2012 State of the Market Report for the ERCOT Wholesale Electricity Markets

Potomac Economics, Ltd.

Independent Market Monitor for the
ERCOT Wholesale Market

June 2013

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

A. Introduction

This report reviews and evaluates the outcomes of the ERCOT wholesale electricity markets in 2012, and is submitted to the Public Utility Commission of Texas ("PUCT") and the Electric Reliability Council of Texas ("ERCOT") pursuant to the requirement in Section 39.1515(h) of the Public Utility Regulatory Act. It includes assessments of the incentives provided by the current market rules and procedures, and analyses of the conduct of market participants. This report also assesses the effectiveness of the Scarcity Pricing Mechanism ("SPM") pursuant to the provisions of PUCT Substantive Rule 25.505(g).

Key findings and statistics from 2012 include the following:

- The ERCOT wholesale market performed competitively in 2012.

- The ERCOT-wide load-weighted average real-time energy price was $28.33 per MWh in 2012, a 47 percent decrease from $53.23 per MWh in 2011. The decrease was primarily driven by more moderate weather and much lower natural gas prices in 2012.

  - The average price for natural gas was 31 percent lower in 2012 than in 2011, decreasing from $3.94 per MMBtu in 2011 to $2.71 per MMBtu in 2012.

  - After the extremes of 2011, loads in 2012 were more moderate with reduced occurrences of shortage conditions. Total ERCOT load in 2012 was 2.7 percent lower than 2011. Peak load decreased by 2.6 percent. Prices at the system-wide offer cap were experienced in dispatch intervals which totaled 1.5 hours in 2012.

- The total congestion revenue generated by the ERCOT real-time market in 2012 was $480 million, a decrease of 9 percent from 2011. This decrease is not as large as might be expected given the much larger decreases in average natural gas prices and real-time energy prices.

  - The Odessa North transformer constraint was the most frequently occurring transmission constraint in 2012. This, and other related localized constraints in west

Texas had significant financial impacts, causing the West zone average price to be higher than the ERCOT average for the first time.

- Even with the increased system-wide offer cap implemented in 2012, net revenues provided by the market were at historic lows as energy prices fell substantially.
  - Net revenues were insufficient to support investment in new generation even though planning reserve margins have fallen to levels that are close to the minimum planning reserve targets.
  - These results underscore the importance of the resource adequacy issues currently under consideration by the Commission, which we discuss in this report.

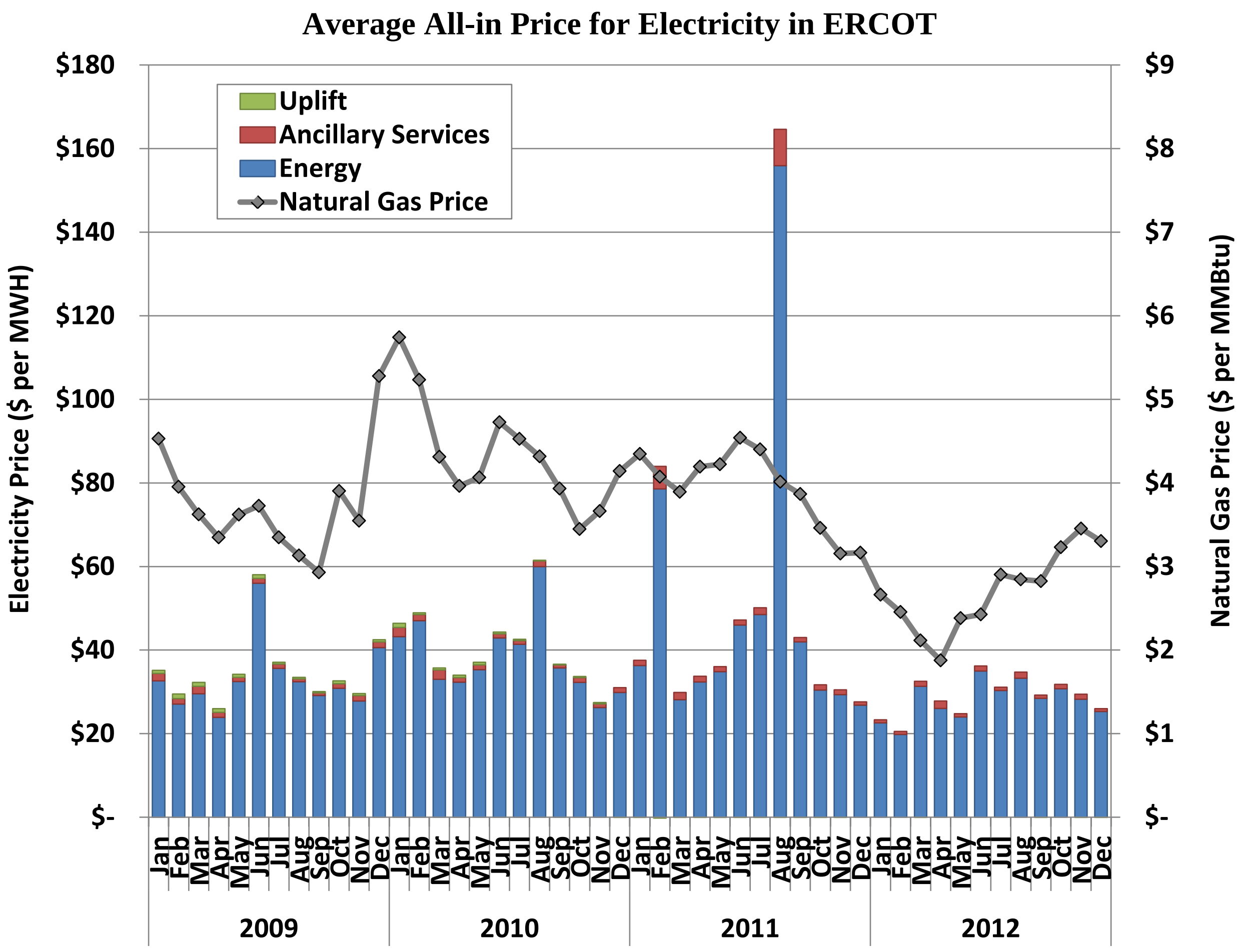
B. Review of Real-Time Market Outcomes

As is typical in other wholesale electricity markets, only a small share of the power produced in ERCOT is transacted in the spot market. However, prices in the real-time energy market are very important because they set the expectations for prices in the forward markets where most transactions take place. Unless there are barriers preventing arbitrage of the prices between the spot and forward markets, the prices in the forward market should be directly related to the prices in the spot market.

The average real-time energy prices by zone in 2009 through 2012 are shown below:

	Average Real-Time Electricity Price ($ per MWh)			
	2009	**2010**	**2011**	**2012**
ERCOT	**$34.03**	**$39.40**	**$53.23**	**$28.33**
Houston	$34.76	$39.98	$52.40	$27.04
North	$32.28	$40.72	$54.24	$27.57
South	$37.13	$40.56	$54.32	$27.86
West	$27.18	$33.76	$46.87	$38.24
Natural Gas ($/MMBtu)	$3.74	$4.34	$3.94	$2.71

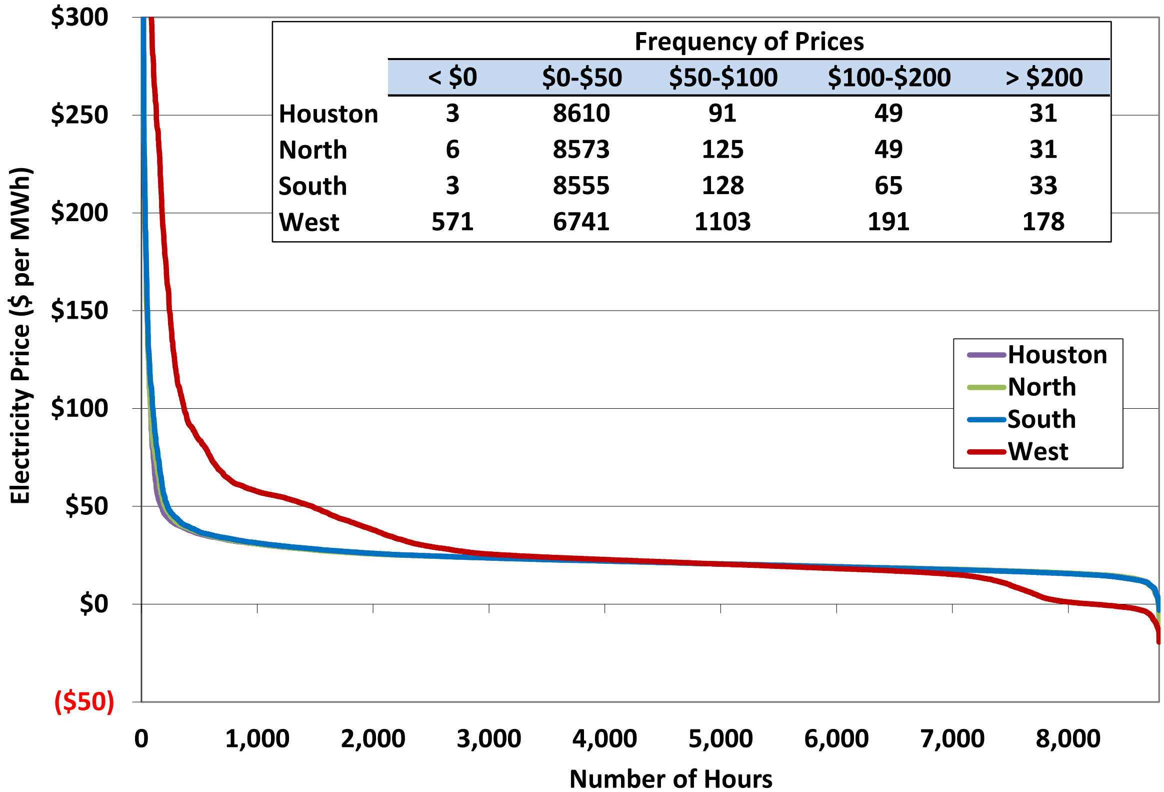
The largest component of the all-in cost of wholesale electricity is the energy cost, which is reflected by the locational marginal prices determined in the real-time energy market. ERCOT average real-time market prices were 47 percent lower in 2012 than in 2011. The ERCOT-wide load-weighted average price was $28.33 per MWh in 2012 compared to $53.23 per MWh in 2011.



The decrease in real-time energy prices was correlated with much lower fuel prices in 2012. The steady decline in natural gas prices from June 2011 to April 2012 resulted in the 2012 average natural gas price of $2.71 per MMBtu, a 31 percent decrease compared to $3.94 per MMBtu in 2011.

To depict how real-time energy prices vary by hour in each zone, the next figure shows the hourly average price duration curve in 2012 for four ERCOT load zones. The Houston, North and South load zones had similar prices over the majority of hours. The price duration curve for the West zone is noticeably different than the other zones, with more hours with prices greater than $50 per MWh and more than 500 hours (6 percent of the time) when the average hourly price was less than zero.

Zonal Price Duration Curves



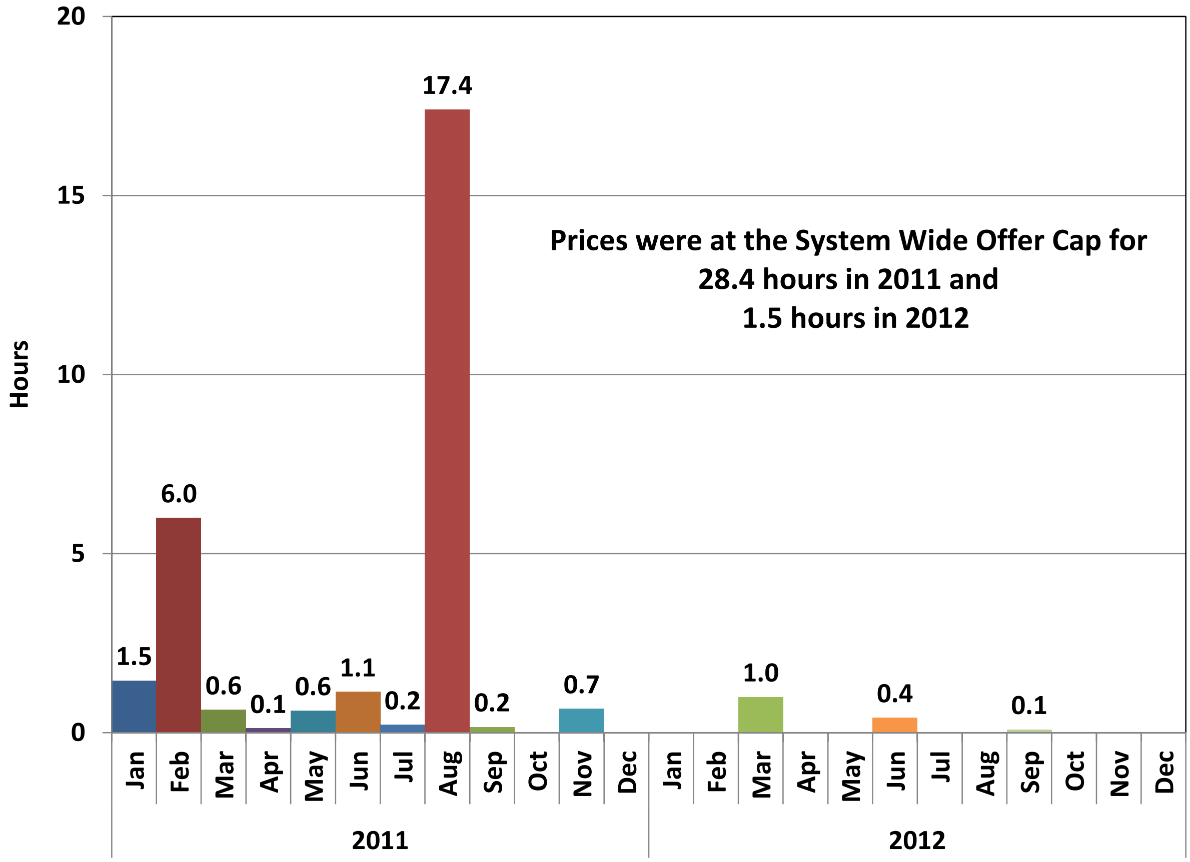
	Frequency of Prices				
	< $0	$0-$50	$50-$100	$100-$200	> $200
Houston	3	8610	91	49	31
North	6	8573	125	49	31
South	3	8555	128	65	33
West	571	6741	1103	191	178

As observed over the past few years, West zone prices are lower than the rest of ERCOT when high wind output in the west results in congested transmission interfaces from the West zone to the other zones in ERCOT. Recently, prices higher than the rest of ERCOT have occurred in the West zone due to local transmission constraints that typically occur under low wind and high load, or outage conditions. Specifics about these transmission constraints are provided in Section III, Transmission and Congestion.

After the extremes of 2011, weather conditions in Texas returned to closer to normal in 2012. As more fully discussed in Section IV, Load and Generation, overall demand for electricity was lower in 2012 than in 2011, resulting in much fewer occasions when the available supply generation capacity was unable to meet customer demands. This resulted in a decreased likelihood that the available generation capacity was not sufficient to meet customer demands for electricity and maintain the required reliability reserves.

As more fully described later in Section V, Resource Adequacy, the nodal market causes energy prices to rise toward the system-wide offer cap as available operating reserves approach minimum required levels to reflect the degradation in system reliability.

Duration of Prices at the System-Wide Offer Cap



Presented in the figure above is the aggregated amount of time represented by all five-minute dispatch intervals where the real-time energy price was at the system-wide offer cap, displayed by month. Prices during 2012 were at the system-wide offer cap for only 1.5 hours, a significant reduction from the 28.4 hours experienced in 2011. Approved during 2012, PUCT SUBST. R.25.508 increased the system-wide offer cap to $4,500 per MWh effective August 1, 2012. As shown in the figure above, there was only a brief period when energy prices rose to the cap after this change was implemented.

Finally, after the multiple protocol revisions implemented in 2012, the non-spinning reserve deployment process remains sub-optimal from a reliability and efficiency perspective. As more fully described in Section I.H, Recommendations, we continue to recommend that ERCOT develop a mechanism that will rationally commit generation and load resources that can start or curtail within 30 minutes. This deficiency in ERCOT's nodal market design should be addressed by implementing a "look ahead" dispatch functionality for the real-time market to produce an energy and ancillary services commitment and dispatch results that are co-optimized and recognize anticipated changes in system demands. This additional functionality represents a major change to ERCOT systems; one we recommend together with improved pricing provisions that will allow offers from load resources to set prices if they are required to meet system demand.

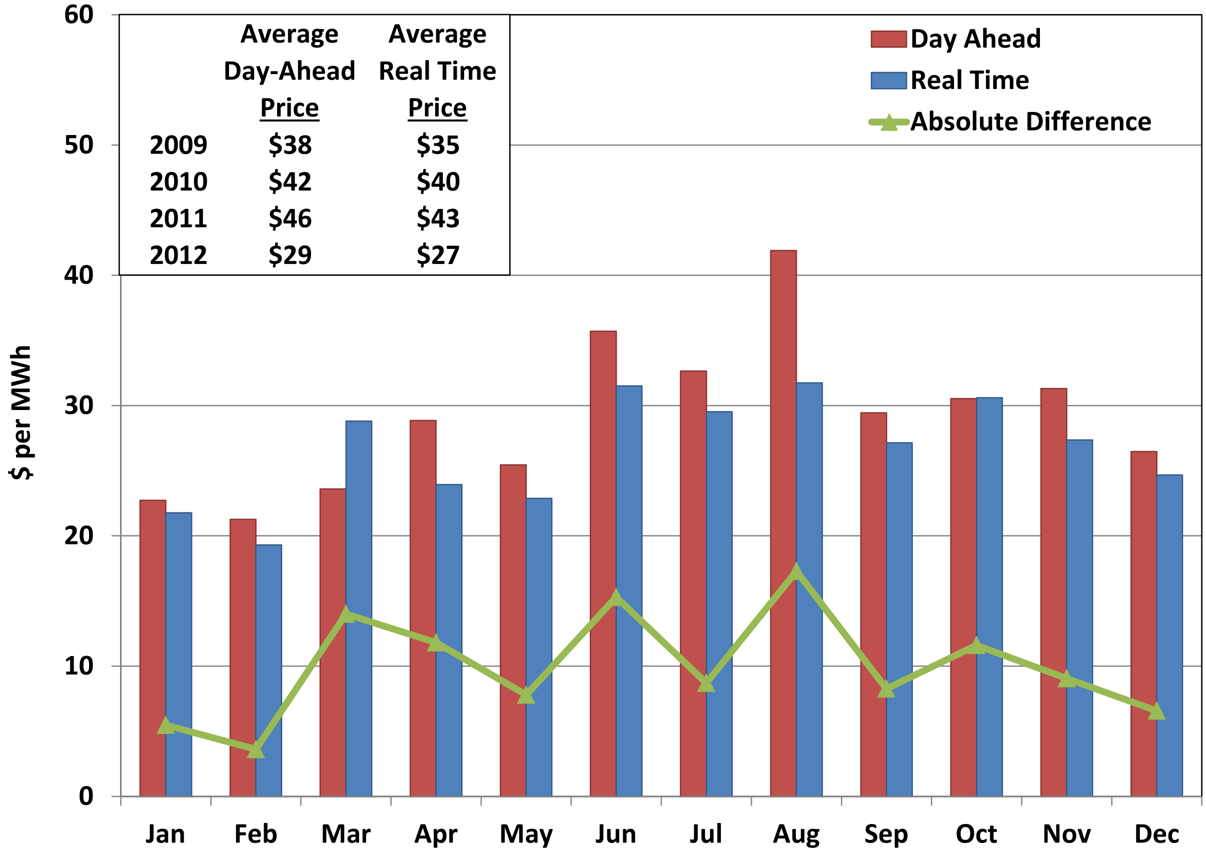
C. Review of Day-Ahead Market Outcomes

ERCOT's centralized day-ahead market allows participants to make financially binding forward purchases and sales of power for delivery in real-time. Although all bids and offers are evaluated in the context of their ability to reliably flow on the transmission network, there are no operational obligations resulting from the day-ahead market clearing. These transactions are made for a variety of reasons, including satisfying the participant's own supply, managing risk by hedging the participant's exposure to the real-time market, or arbitraging with the real-time markets. For example, load serving entities can insure against volatility in the real-time market by purchasing in the day-ahead market. Finally, the day-ahead market plays a critical role in coordinating generator commitments. For all of these reasons, the performance of the day-ahead market is essential.

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

The figure below shows the price convergence between the day-ahead and real-time market, summarized by month. The simple average of day-ahead prices in 2012 was $29 per MWh, compared to the simple average of $27 per MWh for real-time prices. The average absolute difference between day-ahead and real-time prices was $9.96 per MWh in 2012; much lower than in 2011 when average of the absolute difference was $24.50 per MWh. This reduction was due to fewer occurrences of shortage intervals and associated high prices in 2012.

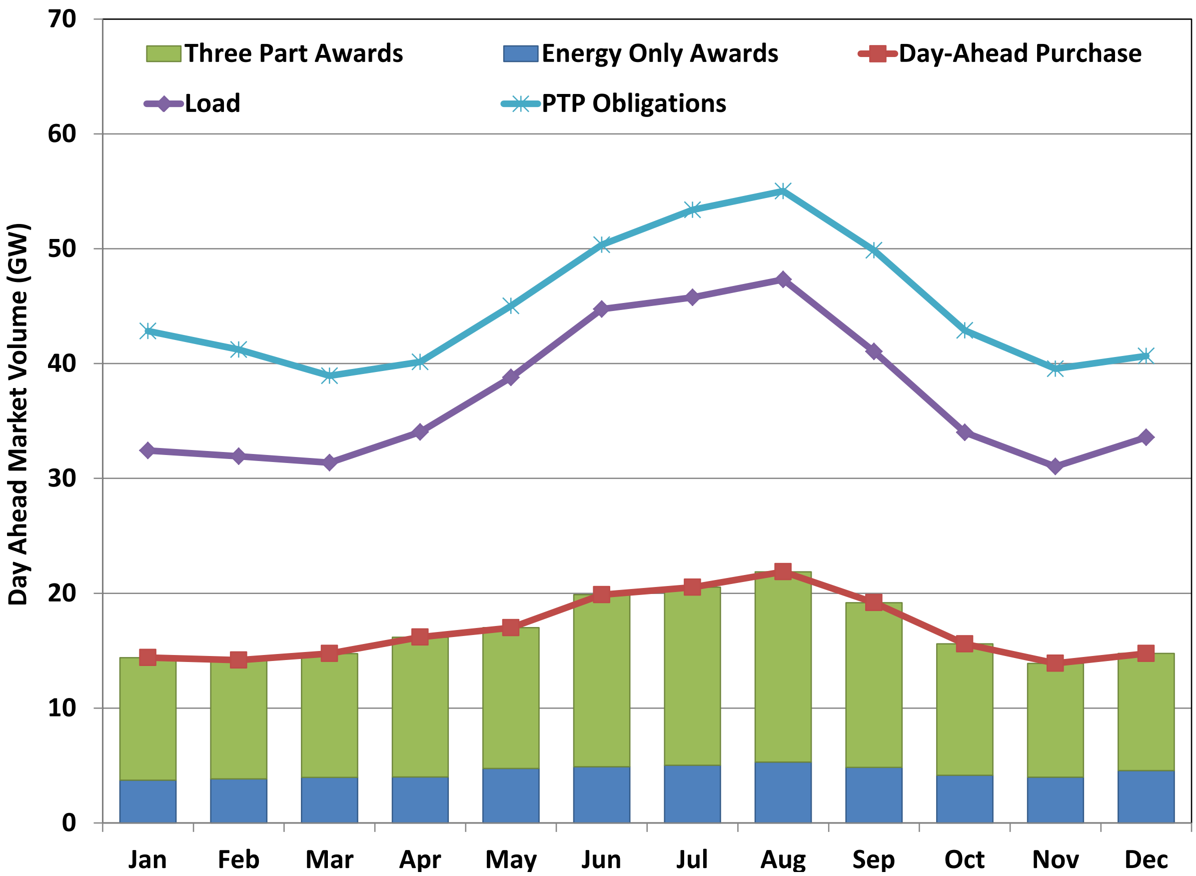
Convergence between Forward and Real-Time Energy Prices



This day-ahead premium is consistent with expectations due to the much higher volatility of real-time prices. Risk is lower for loads purchasing in the day-ahead and higher for generators selling day-ahead. The higher risk for generators is associated with the potential of incurring a forced outage and as a result, having to buy back energy at real-time prices. This explains why the highest premiums tend to occur during the months with the highest demand and highest prices. Overall, the day-ahead premiums were very similar to the differences observed in 2010 and 2011, but remain higher than observed in other organized electricity markets. Real-time energy prices in ERCOT are allowed to rise to levels that are much higher than what is allowed in other organized electricity markets, which increases risk and helps to explain the higher day-ahead premiums regularly observed in ERCOT. Although most months experienced a day-ahead premium (*e.g.*, $10 per MWh in August), it should not be expected over time that every month will always produce a day-ahead premium as the real-time risks that lead to the premiums will materialize unexpectedly on occasion, resulting in real-time prices that exceed day-ahead prices (*e.g.*, in March).

Summarized in the figure below is the volume of day-ahead market activity by month. It shows that day-ahead purchases are approximately 45 percent of real-time load.

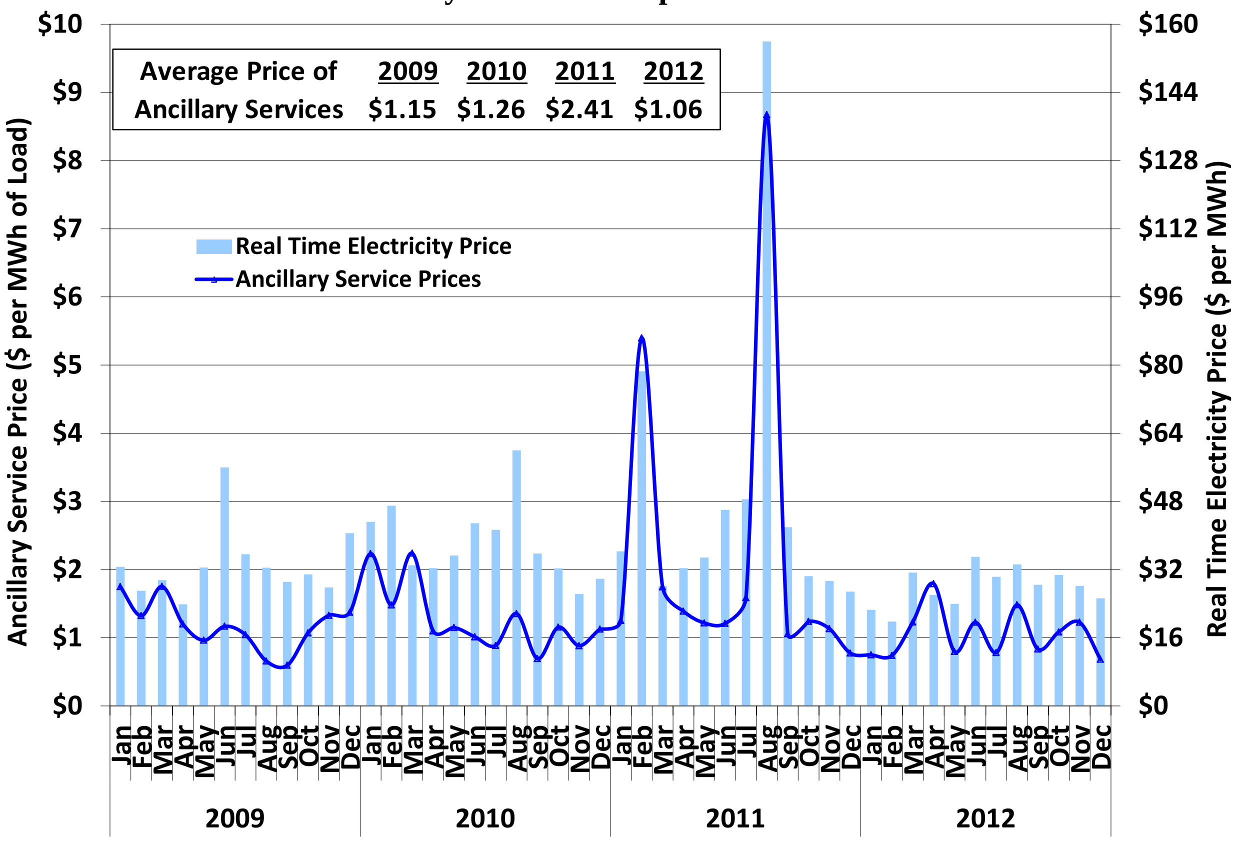
Volume of Day-Ahead Market Activity by Month



Point to Point Obligations are financial instruments purchased in the day-ahead market. Although they do not provide any energy supply themselves, they do provide the ability to avoid the congestion costs associated with transferring the delivery of energy from one location to another. To provide a volume comparison we aggregate all of these "transfers", netting location specific injections against withdrawals. By adding the aggregated transfer capacity associated with purchases of PTP Obligations, we find that on average total volumes transacted in the day-ahead market are greater than real-time load.

