wind generation is evident in Figure 58, which shows the percentage of annual generation from each fuel type for the years 2007 through 2015.

**Figure 58: Annual Generation Mix**



The generation share from wind has increased every year, reaching 12 percent of the annual generation requirement in 2015, up from 3 percent in 2007. The 2015 generation share saw a record high for natural gas and a record low for coal. In 2015 the percentage of generation from natural gas was 48 percent, a significant increase from the 2014 level and the highest share during this time period of 2007-2015.[19] Corresponding with the increase in natural gas share was a significant decrease in the coal share from 36 percent in 2014 to its lowest observed level of 28 percent in 2015.

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 is 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.

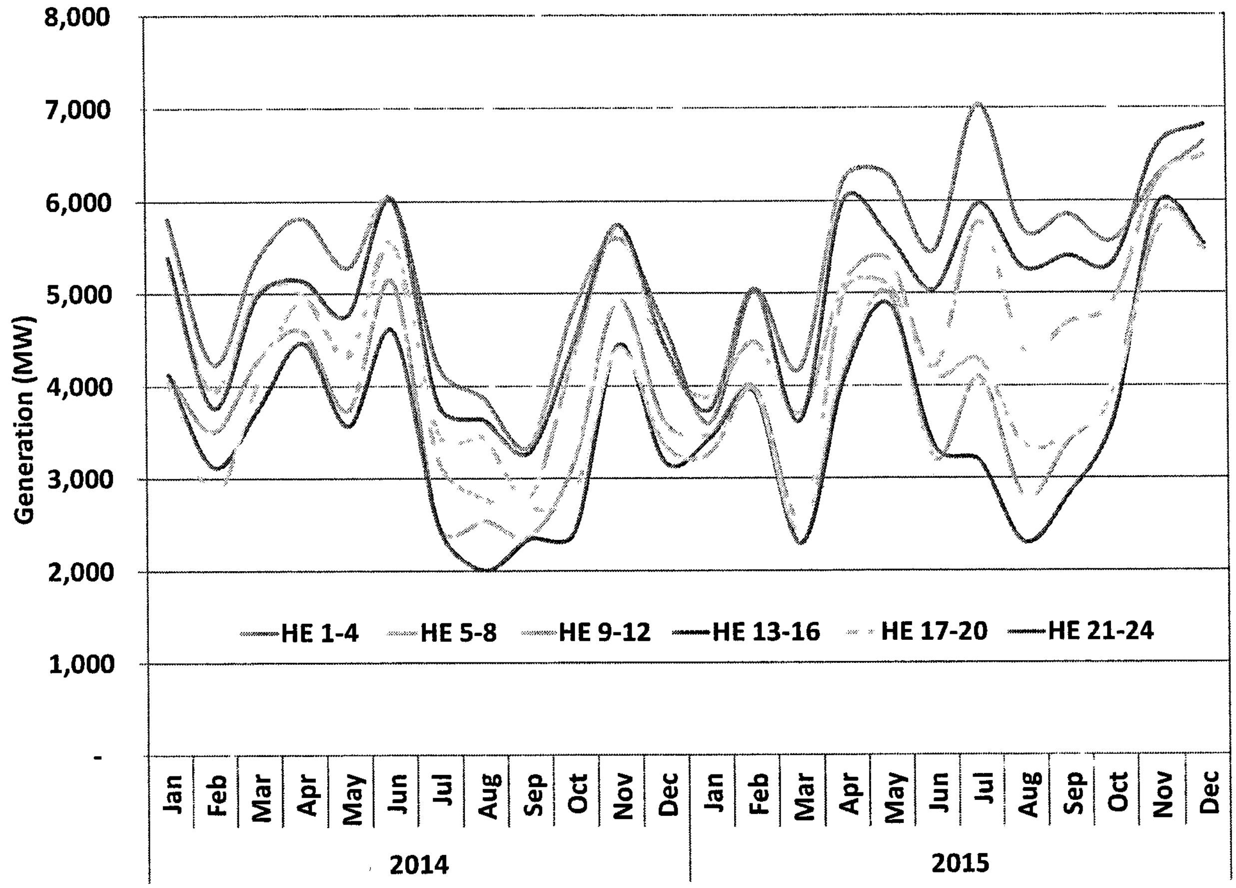
### 1. Wind and Solar Generation

The amount of wind generation installed in ERCOT was approximately 16 GW by the end of 2015. 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. This subsection will more fully describe the characteristics of wind generation in 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 59 shows average wind production for each month in 2014 and 2015, 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 approximately 3 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.

[19] Natural gas provided 40.5 percent of total generation in 2013, and 41.1 percent in 2014.

**Figure 59: Average Wind Production**




ERCOT continued to set new records for peak wind output in 2015. On December 20, wind output exceeded 13 GW, setting the record for maximum output and providing nearly 45 percent of hourly generation.

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 60 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 other areas in ERCOT across various load levels. The "Others" category is primarily composed of wind generators in West Texas and some in the northern part of the state. There is a strong negative relationship between wind output in the "Others" category and increasing load levels. It further shows that the output from wind generators located in the coastal area of the South zone is much more highly correlated with peak electricity demand. Other than at loads greater than 65 GW, Panhandle wind shows a more stable output across the load levels.

**Figure 60: Summer Wind Production vs. Load**

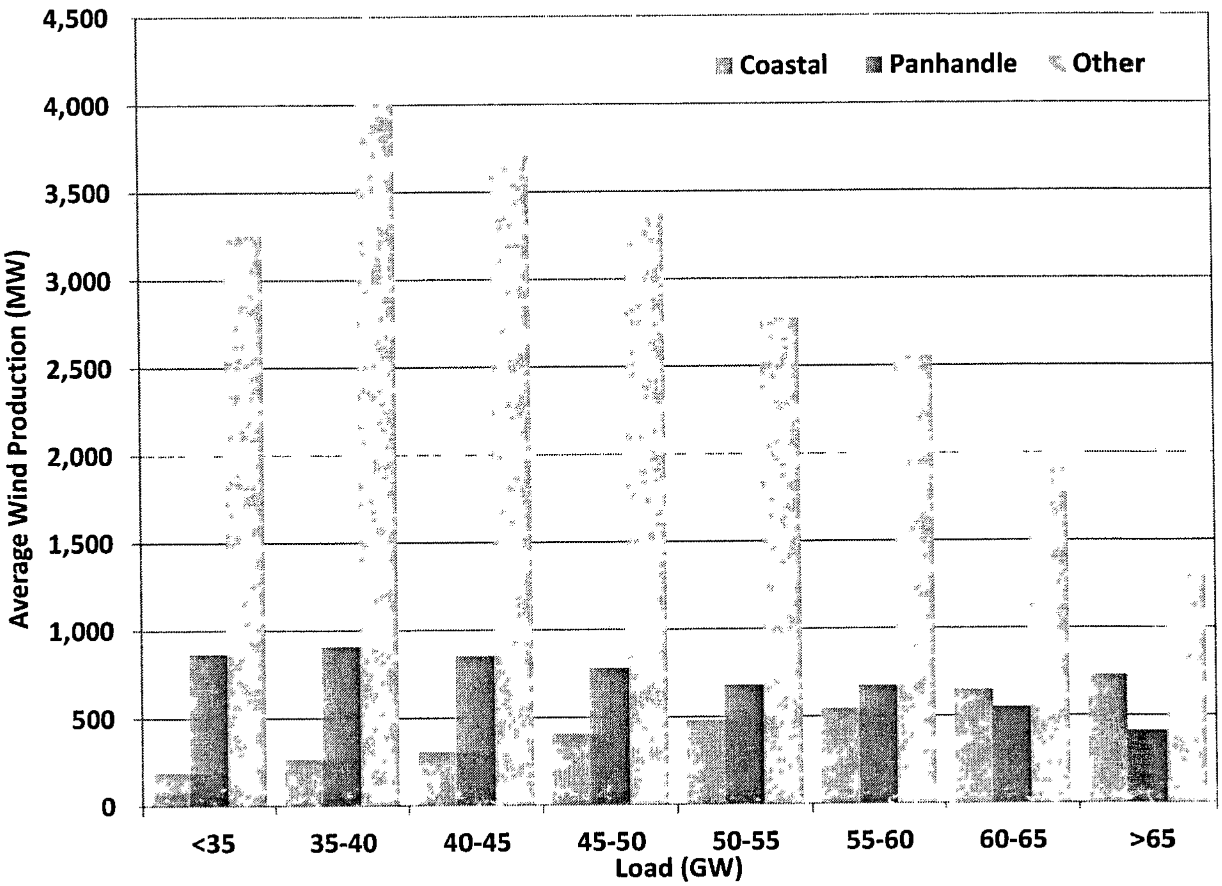



Figure 61 shows the wind production and estimated curtailment quantities for each month of 2012 through 2015. This figure reveals that the total production from wind resources continued to increase, while the quantity of curtailments was up slightly from 2014. The volume of wind actually produced in 2015 was estimated as 99 percent of the total available wind, compared with 99.5 percent in 2014, 98.9 percent in 2013 and 96 percent in 2012.

Figure 61: Wind Production and Curtailment

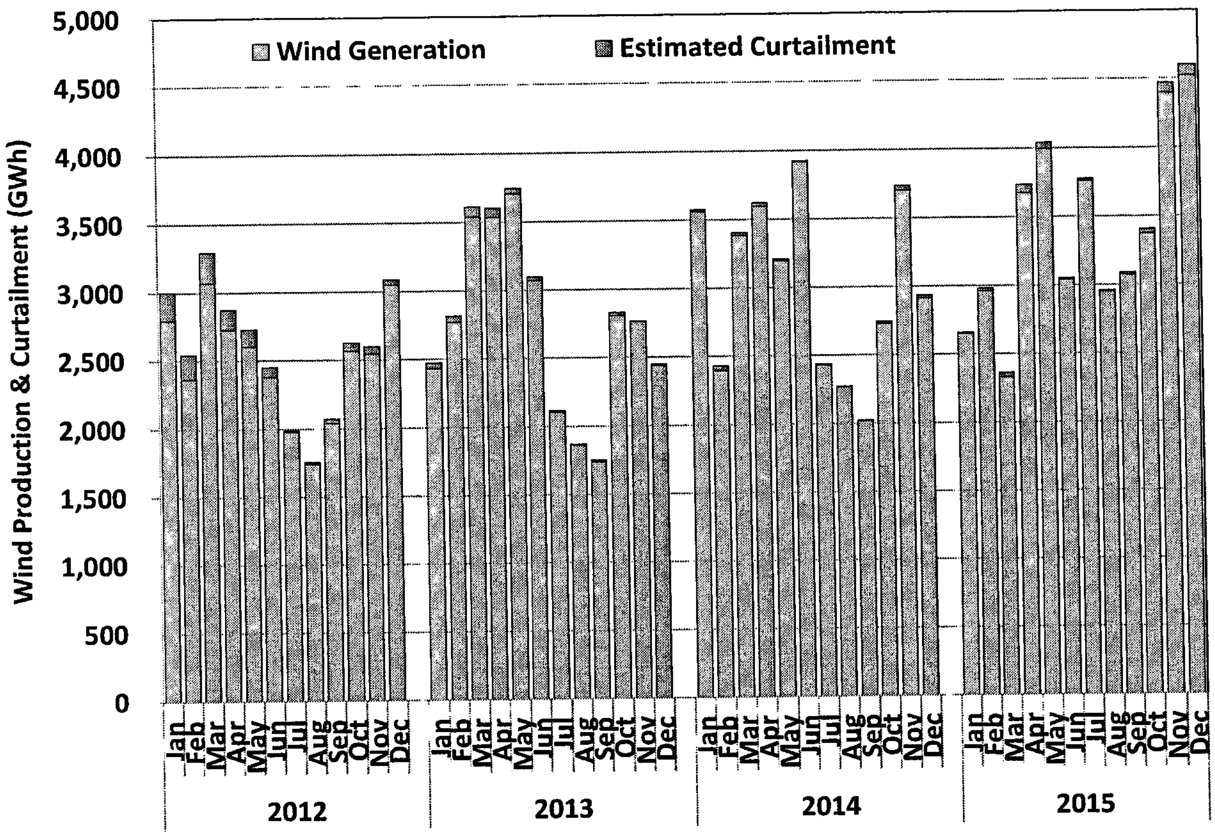
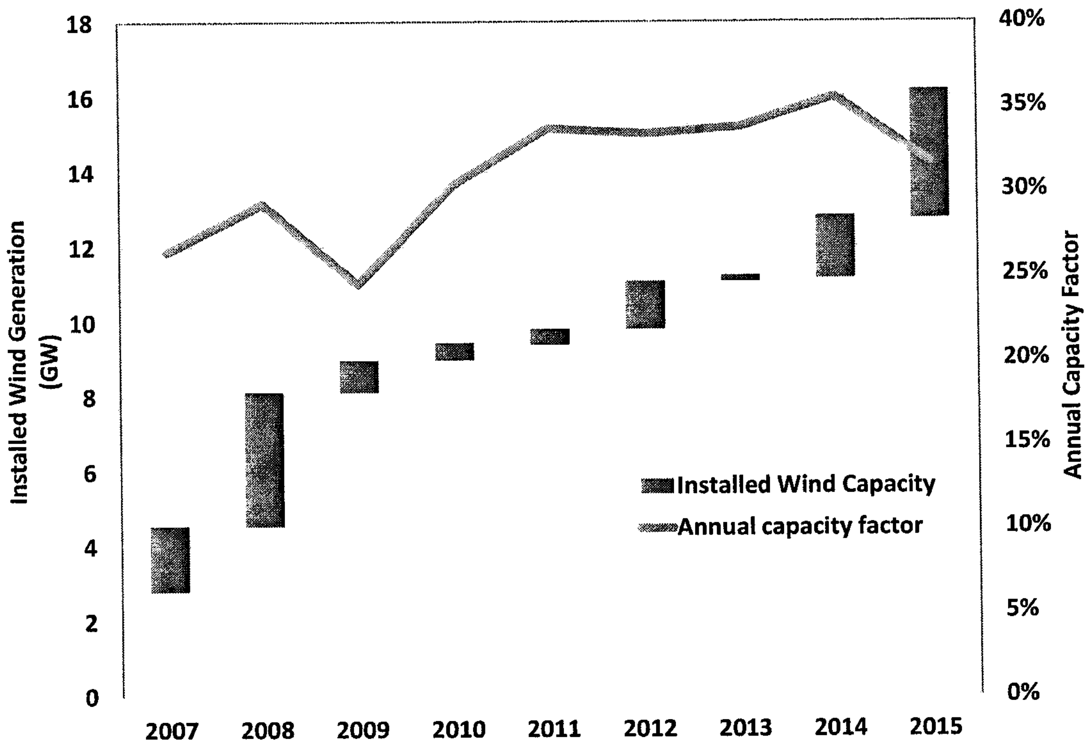