Ancillary Service capacity is procured as part of the day-ahead market clearing. The figure below shows the monthly total ancillary service costs per MWh of ERCOT load and the average real-time energy price for 2009 through 2012. Total ancillary service costs are generally correlated with real-time energy price movements, which in turn are highly correlated with natural gas price movements. The average ancillary service cost per MWh of load decreased to $1.06 per MWh in 2012 compared to $2.41 per MWh in 2011, a decrease of 56 percent. Total ancillary service costs decreased from 4.5 percent of the load-weighted average energy price in 2011 to 3.7 percent in 2012.

Ancillary Service Costs per MWh of Load



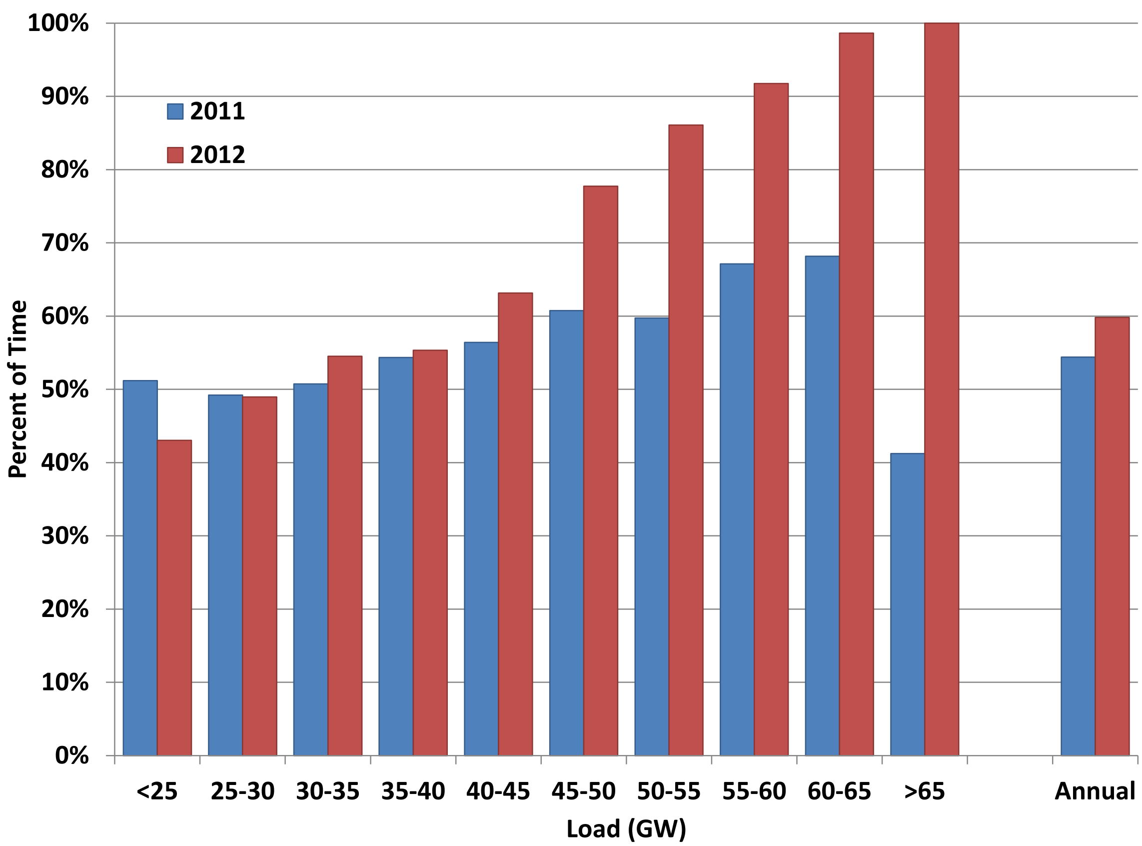
D. Transmission and Congestion

There was a marked change in the nature of real-time transmission congestion during 2012 when compared to previous years. For the past several years a significant portion of real-time transmission congestion could be described as limiting the export of wind generation ***from*** the West zone to the load centers across the rest of ERCOT. Transmission congestion in 2012 was more significantly the result of limitations on the ability to get generation ***to*** loads in the West zone. Some portion of the limitation can be attributed to transmission outages taken to enable the construction of new CREZ transmission lines. Another factor is that loads in far west Texas have increased more than the system-wide average.

The total congestion revenue generated by the ERCOT real-time market in 2012 was $480 million, a decrease of 9 percent from 2011. This decrease is not as large as might be expected given the much larger decreases in average natural gas prices real-time energy prices. Two factors influencing the overall costs of congestion in 2012 were the significant financial impact of several localized transmission constraints in far west Texas and the higher frequency of active transmission constraints.

Shown below is a comparison of the amount of time transmission constraints were active at various load levels in 2012 and 2011.

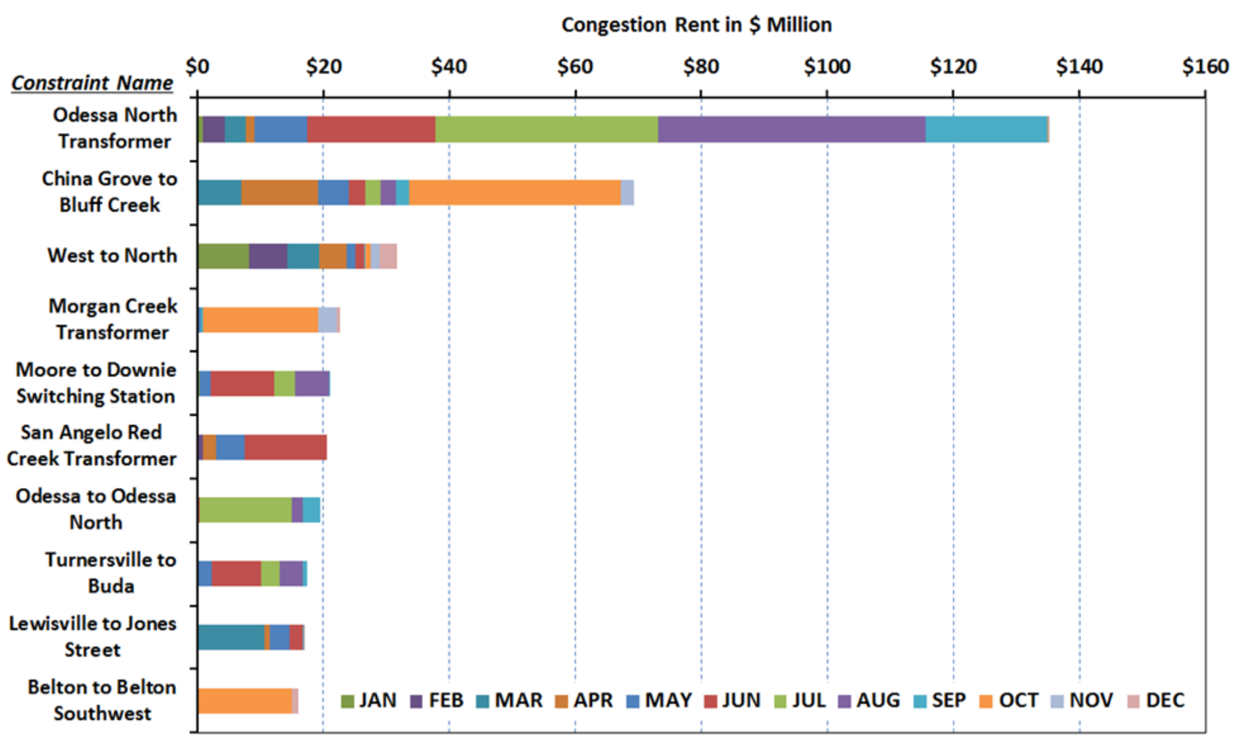
Frequency of Active Constraints



We observe that in 2012 the likelihood of having an active transmission constraint was higher than it was in 2011 and that for loads above 45 GW the frequency was much higher. During 2011, we observed that at higher system loads ERCOT operators did not always activate (or sometimes de-activated) transmission constraints. This was due to a concern that by having a constraint active during periods of high demand the total capacity available to serve load may be limited. However, ERCOT's dispatch software contains parameters that allow it to automatically make the correct decision about when to violate transmission constraints when necessary to serve total system load. Therefore, ERCOT modified their practice in 2012 to retain active transmission constraints even during periods of high demand.

The figure below displays the ten most costly real-time constraints and indicates that the Odessa North 138/69 kV transformer constraint was by far the most highly valued during 2012. This constraint became more pronounced from 2011 to 2012 and is mainly attributed to load growth in far west Texas driven by increased oil and natural gas activity.

Top Ten Real-Time Constraints

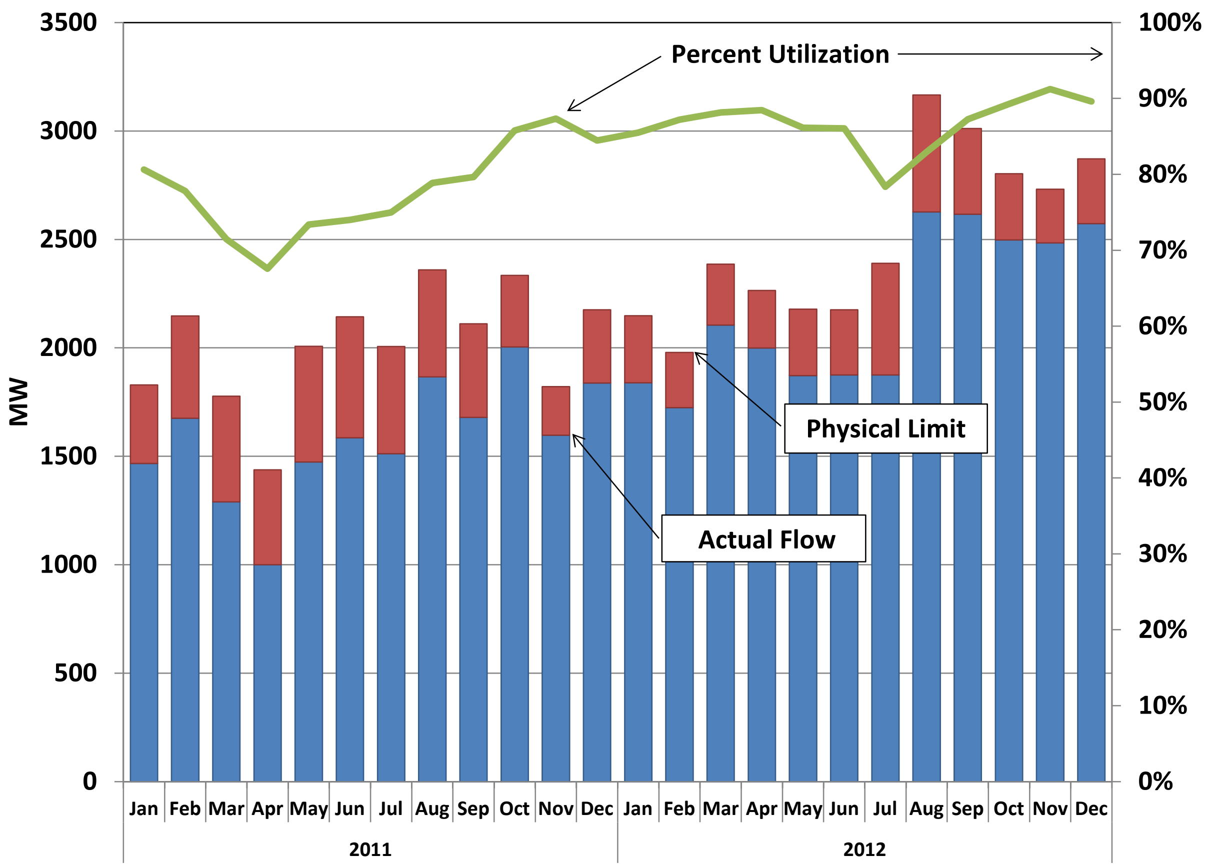


The Odessa North 138/69 kV transformer typically overloads during low wind conditions. The characteristics that load the Odessa North 138/69 kV transformer are the same conditions that also affect Odessa to Odessa North 138 kV line, which is shown as the seventh constraint on the list. Not only did this constraint have nearly twice the financial impact of the second constraint on the list, its impact was more than 40 percent greater than the top constraint from 2011. Its magnitude is even more significant given the overall lower costs of energy in 2012.

Not surprisingly, much public attention was focused on this constraint; much of it questioning its causes and the potential for short-term remedies. ERCOT and the local transmission provider were able to identify two transmission lines, which when opened greatly reduced congestion around the Odessa North station without causing other reliability concerns. After the lines were opened in mid-September, congestion around Odessa North was almost eliminated for the rest of the year.

The figure below presents a summary of the utilization of the West to North interface. The West to North constraint continued to be active at some point during every month of 2012 but with far less impact than in 2011.

Utilization of the West to North Interface Constraint



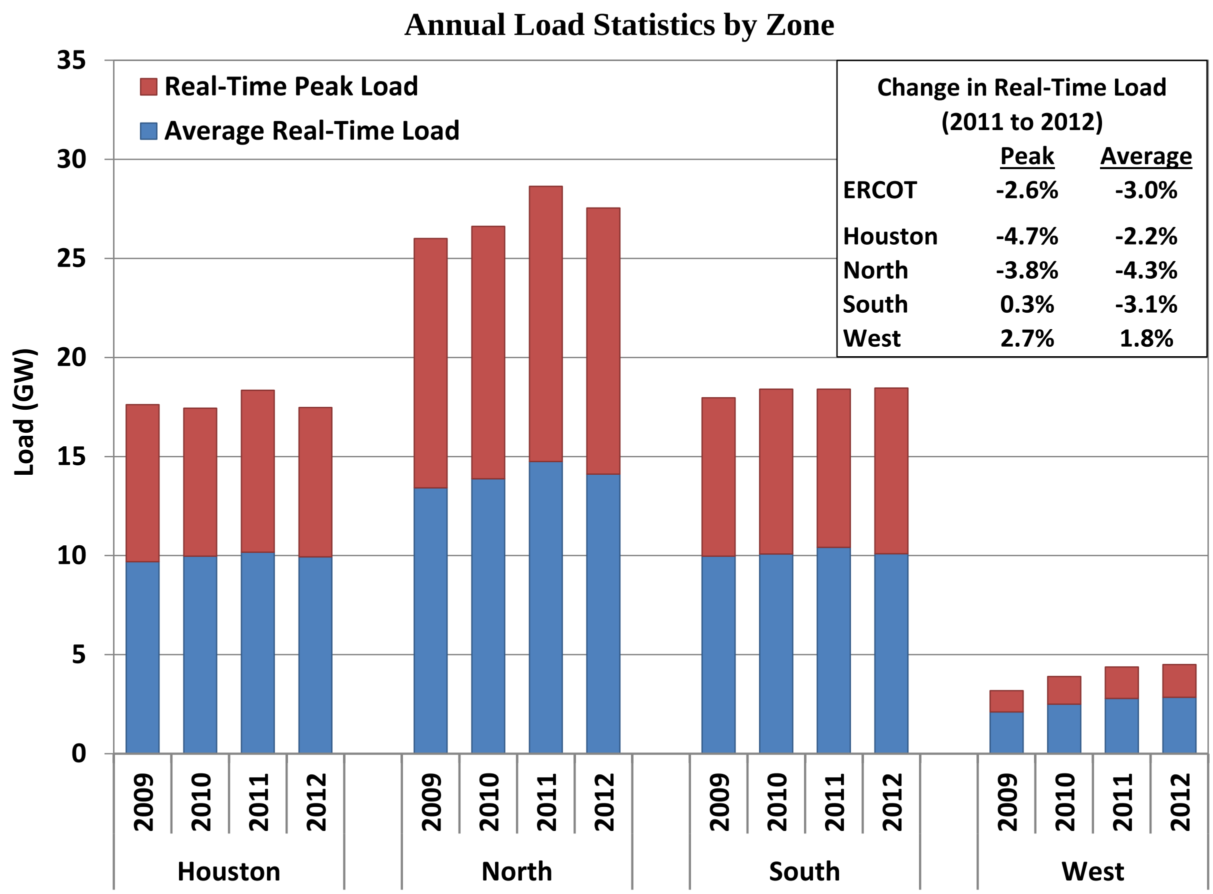
Through the years this constraint has been a major impediment to delivering all the wind generation located and produced in the western reaches of ERCOT to the load centers. This constraint had the highest financial impact of all real-time transmission constraints during 2011, but in 2012 its impact dropped to one third that level. The reduction was a result of the combined impact of higher loads in the west and increased transfer capability due to the first CREZ transmission lines being placed in service.

Through July 2012, the average physical limit was approximately 2,200 MW and the average actual flow during constrained intervals was approximately 1,900 MW. After July 2012, the physical limit increased to an average of 2,900 MW and the actual flow increased to approximately 2,500 MW. In March 2012, a new real-time analysis tool was implemented to better track the dynamic nature of the transient stability limit of the West to North interface. However, there was not a noticeable increase to the transfer limit corresponding to its implementation due to the effects of transmission outages occurring to accommodate maintenance activities and the installation of CREZ lines. Many of these outages were complete by July 2012 which accounts for the increase in the West to North limit.

The average annual utilization of the West to North constraint was 87 percent in 2012, which compares favorably to 78 percent utilization experienced in 2011. Over the long term, the physical limit will continue to increase as CREZ transmission projects are completed.

E. Load and Generation

The figure below shows peak load and average load in each of the ERCOT zones from 2009 to 2012. In each zone, as in most electrical systems, peak demand significantly exceeds average demand. The North zone is the largest zone (with about 38 percent of the total ERCOT load); the South and Houston zones are comparable (27 percent); while the West zone is the smallest (8 percent of the total ERCOT load). The figure also shows the annual non-coincident peak load for each zone. This is the highest load that occurred in a particular zone for one hour during the year; however, the peak can occur in different hours for different zones. As a result, the sum of the non-coincident peaks for the zones was greater than the annual ERCOT peak load.

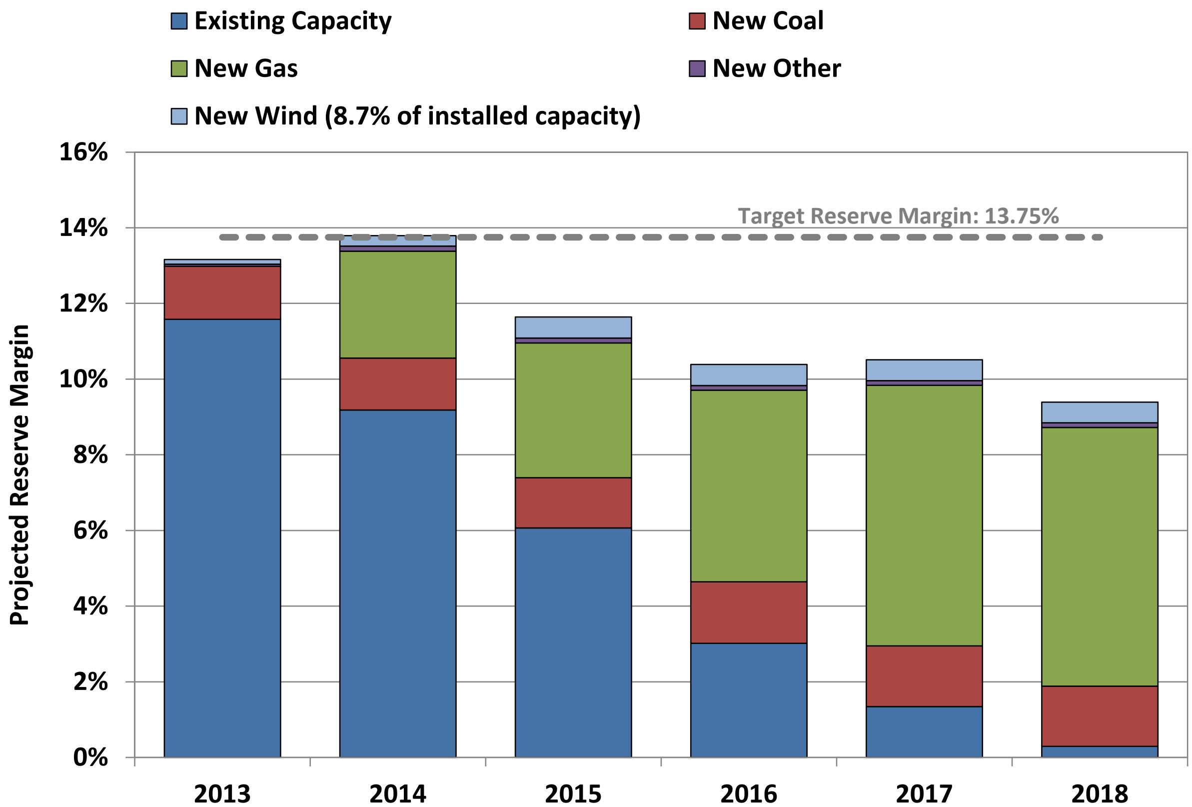


Total ERCOT load decreased from 334 TWh in 2011 to 325 TWh in 2012, a decrease of 2.7 percent or an average of 1,130 MW every hour. Similarly, the ERCOT coincident peak hourly demand decreased from 68,311 MW in 2011 to 66,559 MW, a decrease of 1,752 MW, or 2.6 percent. The results at the zonal level are not consistent. Average load decreased in three of the four zones, but grew by 1.8 percent in the West zone.

New generation resources in 2012 totaled approximately 1 GW; most of which were wind units and the remainder were solar and biomass. Comparing the current mix of installed generation capacity to that in 2007, we find that over these six years wind and coal generation are the two categories with the most increased capacity. The sizable additions in these two categories have been more than offset by retirements of natural gas-fired steam units, resulting in less installed capacity in 2012 than there was in 2007.

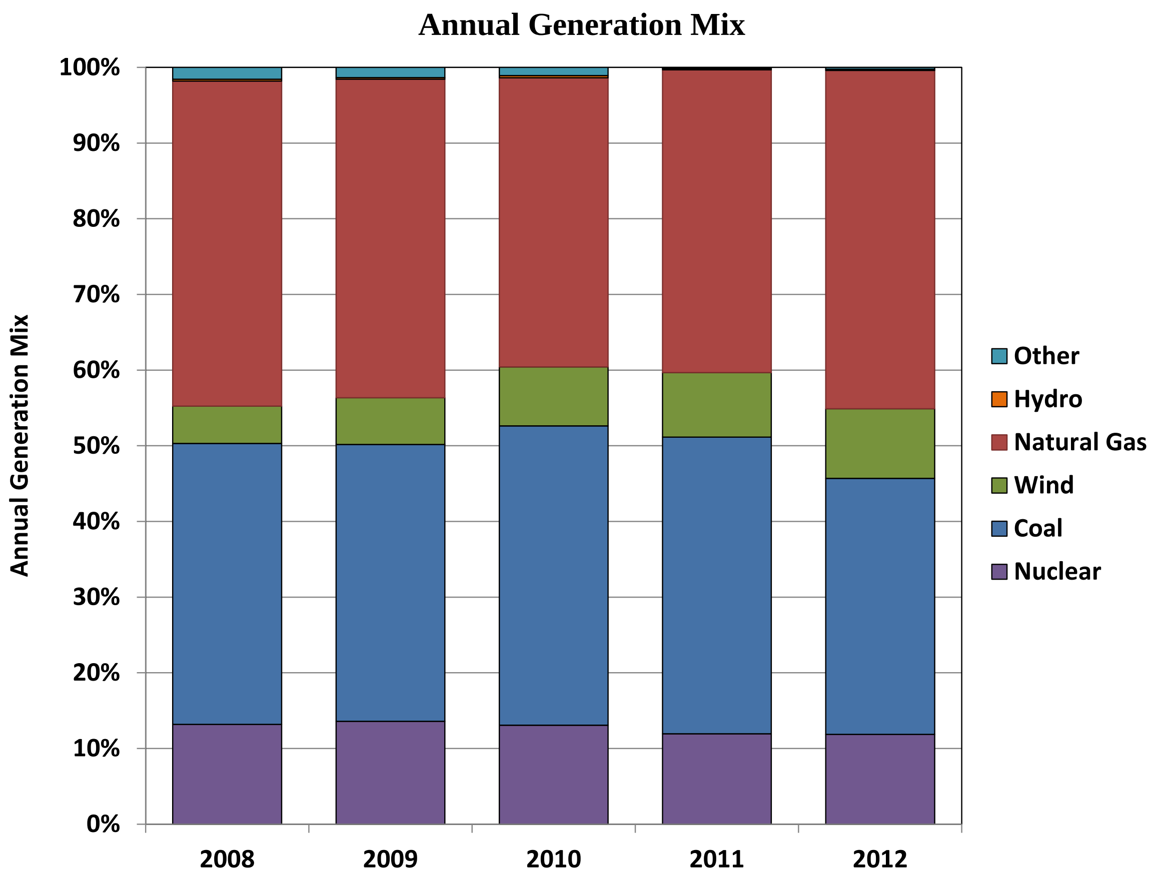
Shown below is ERCOT's most current projection of reserve margins. It indicates that the region will have a 13.2 percent reserve margin heading into the summer of 2013. With the addition of recently announced generation additions, in 2014 the reserve margin is expected to reach 13.8 percent -- just barely above the current target. The bulk of the new capacity being added is natural gas-fired generation, approximately a quarter of which are expansions at existing facilities.