Figure 62 below, shows the quantities of wind generation installed every year from 2007 and the annual capacity factor of the wind output for each year. The amount of wind generation installed in 2015 was almost as large as in 2008. This was likely driven by two factors, the completion of the CREZ transmission lines and the scheduled expiration of the federal production tax credits at the end of 2015. The federal production tax credits have been extended for another four years, which should reduce the pressure on developers to quickly complete all future projects.

Figure 62 also provides the annual wind generation capacity factor. Prior to 2011, annual capacity factors were less than 30 percent, which reflected the large amounts of curtailments incurred due to transmission limitations. As completed CREZ lines allowed more wind to be produced, curtailments were reduced to approximately 1 percent. Capacity factors were 34, 36, and 32 percent in 2013, 2014, and 2015, respectively. These differences are now the result of natural variations in wind availability. So even though wind generation provided 12 percent of annual generation requirements, which was a new record, it occurred in a relatively low wind year.

**Figure 62: Wind Generation Capacity Factor**




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 63 shows the net load duration curves for the years 2015, 2011, and 2007, normalized as a percentage of peak load.

**Figure 63: Net Load Duration Curves**

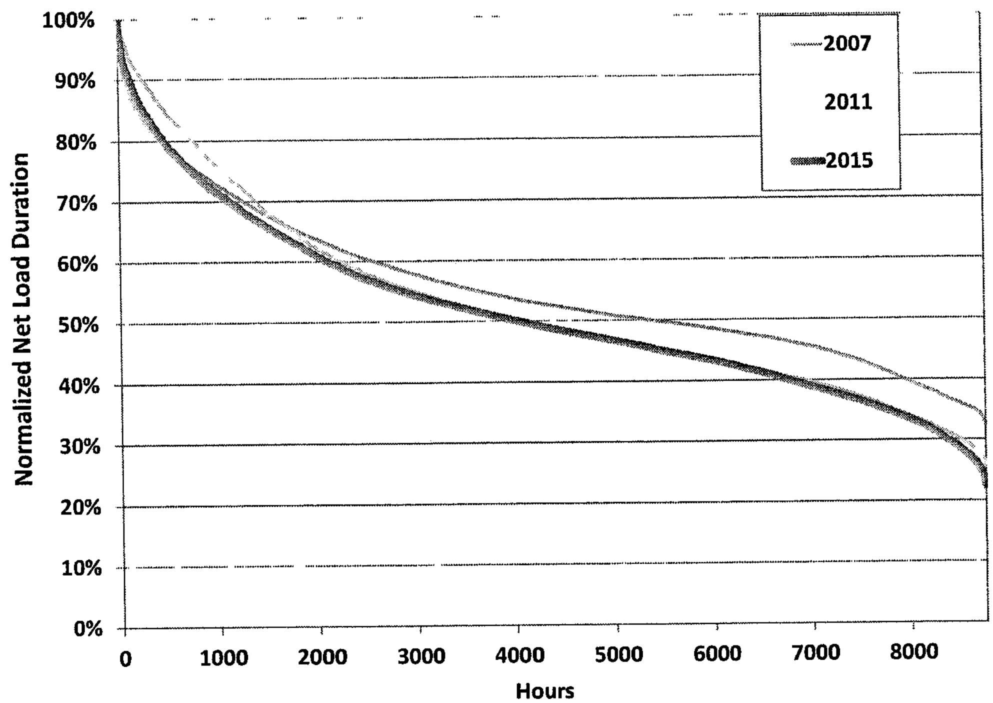
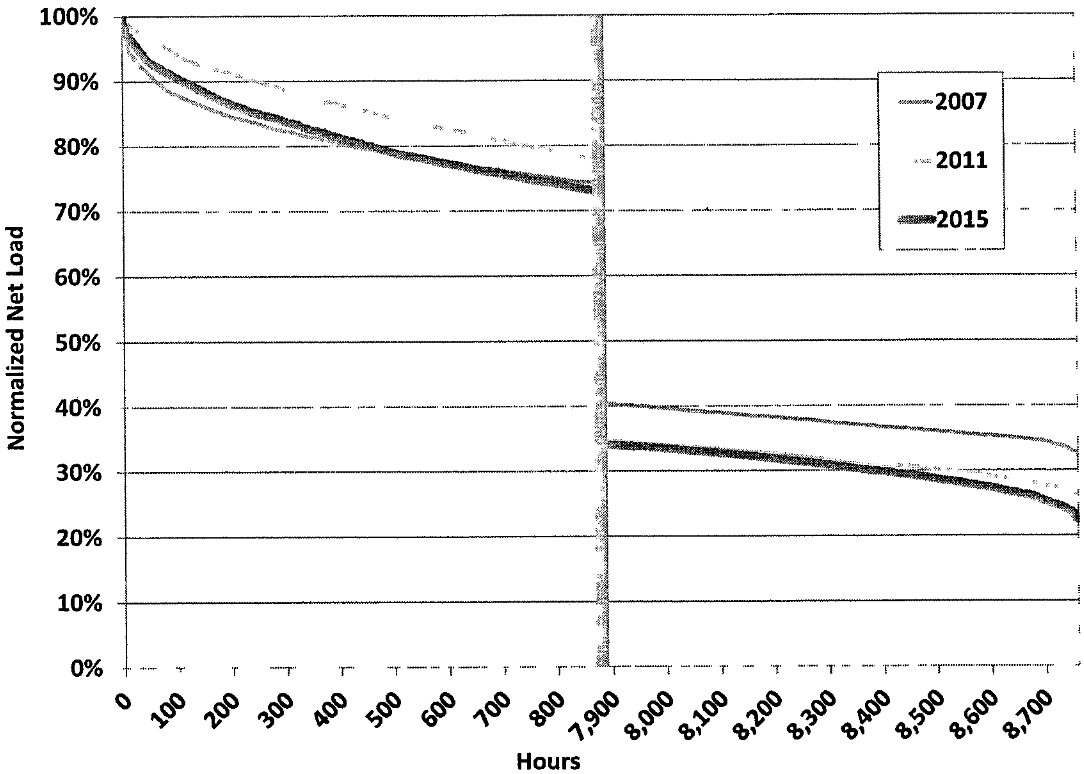



Figure 63 shows the reduction of remaining energy available for non-wind units to serve during most hours of the year, even after factoring in several years of load growth. The impact of wind on the highest net load values is much smaller.

Figure 64 shows net load in the highest and lowest hours. Even with the increased development activity in the coastal area of the South zone, 74 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 hours of highest demand, 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 baseload coal units, while doing very little to reduce the amount of capacity necessary to reliably serve peak load.

**Figure 64: Top and Bottom Ten Percent of Net Load**



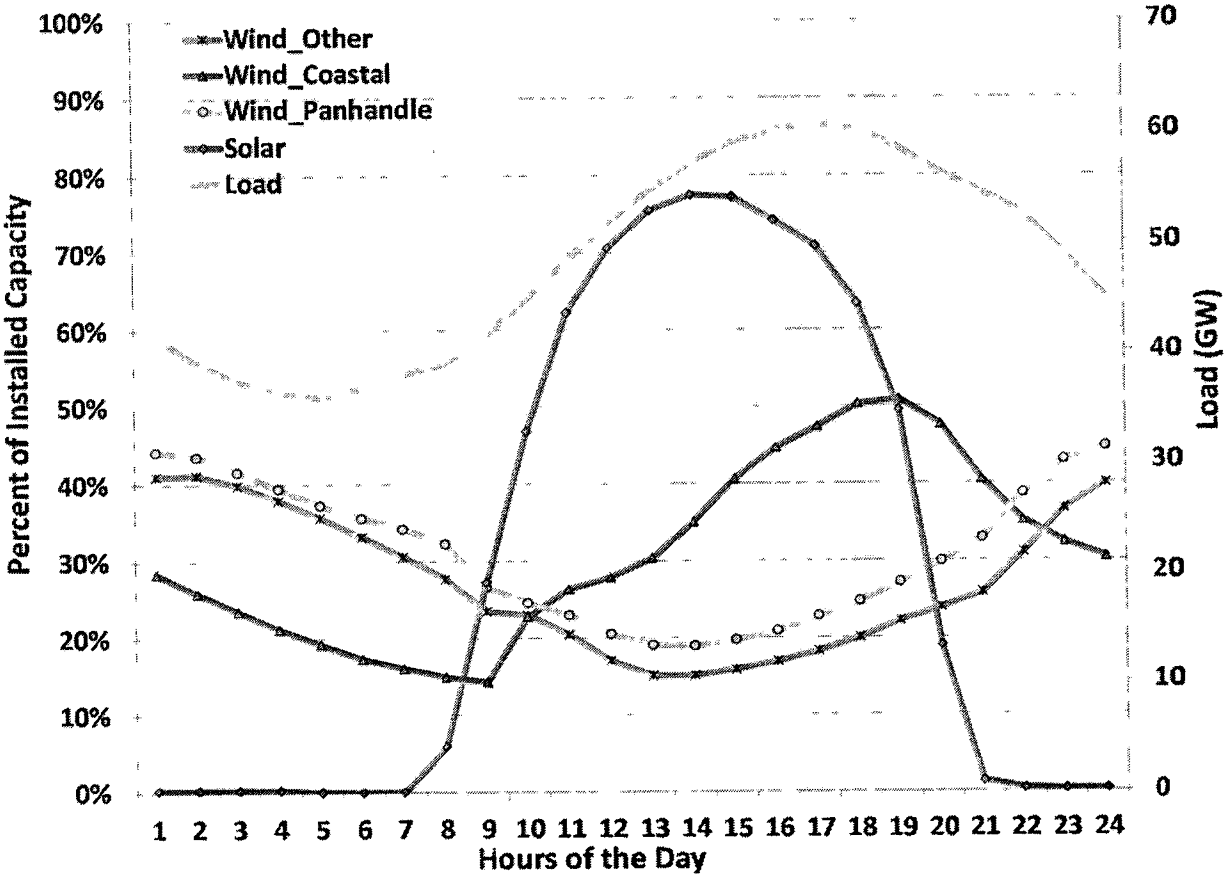

In the hours with the highest net load (left side of the figure above), the difference between peak net load and the 95th percentile of net load has averaged 12.3 GW the past three years. This means that 12.3 GW of non-wind capacity is needed to serve load less than 440 hours per year.