Projected Reserve Margins



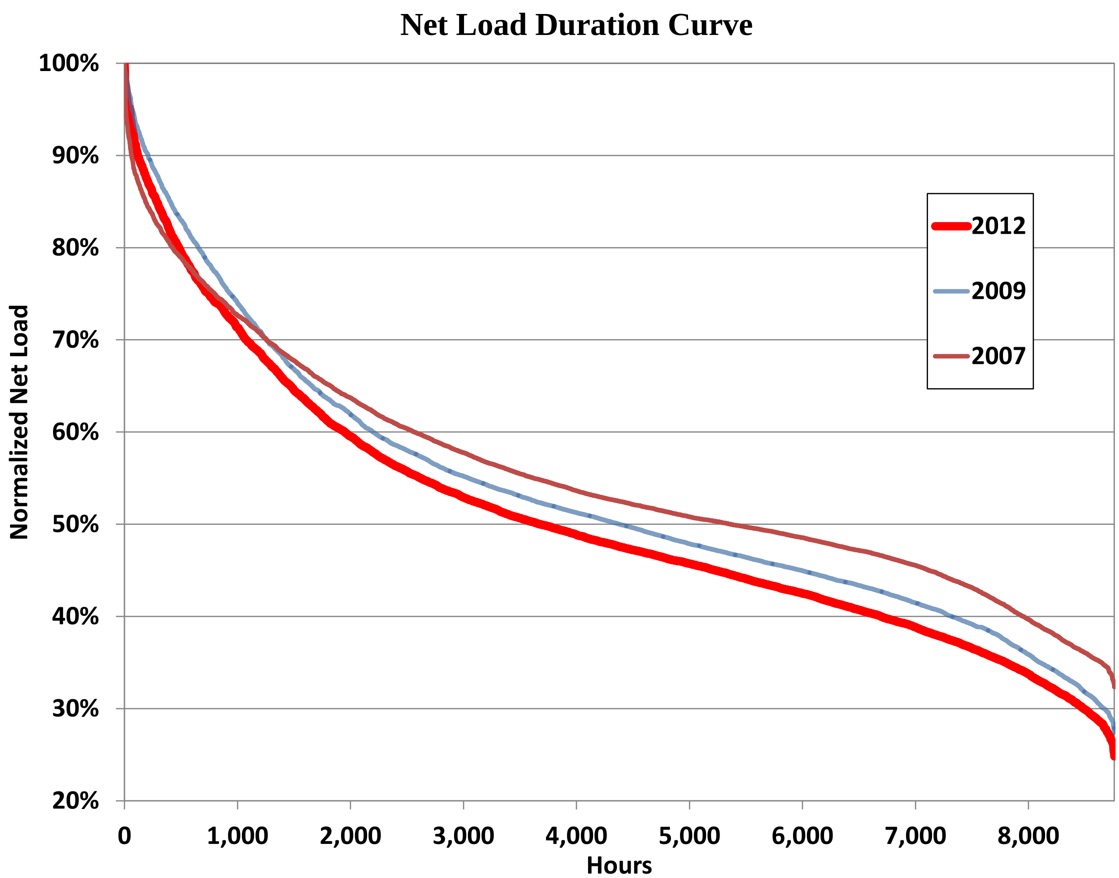
Source: ERCOT Capacity Demand Reserve Reports / 2013 data from Winter 2012, 2014 - 2018 from May 2013

The figure below shows the percent of annual generation from each fuel type for the years 2008 through 2012. The generation share from wind has increased every year, reaching 9 percent of the annual generation requirement in 2012, up from 5 percent in 2008. During the same period, the percentage of generation provided by natural gas decreased from 43 percent in 2008 to 38 percent in 2010, before increasing to 45 percent in 2012. Correspondingly, the percentage of generation produced by coal units increased from 37 percent to 40 percent in 2010 before decreasing to 34 percent in 2012. The increase in the share of generation produced by natural gas, and corresponding reduction in coal generation is due to historically low price of natural gas in 2012.



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 23 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. However, due to the low price of natural gas in 2012, we observe that the share of generation produced from coal-fired and nuclear units decreased to less than half of the energy in ERCOT, with the reduction coming from decreased coal generation.

Increasing levels of wind resource in ERCOT also has important implications for the net load duration curve faced by the non-wind fleet of resources. Net load is defined as the system load minus wind production. The figure below shows the net load duration curves for selected years since 2007, normalized as a percent of peak load. This figure shows the continued erosion of remaining energy available for non-wind units to serve during most hours of the year, with much less impact during the highest loads.

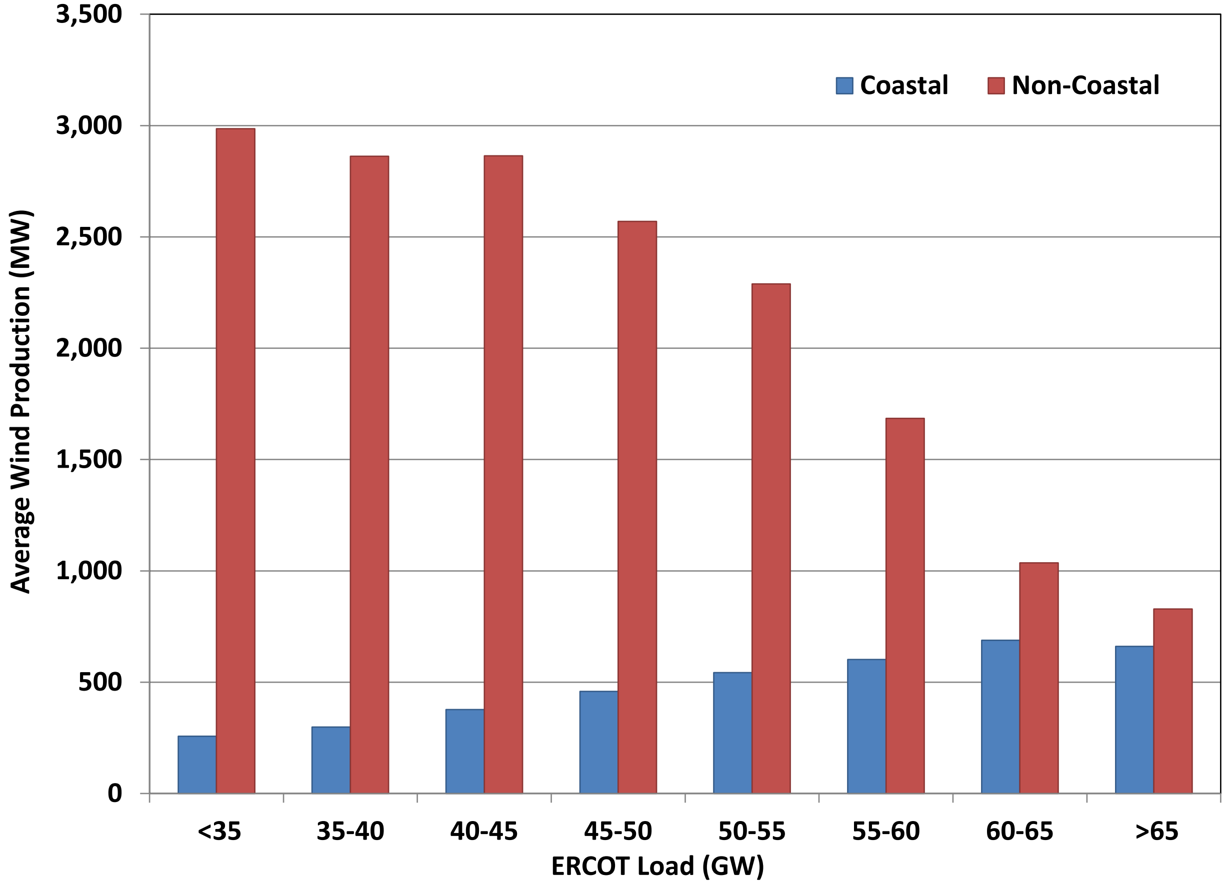


Thus, although the peak net load and reserve margin requirements are projected to continue to increase and create an increasing need for non-wind capacity to meet net load and reliability requirements, the non-wind fleet is expected to operate for fewer hours as wind penetration continues to increase. 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.

Even with the increased development activity in the coastal area of the South zone, more than 80 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 relatively 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 more significant reductions in the net load relative to the actual load in the other hours of the year. The trend shown from 2007 in the figure above may continue with the addition of new wind resources and the reduction in the curtailment of existing wind resources after completion of new transmission facilities.

The next figure compares the output during the summer months of June through August from wind units located in the coastal area of the South zone with those located elsewhere in ERCOT.

Summer Wind Production vs. Load



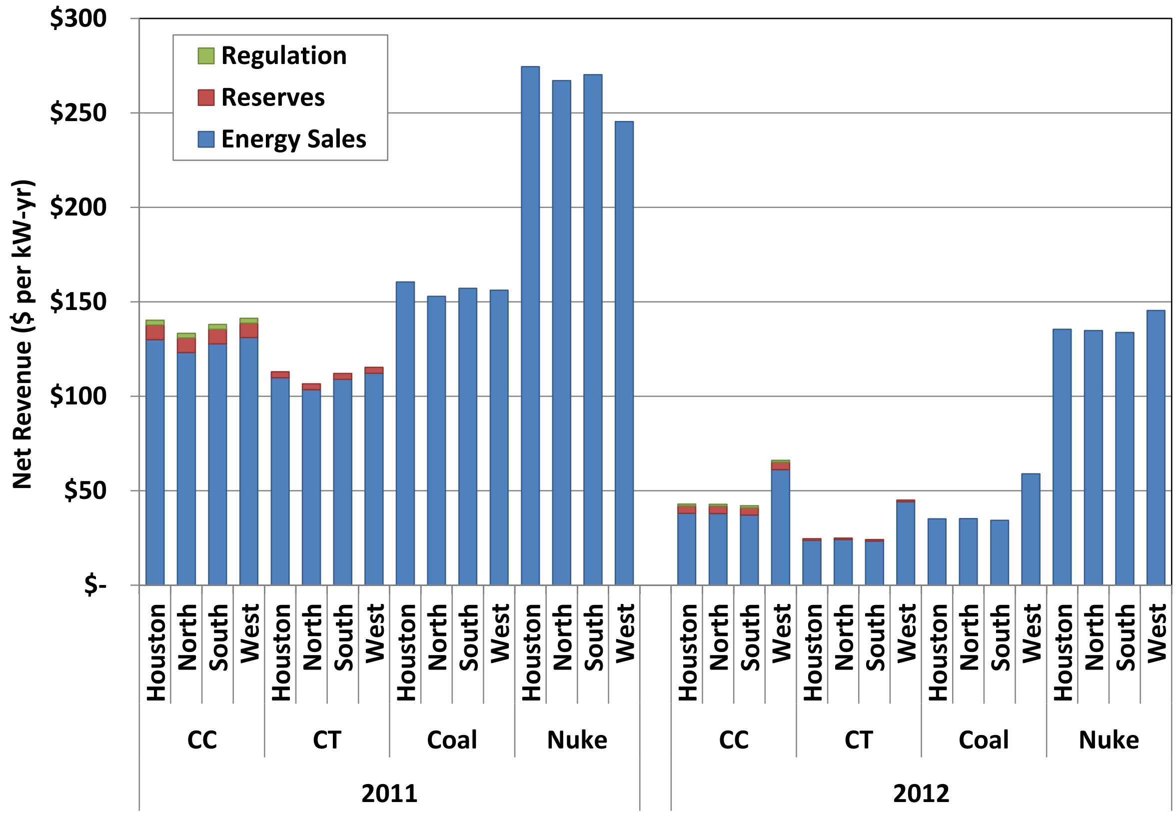
It shows a strong negative relationship between non-coastal wind output and increasing load levels. It further shows that the output from wind generators located in the coastal area of the South zone is much more highly correlated with peak electricity demand.

F. Resource Adequacy

Long-Term Economic Signals: Net Revenue

One of the primary functions of the wholesale electricity market is to provide economic signals that will encourage the investment needed to maintain a set of resources that are adequate to satisfy the system's demands and reliability needs. These economic signals are evaluated by estimating the "net revenue" new resources would receive from the markets. Net revenue is the total revenue that can be earned by a new 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 energy and ancillary services markets together provide the economic signals that inform suppliers' decisions to invest in new generation or retire existing generation. In long-run equilibrium, markets should provide sufficient net revenue to allow an investor to receive a return of, and on an investment in a new generating unit.

Estimated Net Revenue



The figure above shows the results of the net revenue analysis for four types of hypothetical new units in 2011 and 2012. These are: (a) natural gas-fired combined-cycle, (b) natural gas-fired combustion turbine, (c) coal-fired generator, and (d) a nuclear unit. For the natural gas-fired technologies, 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 in other hours that it is available. For coal and nuclear technologies, net revenue is calculated by assuming that the unit will produce at full output.

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 masks what may be very high locational values for a specific generator location. Some generators may also receive uplift payments because of their specific reliability contributions, either as a reliability must run, or through the reliability unit commitment. This source of revenue is not considered in this analysis. The analysis also includes simplifying assumptions that can lead to over-estimates of the profitability of operating in the wholesale market. Start-up costs and minimum run times and ramp restriction, which can prevent the natural gas generators from profiting during brief price spikes, are not explicitly accounted for in the net revenue analysis. Despite these limitations, the net revenue analysis provides a useful summary of signals for investment in the wholesale market.

The figure above shows that the net revenue for every generation technology type decreased substantially in 2012 compared to each zone in 2011.

- For a new coal-fired unit, the estimated net revenue requirement is approximately $210 to $270 per kW-year. The estimated net revenue in 2012 for a new coal unit was approximately $35 per kW-year.

- For a new nuclear unit, the estimated net revenue requirement is approximately $280 to $390 per kW-year. The estimated net revenue in 2012 for a new nuclear unit was approximately $134 per kW-year.

- For a new natural gas-fired combustion turbine, the estimated net revenue requirement is approximately $80 to $105 per kW-year. The estimated net revenue in 2012 for a new gas turbine was approximately $25 per kW-year.

- For a new combined cycle unit, the estimated net revenue requirement is approximately $105 to $135 per kW-year. The estimated net revenue in 2012 for a new combined cycle unit was approximately $42 per kW-year.

These results indicate that the ERCOT markets would not have provided sufficient revenues to support profitable investment in any of the types of generation technology evaluated. Higher energy prices in the West zone during 2012 resulted in higher net revenues in that zone, but they were still not high enough to support new entry there. The net revenues in 2012 were much lower than in 2011. However, it is important to recognize that 2011 was highly anomalous, with some of the hottest summer temperatures on record. Net revenues may have been sufficient to cover the costs of a new combined cycle or new combustion turbine in 2011, however, we would not expect this to be consistently true in years with comparable reserve margins absent the extreme weather conditions, as evidenced by the 2012 net revenue results.

Shortage Pricing, Capacity Markets, and Resource Adequacy

Efficient electricity markets allow energy prices to rise substantially at 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 expected 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.

The nodal market implementation brought about more reliable and efficient shortage pricing. Modifications implemented during 2012, which introduced offer floors associated with the deployment of generator-provided responsive reserves and non-spinning reserves, further improved pricing outcomes. However, ERCOT's real-time energy pricing can be improved by ensuring the value of curtailed load is fully reflected in prices when load resources are deployed and further improving its shortage pricing as recommended in this report.

The PUCT has devoted considerable effort over the past two years deliberating issues related to resource adequacy. These deliberations have resulted in changes to the rules governing the system-wide offer cap and the peaker net margin ("PNM") mechanism. The system-wide offer cap was increased to $4,500 per MWh effective August 1, 2012 and is scheduled to increase every year up to $9,000 per MWh on June 1, 2015. This is intended to raise market revenues to help address resource adequacy concerns. However, inflating the system-wide offer cap may raise efficiency concerns if the resulting energy prices are set at or above levels consistent with the value of lost load during periods when the system is only slightly short of operating reserves and involuntary load curtailment is not imminent. This is a concern because setting prices substantially higher than the expected value of lost load can cause market participants to take inefficient actions, resulting in higher overall market costs. To address this concern, we recommend that ERCOT:

- Modify the slope of the existing power balance penalty curve and the offer floor for responsive reserve service to provide a more gradual slope up to the system-wide offer cap; and
- Modify the automatic pricing of unoffered capacity such that it is not all priced at the system-wide offer cap to avoid the inefficiencies associated with the automated economic withholding of such capacity.

Regardless of the means by which revenues are produced in a wholesale electricity market, it is fundamental that investment will only occur when the total net revenues expected by the investor are greater than its entry costs (including profit on its investment). Additionally, these sources of revenue must be available to all resources, both new and existing, in order to facilitate efficient investment, maintenance, and retirement decisions by all suppliers.

In an energy only market, the primary source of such revenue is the net revenues received during periods of shortage. Expectations about both the magnitude of the energy price during shortage conditions and the frequency of shortage conditions are the primary means to attract new investment in an energy-only market. If the expected revenues are not high enough to facilitate enough investment to satisfy the planning reserve target, one option is to increase the shortage pricing levels to levels that substantially exceed the expected value of lost load. As the planning reserve levels grow, however, the frequency of shortages will tend to drop sharply, which can make it difficult to use this means to meet planning reserve targets.[1] Additionally, as discussed below, such approaches introduce costly operational inefficiencies into the ERCOT energy markets.

[1] The difficulty of relying primarily on shortage pricing will depend on how high the planning reserve target is relative to the planning reserve levels any energy-only market priced at the expected value of lost load would provide. See the discussion of the Brattle Report below.

Most other competitive electricity markets do not rely solely on shortage pricing to generate sufficient revenue to support the capacity additions necessary to satisfy their planning reserve requirements. They employ capacity markets to competitively generate capacity payments over the year that are made to suppliers in return for meeting defined capacity obligations. Capacity prices and associated payments vary monthly or annually based on long-term planning reserve levels, independent of the real-time supply and demand conditions. These capacity markets are designed to ensure that a specified planning reserve margin is achieved.

In 2012, ERCOT engaged The Brattle Group to assess its resource adequacy outlook by evaluating a number of market design scenarios.[2] Brattle also supplemented its report with a comparison of costs and reliability for the energy-only market and two capacity market scenarios with 10% and 14% reserve margin requirements.[3] The results of this analysis are summarized in the table below.

COMPARISON OF COSTS AND RELIABILITY

	Energy-Only Equilibrium	10% Reserve Margin Requirement	14% Reserve Margin Requirement
Reliability			
Reserve Margin	8%	10%	14%
Reserve Margin Certainty	Uncertain	More Certain	More Certain
Annual Avg. Loss of Load Hours	4.1	2.2	0.3
Customer Costs			
Energy Costs ($billions)	$18.3	$16.3	$14.0
Capacity Costs ($billions)	$0	$2.1	$4.7
Total Costs ($billions)	$18.3	$18.4	$18.7
Cost Increase over Energy-Only Equilibrium (%)	NA	0.7%	2.4%
Rate Increase over Energy-Only Equilibrium (%)	NA	0.4%	1.4%
Combustion Turbine Energy Margins and Capacity Revenues			
Energy Margins ($/kW-y)	$105	$75	$41
Capacity Revenues ($/kW-y)	$0	$30	$64
Total Margins ($/kW-y)	$105	$105	$105

Notes: 8% energy-only equilibrium reserve margin based on The Brattle Group's simulations with a $9,000 price cap and gradually sloping scarcity pricing function. Rate impacts assume generation costs comprise 60% of total retail rates.

[2] *ERCOT Investment Incentives and Resource Adequacy*, The Brattle Group, PUCT Docket No. 37987 (June 1, 2012).

[3] *Customer Cost Comparison*, The Brattle Group, PUCT Docket No. 40000 (Sept. 4, 2012).

As indicated in the table above, Brattle estimates that even with $9,000 per MWh system-wide offer caps, economic equilibrium for the ERCOT energy-only market is achieved at an 8 percent planning reserve margin, although the actual reserve margin outcomes will be uncertain. Brattle further estimates that in the energy-only market at annual equilibrium, wholesale generation costs will be $18.3 billion. In contrast, Brattle's assessment of a capacity market with a more certain 14 percent reserve margin expectation, estimates generation costs at annual equilibrium to be $18.7 billion.

It is important to recognize that this increase in cost is not due to the introduction of the capacity market, it is due to the requirement to sustain a planning reserve margin greater than 8 percent. In fact, the Brattle analysis indicates that a capacity market would deliver the higher planning reserve margin at a relatively low incremental cost with much more certainty. Further, recent studies have indicated that to maintain the same small level of risk of having an involuntary curtailment of firm load, the planning reserve target should be increased from 13.75 percent to approximately 16 percent. Hence, the difficulty of satisfying ERCOT's planning needs with shortage pricing alone will grow if this recommendation is adopted.

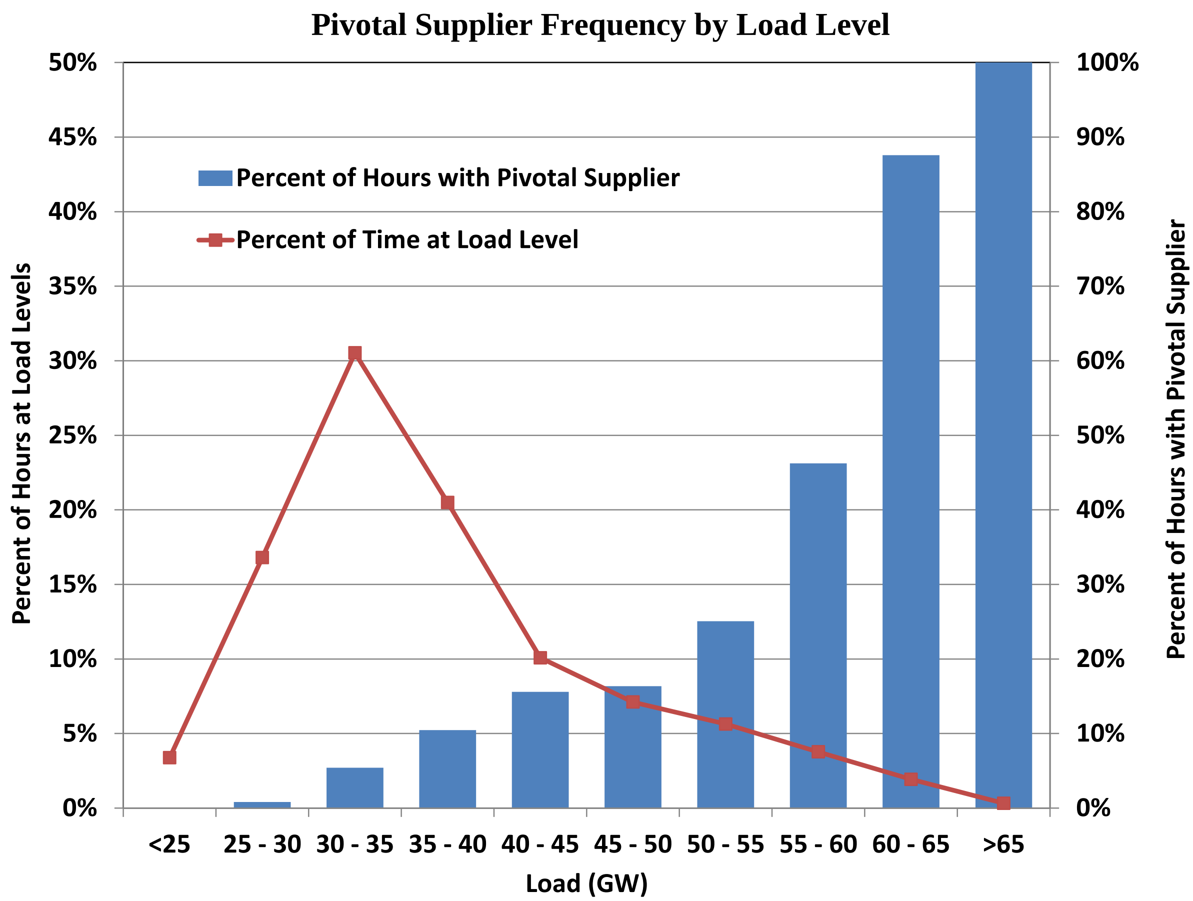
In response to these observations, proposals have been put forth that would introduce significant operational inefficiencies into the ERCOT energy markets, such as a requirement to substantially increase the quantity of operating reserves ERCOT procures and to, by rule, economically withhold these surplus reserves from the market. Such approaches would introduce significant inefficiencies into ERCOT day ahead and real time operations in an effort to manufacture more frequent shortage pricing and a higher planning reserve margin than would be achieved in a pure energy-only market framework. However, such approaches will not guarantee that the planning reserve targets will be satisfied and, because of the resulting operational inefficiencies, will be more costly for ERCOT's consumers. Hence, consistent with Brattle's findings, it is our view that if the planning reserve margin is viewed as a minimum requirement, implementation of a capacity market is the most efficient mechanism to achieve this objective. As observed by Brattle, a well-designed capacity market can efficiently meet a planning reserve requirement without impairing the efficiency of energy market operations. However, there are many determinations required in the design, implementation and maintenance of a capacity market construct.[4]

G. Analysis of Competitive Performance

The report evaluates market power from two perspectives, structural (does market power exist) and behavioral (have attempts been made to exercise it). The Residual Demand Index ("RDI") is used to as the primary indicator of potential structural market power. The RDI measures 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. When the RDI is greater than zero the largest supplier is pivotal; that is, 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, it does identify conditions under which a supplier would have the ability to raise prices significantly by withholding resources.

[4] *ERCOT Investment Incentives and Resource Adequacy*, The Brattle Group, at 115-119, PUCT Docket No. 37987 (June 1, 2012).



The figure above summarizes the results of the RDI analysis by displaying the percent of time at each load level there was a pivotal supplier. At loads greater than 65 GW there was a pivotal supplier 100 percent of the time. The figure also displays the percent of time each load level occurs. Combining these values we find that there was a pivotal supplier in approximately 12 percent of all hours of 2012. As a comparison, the same system-wide measure for the Midwest ISO showed less than 1 percent of all hours with a pivotal supplier.

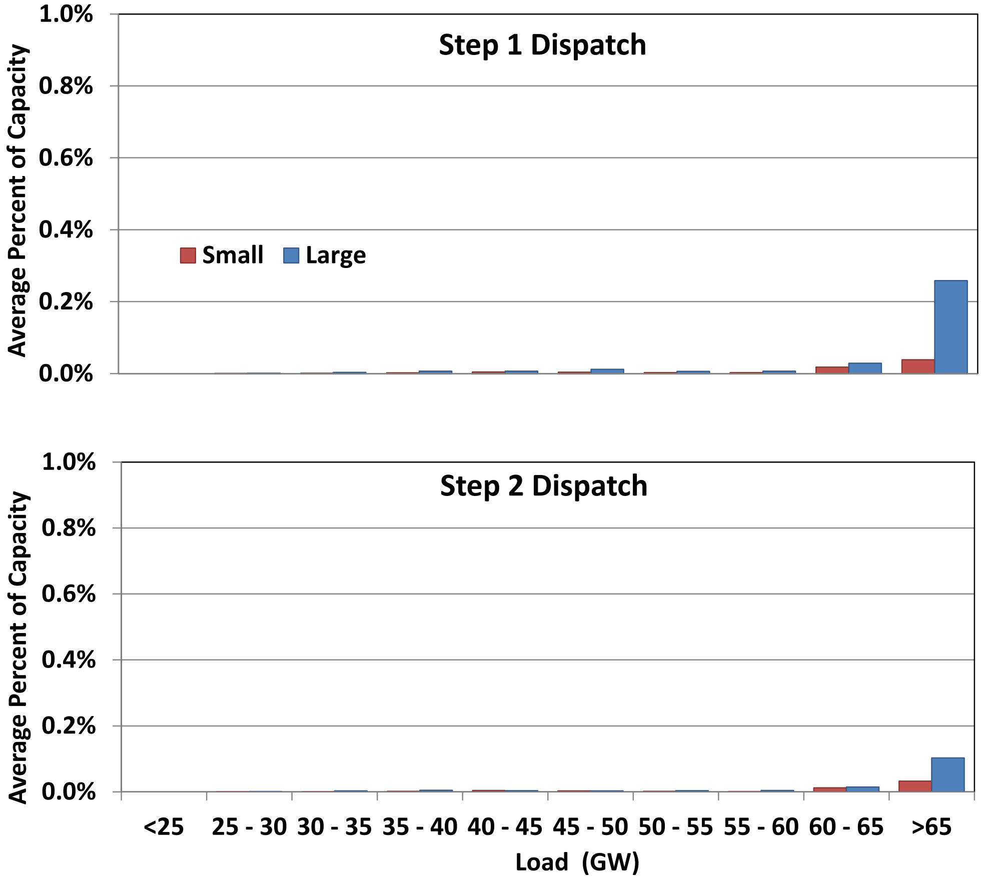
The behavioral aspects of market power abuse are evaluated by calculating an "output gap." The output gap is defined as the quantity of energy that is not being produced by in-service capacity even though the in-service capacity is economic by a substantial margin given the real-time energy price. A participant can economically withhold resources, as measured by the output gap, by raising its energy offers so as not to be dispatched.

Resources are considered for inclusion in the output gap when they are committed and producing at less than full output. Energy not produced from committed resources is included in the output gap if the real-time energy price exceeds by at least $50 per MWh that unit's mitigated offer cap, which serves as an estimate of the marginal production cost of energy from that resource.

The output gap is measured at both steps in ERCOT's two-step dispatch because 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 of ERCOT's dispatch process. Although in the second step, the offer appears to be mitigated, the market participant has still influenced the market price. This output gap is measured by the difference between the capacity level on a generator's original offer curve at the first step reference price and the capacity level on the generator's cost curve at the first step reference price. However, this output gap is only indicative because no output instructions are sent based on the first step. It is only used to screen out whether a market participant is withholding in a manner that may influence the reference price.

The ultimate output gap is 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. 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 due to a market participant exerting market power.

Incremental Output Gap by Load Level and Participant Size



The figure above shows the magnitude of the output gap to be very small, even at the highest load levels, for both steps in the dispatch process, and for both small and large generators. These small quantities raise no competitive concerns. In addition to this metric, we also evaluate outages, deratings, and economic units that were not committed to identify other means suppliers may have used to withhold resources. Based on the analysis above and our other monitoring screens, we find that the ERCOT nodal wholesale market performed competitively in 2012.

H. Recommendations

Last year we recommended changes to the automated mitigation procedures that are part of the real-time dispatch to eliminate the occurrences of over-mitigation we have observed. As more fully described in Section I.E, Mitigation at page 20, we support the introduction of a test to determine whether a unit is either contributing to, or helping to resolve a transmission constraint and only subject the relieving units to mitigation. These changes were included in NPRR520 and should substantially reduce the occurrence of mitigating resources that are not in a position to exert market power related to the relief of transmission constraints. Various parameters will be approved related to the implementation of these changes, currently scheduled for summer 2013, and performance should be closely monitored to determine if any adjustments are required.

Last year we also recommended a change to the real-time market software to allow it to "look ahead" a sufficient amount of time to better commit load and generation resources that can be online within 30 minutes. More discussion of this topic can be found starting on page 84 in Section V.B, Effectiveness of the Scarcity Pricing Mechanism. We still believe this functionality would enhance the performance of the ERCOT market. However, this will need to be coordinated with the other fundamental market design changes still currently under consideration.

Whatever these future market design changes may entail, we recommend stakeholder consideration of the following three modifications, particularly as the system-wide offer cap rises above $5,000 per MWh.

1. Modify the slope of the existing power balance penalty curve and the offer floors for responsive reserve service to provide a more gradual slope up to the system-wide offer cap such that the cap is reached when operating reserves are down to the level that ERCOT would initiate involuntary curtailment of firm load.

    - If system-wide offer caps of $5,000, $7,000 or $9,000 are intended to reflect the value of lost load, then real-time energy prices should only get to those levels once firm load is being involuntarily curtailed.

- A more “well-behaved” reserve shortage pricing function could be achieved in the following ways:
  - Implement real-time co-optimization of energy and reserves with an operating reserve demand curve;
  - Introduce an operating reserve demand curve but not include real-time co-optimization.
  - Adjust the current operating reserve offer floors to better reflect the loss of load probability and the value of lost load at various levels of operating reserves.

2. Modify the Protocols related to proxy offer curve provisions such that all unoffered capacity is not automatically priced at the system-wide offer cap. Currently, if available capacity does not have an associated energy offer, ERCOT’s dispatch software “fills in” with an offer priced at the system-wide offer cap. During 2012, the average amount of capacity priced in this manner exceeded 100 MW.

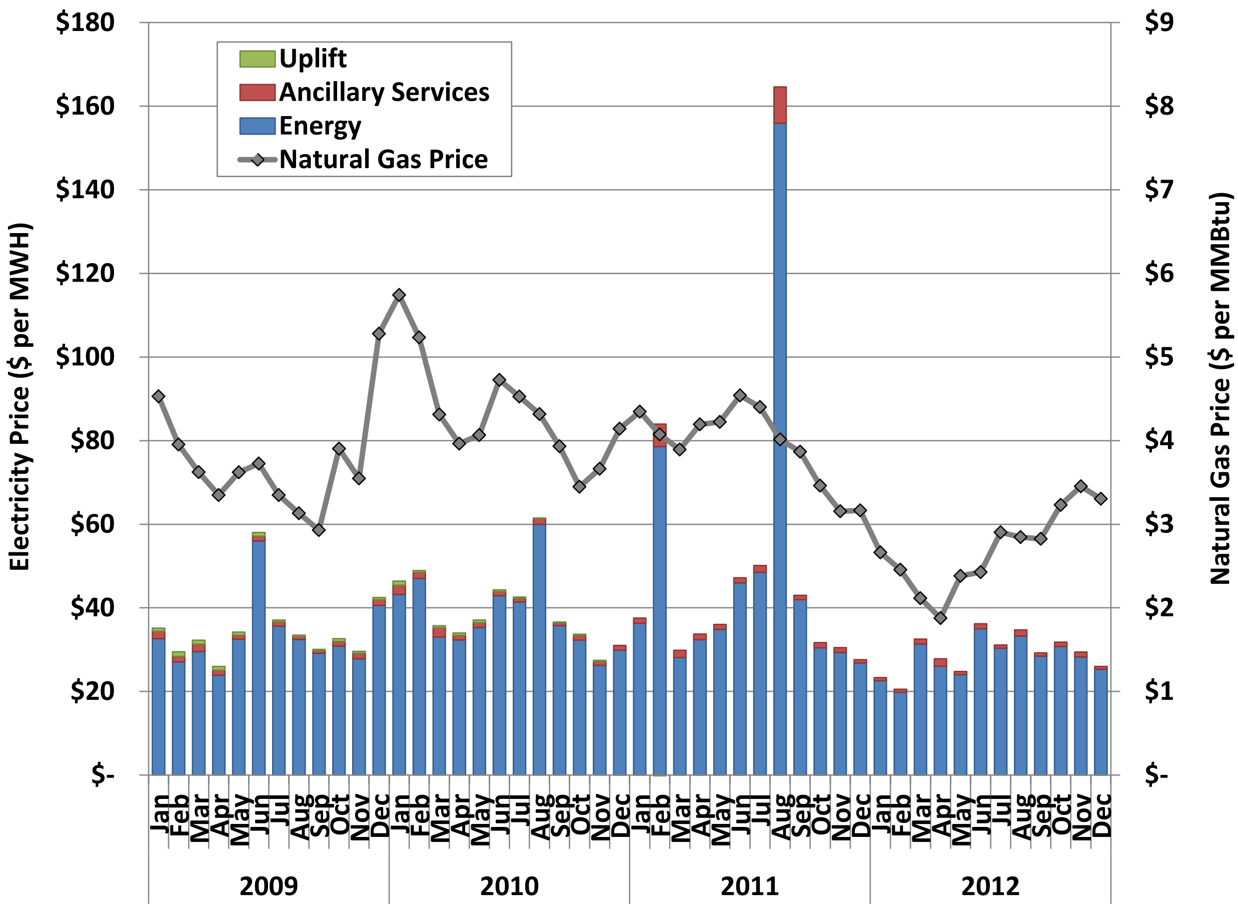
3. Implement changes that ensure ERCOT deployments of load resources, Emergency Response Service (ERS), or the involuntary curtailment of firm load are reflected in the real-time dispatch energy and reserve prices. This may be achieved through various means, either as a component of integrating load bids in the real-time dispatch software, or through simple administrative shortage pricing rules.

I. REVIEW OF REAL-TIME MARKET OUTCOMES

A. Real-Time Market Prices

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

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



Energy, ancillary services and uplift costs are the three components in the all-in price of electricity. The ERCOT wide price is the load weighted average of the real-time market prices from all load zones. Prior to ERCOT's conversion to the nodal market in December 2010, energy costs were determined from the zonal balancing energy market. Ancillary services costs are estimated based on total system demand and prices in the ERCOT markets for regulation, responsive reserves, and non-spinning reserves. Uplift costs are assigned market-wide on a load-ratio share basis to pay for charges associated with additional reliability unit commitment and any reliability must run contracts.[5]

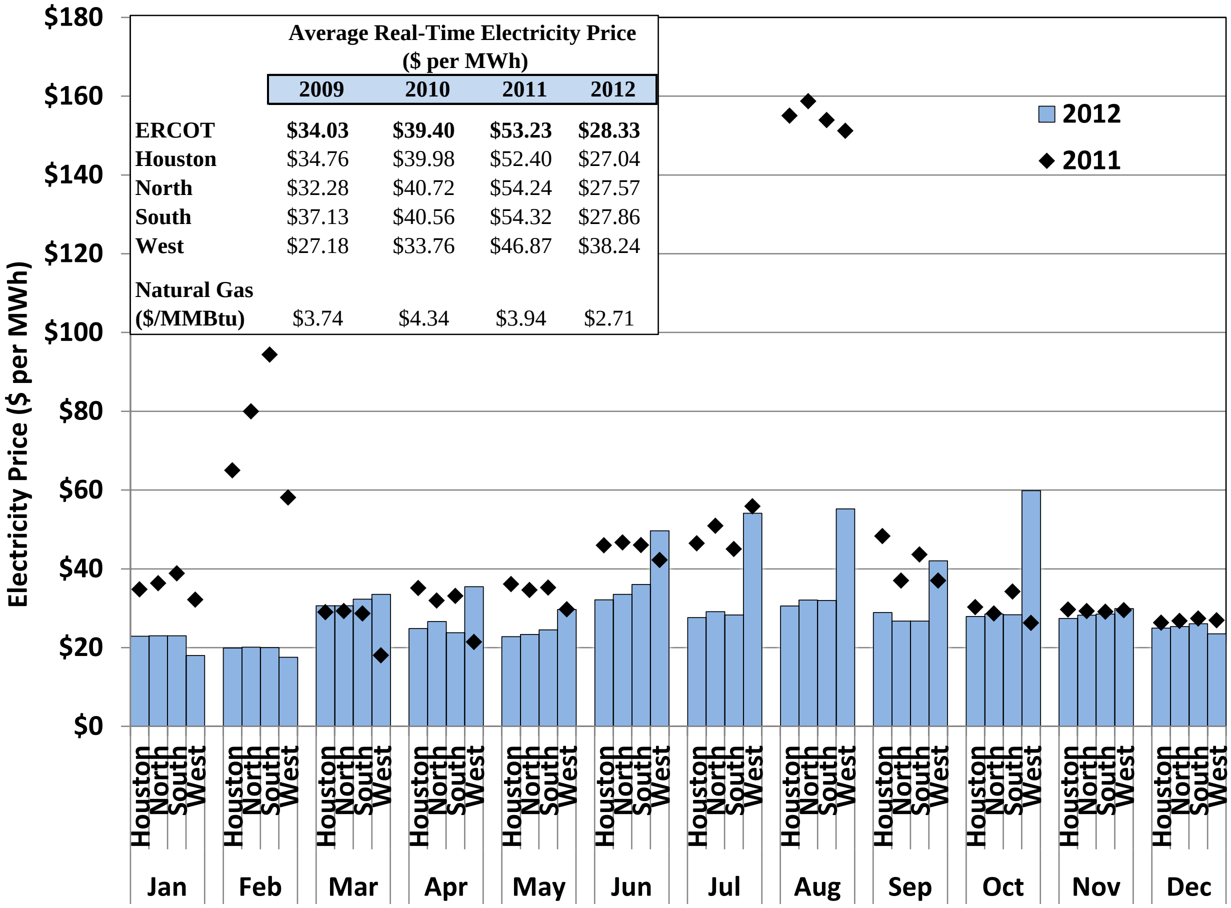
Figure 1 shows the monthly average all-in price for all of ERCOT from 2009 to 2012 and the associated natural gas price. This figure indicates that natural gas prices were a primary driver of the trends in electricity prices from 2009 to 2012. Again, this is not surprising given that natural gas is a widely-used fuel for the production of electricity in ERCOT, especially among generating units that most frequently set locational marginal prices in the nodal market.

The largest component of the all-in cost of wholesale electricity is the energy cost, which is reflected by the locational marginal prices. As is typical in other wholesale electricity markets, only a small share of the power produced in ERCOT is transacted in the spot market. However, prices in the real-time energy market are very important because they set the expectations for prices in the forward markets (including bilateral markets) where most transactions take place. Unless there are barriers preventing arbitrage of the prices between the spot and forward markets, the prices in the forward market should be directly related to the prices in the spot market (*i.e.*, the spot prices and forward prices should converge over the long-run). Hence, artificially low prices in the real-time energy market will translate to artificially-low forward prices. Likewise, price spikes in the real-time energy market will increase prices in the forward markets. This section evaluates and summarizes electricity prices in the real-time market during 2012.

To summarize the price levels during the past four years, Figure 2 shows the monthly load-weighted average prices in the four geographic ERCOT load zones. These prices are calculated by weighting the energy price for each interval and each zone by the total zonal load in that interval. Since December 2010 these prices were determined by the nodal real-time energy market. Prior prices were derived from the zonal balancing energy market. Load-weighted average prices are the most representative of what loads are likely to pay, assuming that real-time energy prices are, on average, generally consistent with bilateral contract prices.

[5] Prior to December 2010 uplift costs included charges for out-of-merit energy and capacity, replacement reserve services and any reliability must run contracts.

Figure 2: Average Real-Time Energy Market Prices



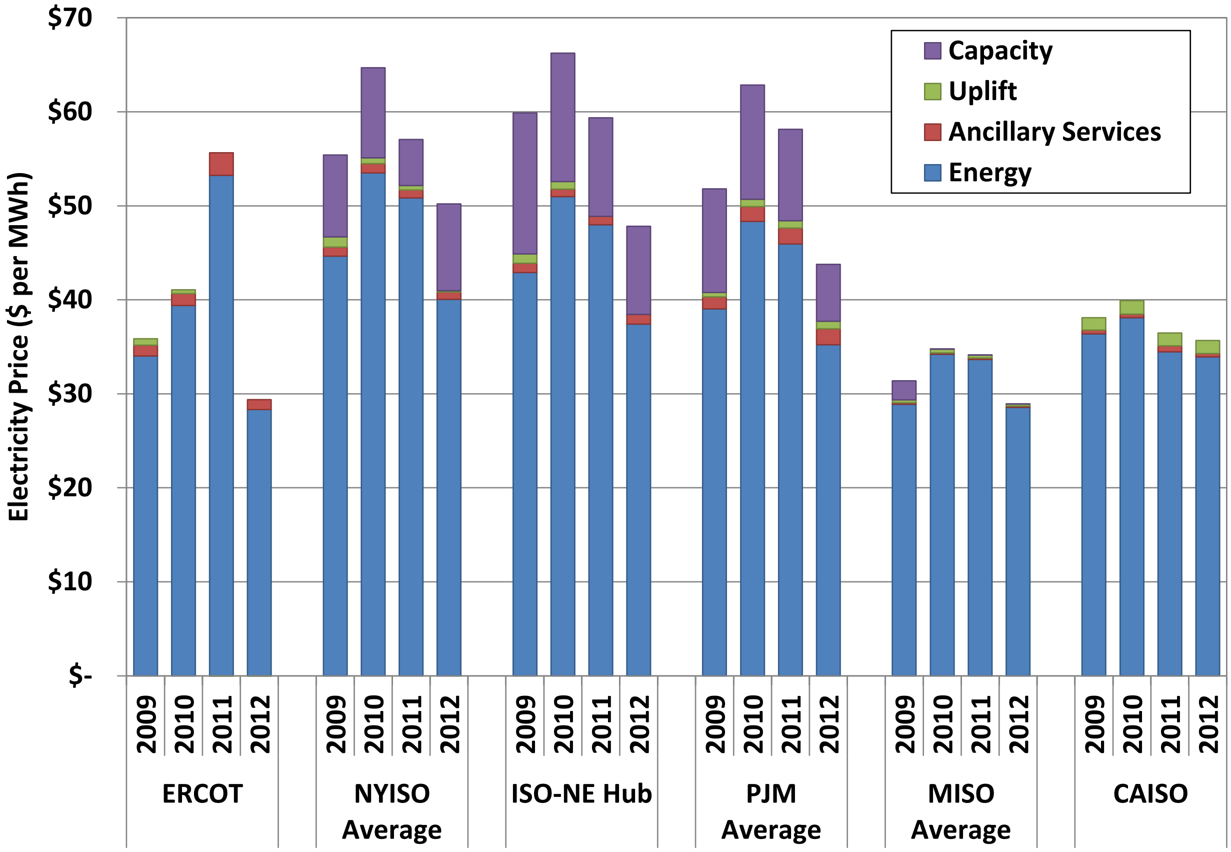
Average Real-Time Electricity Price ($ per MWh)

	2009	2010	2011	2012
ERCOT	**$34.03**	**$39.40**	**$53.23**	**$28.33**
Houston	$34.76	$39.98	$52.40	$27.04
North	$32.28	$40.72	$54.24	$27.57
South	$37.13	$40.56	$54.32	$27.86
West	$27.18	$33.76	$46.87	$38.24
Natural Gas ($/MMBtu)	$3.74	$4.34	$3.94	$2.71

ERCOT average real-time market prices were 47 percent lower in 2012 than in 2011. The ERCOT-wide load-weighted average price was $28.33 per MWh in 2012 compared to $53.23 per MWh in 2011. The decrease in real-time energy prices was correlated with much lower fuel prices in 2012. The steady decline in natural gas prices from June 2011 to April 2012 resulted in the 2012 average natural gas price of $2.71 per MMBtu, a 31 percent decrease compared to $3.94 per MMBtu in 2011.

To provide additional perspective on the outcomes in the ERCOT market, our next analysis compares the all-in price metrics for ERCOT and other electricity markets. The following figure compares the all-in prices in ERCOT with other organized electricity markets in the U.S.: New York ISO, ISO New England, PJM, Midwest ISO, and California ISO.

Figure 3: Comparison of All-in Prices across Markets

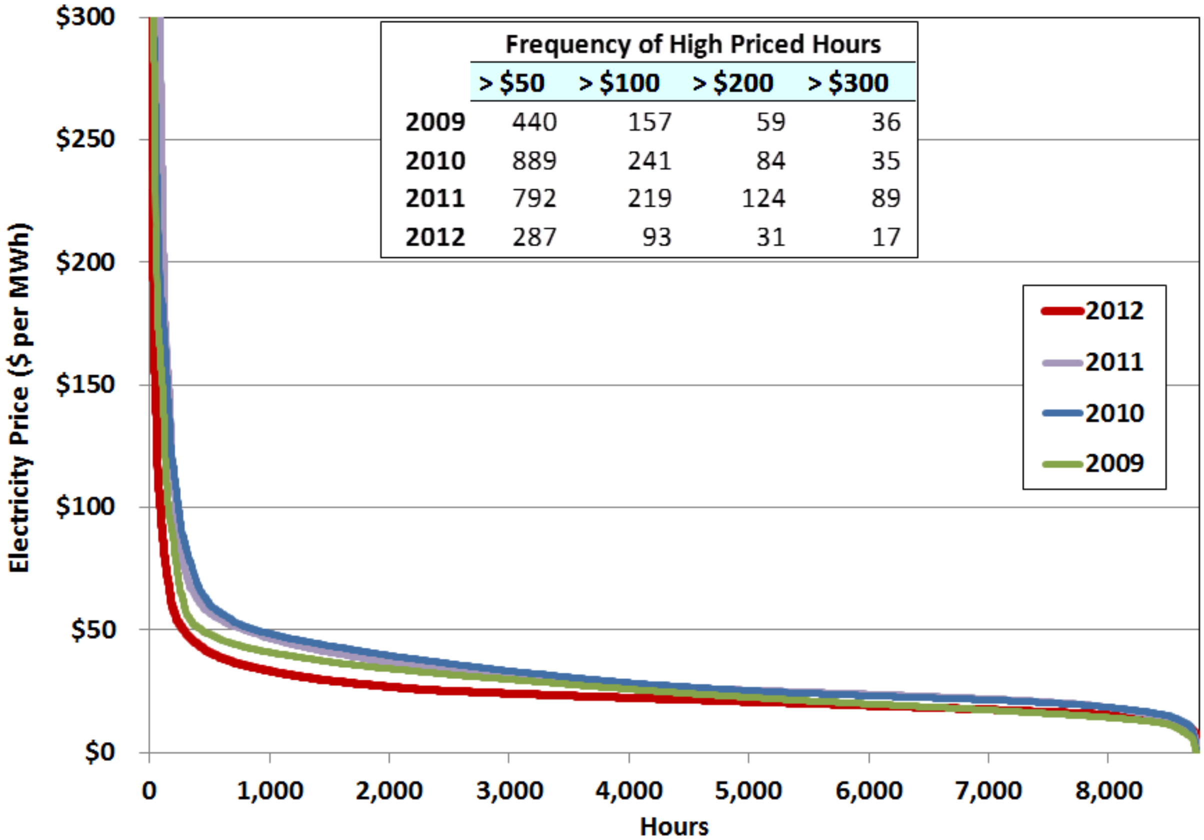


For each region, the figure reports the average cost (per MWh of load) for energy, ancillary services (reserves and regulation), capacity markets (if applicable), and uplift for economically out-of-merit resources. Figure 3 shows that ERCOT all-in prices in 2012 were roughly equivalent to the Midwest ISO and significantly lower than all other regions. Although prices in all markets declined from 2011 to 2012, no other region experienced anything close to the magnitude of reduction seen in ERCOT.

Figure 4 presents price duration curves for ERCOT energy markets in each year from 2009 to 2012. A price duration curve indicates the number of hours (shown on the horizontal axis) that the price is at or above a certain level (shown on the vertical axis). The prices in this figure are the hourly load-weighted zonal balancing energy price for the zonal market and hourly load-weighted nodal settlement point price for the nodal market.[6]

6 ERCOT switched to a nodal market on December 1, 2010. The December nodal prices are included in the 2010 price duration curve.

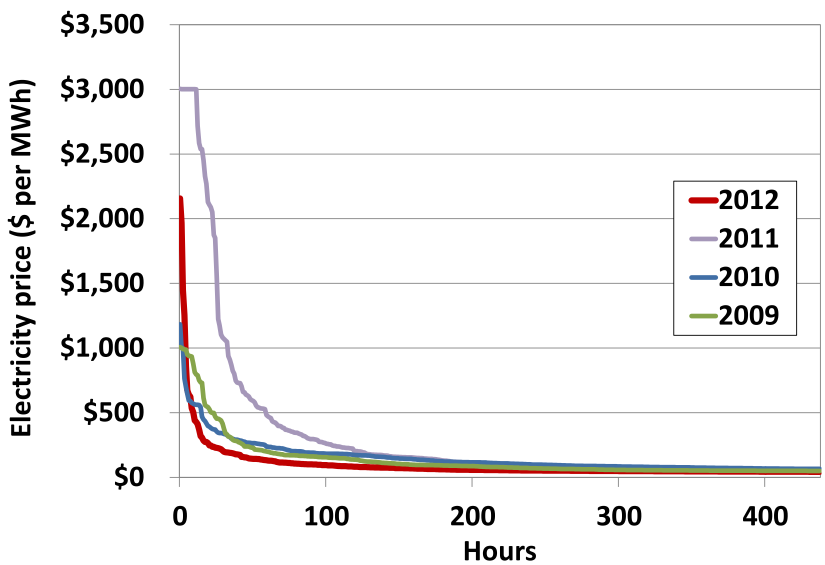
Figure 4: ERCOT Price Duration Curve



Due to the lowest natural gas prices seen in ten years, the 2012 price duration curve is below the duration curve of other years in most hours.

To see where the prices during 2012 were much different than in the previous three years, we present Figure 5, which compares prices for the highest five percent of hours. In 2011, energy prices for the top 100 hours were significantly higher due to higher loads leading to more shortage conditions coupled with a more effective shortage pricing mechanism

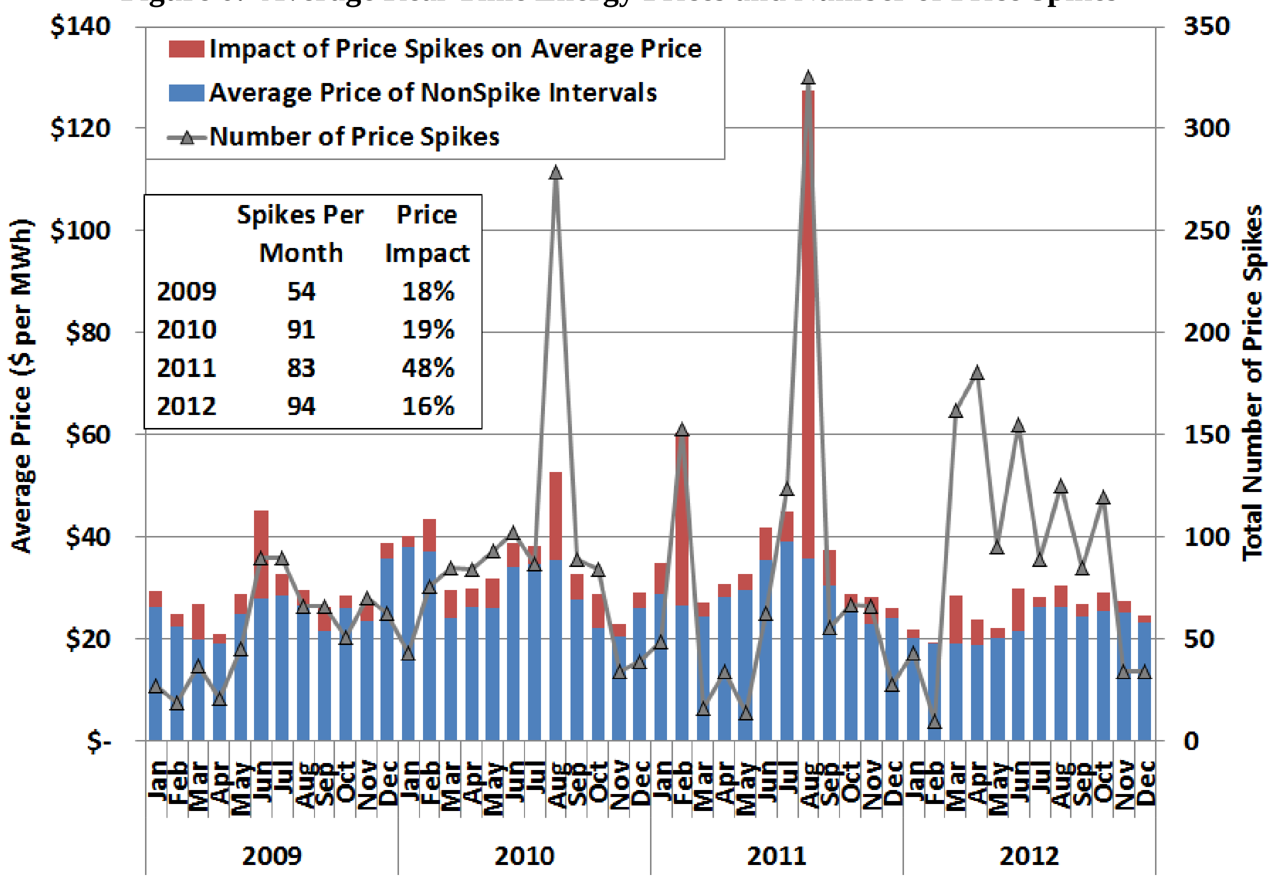
Figure 5: ERCOT Price Duration Curve
Top 5% of Hours



implemented as part of the nodal market design. In 2012, the energy duration curve for the top five percent of hours is lower than the past three years for the majority of hours, reflecting lower loads and resulting fewer occassions of shortage conditions. However, during the brief periods of shortage that were experienced in 2012 prices rose to levels higher than those experienced in 2009 and 2010 when all prices in the zonal market were the result of actual submitted offers.