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 last year, even with sizable growth in total annual load. This continues to put operational pressure on the 23.5 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 meet net load and 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 within 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 65 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 65: Summer Renewable Production**




This figure shows that the total installed capacity of solar generation is much smaller than that of wind generation. However, its production as a percentage of installed capacity is the highest in the early afternoon, nearing 80 percent, and producing more than 60 percent of its installed capacity during peak load hours.

The contrast between coastal wind and non-coastal wind is also clearly displayed in Figure 65. Coastal wind produced nearly 50 percent of its installed capacity during summer peak hours. Output from Panhandle wind exceeded 20 percent, while output from non-coastal wind (primarily West Zone) was less than 20 percent during summer peak hours.

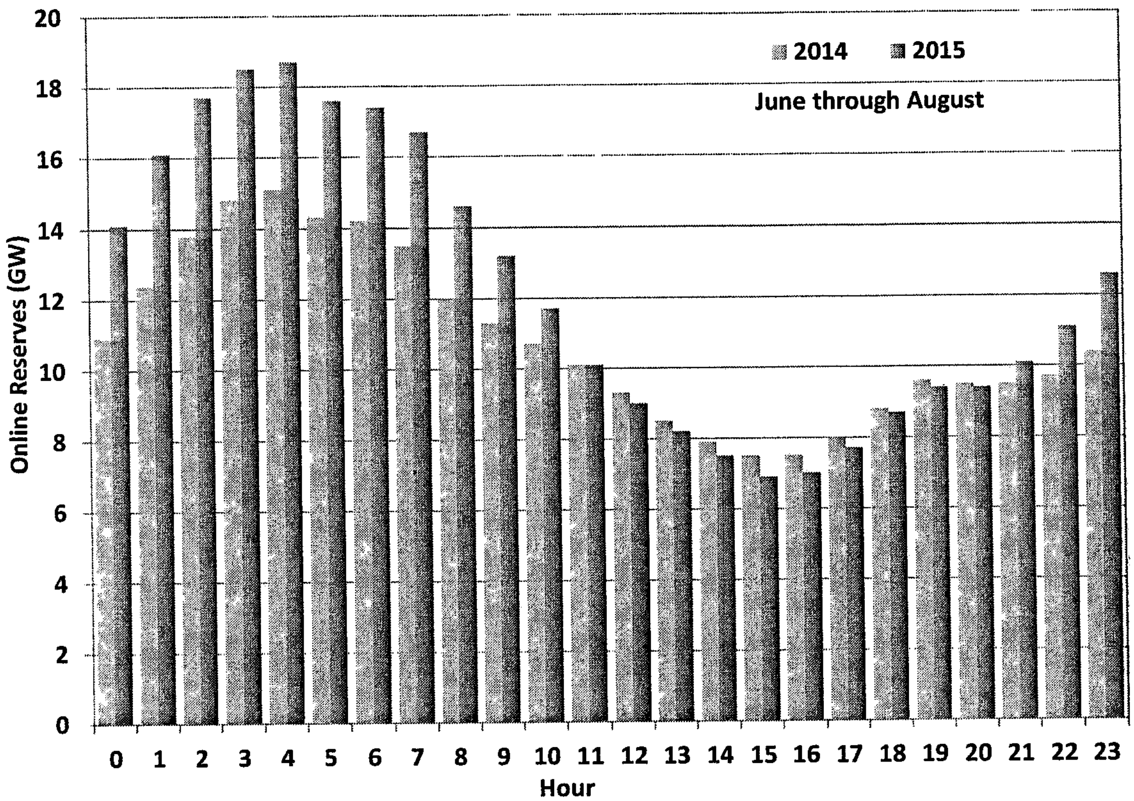
### 2. Resource Commitments for Reliability

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 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 it is important to note that 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.

The following figure compares the amount of on-line reserves, by hour, for the summer months of June through August in 2015 and 2014. The amount of on-line reserves is equal to the amount of capacity committed in excess of expected demand. Figure 66 displays available online reserves by operating hour and shows the expected pattern of declining reserves as system load increases during peak demand hours. Two interesting patterns emerge from this data. First, there were significantly more online reserves during overnight hours in 2015. Second, the online reserves during peak operating hours in summer 2015 were generally lower than in 2014.

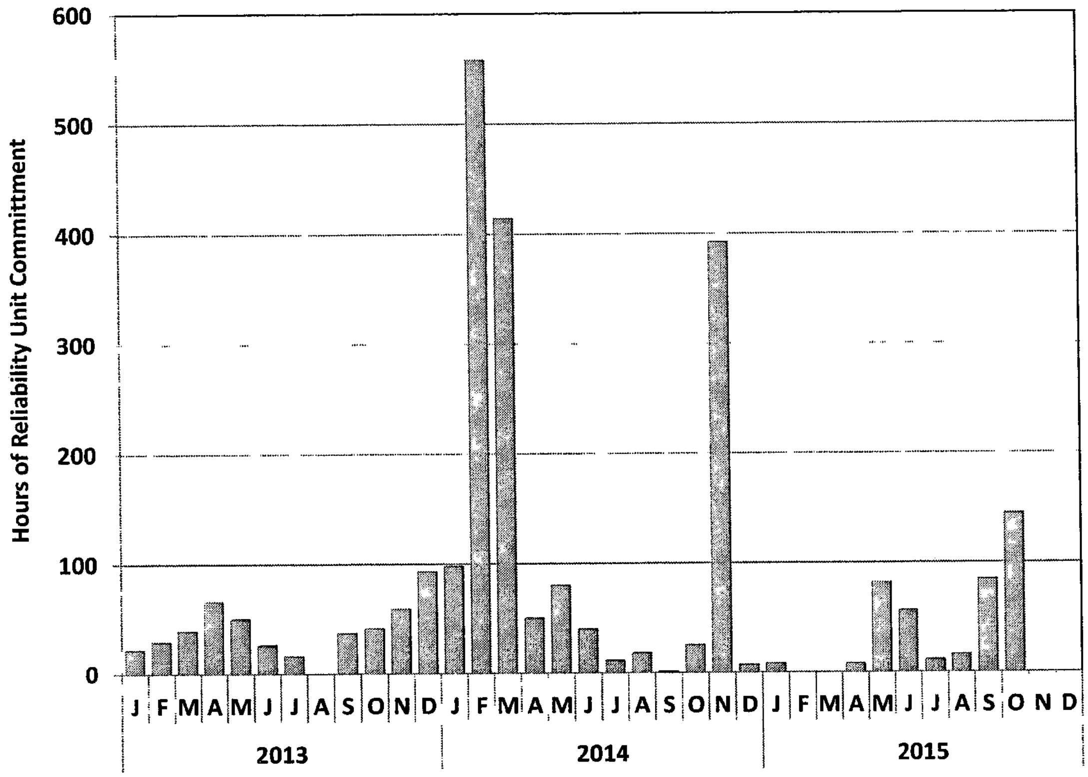
**Figure 66: Average On-line Summer Reserves**




One possible explanation for the increased off-peak capacity commitments in 2015 is that the low natural gas prices reduced minimum load operating costs, leading many generators to keep their units online overnight rather than incurring the risk and cost of shutting down and starting the unit every day. The small (500-600 MW) reductions to online reserves during peak hours in 2015 are explained by much higher (2500-2600 MW) average loads in those hours.

Once ERCOT assesses the unit commitments resulting from the day-ahead market, additional capacity commitments are made, if needed, using a reliability unit commitment (RUC) process that executes both on a day-ahead and hour-ahead basis. These additional unit commitments may be made for one of two reasons. Either additional capacity is required to ensure forecasted total demand will be met, or a specific generator is required to resolve a transmission constraint. The constraint may be either a thermal limit or to support a voltage concern. Figure 67 shows how frequently these reliability unit commitments have occurred over the past three years, measured in unit-hours.

**Figure 67: Frequency of Reliability Unit Commitments**




There was a significant decrease in the frequency of reliability unit commitments in 2015. During 2015, five percent of hours had at least one unit receiving a reliability unit commitment instruction. This is the same as the percent of hours in 2013 and down from 2014 when 19 percent of hours had RUC instructions. Most of the unusually high RUC activity in 2014 occurred during cold winter weather. In 2015, RUC commitments were most frequent in the fall due to congestion in Dallas and the Rio Grande Valley.

Table 4 provides the units most frequently called upon for RUC. Also provided are the hours of RUC instruction and the number of hours in which the unit opted out. A unit that receives a RUC instruction is guaranteed payment of its start-up and minimum energy costs (RUC Make-Whole). 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 Claw-Back). Beginning in January 2014, a unit receiving a RUC instructions had the choice to "Opt Out," meaning it would forgo all RUC Make-Whole in return for not being subject to RUC Claw-Back. In 2015, units receiving RUC instructions elected to opt out 34 percent of unit-hours.

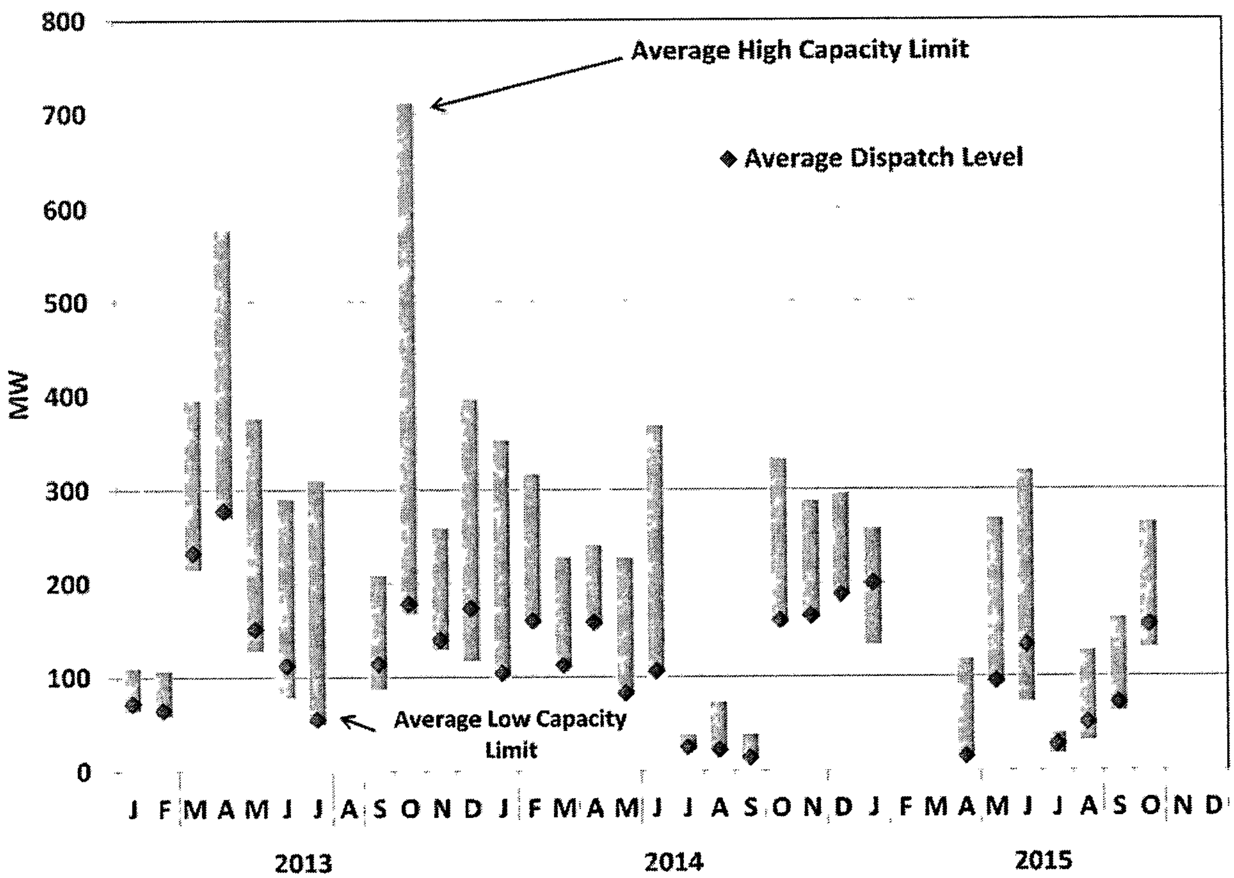
**Table 4: Most Frequent Reliability Unit Commitments**

| Resource | Location | Reason | Unit RUC Hours | Unit OPTOUT Hours |
|---|---|---|---|---|
| Silas Ray CC1 | Valley | Congestion | 54 | 44 |
| Morgan Creek Unit 8 | Dallas | Congestion | 58 | - |
| Silas Ray 10 | Valley | Congestion | 15 | 23 |
| Morgan Creek Unit 7 | Dallas | Congestion | 30 | - |
| Decker DPG2 | Austin | Congestion | - | 22 |
| Barney Davis CC1 | Corpus Christi | Congestion | - | 16 |
| Decker DPG1 | Austin | Congestion | - | 16 |
| Nueces Bay CC1 | Corpus Christi | Congestion | 16 | - |
| North Edinburg CC1 | Valley | Congestion & Voltage | 14 | - |
| Frontera CC1 | Valley | Congestion | 9 | - |
| Lake Hubbard Unit 1 | Dallas | Congestion | 3 | 6 |
| Mountain Creek Unit 6 | Dallas | Congestion | 8 | - |
| Midlothian CT 4 | Dallas | Capacity | - | 8 |
| Forney CC1 | Dallas | Voltage | 7 | - |
| Midlothian CT 5 | Dallas | Voltage | 7 | - |