To better observe the effect of the highest-priced hours, the following analysis focuses on the frequency of price spikes in the real-time energy market. Data prior to December 2010 is from the zonal balancing energy market. Figure 6 shows the average price and the number of price spikes in each month. For this analysis, price spikes are defined as intervals where the load-weighted average energy price in ERCOT is greater than 18 MMBtu per MWh times the prevailing natural gas price. Prices at this level have historically exceeded the marginal costs of virtually all of the on-line generators in ERCOT.

Figure 6: Average Real-Time Energy Prices and Number of Price Spikes

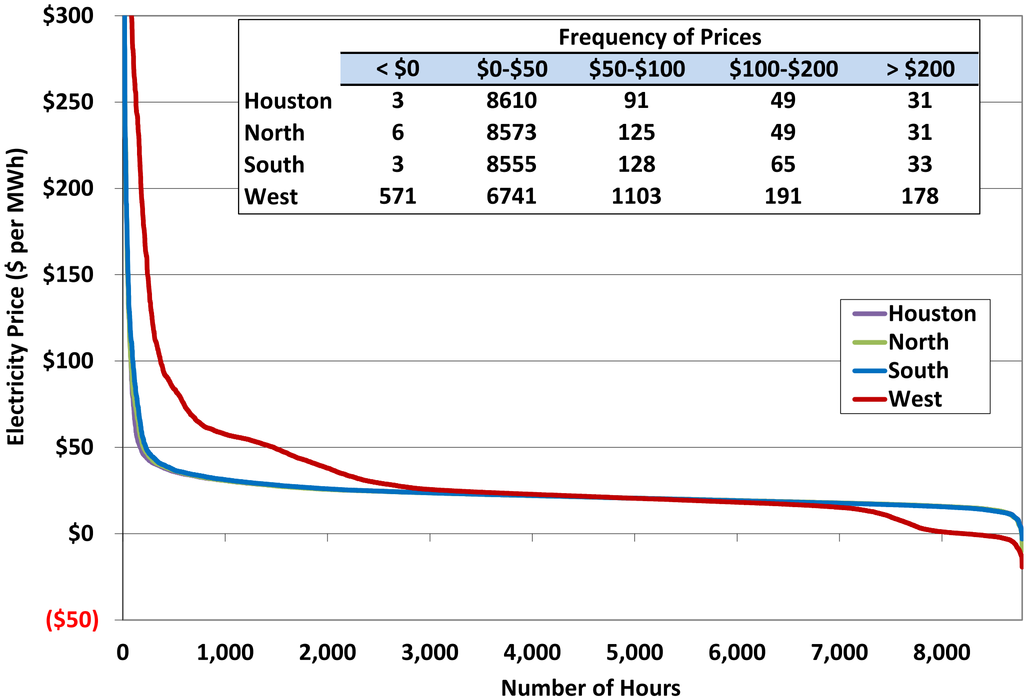


The number of price spike intervals during 2012 was 94 per month, an increase from the 83 per month in 2011. However, just looking at the average can be misleading. Due to extreme weather in February and August 2011, there were only two months with very high numbers of price spikes in 2011. In contrast, all months from March to October 2012 had at least 85 price spikes. As discussed later in this section, the high number of price spikes in 2012 is likely related to the very low price of natural gas and resulting ‘overlap’ of offers from natural gas and coal-fired units.

To measure the impact of these price spikes on average price levels, the figure also shows average prices with and without the price spike intervals. The top portions of the stacked bars show the impact of price spikes on monthly average price levels. Prior to 2012, the impact grew with the frequency of the price spikes, averaging $4.67, $5.53 and $14.09 per MWh during 2009, 2010 and 2011, respectively. Although the frequency of price spikes increased in 2012, the magnitude of their price impact decreased. The impact on average energy price in 2012 declined to $3.63 per MWh, or 16 percent of the annual average price. This is explained by much lower natural gas prices in 2012, resulting in a much lower threshold level for the definition of a “price spike”.

To depict how real-time energy prices vary by hour in each zone, Figure 7 below shows the hourly average price duration curve in 2012 for four ERCOT load zones. The Houston, North and South load zones had similar prices over the majority of hours. The price duration curve for the West zone is noticeably different than the other zones, with more hours with prices greater than $50 per MWh and more than 500 hours (6 percent of the time) when the average hourly price was less than zero. As observed over the past few years, West zone prices are lower than the rest of ERCOT when high wind output in the West results in congested transmission interfaces from the West zone to the other zones in ERCOT. Recently, prices higher than the rest of ERCOT have occurred in the West zone due to local transmission constraints that typically occur under low wind and high load, or outage conditions.

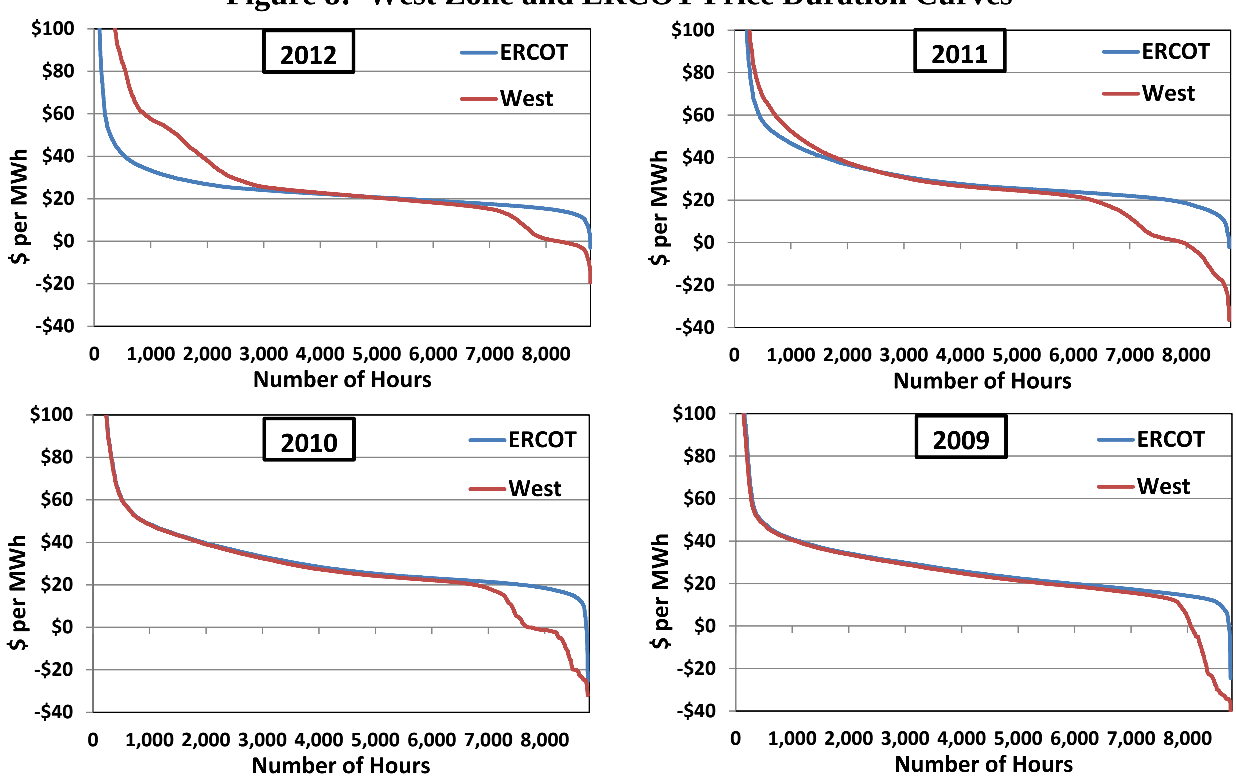
Figure 7: Zonal Price Duration Curves



	Frequency of Prices				
	< $0	$0-$50	$50-$100	$100-$200	> $200
Houston	3	8610	91	49	31
North	6	8573	125	49	31
South	3	8555	128	65	33
West	571	6741	1103	191	178

Figure 8 below shows the relationship between West zone and ERCOT average prices for the 2009 through 2012.

Figure 8: West Zone and ERCOT Price Duration Curves



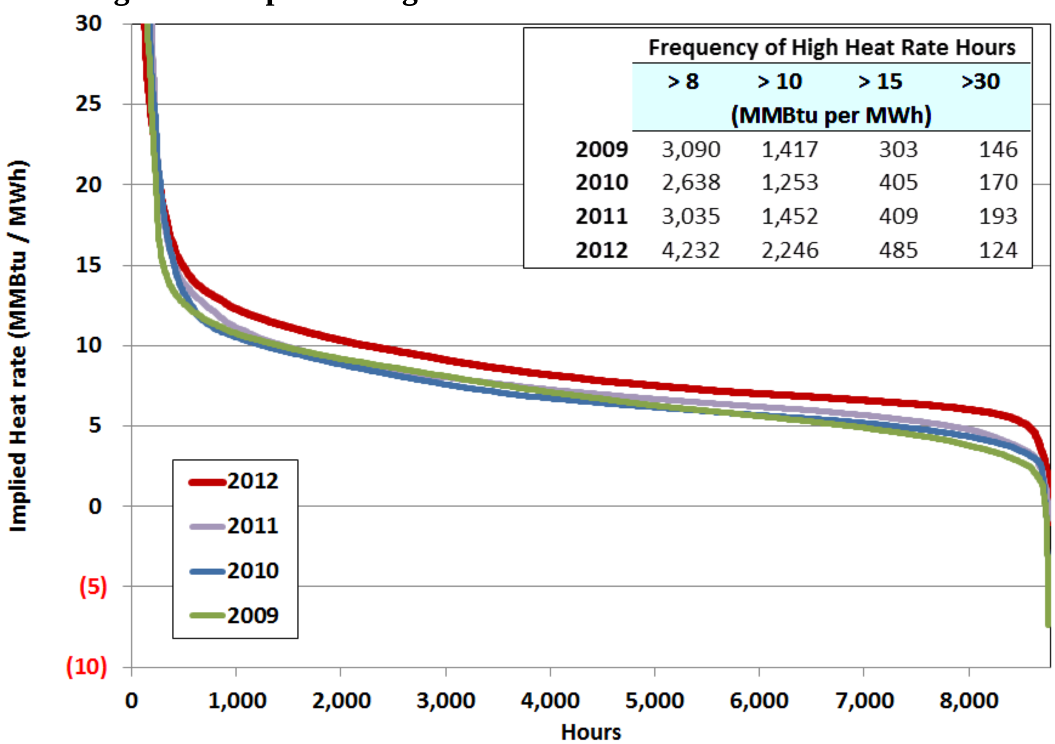
On the low price end, we observe a reduction in the number of hours when West zone prices were below the ERCOT average. We also note that minimum West zone prices have increased; that is, become “less negative”. During 2012, for the first time, West zone prices were much higher than ERCOT average for a significant number of hours. The combination of more hours with higher prices, and fewer hours with less negative prices resulted in the average real-time energy price in the West zone being greater than the ERCOT average.

More details about the transmission constraints influencing energy prices in the West zone are provided in Section III, Transmission and Congestion.

B. Real-Time Prices Adjusted for Fuel Price Changes

Although real-time electricity prices are driven to a large extent by changes in fuel prices, natural gas prices in particular, they are also influenced by other factors.

Figure 9: Implied Marginal Heat Rate Duration Curve – All hours



To clearly identify changes in electricity prices that are not driven by changes in natural gas prices, Figure 9 and Figure 10 show the load weighted, hourly average real-time energy price adjusted to remove the effect of natural gas price fluctuations. The first chart shows a duration curve where the real-time energy price is replaced by the marginal heat rate that would be implied if natural gas was always on the margin.[7]

The second chart shows the same duration curves for the five percent of hours in each year with the highest implied heat rate. Both figures show duration curves for the implied marginal heat rate for 2009 to 2012. In contrast to Figure 4 where the 2012 price duration curve lies below the curves of other years, Figure 9 shows that the implied marginal heat rates were higher in 2012 as compared to the three prior years. This can be explained by the much lower natural gas prices in 2012.

Figure 10: Implied Marginal Heat Rate Duration Curve – Top five percent of hours

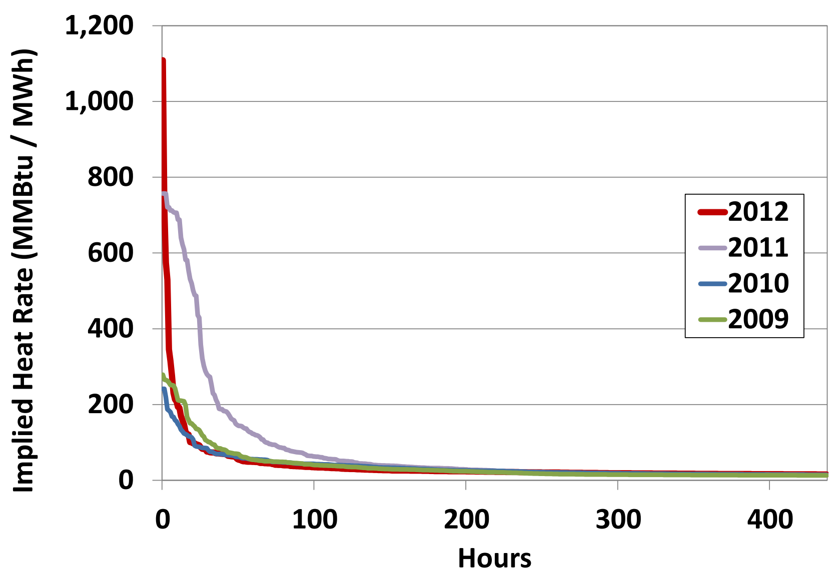
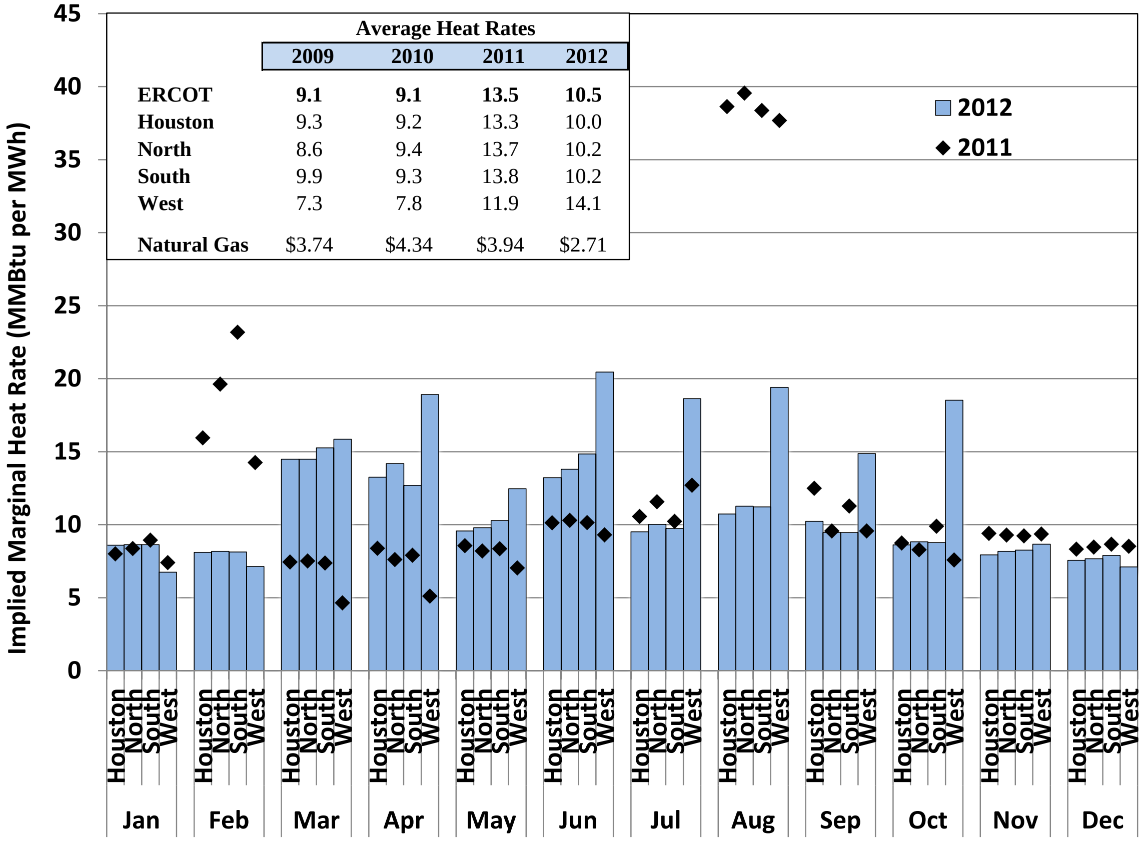


Figure 10 shows the implied marginal heat rates for the top five percent of hours in 2009 through 2012 and highlights that the implied heat rate in 2012 at the top 5 percent of hours is consistent with other years, except for 2011, where the heat rates were higher at top hours.

To further illustrate these differences, the next figure shows the implied marginal heat rates on a monthly basis in each of the ERCOT zones in 2011 and 2012, with annual average heat rate data for 2009 through 2012. This figure is the fuel price-adjusted version of Figure 2 in the prior sub-section. Adjusting for natural gas price influence, Figure 11 shows that the annual, system-wide average implied heat rate decreased in 2012 compared to 2011. However, it was still higher than 2009 and 2010.

[7] The *Implied Marginal Heat Rate* equals either the *Balancing Energy Price* (zonal) or the *Real-Time Energy Price* (nodal) divided by the *Natural Gas Price*. This methodology implicitly assumes that electricity prices move in direct proportion to changes in natural gas prices.

Figure 11: Monthly Average Implied Heat Rates

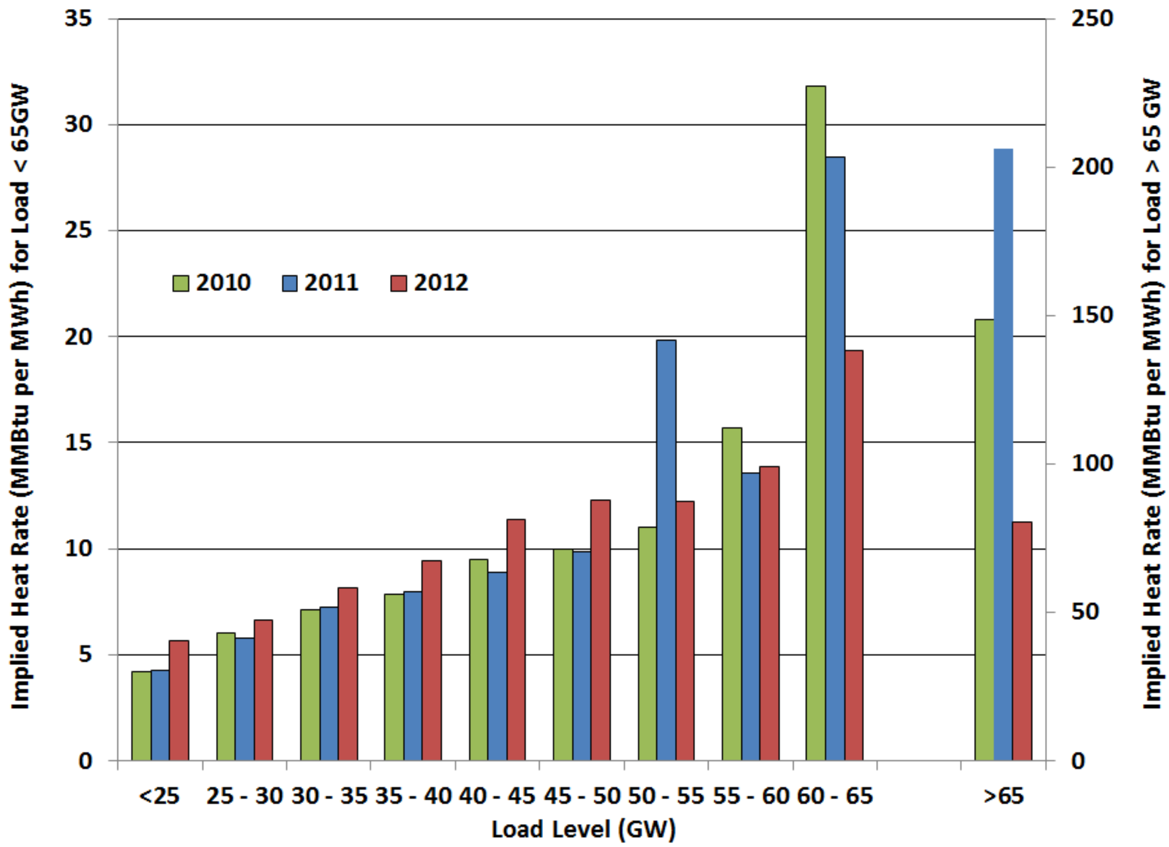


The monthly average implied heat rates in 2012 are generally consistent with 2011, with notable exceptions in February and August 2011. Higher heat rates in February can be explained by the extended period when real-time prices were $3,000 per MWh due to extreme cold weather and the resulting unplanned outages of numerous generators. Extended hot, dry weather resulted in record system peak demands in August, and another extended period of energy prices reflecting shortage conditions. The differences in the average annual implied heat rates observed at the zonal level can be attributed to the continued significant congestion related to wind generation exports from the West zone.

We conclude our examination of implied heat rates from the real-time energy market by evaluating them at various load levels. Figure 12 below, provides the average heat rate at various system load levels from 2010 through 2012.[8]

8 To appropriately compare twelve months of data under each market design, data labeled as 2010 in Figure 12 are from December 1, 2009 through November 30, 2010.

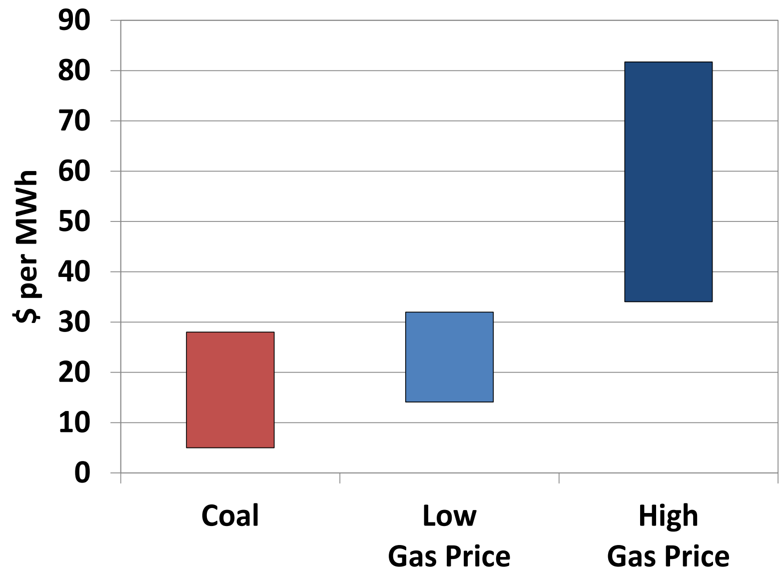
Figure 12: Heat Rate and Load Relationship



In a well performing market, a clear positive relationship between these two variables is expected since resources with higher marginal costs must be dispatched to serve higher loads. Although we do see a generally positive relationship, there is a noticeable disparity for loads between 50 and 55 GW. During the extreme cold weather event in early February 2011, loads were at this level while prices reached $3,000 per MWh for a sustained period of time. Focusing on 2012 data, we observe the desired positive relationship between load and implied heat rates.

The higher heat rates observed at lower loads in 2012 are likely due to the interplay between coal and natural gas prices because of the low natural gas prices experienced in 2012. This interaction warrants more explanation. The price of energy offered from coal units is generally very stable, due to the long term nature of contracts for both fuel and transportation. The large majority of energy offered from coal units is generally priced between $5 and $30 per MWh. The price of energy from natural gas-fired units in ERCOT is much more variable but closely tied to the price of natural gas fuel. In fact, the implied heat rate (the measure of conversion efficiency from MMBtu of fuel to MWh of electricity) of natural gas based offers has remained within a fixed range of 7.5 MMBtu per MWh to 18 MMBtu per MWh. Focusing on the past two years, natural gas prices peaked in June 2011 at $4.54 per MMBtu and declined steadily to a nadir of $1.88 per MMBtu in April 2012. At these low prices, energy offers from natural gas units competed directly with offers from coal units. Figure 13 compares the typical ranges of energy offers from coal units with those from natural gas units, under both high and low natural gas prices. As discussed later, one of the effects of this price overlap was reduced generation from coal units during 2012. At natural gas prices above $3 per MMBtu the amount of overlap between energy offers from coal and natural gas units decreases.

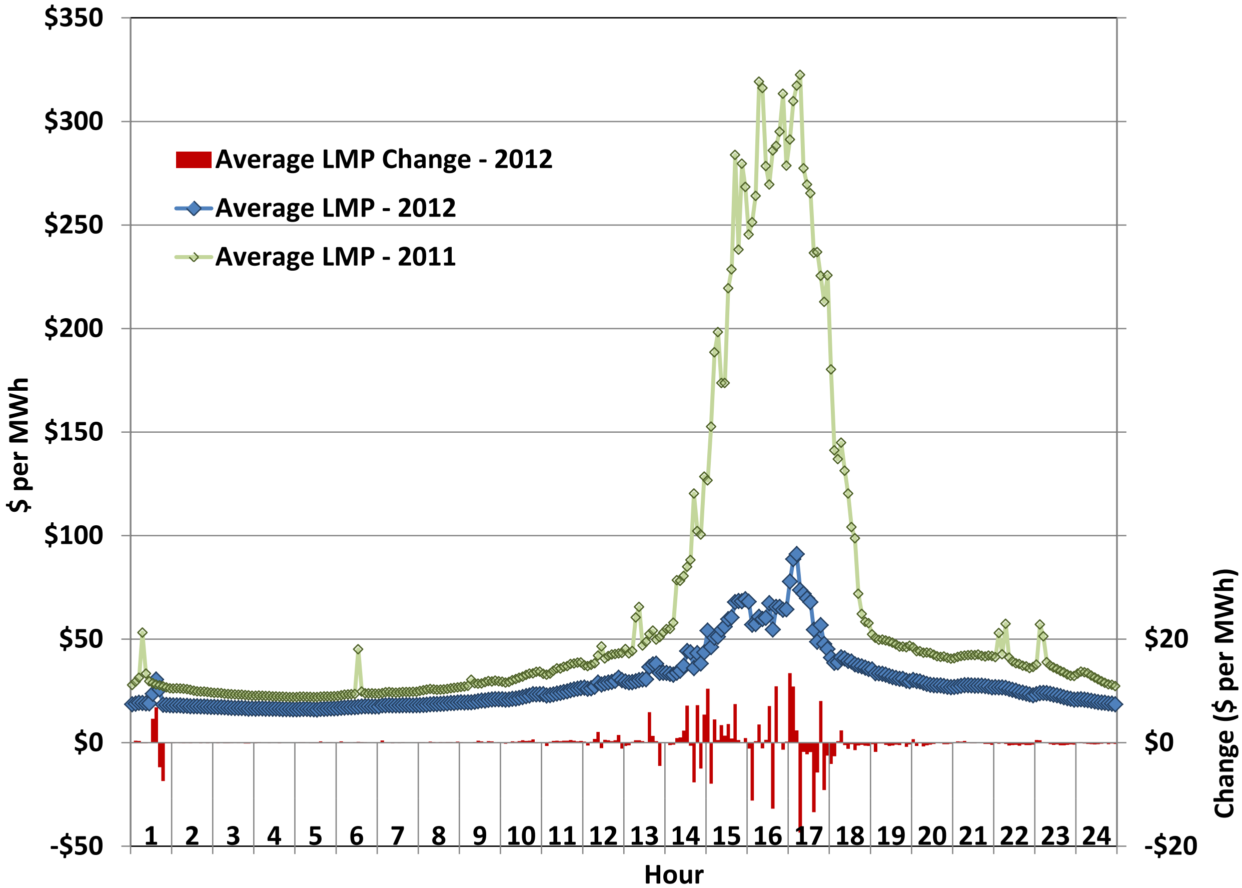
Figure 13: Typical Energy Offers



C. Real-Time Price Volatility

Volatility in real-time wholesale electricity markets is expected because system load can change rapidly and the ability for supply to adjust can be restricted by physical limitations of the resources and the transmission network. Figure 14 below presents a view of the price volatility experienced in ERCOT's real-time energy market during the summer months of May through August. Average five-minute real-time energy prices are presented along with the magnitude of change in price for each five-minute interval. Average real-time energy prices from the same period in 2011 are also presented. Comparing average real-time energy prices for 2012 with those from 2011, the effects of lower natural gas prices on average prices during non-peak hours and the effects of fewer shortage intervals during peak hours are observed. Outside of the hours from 15 to 18 (2:00 pm to 6:00 pm), short-term increases in average real-time energy prices are typically due to high prices resulting from generator ramp rate limitations occurring at times when significant amounts of generation is changing its online status. Factoring current market conditions into generators' daily operational decisions about the specific timing of startup and shutdown may have led to the reduced the effects of this type of ramp rate limitation on price spikes during 2012. The average of the absolute value of changes in five-minute real-time energy prices, expressed as a percentage of average price was less than 4 percent in 2012 compared to approximately 6 percent for the same period in 2011.

Figure 14: Real-Time Energy Price Volatility (May – August)



Reduced price volatility in 2012 is also observed in 15 minute settlement point prices for the four geographic load zones, as shown below in Table 1.

Table 1: 15 Minute Price Changes as a Percent of Annual Average Price

Load Zone	*2011*	*2012*
Houston	21.4%	13.0%
South	19.9	13.1
North	22.5	13.9
West	26.2	19.4

The table shows that the price volatility fell substantially from 2011 to 2012. This was primarily due to the sharply reduction in shortages that occurred in 2012, which exhibit relatively normal summer weather conditions. In contrast, 2011 exhibited the hottest summer temperatures in more than 100 years, leading to frequent shortages and associated higher price volatility. The table also shows that price volatility in the West zone has continued to be higher than in the other zones, which is expected given the very high penetration of variable output wind generation located in that area.

D. Prices at the System-Wide Offer Cap

After the extremes of 2011, weather conditions in Texas returned to closer to normal in 2012. As more fully discussed in Section IV Load and Generation, overall demand for electricity was lower in 2012 than in 2011, resulting in much fewer occasions when the available supply generation capacity was unable to meet customer demands. This resulted in a decreased likelihood that the available generation capacity was not sufficient to meet customer demands for electricity and maintain the required reliability reserves.

As more fully described later in Section V, Resource Adequacy, the nodal market causes energy prices to rise toward the system-wide offer cap as available operating reserves approach minimum required levels to reflect the degradation in system reliability.

Figure 15: Duration of Prices at the System-Wide Offer Cap

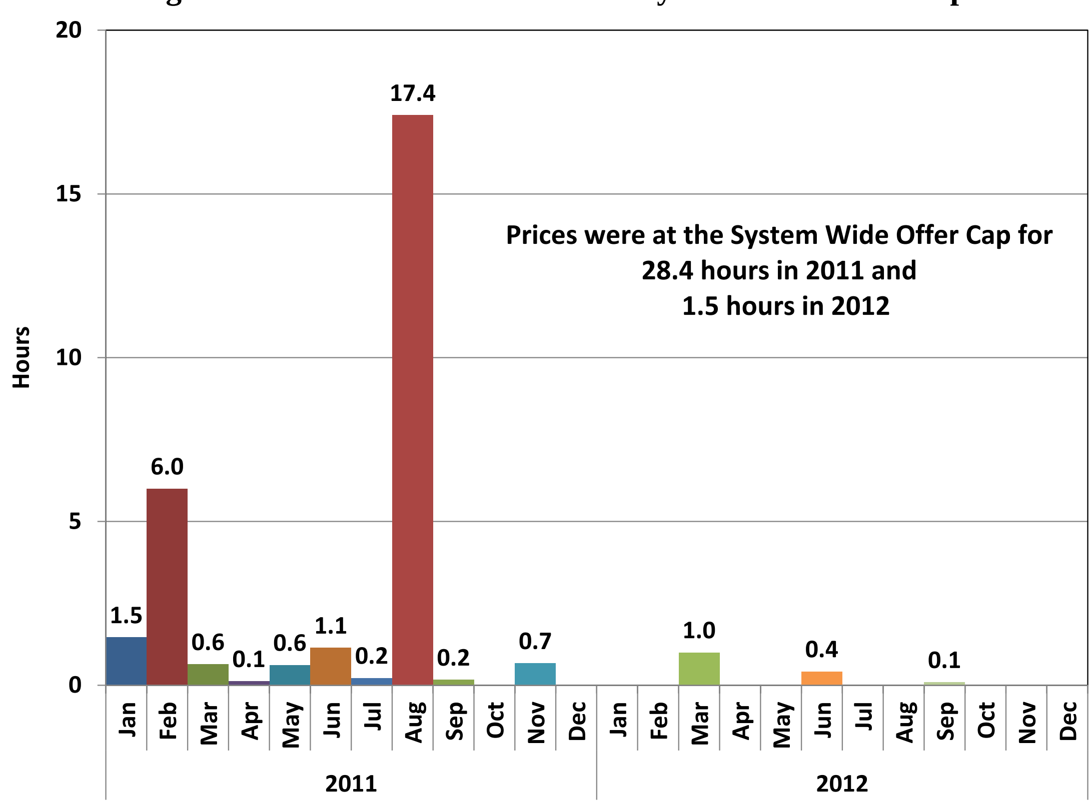
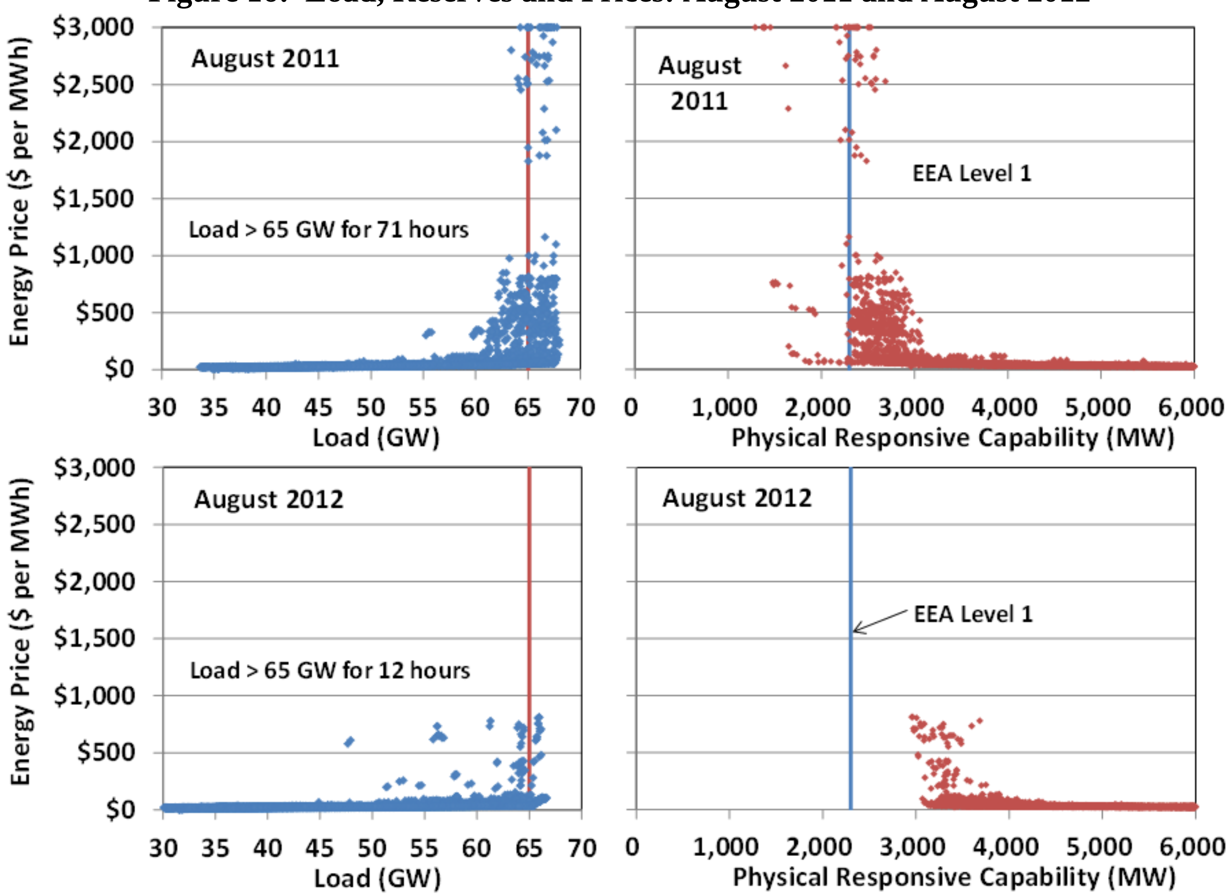


Figure 15 above shows the aggregated amount of time represented by all five-minute dispatch intervals where the real-time energy price was at the system-wide offer cap, displayed by month. Prices during 2012 prices were at the system-wide offer cap for only 1.5 hours, a significant reduction from the 28.4 hours experienced in 2011. Approved during 2012, PUCT SUBST. R.25.508 increased the system-wide offer cap to $4,500 per MWh effective August 1, 2012. As shown in Figure 15 above, there was only a brief period when energy prices rose to the cap after this change was implemented

The next figure provides a detailed comparison of August's load, required reserve levels, and prices for 2011 and 2012. As expected, the weather in ERCOT was extremely hot and dry during both months, but there were very few dispatch intervals when real-time energy prices reached the system-wide offer cap in 2012 compared to the relatively high frequency it occurred in 2011. Although the weather may have been similar, there were significant differences in load and available operating reserve levels, resulting in much higher prices in August 2011.

Figure 16: Load, Reserves and Prices: August 2011 and August 2012



Shown on the left side of Figure 16 is the relationship between real-time energy price and load level for each dispatch interval for the month of August in 2011 (top) and 2012 (bottom). ERCOT loads were greater than 65 GW for 71 hours in 2011, whereas loads reached that level for less than 12 hours during August 2012. As previously discussed, a strong positive correlation between higher load and higher prices is expected in a well-functioning energy market. We observe such a relationship between higher prices and higher loads in both months.

Although load levels are strong predictors of energy prices, an even more important predictor is the level of operating reserves. Simply put, operating reserves are the difference between the total capacity of operating resources and the current load level. As load level increases against a fixed quantity of operating capacity, the amount of operating reserves diminishes. The minimum required operating reserves prior to the declaration of Energy Emergency Alert ("EEA") Level 1 by ERCOT is 2,300 MW. As the available operating reserves approach the minimum required amount, energy prices should rise toward the system-wide offer cap to reflect the degradation in system reliability.

On the right side of Figure 16 are data showing the relationship between real-time energy prices and the quantity of available operating reserves for each dispatch interval during August of 2011 (top) and 2012 (bottom). This figure shows a strong correlation between diminishing operating reserves and rising prices. With the lower loads in August 2012, available operating reserves were well above minimum levels for the entire month, and there were no occurrences where the energy price reached the system-wide offer cap. In contrast, there were numerous dispatch intervals in August 2011 when the minimum operating reserve level was approached or breached, with 17.4 hours where prices reached $3,000 per MWh.[9] It should be noted that during August 2011 there were a number of dispatch intervals where operating reserves were below minimum requirements with prices well below the level that would be reflective of the reduced state of reliability. In Section IV, Load and Generation at page 93, we provide an example explaining why this can occur and offer a recommendation for improvement.

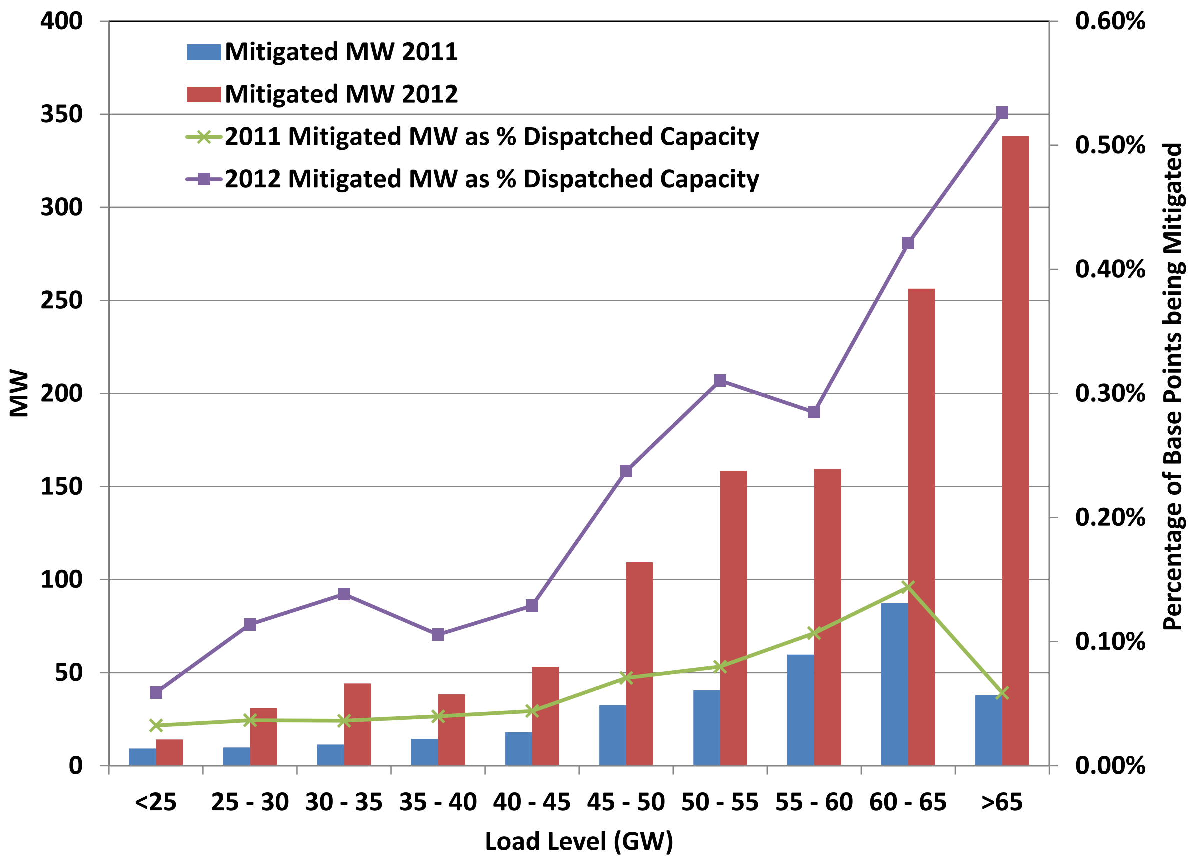
[9] The system-wide offer cap during August 2011 was set at $3,000 per MWh. It was increased to $4,500 per MWh effective August 1, 2012.

E. Mitigation

The 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 only considering 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 specific generator to raise prices in the event of a transmission constraint that requires their output to resolve. In this section we analyze the quantity of capacity affected by this mitigation process.

Our first analysis computes how much capacity, on average, is actually mitigated during each dispatch interval. The results, shown in Figure 17, are provided by load level.

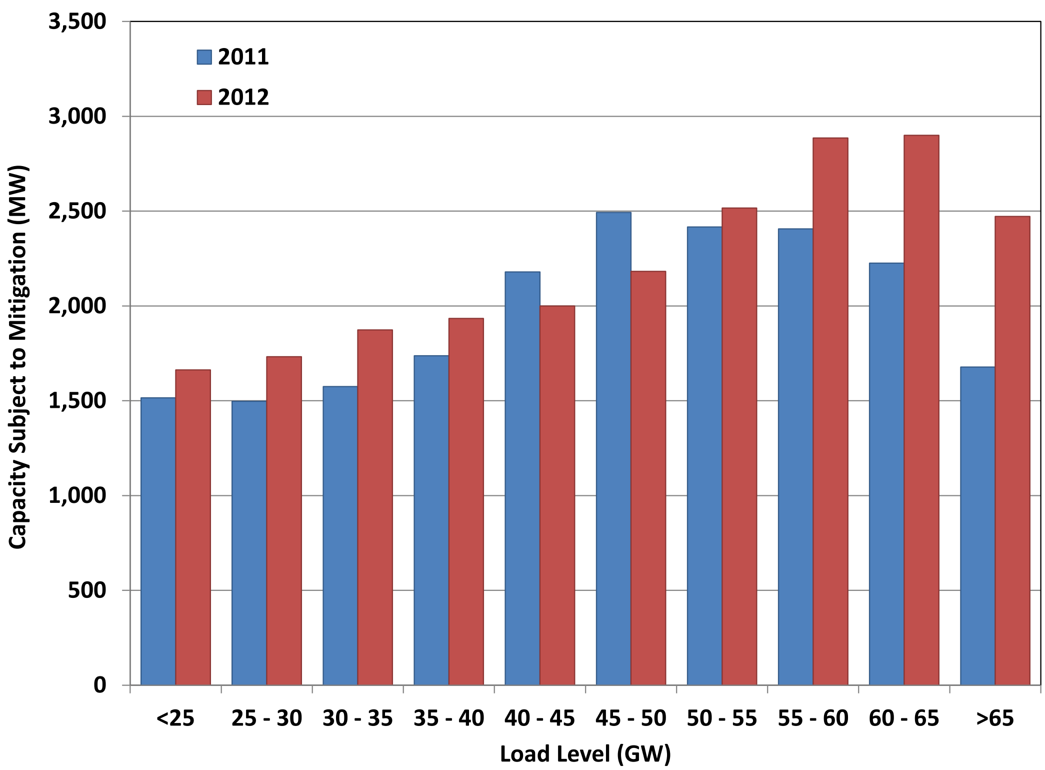
Figure 17: Mitigated Capacity by Load Level



The quantities of capacity actually mitigated in 2012 were much larger than during 2011, averaging 14 MW at low loads and increasing to 338 MW at loads above 65 GW. Although the quantities of mitigated capacity were greater in 2012, they were less than one-half of one percent of the total dispatched capacity at all but the very highest load levels. The decrease in mitigated capacity at high loads observed in 2011 due to the reluctance by ERCOT operators to activate certain transmission constraints during very high system load conditions is not present in 2012.

In the previous figure only the amount of capacity that can be dispatched within one interval is counted as mitigated. In our next analysis we compute the total capacity subject to mitigation. These values are determined by comparing a generator's mitigated and unmitigated (as submitted) offer curves and determining the point at which they diverge. We then take the difference between the total unit capacity and the capacity at the point the curves diverge. This calculation is performed for all units and aggregated by load level, as shown in Figure 18. From this figure we observe that at most 7 percent of capacity necessary to serve load is subject to mitigation. An important note about this capacity measure is that it includes all capacity above the point at which a unit's offers become mitigated, without regard for whether that capacity is actually required to serve load.

Figure 18: Capacity Subject to Mitigation



Although executing all the time, the automatic price mitigation aspect of the two step dispatch process only has an effect when a non-competitive transmission constraint is active. This can result in mitigating certain units inappropriately. The mitigation process is intended to limit the ability of a generator to affect price when their output is required to manage congestion. The process does not currently identify a situation where there are a competitively sufficient number of generators on the other side of the constraint and mitigates all their offers. This unnecessary mitigation will be addressed with the implementation of changes described in NPRR520. This change will introduce an impact test to determine whether units are relieving or contributing to a transmission constraint, and only subject the relieving units to mitigation.

II. REVIEW OF DAY-AHEAD MARKET OUTCOMES

ERCOT's centralized day-ahead market allows participants to make financially binding forward purchases and sales of power for delivery in real-time. Offers to sell can take the form of either a three-part supply offer, which allow sellers to reflect the unique financial and operational characteristics of a specific generation resource, or an energy only offer, which is location specific but is not associated with a generation resource. Bids to buy are also location specific. Ancillary services are also procured as part of the day-ahead market clearing. The third type of transaction included in the day-ahead market is bids to buy Point to Point ("PTP") Obligations, which allow parties to hedge the incremental cost of congestion between day-ahead and real-time operations.

With the exception of the acquisition of ancillary service capacity, the day-ahead market is a financial market. Although all bids and offers are evaluated in the context of their ability to reliably flow on the transmission network, there are no operational obligations resulting from the day-ahead market clearing. These transactions are made for a variety of reasons, including satisfying the participant's own supply, managing risk by hedging the participant's exposure to the real-time market, or arbitraging with the real-time markets. For example, load serving entities can insure against volatility in the real-time market by purchasing in the day-ahead market. Finally, the day-ahead market plays a critical role in coordinating generator commitments. For all of these reasons, the performance of the day-ahead market is essential.

In this section we review energy pricing outcomes from the day-ahead market and compare their convergence with real-time energy prices. We will also review the volume of activity in the day-ahead market, including a discussion of PTP Obligations. We conclude this section with a review of the ancillary service markets.

A. Day-Ahead Market Prices

One indicator of market performance is the extent to which forward and real-time spot prices converge over time. Forward prices will converge with real-time prices when two main conditions are in place: a) there are low barriers to shifting purchases and sales between the forward and real-time markets; and b) sufficient information is available to market participants to allow them to develop accurate expectations of future real-time prices. When these conditions are met, market participants can be expected to arbitrage predictable differences between forward prices and real-time spot prices by increasing net purchases in the lower-priced market and increasing net sales in the higher-priced market. These actions will tend to improve the convergence of forward and real-time prices.

In this section, we evaluate the price convergence between the day-ahead and real-time markets. This analysis reveals whether persistent and predictable differences exist between day-ahead and real-time prices, which participants should arbitrage over the long-term. To measure the short-term deviations between real-time and day-ahead prices, we also calculate the average of the absolute value of the difference between the day-ahead and real-time price on a daily basis.

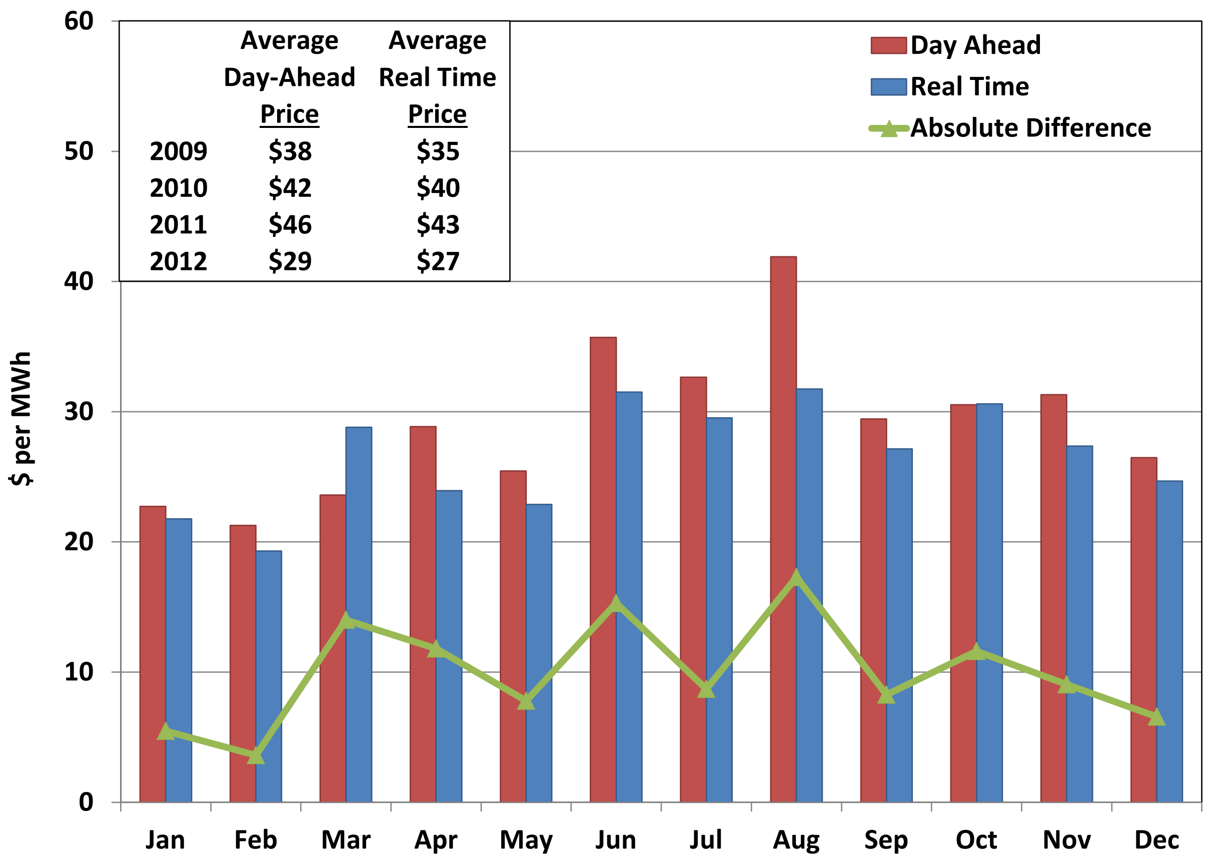
This measure captures the volatility of the daily price differences, which may be large even if the day-ahead and real-time energy prices are the same on average. For instance, if day-ahead prices are $70 per MWh on two consecutive days while real-time prices are $40 per MWh and $100 per MWh on the two days, the absolute price difference between the day-ahead market and the real-time market would be $30 per MWh on both days, while the difference in average prices would be $0 per MWh.