The vast majority of the 411 unit-hours with RUC instructions during 2015 were to resolve localized thermal transmission constraints (93 percent), and of those the majority were to units located in the Rio Grande Valley (37 percent) and in Dallas (27 percent). A small number of commitments, 20 unit-hours or 5 percent, were made due to voltage concerns in the Rio Grande Valley and in Dallas during the off-peak hours during a planned outage of one of the Comanche Peak units. There were eight unit-hour commitments (2 percent) given to a Midlothian Combustion Turbine for system-wide capacity requirements. This compares to 2014 when 18 percent of the unit hours of RUC instructions were for system-wide capacity requirements, primarily during the period from January through March.

The next analysis compares the average dispatched output of the reliability committed units with the operational limits of the units. Figure 68 below shows that the quantity of reliability unit commitment generation also decreased in 2015 compared to 2014. This figure shows that the average quantity dispatched during any month in 2015 was always less than 200 MW, and less than 100 MW in five of the eight months with reliability unit commitments. Therefore, the energy produced from reliability committed units does not generally displace a large quantity of energy from market committed units.

**Figure 68: Reliability Unit Commitment Capacity**




Additionally, this figure shows a decreasing trend in reliability unit commitments since 2013, which is good because such commitments and undermine real-time pricing.

## C. Demand Response Capability

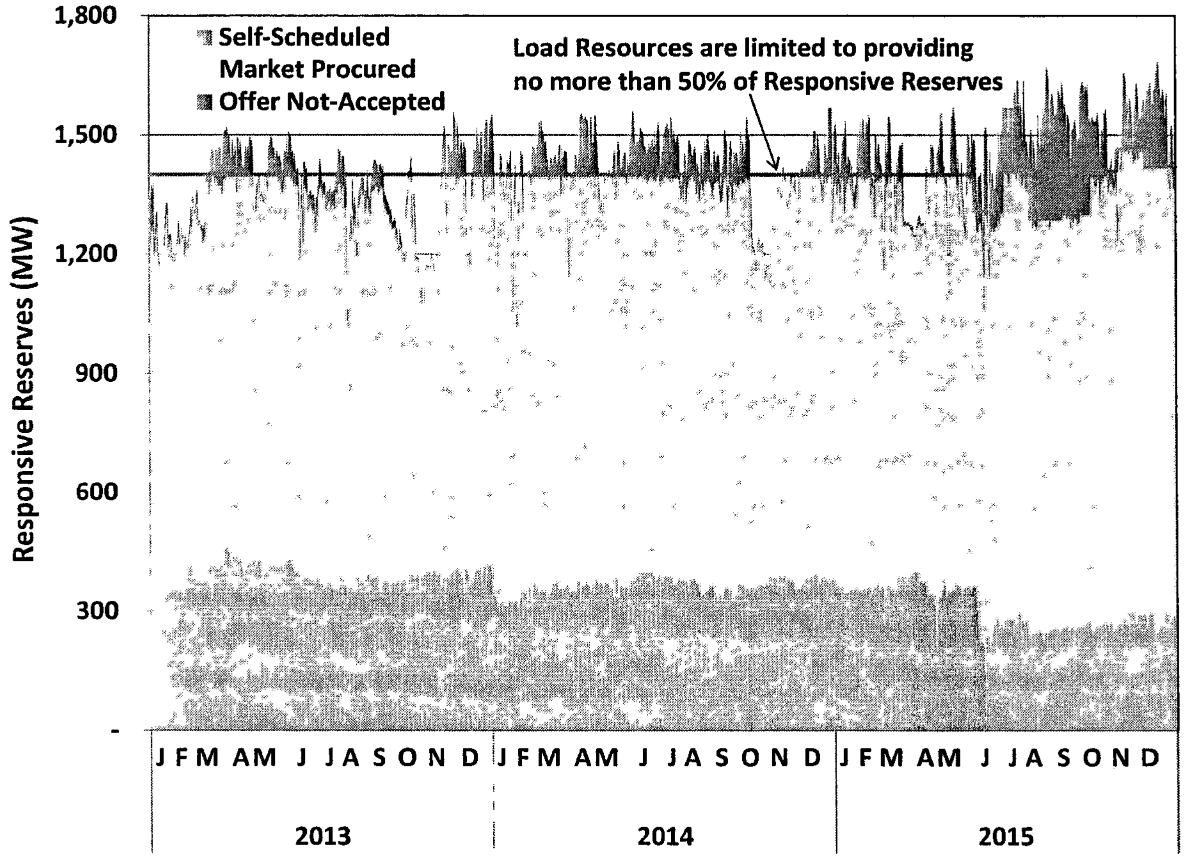
Demand response is a term that broadly refers to actions that can be taken by end users of electricity to reduce load in response to instructions from ERCOT or in response to certain market or system conditions. The ERCOT market allows participants with demand-response capability to provide energy and 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. Unlike active participation in ERCOT-administered markets, self-dispatch by demand is not directly tracked by ERCOT.

### 1. 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 2015, approximately 3,413 MW were qualified as Load Resources.

Figure 69 shows the amount of responsive reserves provided from load resources on a daily basis for the past three years. For reliability reasons the maximum amount of responsive reserves that can be provided by load resources is limited to 50 percent of the total.

**Figure 69: Daily Average of Responsive Reserves Provided by Load Resources**



For the first five months of 2015, the RRS procurement amount was constant at 2,800 MW, of which 1,400 MW could be provided by load. On June 1, 2015, ERCOT began procuring a variable amount of RRS based on season and time of day. The total amount of RRS varies between 2,300 to 3,000 MW. During 2015, there were no system-wide manual deployments of load resources providing RRS and only one automatic deployment of a small portion of frequency responsive load.

Figure 69 shows amounts of responsive reserves that were either self-scheduled or offered by load resources. The quantity of offers submitted by load resources exceeds the 50 percent limit most of the time. This is only generally not the case 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.

### 2. Reliability Programs

There are two main reliability programs in which demand can participate in ERCOT – Emergency Response Service and transmission provider load management programs. The Emergency Response Service (ERS) product is defined by a PUCT Rule enacted in March of 2012 setting a program budget of $50 million.[20] The amount of ERS procured ranged from just over 400 MW to almost 900 MW across the various periods in the 2015 program year. 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. The average price paid for ERS over the contract periods from February 2015 through January

[20] *See* 16 TEX. ADMIN. CODE § 25.507.

2016 was $10.50 per MW-hour, higher than the average price of $6.92 per MW-hour paid for non-spinning reserves in 2015. ERS was not deployed in 2015.

A load has to make a choice between participating in ERS, providing Ancillary Services, or simply choosing to curtail in response to high wholesale energy prices. A specific load cannot provide more than one of these functions at the same time. Given the high budget allotted and the low risk of deployment, ERS is a very attractive program for loads. Because the ERS program is so lucrative, there is concern that it is limiting the motivation for loads to actively participate in SCED and contribute to price formation. We suggest that the PUCT evaluate the need for and structure of this program.[21]

Beyond ERS there are slightly less than 200 MW of load participating in load management programs administered by transmission providers. 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).

### 3. 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 to all loads in ERCOT 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. Over the last three years, transmission costs have risen by more than 60 percent, thus

[21] On May 4, 2016, the PUCT opened Docket 45927, *Rulemaking Regarding Emergency Response Service.*

significantly increasing an already substantial incentive to reduce load during probable peak intervals in the summer. ERCOT estimates that over 800MW of load is actively pursuing reduction during these intervals.[22]

Two recent changes in the ERCOT market have made advances in appropriately pricing actions taken by load during 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, performs a second pricing run of SCED to account for the amount of load deployed, including ERS.

22 *See* ERCOT, *2015 Annual Report of Demand Response in the ERCOT Region* (Mar. 2016) at 6, available at http://www.ercot.com/services/programs/load.

# V. 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

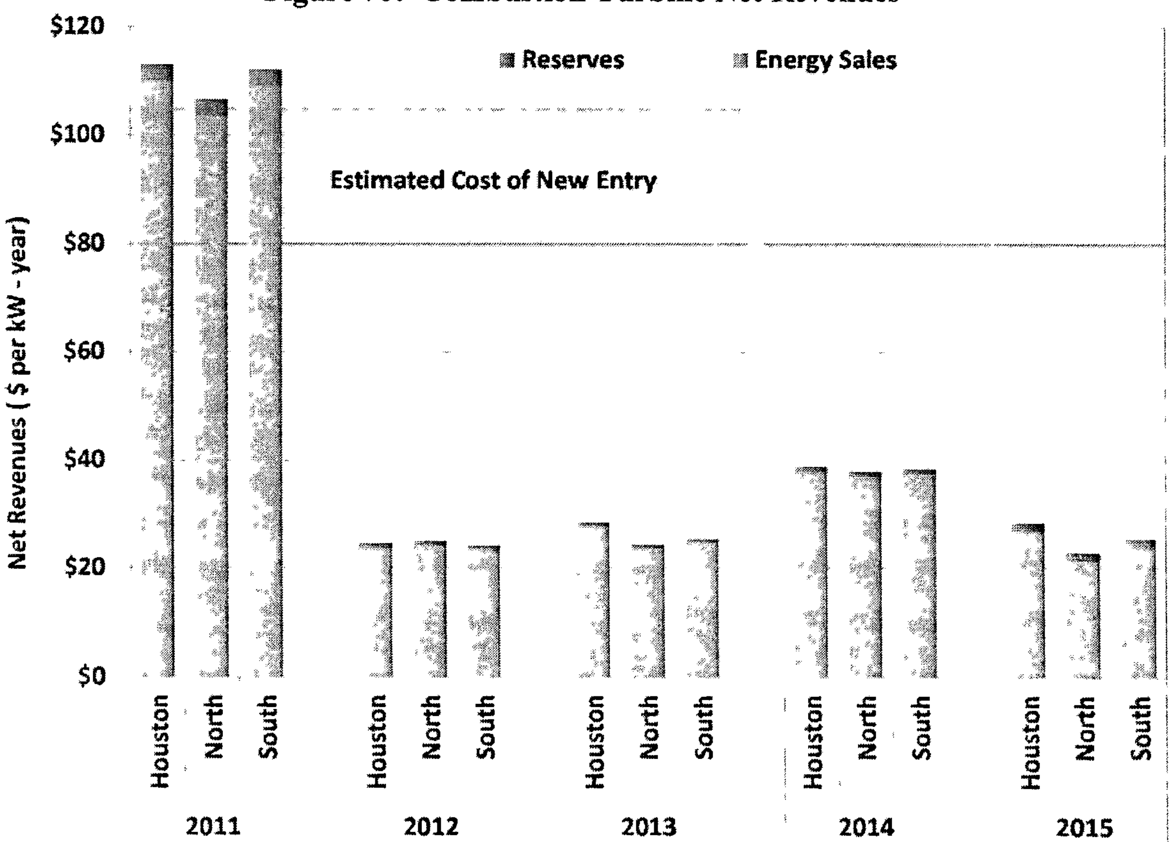
Net revenue is calculated by determining the total revenue that could have been earned by a generating unit less its variable production costs. Put another way, it is the revenue in excess of short-run operating costs that is available to recover a unit's fixed and capital costs, including a return on the investment. Net revenues from the real-time energy and ancillary services markets provide economic signals that help inform suppliers' decisions to invest in new generation or retire existing generation. 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 new natural gas combustion turbine (Figure 70) and combined cycle generation (Figure 71).