Figure 19 shows the price convergence between the day-ahead and real-time market, summarized by month. Day-ahead prices averaged $29 per MWh in 2012 compared to an average of $27 per MWh for real-time prices.[10] The average absolute difference between day-ahead and real-time prices was $9.96 per MWh in 2012; much lower than in 2011 when average of the absolute difference was $24.50 per MWh. This reduction was due to fewer occurrences of shortage intervals and associated high prices in 2012. This day-ahead premium is consistent with expectations due to the much higher volatility of real-time prices. Risk is lower for loads purchasing in the day-ahead and higher for generators selling day-ahead. The higher risk for generators is associated with the potential of incurring a forced outage and as a result, having to buy back energy at real-time prices. This explains why the highest premiums tend to occur during the months with the highest demand and highest prices. Overall, the day-ahead premiums

[10] These values are simple averages, rather than load-weighted averages presented in Figure 1and Figure 2.

were very similar to the differences observed in 2010 and 2011, but remain higher than observed in other organized electricity markets.[11] Real-time energy prices in ERCOT are allowed to rise to levels that are much higher than what is allowed in other organized electricity markets, which increases risk and helps to explain the higher day ahead premiums regularly observed in ERCOT Although most months experienced a day-ahead premium (*e.g.*, $10 per MWh in August), it should not be expected over time that every month will always produce a day-ahead premium as the real-time risks that lead to the premiums will materialize unexpectedly on occasion, resulting in real-time prices that exceed day-ahead prices (*e.g.*, in March).

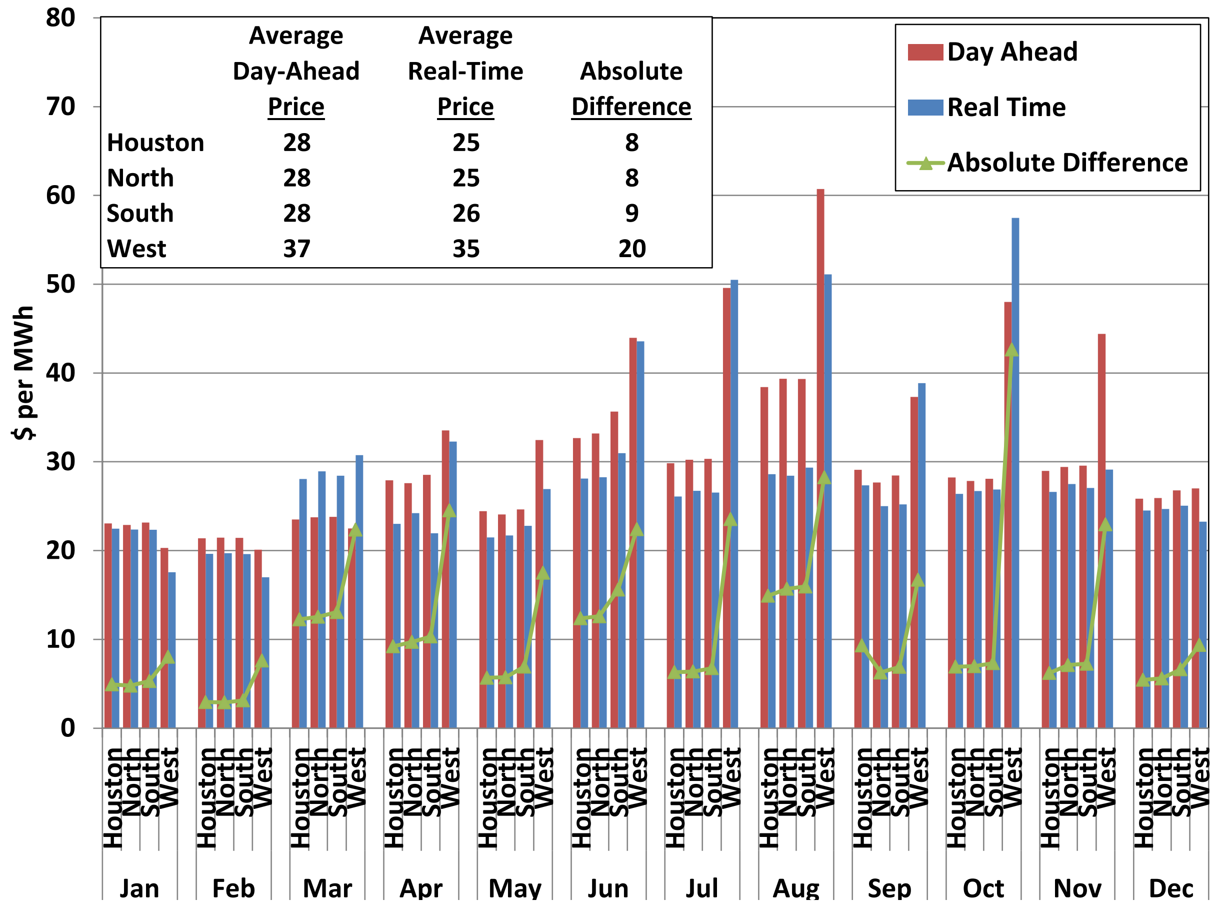
Figure 19: Convergence between Forward and Real-Time Energy Prices



[11] In 2009 and 2010 under the zonal market the comparison was made between on-peak forward prices and prices for the same on-peak period in the balancing energy market.

In Figure 20 below, monthly day-ahead and real-time prices are shown for each of the geographic load zones. Of note is the difference in the West zone data compared to the other regions. The higher volatility in West zone pricing is likely associated with the uncertainty of forecasting wind generation output and the resulting price differences between day-ahead and real-time.

Figure 20: Day-Ahead and Real-Time Prices by Zone



B. Day-Ahead Market Volumes

Our next analysis summarizes the volume of day-ahead market activity by month. In Figure 21 below, we find that day-ahead purchases are approximately 45 percent of real-time load. These energy purchases are met through a combination of generator specific and virtual offers.

As discussed in more detail in the next sub-section, Point to Point Obligations are financial instruments purchased in the day-ahead market. They do not provide any energy supply themselves, but they do provide the ability to avoid the congestion costs associated with transferring the delivery of energy from one location to another. To provide a volume comparison we aggregate all of these "transfers", netting location specific injections against withdrawals. By adding the aggregated transfer capacity associated with purchases of PTP Obligations, we find that on average, total volumes transacted in the day-ahead market are greater than real-time load.

Figure 21: Volume of Day-Ahead Market Activity by Month

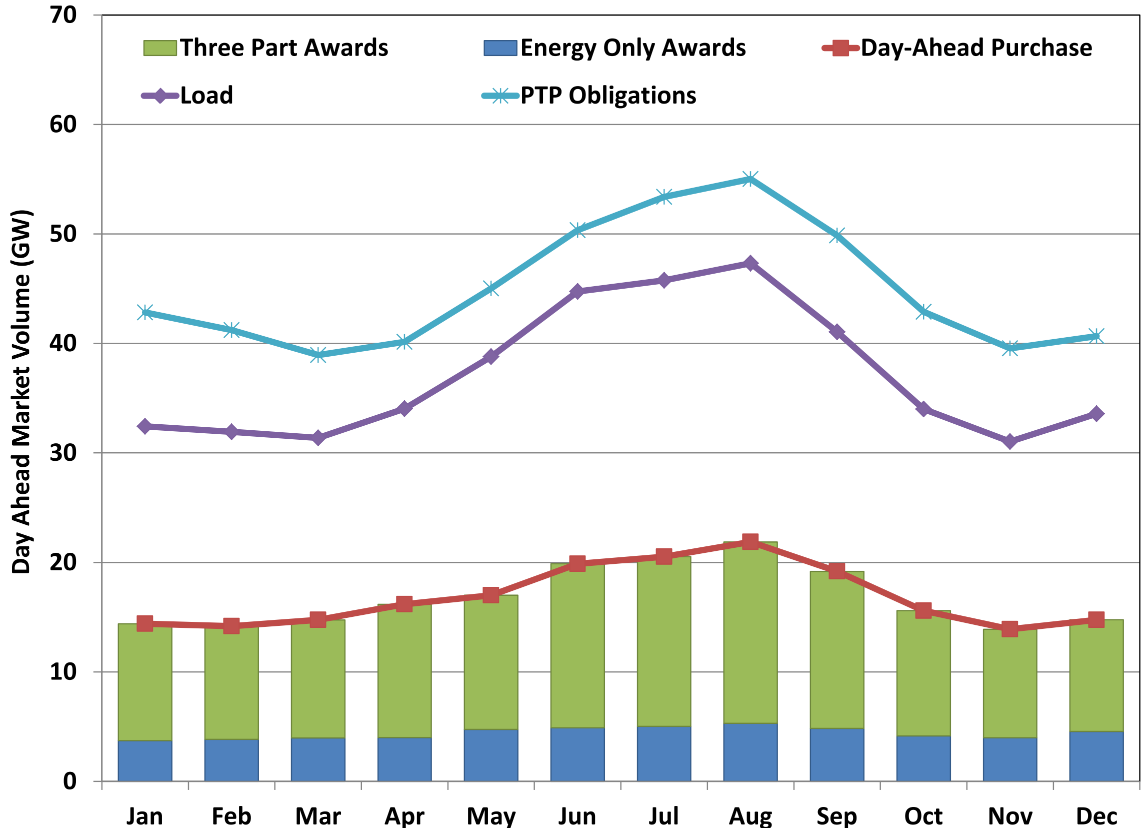
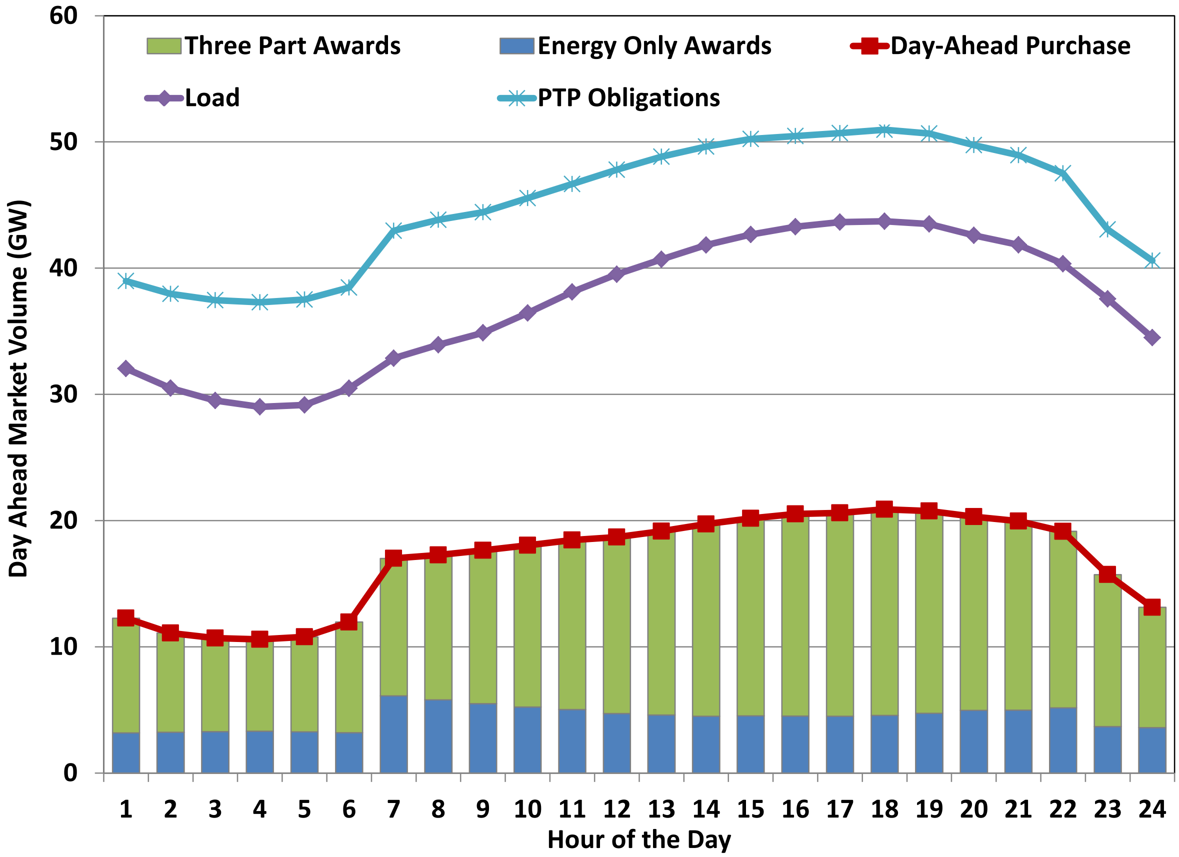


Figure 22 below, presents the same day-ahead market activity data summarized by hour of the day. In this figure the volume of day-ahead market transactions is disproportionate with load levels between the hours of 6 and 22. Since these times align with common bilateral transaction terms, it appears that market participants are using the day-ahead market to trade around those positions.

Figure 22: Volume of Day-Ahead Market Activity by Hour



C. Point to Point Obligations

Purchases of Point to Point ("PTP) Obligations comprise a significant portion of day-ahead market activity. These instruments are similar to, and can be used to complement Congestion Revenue Rights (CRRs). CRRs, as more fully described in Section III, are acquired via monthly and annual auctions and allocations. CRRs accrue value to their owner based on locational price differences as determined by the day-ahead market. PTP Obligations are instruments that are purchased as part of the day-ahead market and accrue value to their owner based on real-time locational price differences. By acquiring a PTP Obligation using the proceeds due to them from their CRR holding, a market participant may be described as rolling their hedge to real-time.

In this subsection we provide additional details about the volume and profitability of these PTP Obligations.

Figure 23: Point to Point Obligation Volume

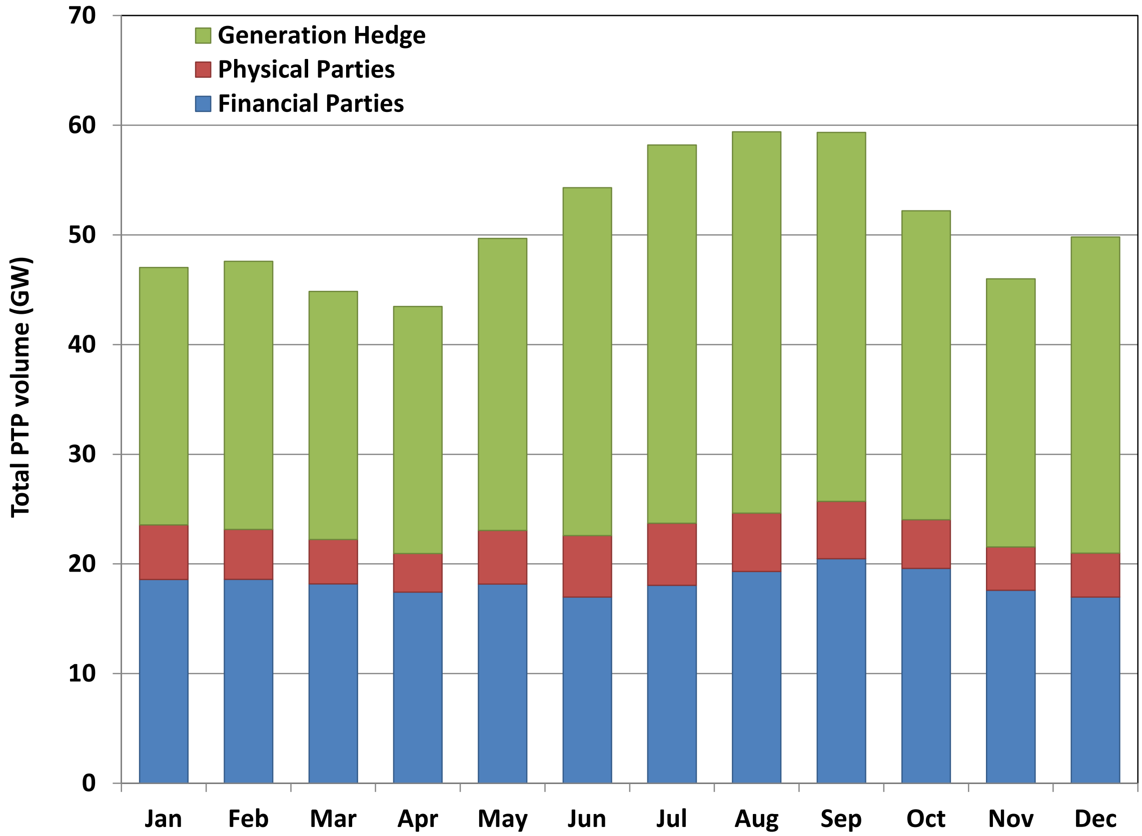
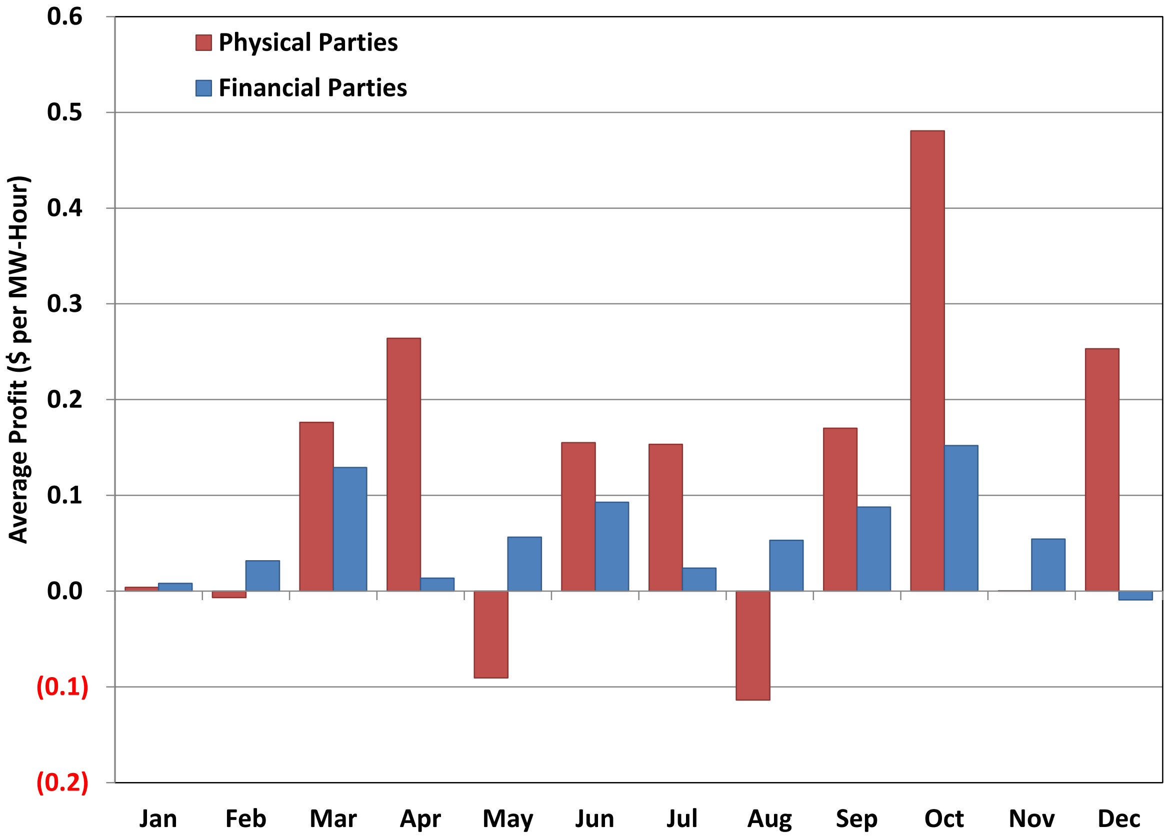


Figure 23 presents the total volume of PTP Obligation purchases divided into three categories. Compared to the previous two figures which showed net flows associated with PTP Obligations, in this figure we examine the total volume. For all PTP Obligations that source at a generator location, we attribute capacity up to the actual generator output as a generator hedge. From the figure above we see that this comprised most of the volume of PTP Obligations purchased. The remaining volumes of PTP Obligations are not directly linked to a physical position and are assumed to be purchased primarily to profit from anticipated price differences between two locations. We further separate this arbitrage activity by type of market participant. Physical parties are those that have actual real-time load or generation, whereas financial parties have neither.

To the extent the price difference between the source and sink of a PTP Obligation is greater in real-time than it was for the day-ahead price, the owner will have made a profit. Conversely, if the price difference does not materialize in real-time, the PTP Obligation may be unprofitable.

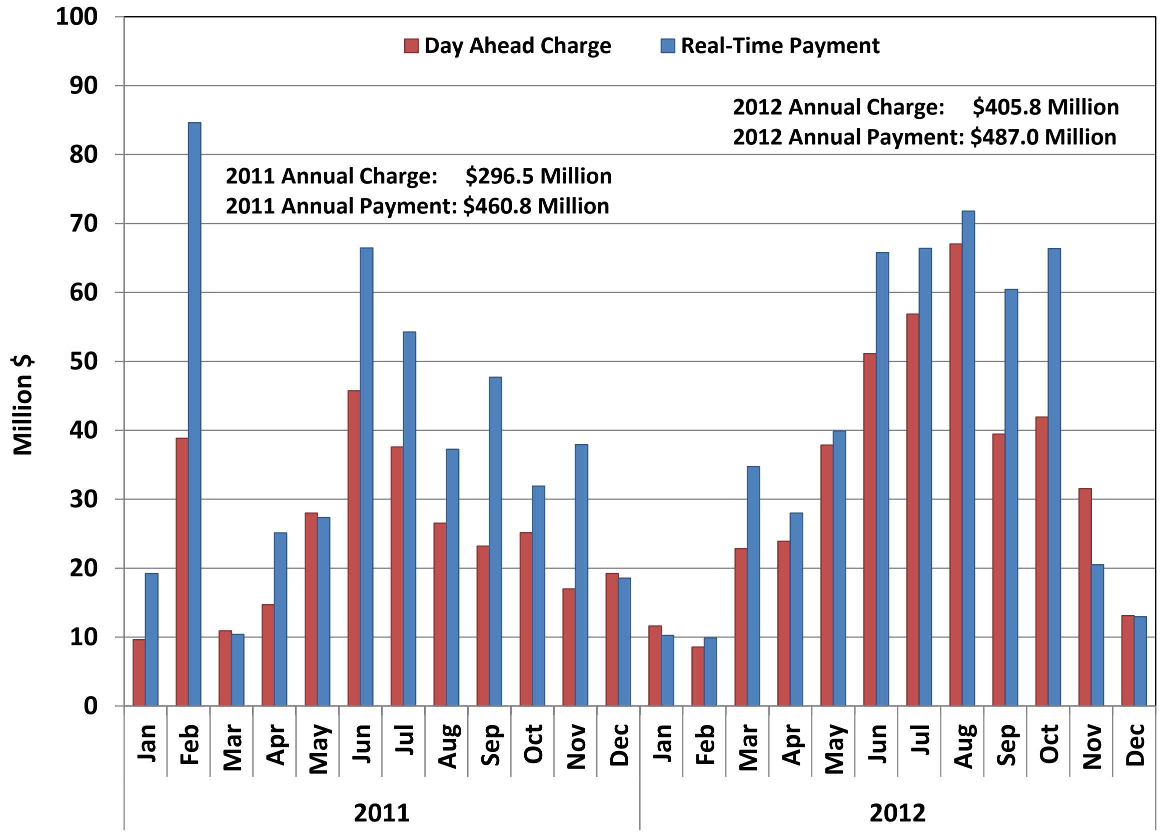
We compare the profitability of PTP Obligation holdings by the two types of participants in Figure 24.

Figure 24: Average Profitability of Point to Point Obligations



From the figure above we can infer different motivations between the two types of participants. Because financial participants have no real-time load or generation they have no other exposure to real-time prices. If a financial participant is not making a profit on their PTP Obligations there is no reason for them to buy any. In fact, their profit seeking action of buying PTP Obligations between points where congestion is expected helps make the day-ahead market converge with real-time market outcomes. On the other hand, physical participants do have exposure to real-time prices. It is reasonable to expect that this type of participant is most interested in limiting that exposure by using PTP Obligations as a hedge.

Figure 25: Point to Point Obligation Charges and Payments

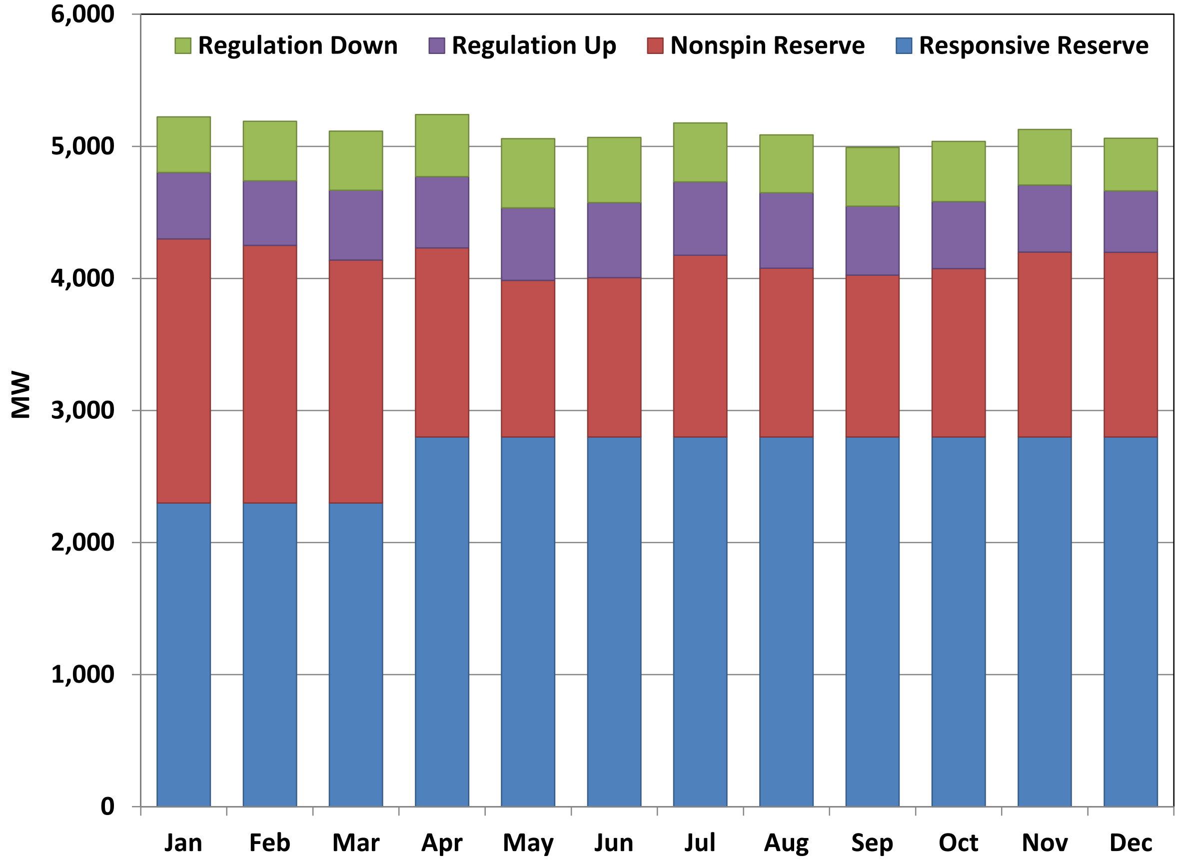


To conclude our analysis of PTP Obligations, in Figure 25 we compare the total amount paid for these instruments day-ahead, with the total amount received by their holders in real-time. In most months owners received, in aggregate, more in payments for their PTP Obligations than they paid to acquire them. This occurs when real-time congestion is greater than what occurred in the day-ahead. The payments made to PTP Obligation owners come from real-time congestion rent. We assess the sufficiency of real-time congestion rent to cover both PTP Obligations and payments to owners of CRRs who elect to receive payment based on real-time prices in Section III, Transmission and Congestion at page 52.