**Figure 70: Combustion Turbine Net Revenues**

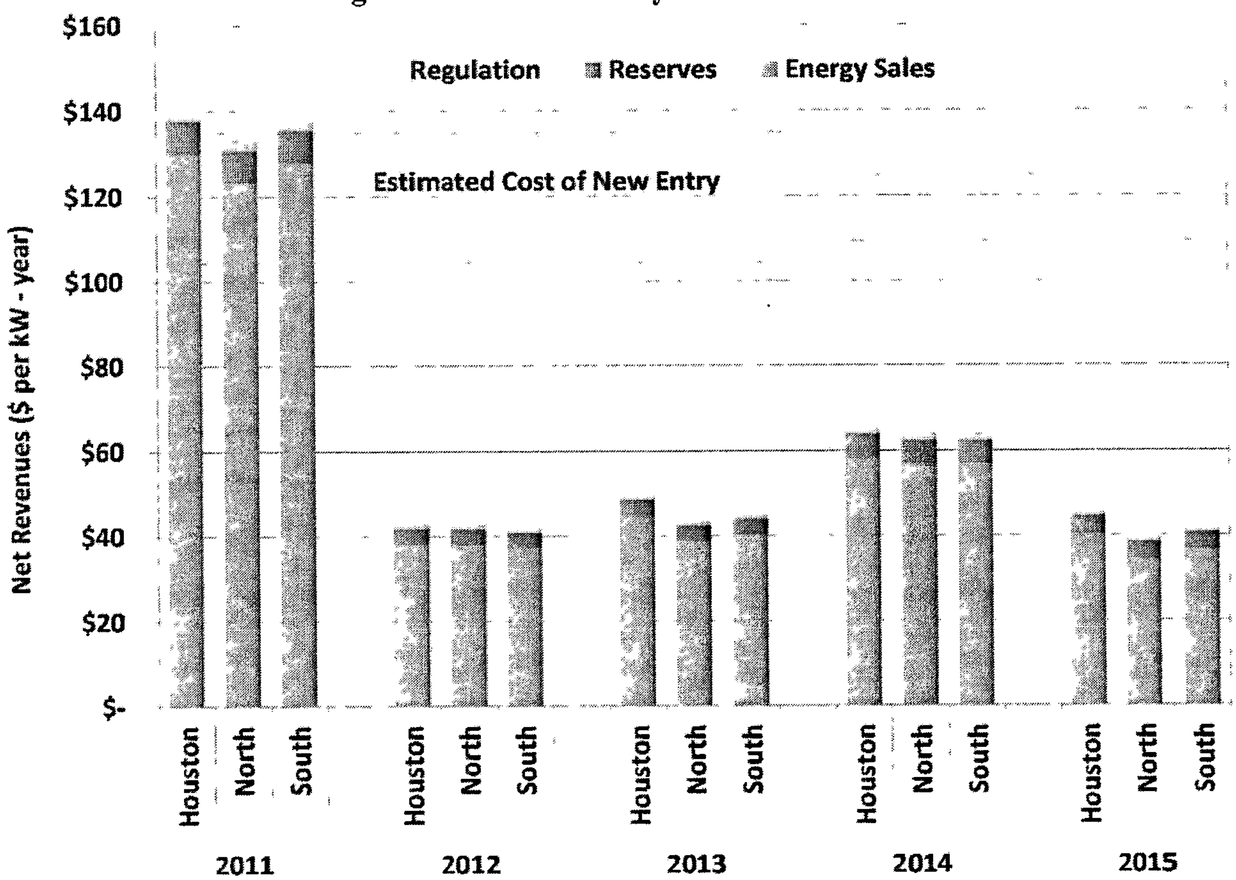





0000000126

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 2015 for a new gas turbine was calculated to be approximately $23 to 29 per kW-year, depending on the zone. These values are well below the estimated cost of new gas turbine generation.

**Figure 71: 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 2015 for a new combined cycle unit was calculated to be approximately $40 to 46 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 that exists over the minimum target level, which contributed to infrequent shortages in 2015. 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.

Given the very low energy prices during 2015 in non-shortage hours, the economic viability of existing coal and nuclear units was evaluated. The prices in these hours, which have been substantially affected by the prevailing natural gas prices, determine the vast majority of net revenues received by these baseload units. As previously described, the load-weighted ERCOT-wide average energy price in 2015 was $26.76 per MWh. The generation-weighted average price for the four nuclear units - approximately 5GW of capacity - was $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 $27.53 per MWh in 2015.[23] Assuming that operating costs in ERCOT are similar to the U.S. average, considering only fuel and operating and maintenance costs indicates that nuclear generation was not profitable in ERCOT during 2015. To the extent nuclear units in ERCOT had any associated capital costs, it is likely those costs were not recovered.

The generation-weighted price of all coal and lignite units in ERCOT during 2015 was $25.94 per MWh. Although specific unit costs may vary, index prices for Powder River Basin coal delivered to ERCOT were approximately $3 per MMBtu in 2015. With a typical heat rate of 10 MMBtu per MWh, the fuel-only operating costs for coal units in 2015 may be inferred to be approximately $30 per MWh. As with nuclear units, it appears that coal units were likely not profitable in ERCOT during 2015. This is significant because the retirement or suspended operation of some of these units could cause ERCOT's capacity margin to fall below the minimum target more quickly than anticipated.

These results indicate that during 2015 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.

[23] NEI Whitepaper, "Nuclear Costs in Context", April 2016, available at http://www.nei.org/CorporateSite/media/filefolder/Policy/Papers/Nuclear-Costs-in-Context.pdf?ext=.pdf.

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 2015. 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 2015, 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 the 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.

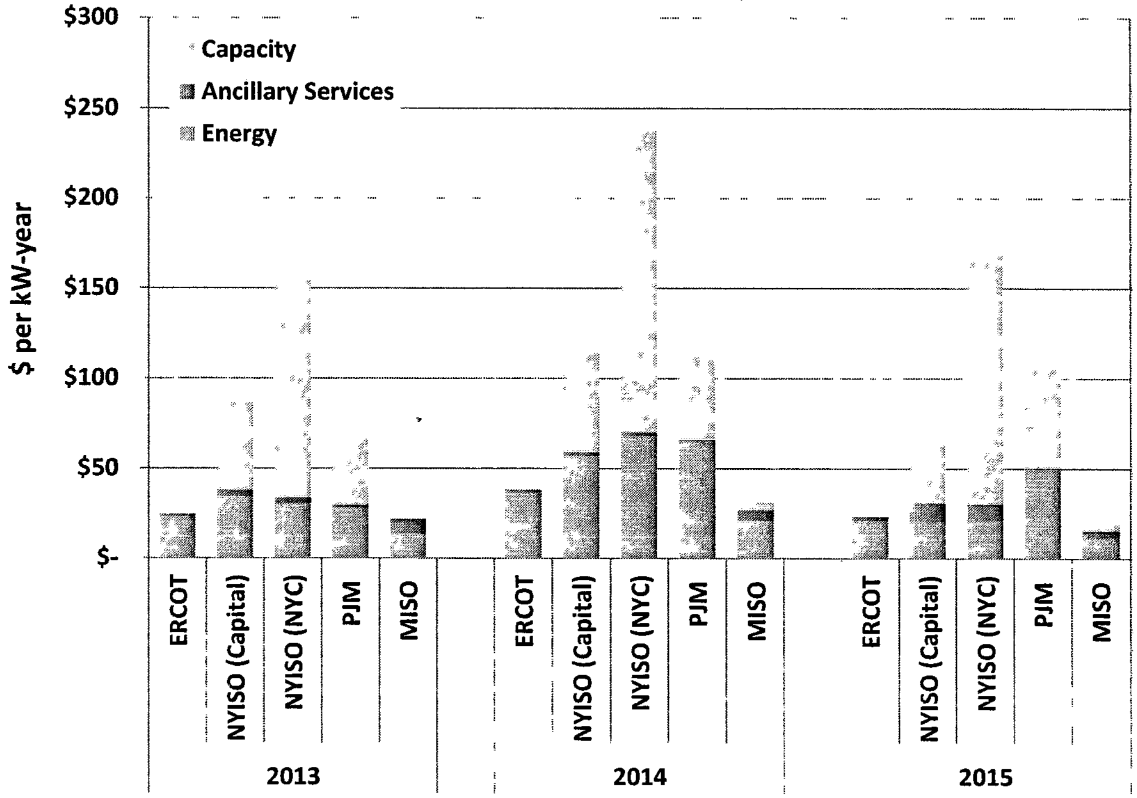
Finally, the costs of new entry used in this report are generic and reflective of the costs of a new unit on an undeveloped greenfield site. They have been reduced somewhat to reflect the lower costs of construction in Texas. However, companies may have opportunities to build generation at much lower 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 lower than 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.

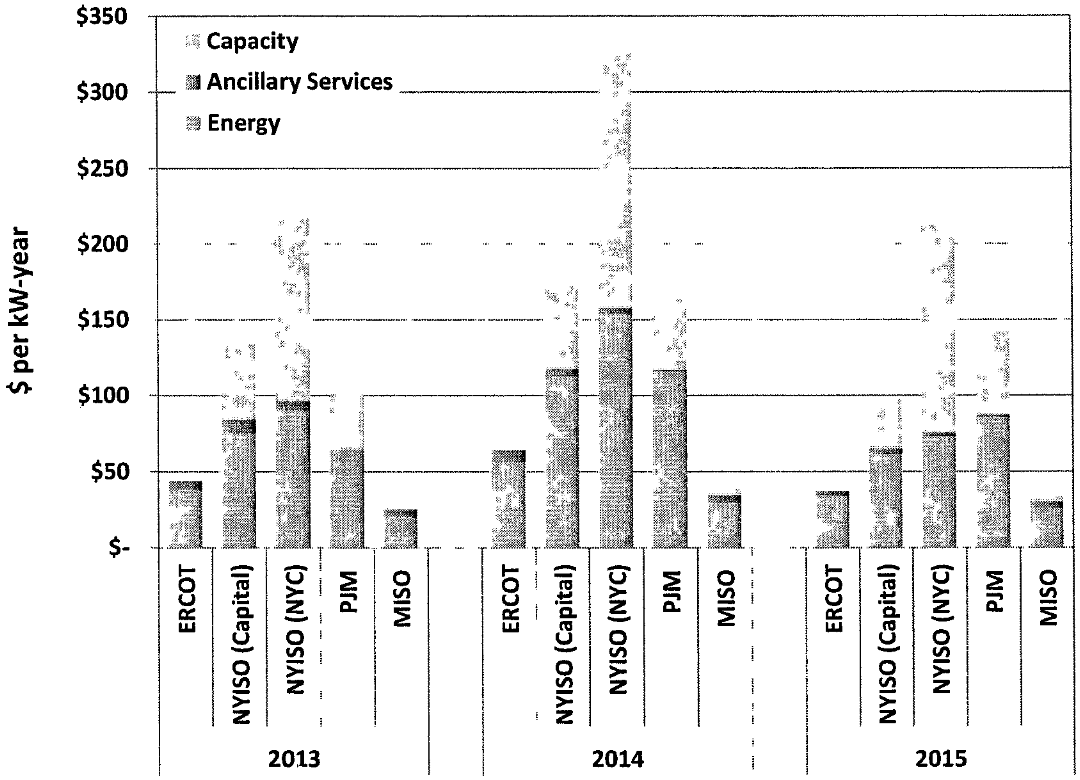
The next two figures compare estimates of net revenue for these two types of natural gas generators for the ERCOT North zone, PJM, two locations within the New York ISO, and the Midcontinent ISO. Figure 72 provides a comparison of net revenues for a combustion turbine and Figure 73 provides the same comparison for a combined cycle unit.

**Figure 72: 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 NYISO 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 73: Combined Cycle Net Revenue Comparison Between Markets**




Both figures indicate that across all markets net revenues decreased substantially in 2015 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,[24] this subsection assesses the Scarcity Pricing Mechanism (SPM) in 2015 under ERCOT's energy-only market structure.

[24] *See* 16 TEX. ADMIN. CODE § 25.505(g)(6)(D).

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 19 on page 20, there have been very brief periods when energy prices rose to the cap since the system-wide offer cap was increased to greater than $3,000 per MWh. There have been no instances of prices rising above $5,000 per MWh.

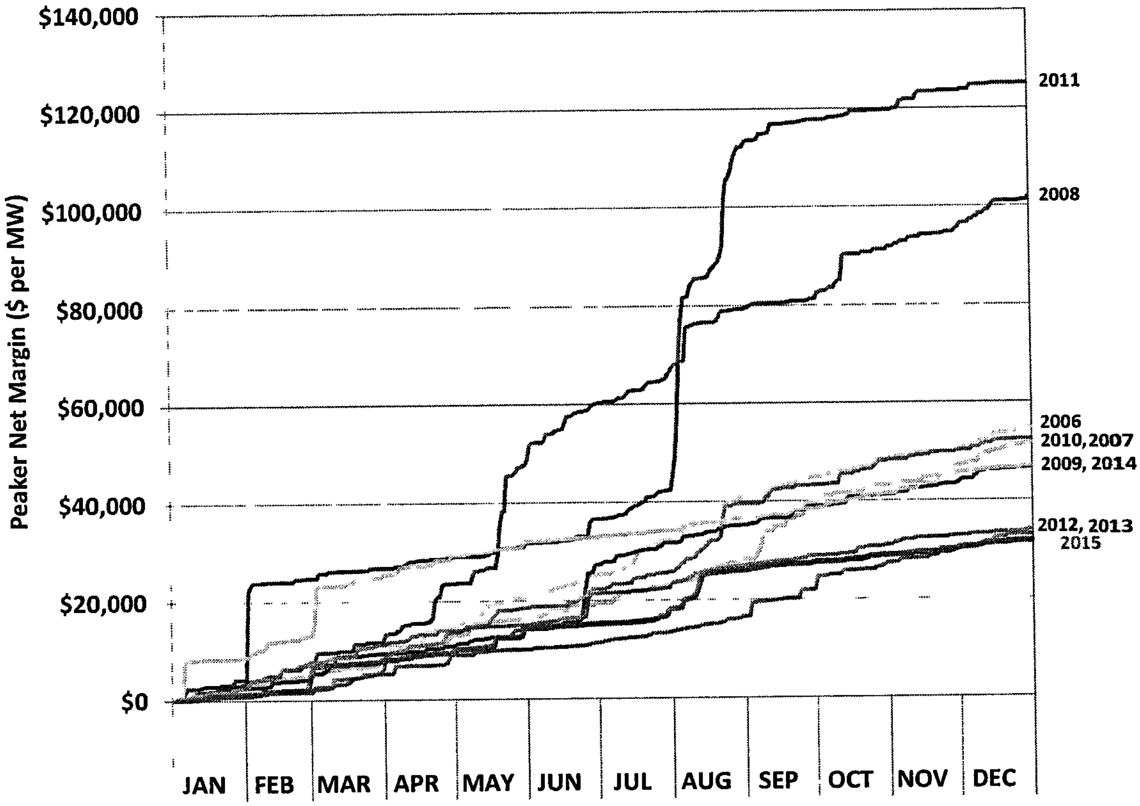
The SPM includes a provision termed the Peaker Net Margin (PNM) that is designed to provide a fail-safe pricing measure, which if exceeded would cause the system-wide offer cap to be reduced. If the PNM for a year reaches a cumulative total of $315,000 per MW, the system-wide offer cap is then reduced to the higher of $2,000 per MWh or 50 times the daily natural gas price index.[25] PNM also serves as a simplified measure of the annual net revenue of a hypothetical peaking unit.[26]

Figure 74 shows the cumulative PNM results for each year from 2006 through 2015 and shows that PNM in 2015 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.

[25] 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.

[26] 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 74: 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.[27] 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.D, 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.

27 Not that the zonal market design was the problem per se, but rather its reliance on high-priced offers to set high prices during periods of shortage.

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."[28] Given how long it has been in place, 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.

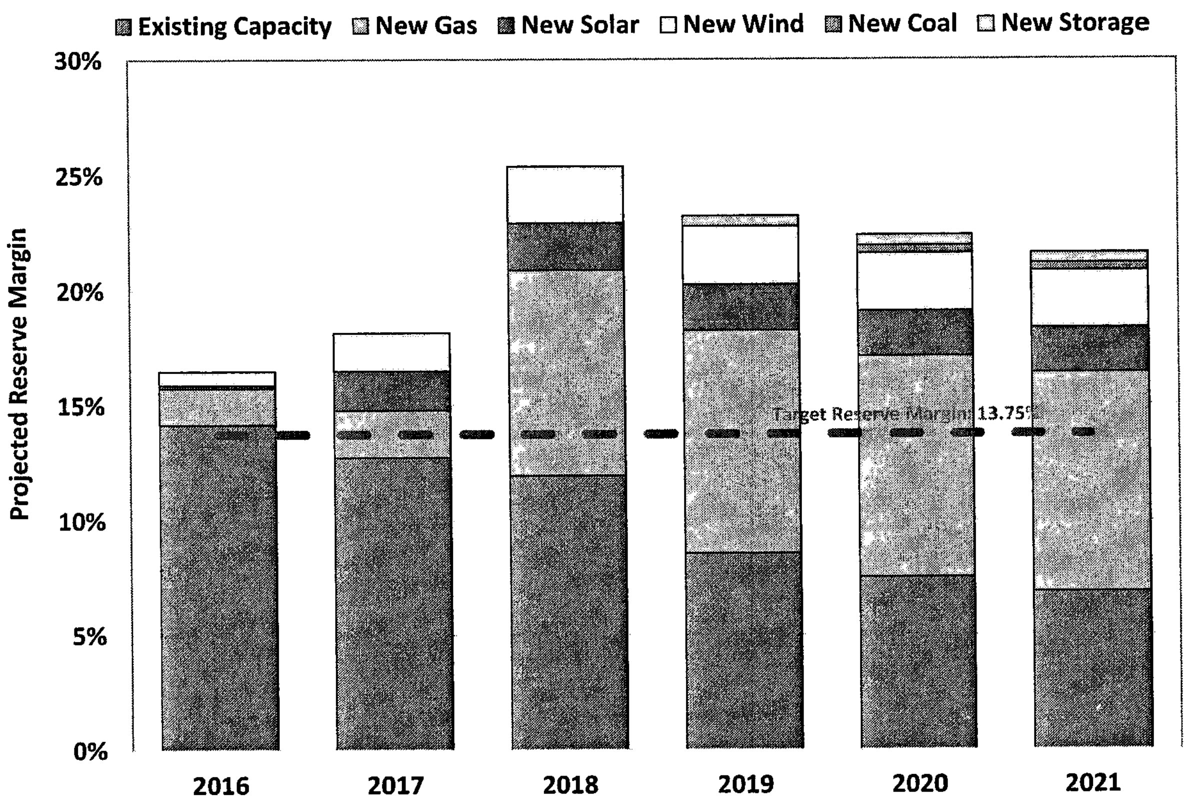
Nonetheless, whatever the outcome of this ongoing review, the fact that responsive and regulating reserves are forced to be maintained (held behind the High Ancillary Service Limit (HASL)) under the current market design will continue to be problematic, regardless of the ORDC parameters that are selected. 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 reserve margins.

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

**Figure 75: Projected Reserve Margins**




Source: ERCOT Capacity Demand Reserve Reports / 2016 from December 2015 and 2017-2021 from May 2016

Figure 75 above indicates that the region will have a 16.5 percent reserve margin heading into the summer of 2016. Reserve margins are now expected to exceed the target level of 13.75 percent for the next several years. These projections are higher than those developed last year, which were higher than in 2013. These increases are due to more new generation capacity expected to be constructed in ERCOT. The current outlook is very different than it was in 2013, when reserve margins were expected to be below the target level of 13.75 percent for the foreseeable future.

This current projection of reserve margins combined with relatively infrequent shortage pricing may raise doubts regarding the likelihood of all announced generation actually coming on line as currently 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 baseload resources casts doubt on the assumption embedded in the CDR that all existing generation will continue to operate.

As previously discussed, the decreased share of total generation provided by coal is evidence of the economic strain that coal units were under in 2015. The share of coal in the overall generation mix has historically dropped whenever natural gas prices have fallen below $3 per MMBtu.[29] This was true again in 2015. 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. In the face of these challenging fuel market economics, increased environmental requirements are also coming to the fore. Although currently under litigation, the interplay of multiple environmental regulations would require additional investment in pollution control technology at many existing coal units, thus further increasing the challenge to economic operations.

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.

ERCOT's installed reserve margin is compared that of other regions below. Figure 76 provides the anticipated reserve margins for the North American Electric Reliability Council (NERC) regions in the United States for the summer of 2016, as of NERC 2015 Long-Term Reliability Assessment.[30] Figure 76 shows that required, or reference level reserve margins center around 15 percent across other regions. These regions run the gamut from traditional bundled, regulated utility service territories to fully competitive, centrally operated wholesale markets. There are differences in the level of planning reserves expected for the summer of 2016. However, reserve margins are lower in nearly every region this year compared to last. ERCOT continues to stand out with its anticipated reserve margin remaining very close to its target level. Even with the

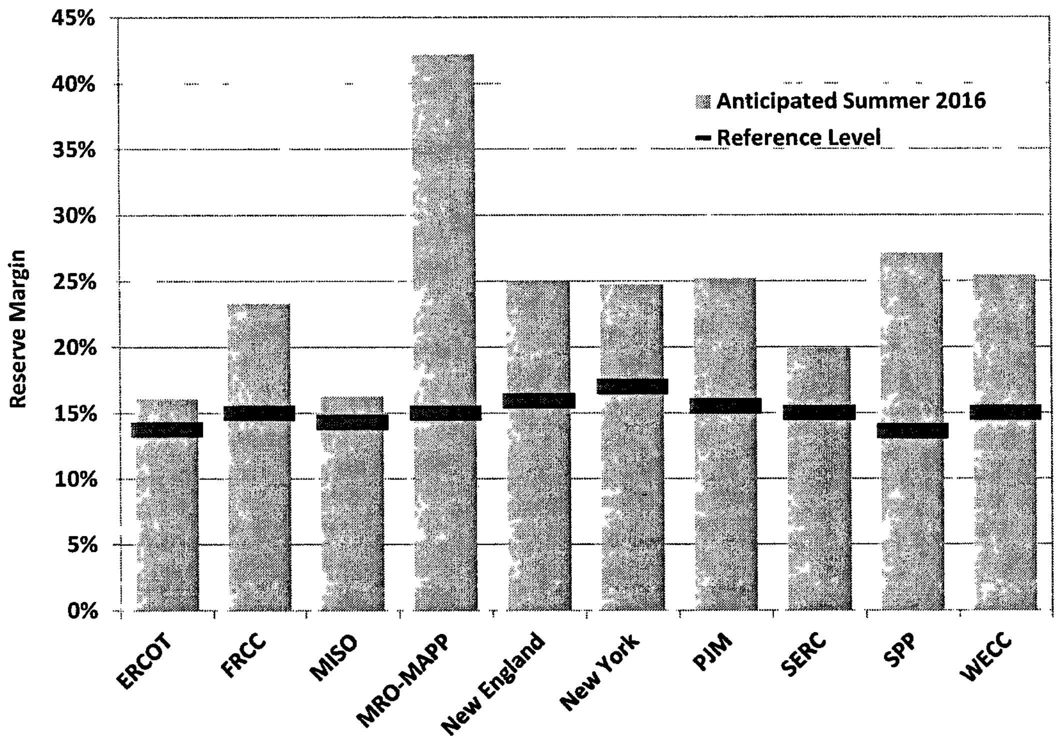
---

29 *See* 2012 State of the Market Report for the ERCOT Wholesale Electricity Markets, Potomac Economics, Ltd.

30 Data from NERC 2015 Long-Term Reliability Assessment (January 2016) available at http://www.nerc.com/pa/RAPA/ra/Reliability%20Assessments%20DL/2015LTRA%20-%20Final%20Report.pdf For the most recent projected reserve margins for ERCOT, please see Figure 75 and the associated discussion

forecasted additions, ERCOT is projected to sustain lower reserve margins than many other regions. This makes it important to ensure that the ERCOT market is designed to provide adequate economic signals to remain near this target, which is discussed below.

**Figure 76: Reserve Margins in Other Regions**




## 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 are continually increasing, 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. As discussed in the net revenue subsection, ancillary service payments are a small contributor: $5 - $10 per kW-year. Setting them 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 produce a planning reserve margin that is less than the planning reserve target.

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 reserve margins should be closely monitored to determine whether shortage pricing alone is leading to the desired level of planning reserves.

# VI. 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 2015. 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 2015.

## 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.[31] 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 as long as 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,

[31] 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.

it does identify conditions under which a supplier would have the ability to raise prices significantly by withholding resources.

Figure 77 shows the ramp-constrained RDI relative to load for all hours in 2015. 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.

**Figure 77: Residual Demand Index**

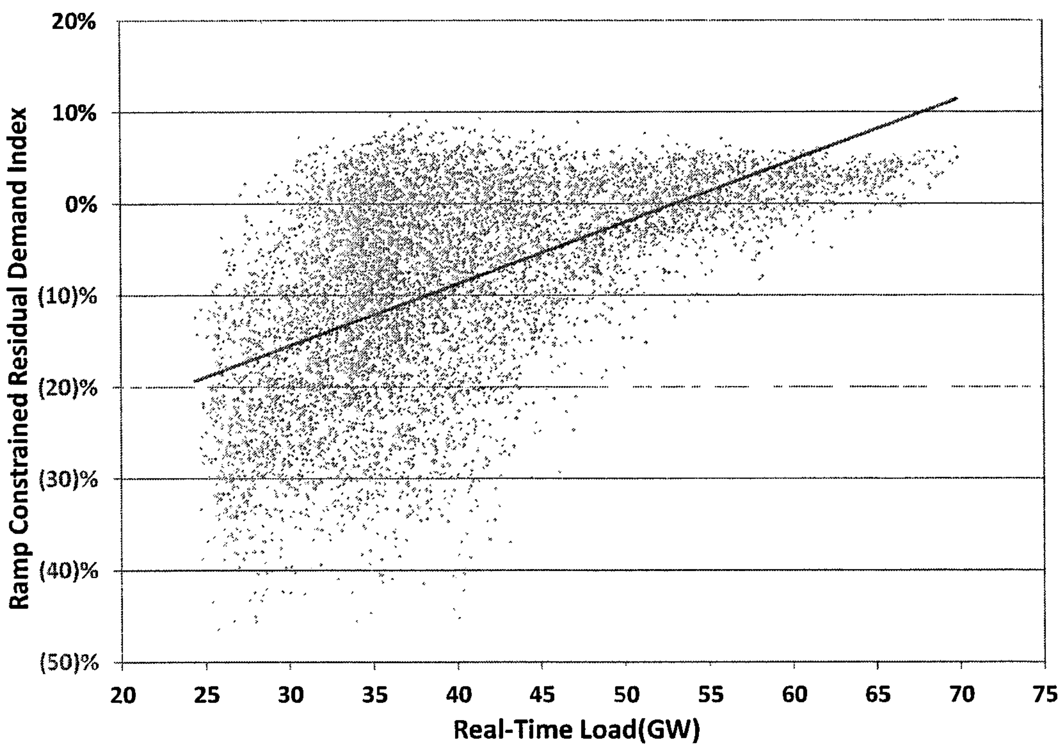
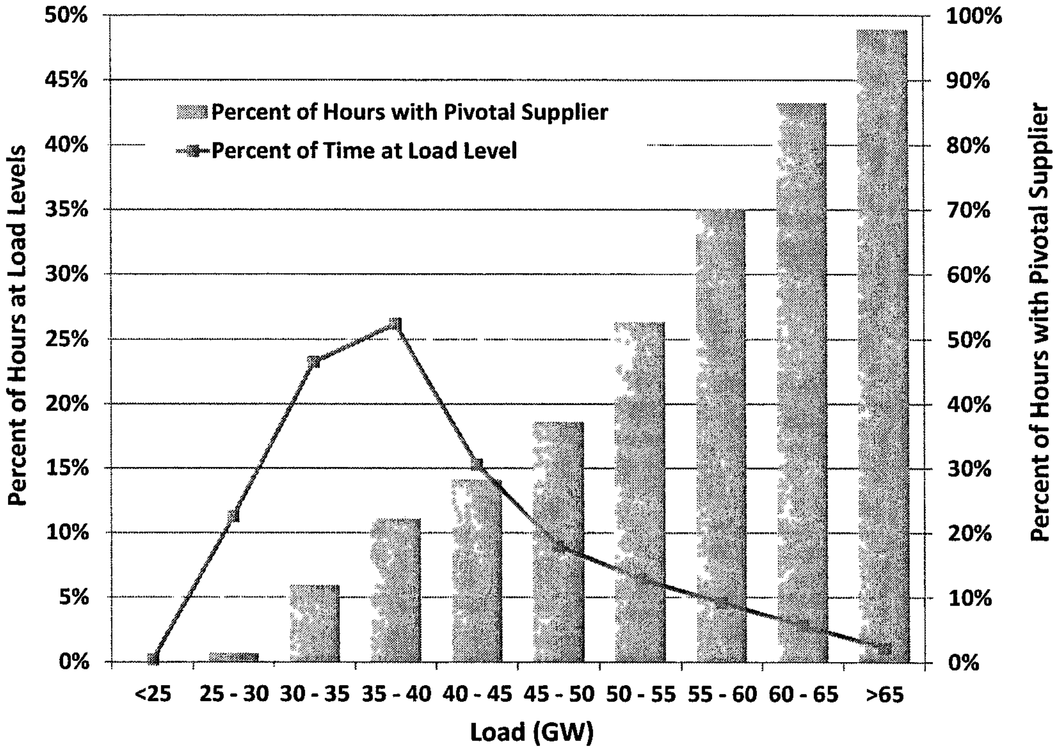



Figure 78 below summarizes the results of the RDI analysis by displaying the percent of time at each load level there was a pivotal supplier. The figure also displays the percentage of time each load level occurs.

**Figure 78: Pivotal Supplier Frequency by Load Level**




At loads greater than 65 GW there was a pivotal supplier 98 percent of the time. The occurrences of higher loads were more frequent in 2015 resulting in a pivotal supplier in approximately 26 percent of all hours of 2015, up from 23 percent of hours in 2014. 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.

### *Voluntary Mitigation Plans*

Voluntary Mitigation Plans (VMPs) existed for two market participants – NRG and Calpine – for the full year in 2015. In addition, the PUCT approved a VMP for Luminant in May 2015.[32] 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 in the ERCOT real-time energy market through economic withholding. The same restrictions are not required in forward energy markets (e.g., the ERCOT day-ahead market) because the price in forward energy markets is 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.

NRG's plan, initially approved in June 2012 and modified in May 2014,[33] 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.

---

32 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).

33 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).

Calpine's VMP was approved in March of 2013.[34] 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 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.

The most recently approved VMP for Luminant 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 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 Luminant's approximately 1000 MW of quick-start qualified combustion turbines 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/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

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

possessing market power that are unreasonably discriminatory or tend to unreasonably restrict, impair, or reduce the level of competition."[35] The exercise of market power may not rise to the level of an abuse of market power if it does not unreasonably impair competition, which would typically involve profitably raising prices significantly above the competitive level for a significant period of time. Thus, although the offer thresholds provided in the VMPs are designed to promote competitive market outcomes, the short termination provision provides additional assurance that any unintended consequences associated with the potential exercise of market power can be addressed in a timely manner rather than persisting and rising to the level of an abuse of market power.

The amount of offer flexibility afforded by the VMPs is small when compared to the offer flexibility that small participants, those with less than 5 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 2015. 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. Over the past five years, there were 13 hours with no surplus capacity, with all but one hour occurring in 2011. These correspond to times when ERCOT was unable to meet load and maintain all operating reserve obligations.

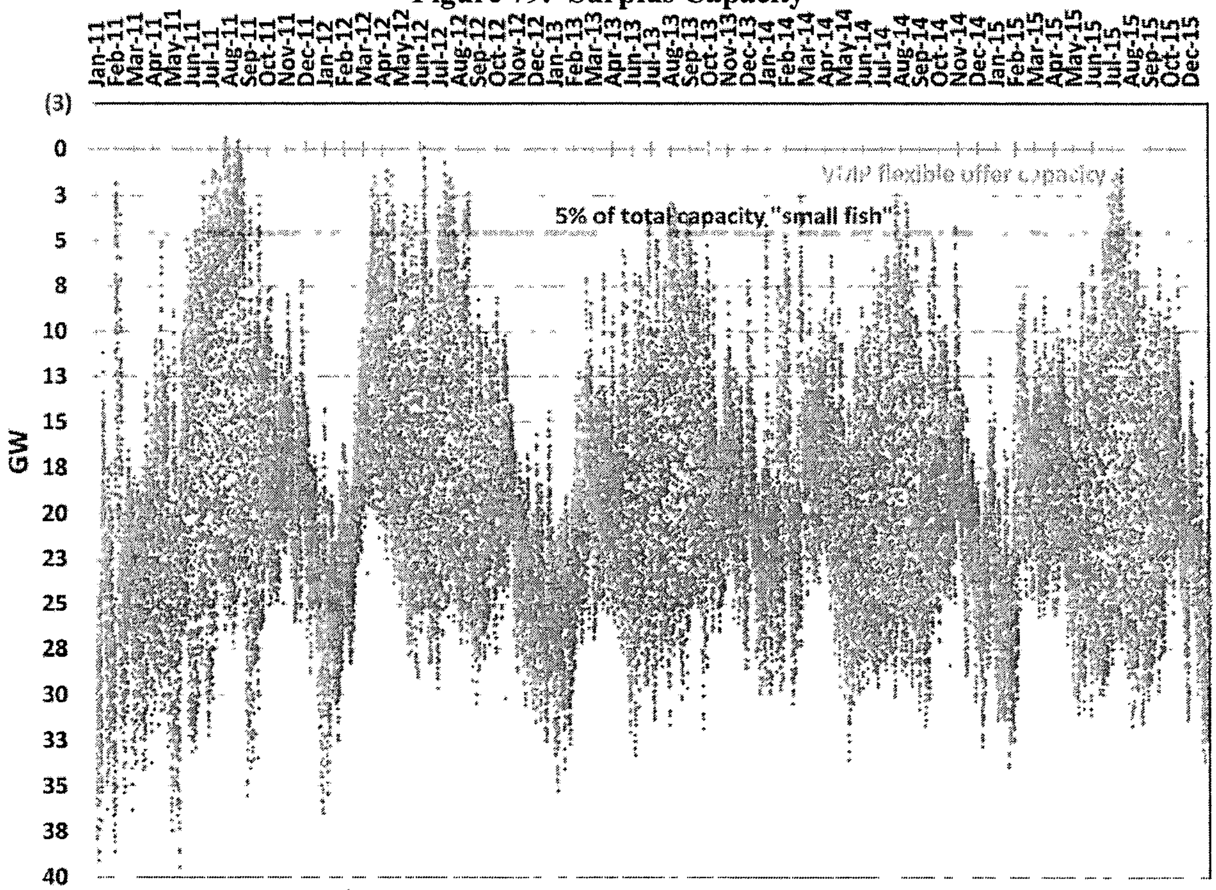
Currently, the 5 percent "small fish" threshold is roughly 4,000 MW, as indicated by the red line in Figure 79. There were 547 hours over the past five years with less than 4,000 MW of surplus capacity.[36] During these times a large "small fish" would have been pivotal and able to increase the market clearing price through its offer, potentially as high as the system-wide offer cap. In

35 PURA § 39.157(a).

36 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, and 56 hours in 2015.

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 five years.

**Figure 79: Surplus Capacity**




The effects of such actions became much more pronounced after June 21, 2013 when changes to real-time mitigation measures went into effect. These changes narrowed the scope of mitigation addressing the concern that mitigation measures were being applied much more broadly than intended or necessary in the ERCOT real-time energy market.[37] Although "small fish" market participants have always been allowed to offer all capacity at prices up to the system-wide offer cap, the effect on market outcomes of a large "small fish" offering substantial quantities at high prices became more noticeable after the scope of mitigation was narrowed.

37 Refer to Section I.F. Mitigation.



0000000145

## B. Evaluation of Supplier Conduct

The previous subsection presented a structural analysis that supports inferences about potential market power. In this subsection actual participant conduct is evaluated 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.

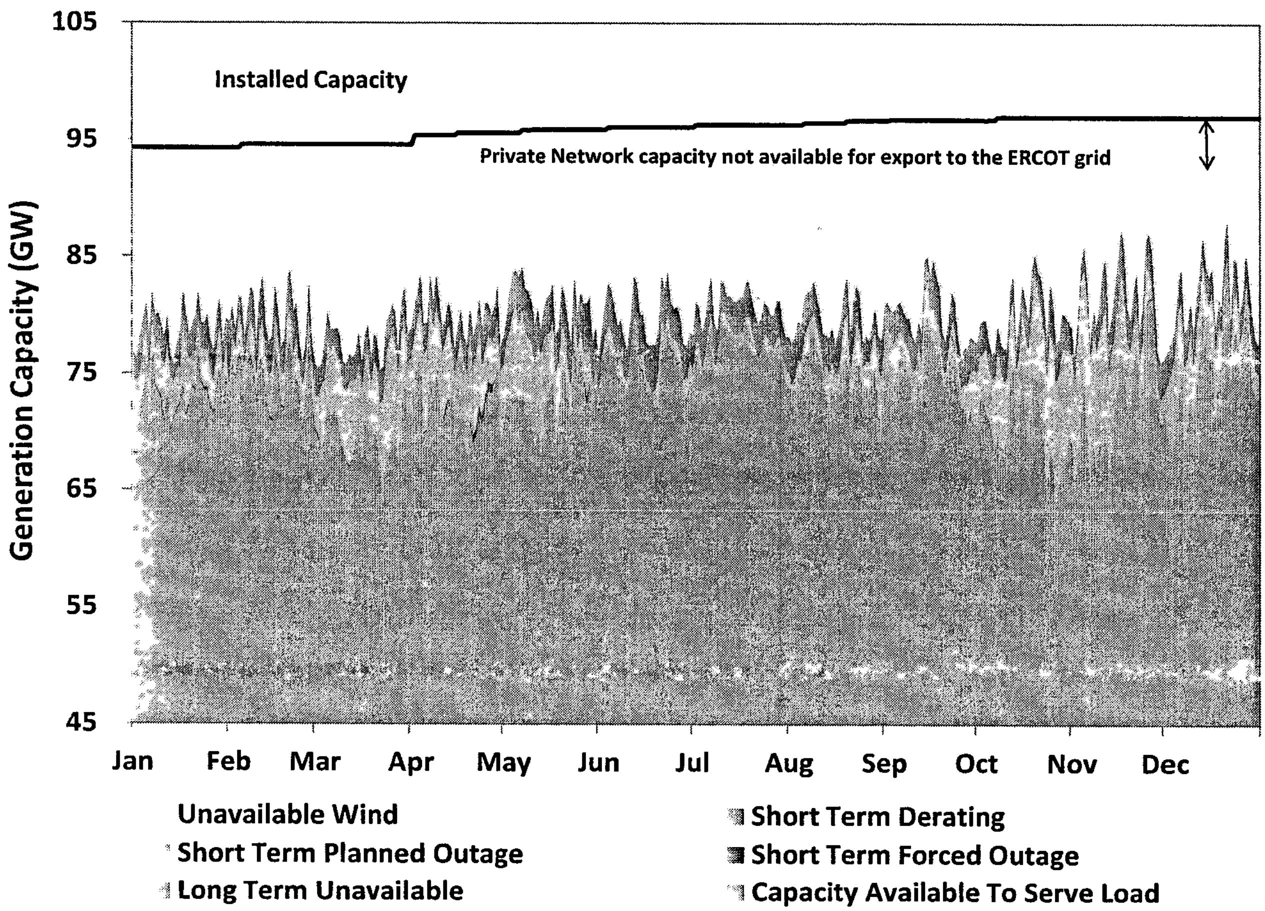
In a single-price auction like the real-time energy market, suppliers may attempt to exercise market power by withholding resources. The purpose of withholding is to cause more expensive resources to set higher market clearing prices, allowing the supplier to profit on its other sales in the real-time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

### 1. 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 80 shows a breakdown of total installed capacity for ERCOT on a daily basis during 2015. 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 80: Reductions in Installed Capacity**




Outages and deratings of non-wind generators fluctuated between 5 and 26 GW, as shown in Figure 80, while wind unavailability varied between 4 and 16 GW. Short-term planned outages were largest between mid-October and mid-November, and smallest during the summer months, which is consistent with expectations. Short-term forced outages also declined during the summer months. Short-term deratings peaked during October.

The quantity of long-term (greater than 30 days) unavailable capacity, peaked in March at 11.2 GW, reduced to less than 1.5 GW during the summer months, and increased to almost 9.2 GW in November. This pattern reflects the choice by some owners to mothball certain units on a seasonal basis, maintaining the units' operational status only during the high load summer season when more costly units have a higher likelihood of operating.

The next analysis focuses specifically on short-term planned and forced outages and deratings 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 81 shows the average magnitude of the outages and deratings lasting less than 30 days for the year and for each month during 2015.

**Figure 81: Short-Term Outages and Deratings**

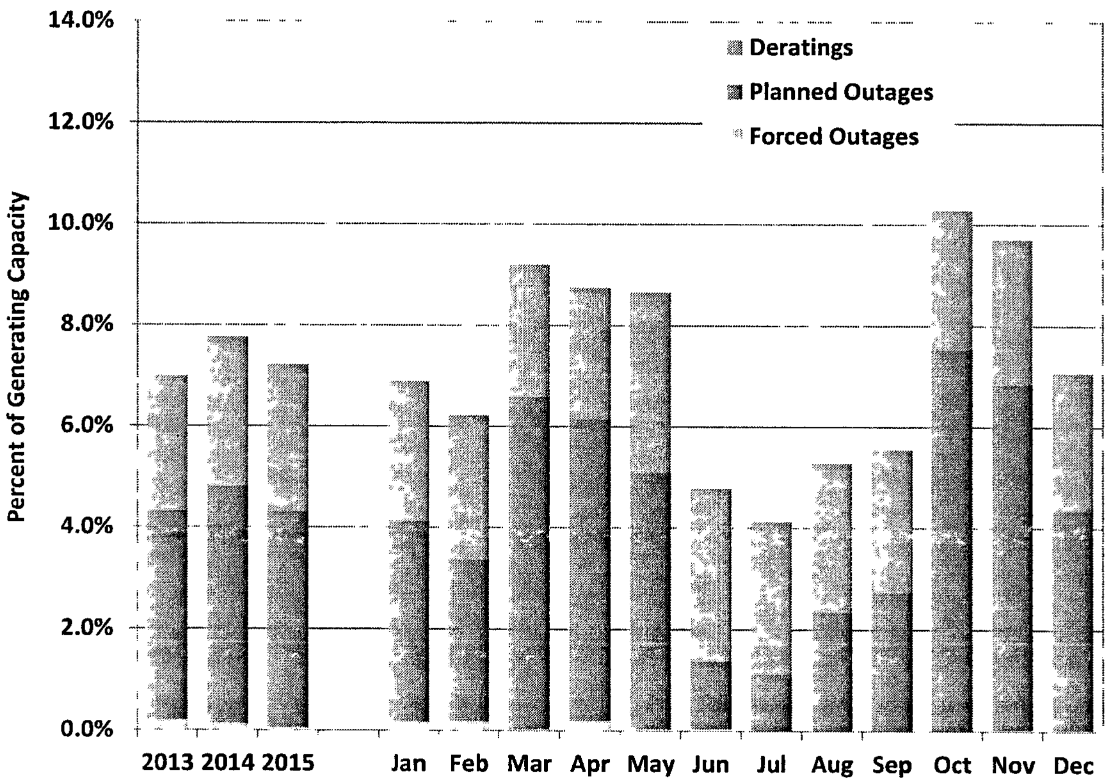



Figure 81 shows that total short-term deratings and outages were as large as 10.3 percent of installed capacity in October, and averaged less than 5 percent during the summer. Most of this fluctuation was due to anticipated planned outages. The amount of capacity unavailable during 2015 averaged 7.2 percent of installed capacity. This is a slight decline from 7.8 experienced in 2014 and an increase from the 7.0 percent experienced in 2013. 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.

### 2. 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 77 and Figure 78 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 82 shows the average relationship of short-term deratings and forced outages as a percentage of total installed capacity to real-time load level for large and small suppliers. 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 82: Outages and Deratings by Load Level and Participant Size June to August, 2015**

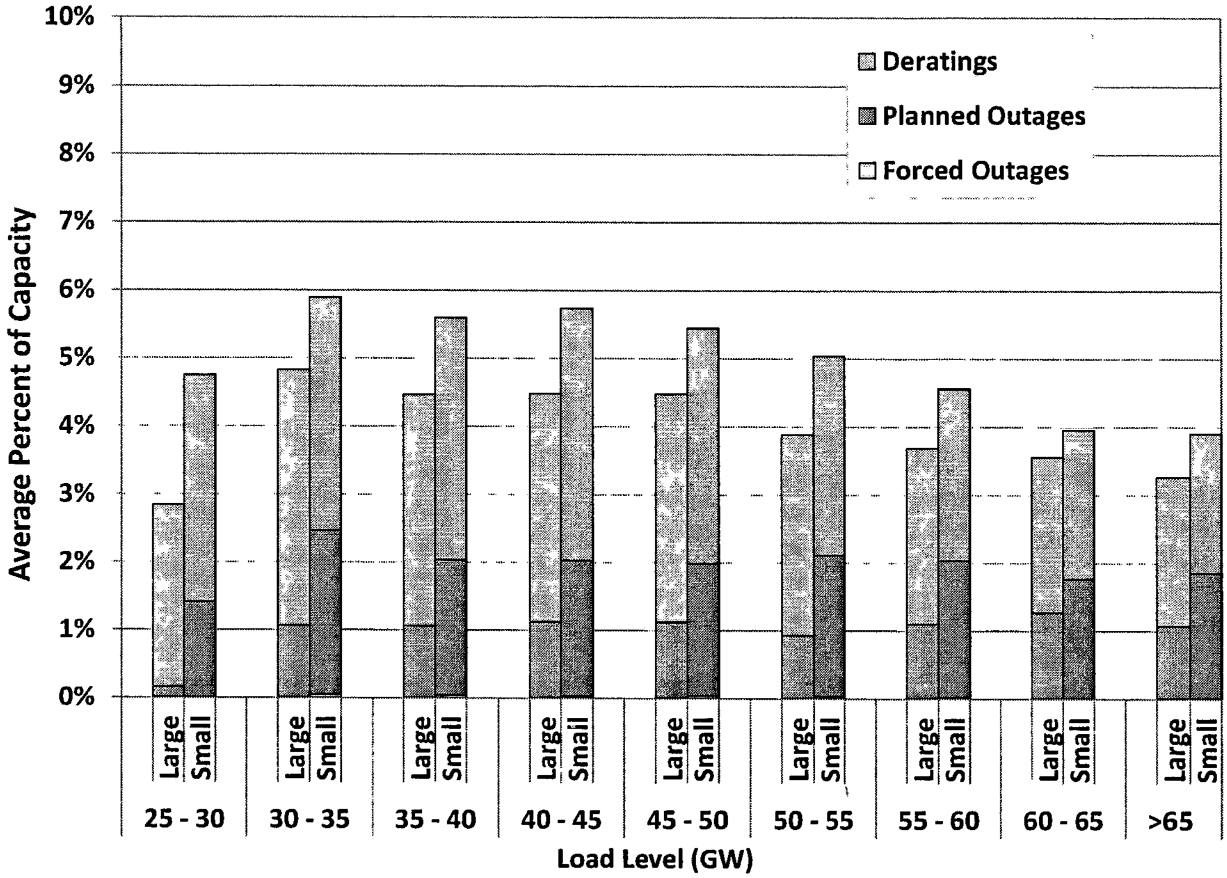



Figure 82 suggests that as demand for electricity increases, all market participants tend to make more capacity available to the market. For large suppliers, the combined short-term derating and forced outage rates decreased from 5 percent at low demand levels to approximately 3.5 percent at load levels above 65 GW. Rates for small participants also declined at higher loads. Outage rates for large participants were lower than those of small participants across all load levels. 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. Hence, these results do not raise potential competitive concerns.

### 3. 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 defined as the quantity of energy that is not being produced by in-service capacity even though the in-service capacity is