D. Ancillary Services Market

The primary ancillary services are up regulation, down regulation, responsive reserves, and non-spinning reserves. Market participants may self-schedule ancillary services or purchase their required ancillary services through the ERCOT markets. This section reviews the results of the ancillary services markets in 2012. We start with a display of the quantities of each ancillary service procured each month shown in Figure 26.

Figure 26: Ancillary Service Capacity

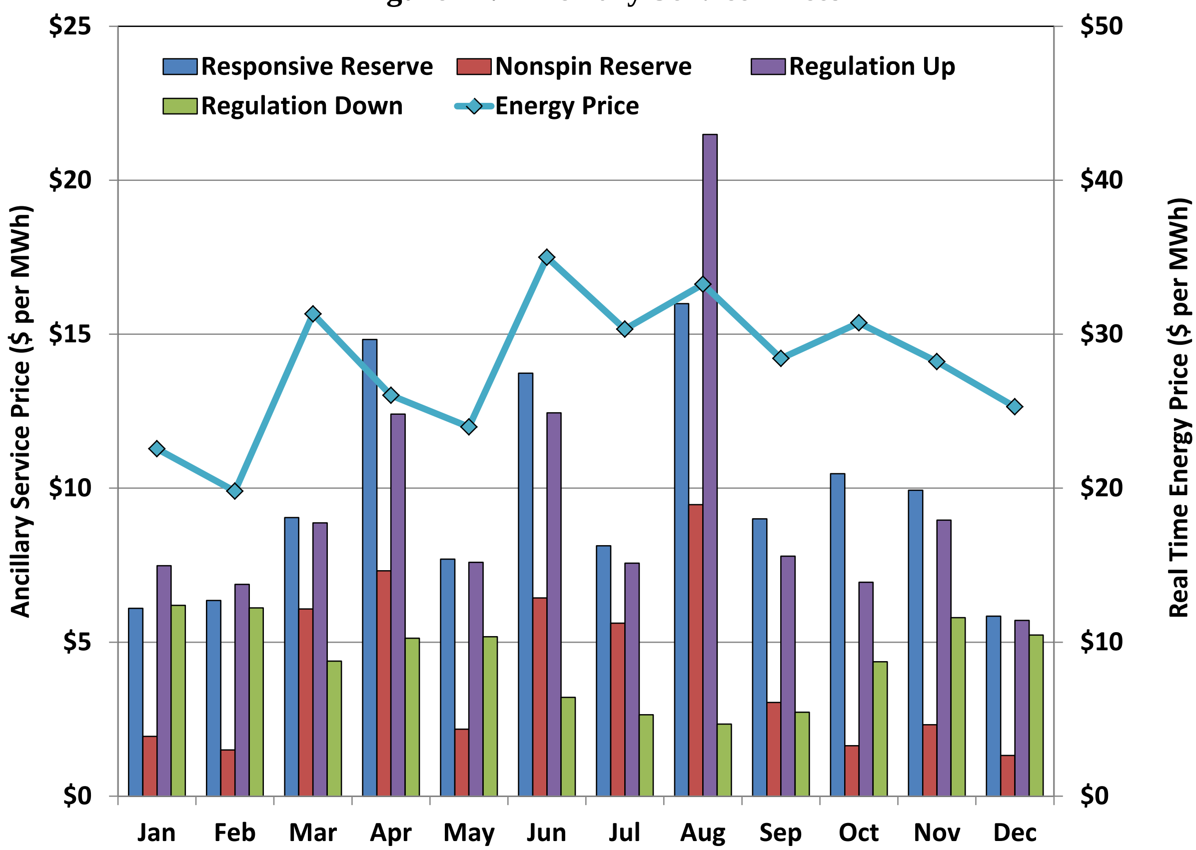


In general, the purpose of responsive and non-spinning reserves is to protect the system against unforeseen contingencies (*e.g.*, unplanned generator outages, load forecast error, wind forecast error), rather than for meeting normal load fluctuations. ERCOT procures responsive reserves to ensure that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

One notable change made to ancillary service procurements during 2012 was the increased amount of responsive reserve procured beginning in April. This 500 MW increase was balanced with the same amount of decrease in the amount of non-spinning reserves procured. Although the minimum level of required responsive reserve remains at 2,300 MW, having the additional 500 MW of responsive reserve provides a higher quality – that is, faster responding capacity available to react to sudden changes in system conditions. As the amount of wind generation in ERCOT continue to increase, this additional responsive reserve should contribute to improved system reliability.

Under the nodal market, ancillary service offers are co-optimized as part of the day-ahead market clearing. This means that market participants no longer have to include their expectations of forgone energy sales in their ancillary services capacity offers. However, because clearing prices for ancillary services capacity will explicitly account for the value of energy, there is a much higher correlation between ancillary services prices and real-time energy prices.

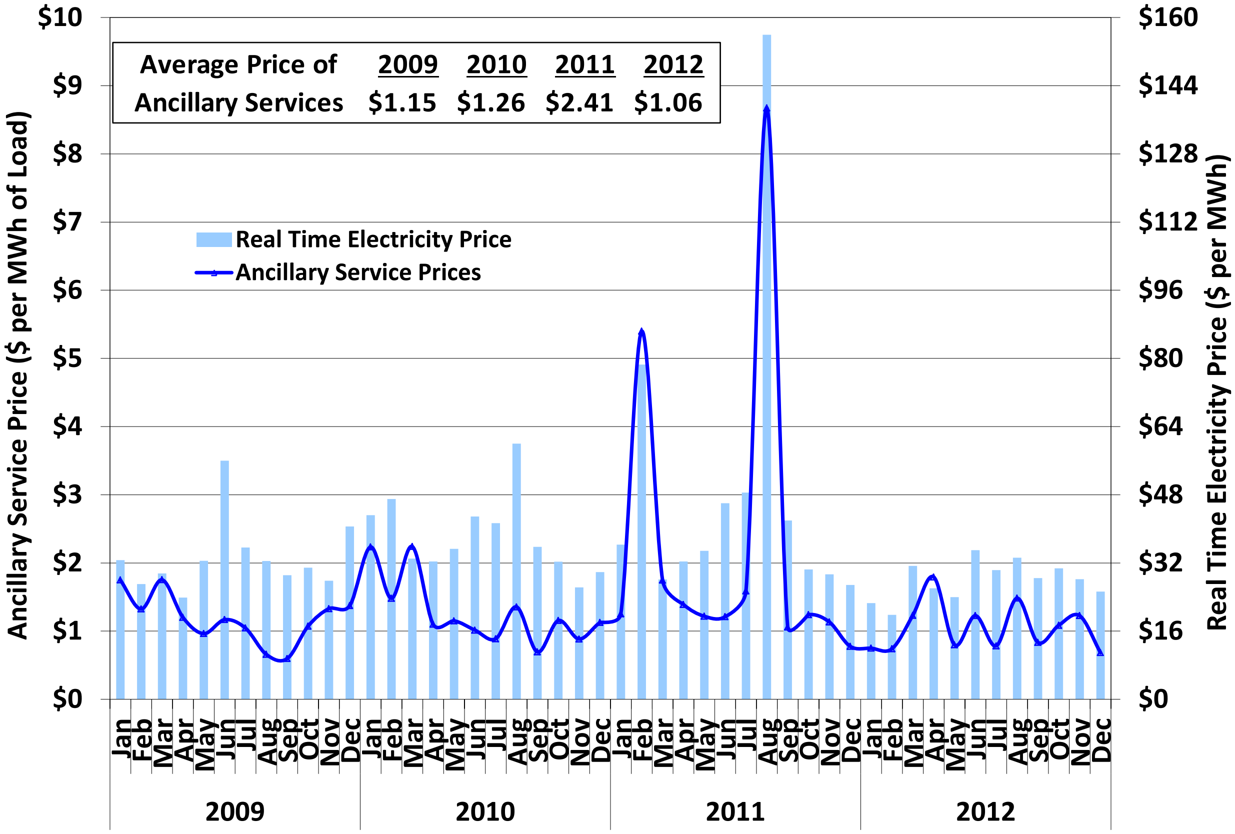
Figure 27: Ancillary Service Prices



With average energy prices varying only between $20 and $35 per MWh, we observe the prices of ancillary services remaining fairly stable throughout the year. However, the average price of responsive reserve was higher in April than in March, even though average energy price declined for the same period. We attribute this pricing outcome to the increased responsive reserve requirement being implemented in April.

In contrast to the previous data that showed the individual ancillary service capacity prices, Figure 28 shows the monthly total ancillary service costs per MWh of ERCOT load and the average real-time energy price for 2009 through 2012. This figure shows that total ancillary service costs are generally correlated with real-time energy price movements, which, as previously discussed, are highly correlated with natural gas price movements.

Figure 28: Ancillary Service Costs per MWh of Load

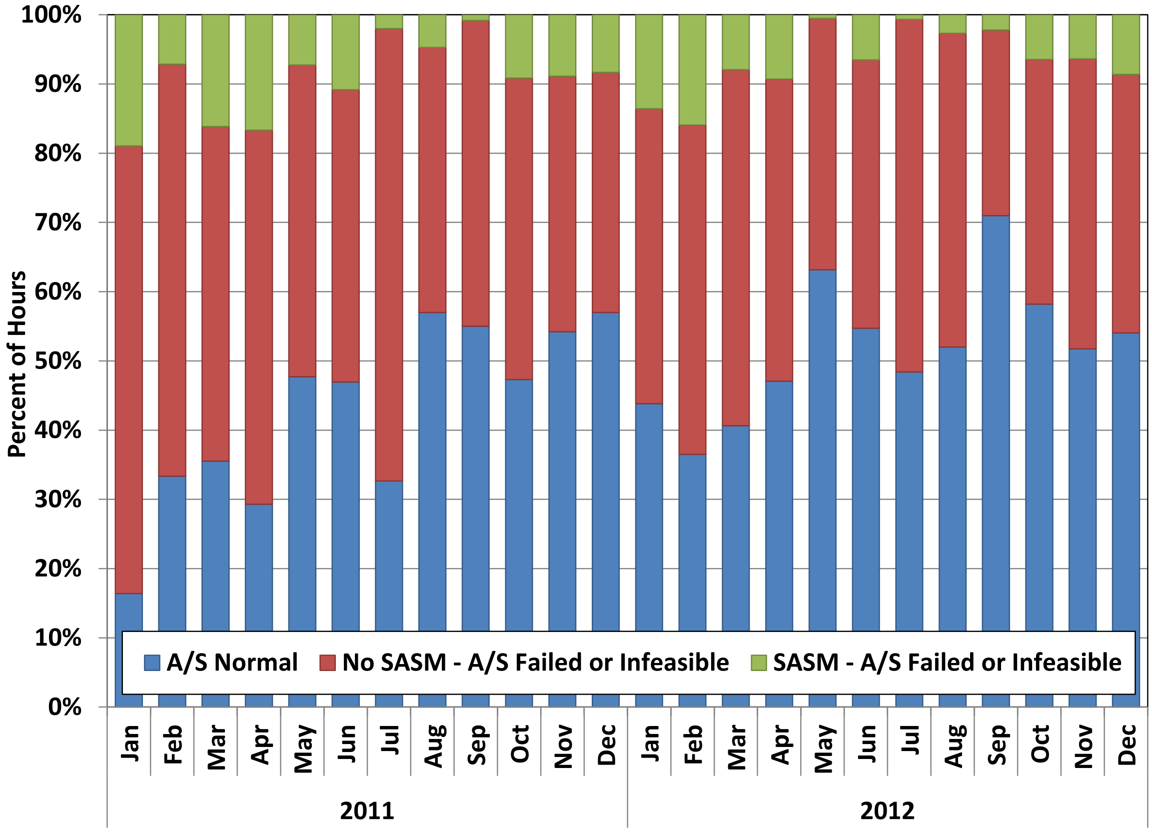


The average ancillary service cost per MWh of load decreased to $1.06 per MWh in 2012 compared to $2.41 per MWh in 2011, a decrease of 56 percent. Total ancillary service costs decreased from 4.5 percent of the load-weighted average energy price in 2011 to 3.7 percent in 2012.

Ancillary Service capacity is procured as part of the day-ahead market clearing. Between the time it is procured and the time that it is required to be provided events can occur which make this capacity unavailable to the system. Transmission constraints can arise which make the capacity undeliverable. Outages or limitations at a generating unit can lead to failures to provide. When either of these situations occurs, ERCOT may open a supplemental ancillary services market ("SASM") to procure replacement capacity.[12]

Figure 29 presents a summary of the frequency with which ancillary service capacity was not able to be provided and the number of times that a SASM was opened.

Figure 29: Frequency of SASM Clearing



The percent of time that capacity procured in the day-ahead was actually able to provide the service in the hour it was procured for continued to increase in 2012. This percentage was 52 percent in 2012 compared to 43 percent in 2011. Even though in more than 40 percent of the hours there were deficiencies in ancillary service deliveries, SASMs were opened to procure replacement capacity only 7 percent of the time, down from 9 percent of the hours in 2011.

12 ERCOT may also open a SASM if they change their ancillary service plan. This did not occur during 2012.

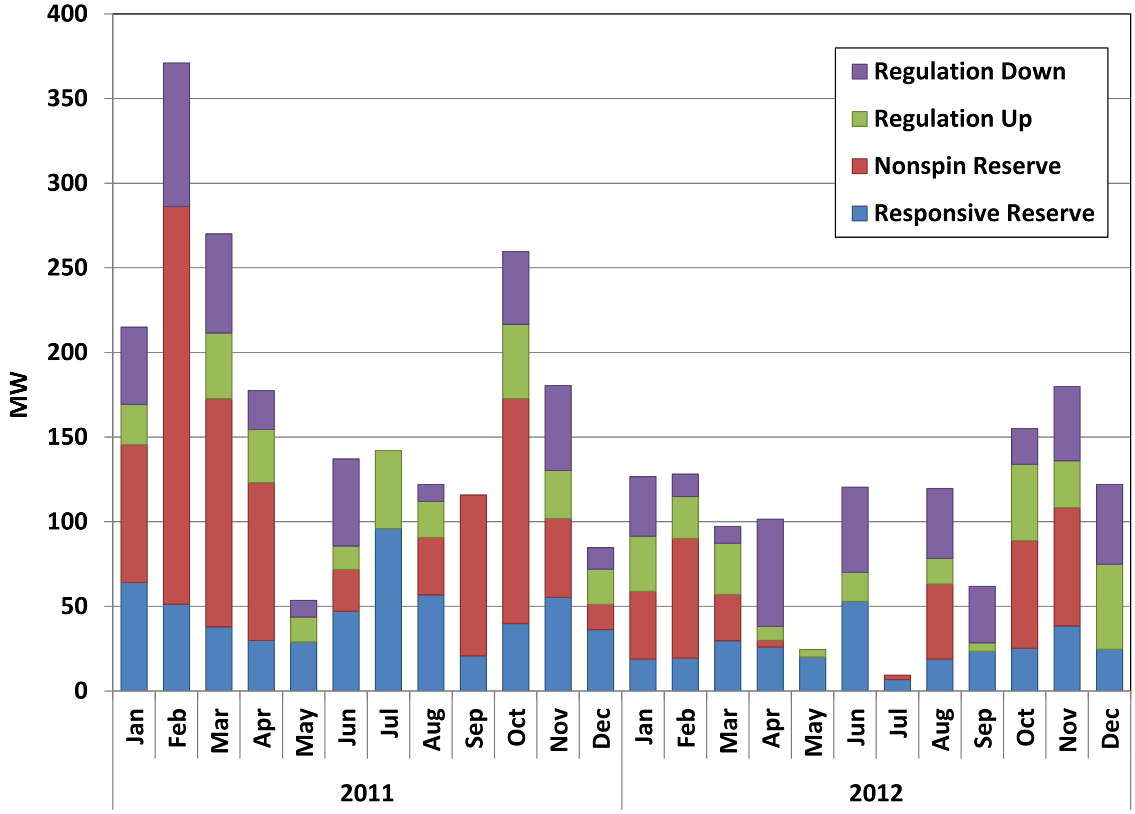
In Table 2 below, we provide an annual summary of the frequency and quantity of ancillary service deficiency, which is defined as either failure-to-provide or as undeliverable.

Table 2: Ancillary Service Deficiency

2012 *Service*	*Hours Deficient*	*Mean Deficiency (MW)*	*Median Deficiency (MW)*
Responsive Reserve	3756	34	15
Non-Spin Reserve	664	36	8
Up Regulation	750	41	25
Down Regulation	522	48	39
2011			
Responsive Reserve	4053	39	20
Non-Spin Reserve	1254	90	39
Up Regulation	1222	27	20
Down Regulation	1235	22	11

The number of hours with deficiency decreased for all types of ancillary services when compared to 2011. Responsive Reserve service was deficient most frequently. Well over 90 percent of the deficiency occurrences were caused by failure to provide by the resource rather than undeliverability related to a transmission constraint. We also note that the average magnitude of non-spin reserve deficiency reduced from 90 MW in 2011 to 36 MW in 2012. On the other hand, the amount of deficiency in regulation services (up and down) increased slightly in 2012 when compared to 2011.

Figure 30: Ancillary Service Quantities Procured in SASM



Our final analysis in this section, shown in Figure 30, summarizes the average quantity of each service that was procured via SASM. Along with reduced occurrences of ancillary service deficiency, the quantity of services procured via SASM also declined in 2012. The average quantities of responsive and non-spin reserves procured via SASM in 2012 were roughly half what they were the previous year.

III. Transmission and Congestion

One of the most important functions of any electricity market is to manage the flows of power over the transmission network by limiting additional power flows over transmission facilities when they reach their operating limits. The action taken to ensure operating limits are not violated is called congestion management. The effect of congestion management is to change generator(s) output level so as to reduce the amount of electricity flowing on the transmission line nearing its operating limit. Because the transmission system is operated such that it can withstand the unexpected outage of any element at any time, congestion management actions most often occur when a transmission element is expected to be overloaded if a particular unexpected outage (contingency) were to occur. Congestion leads to higher costs due to lower cost generation being reduced and higher priced generation increased. Different prices at different nodes are the result. The decision about which generator(s) will vary their output is based on generating unit specific offer curves and the relative shift factors to the contingency and constraint pair. This leads to a dispatch of the most efficient resources available to reliably serve demand.

After summarizing congestion activity in 2012, this section of the report provides a review of the costs and frequency of transmission congestion in both the day-ahead and real-time markets. We will then provide a review the activity in the congestion rights market.

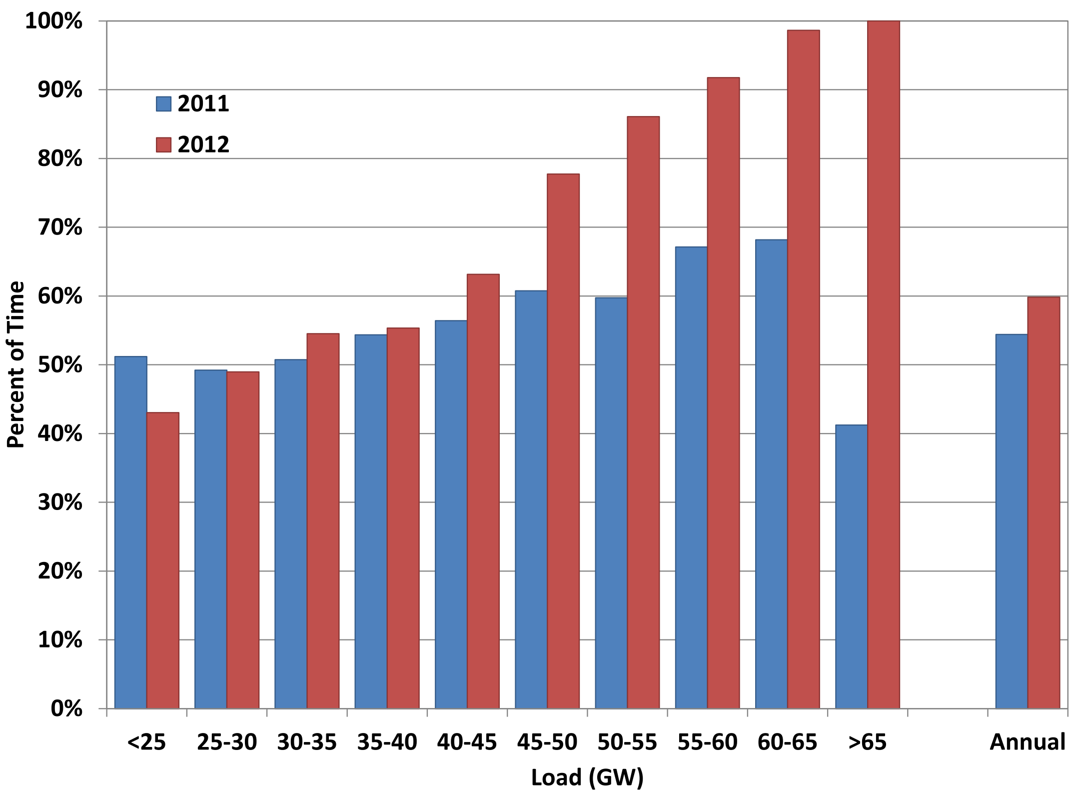
A. Summary of Congestion

There was a marked change in the nature of real-time transmission congestion during 2012 when compared to previous years. For the past several years a significant portion of real-time transmission congestion could be described as limiting the export of wind generation ***from*** the West zone to the load centers across the rest of ERCOT. Transmission congestion in 2012 was more significantly the result of limitations on the ability to get generation ***to*** loads in the West zone. Some portion of the limitation can be attributed to transmission outages taken to enable the construction of new CREZ transmission lines. Another factor is that loads in far west Texas have experienced much greater growth than the system-wide average.

The total congestion revenue generated by the ERCOT real-time market in 2012 was $480 million, a decrease of 9 percent from 2011. This decrease was not as large as might be expected given the much larger decreases in average natural gas prices real-time energy prices. Two factors influencing the overall costs of congestion in 2012 were the significant financial impact of several localized transmission constraints in far west Texas and the higher frequency of active transmission constraints.

Figure 31 provides a comparison of the amount of time transmission constraints were active at various load levels in 2012 and 2011. Active transmission constraints are those for which generators are being dispatched to a less efficient output level in order to maintain transmission flows at reliable levels.

Figure 31: Frequency of Active Constraints



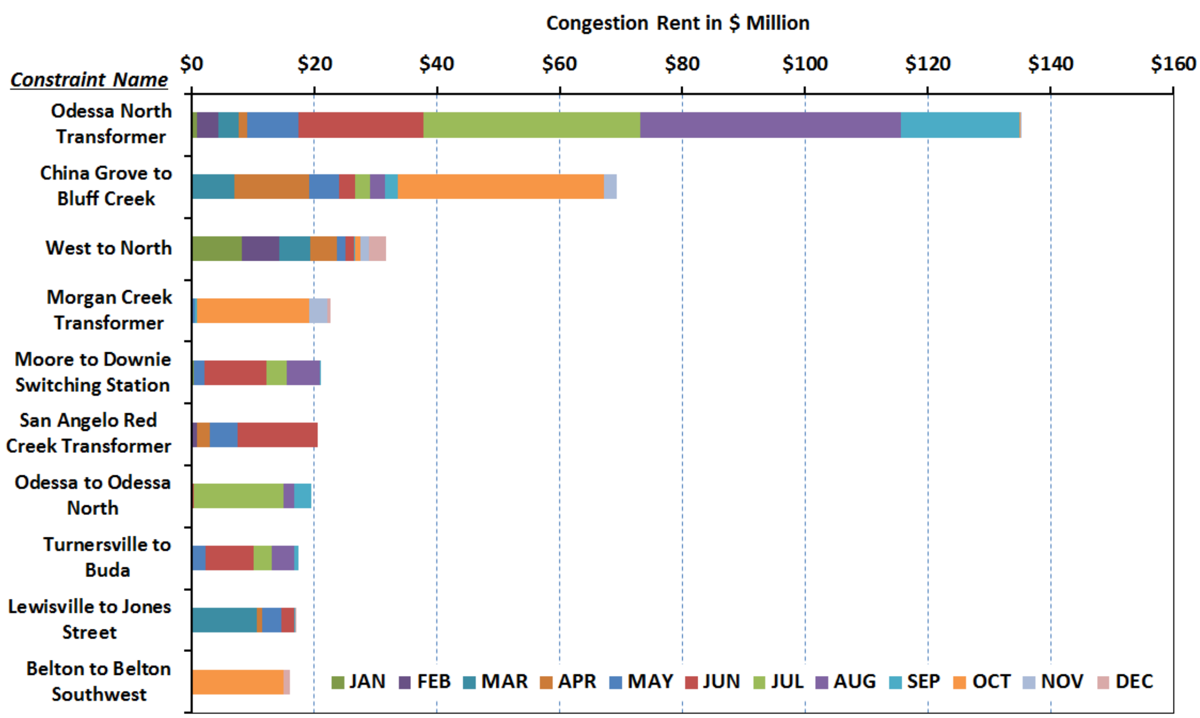
We observe that in 2012 the likelihood of having an active transmission constraint was higher than it was in 2011 and that for loads above 45 GW the frequency was much higher. During 2011, we observed that at higher system loads ERCOT operators did not always activate (or sometimes de-activated) transmission constraints. This was due to a concern that by having a constraint active during periods of high demand the total capacity available to serve load may be limited. However, ERCOT's dispatch software contains parameters that allow it to automatically make the correct decision about when to violate transmission constraints when necessary to serve total system load. Therefore, ERCOT modified their practice in 2012 to retain active transmission constraints even during periods of high demand.

B. Real-Time Constraints

We begin our review by examining the real-time transmission constraints with the highest financial impact as measured by congestion rent. There were more than 350 different constraints active at some point during 2012, a slight increase from 2011. The median financial impact, as measured by congestion rent, was approximately $200,000 during 2012. This is a significant decrease from 2011, when the median impact was approximately $300,000. Given the significant reduction in average energy costs from 2011 to 2012, reduced financial impact of congestion is expected.

Figure 32 below displays the ten most highly valued real-time constraints as measured by congestion rent and indicates that the Odessa North 138/69 kV transformer constraint was by far

Figure 32: Top Ten Real-Time Constraints



the most highly valued during 2012. This constraint became more pronounced from 2011 to 2012 and is mainly attributed to load growth in far west Texas driven by increased oil and natural gas activity.

The Odessa North 138/69 kV transformer typically overloads under low wind conditions. The characteristics that load the Odessa North 138/69 kV transformer are the same conditions that also affect the Odessa to Odessa North 138 kV line, which is shown as the seventh constraint on the list. Not only did this constraint have nearly twice the financial impact of the second constraint on the list, its impact was more than 40 percent greater than the top constraint from 2011. Its magnitude is even more significant given the overall lower costs of energy in 2012. Not surprisingly, much public attention was focused on this constraint; much of it questioning its causes and the potential for short-term remedies. ERCOT and the local transmission provider were able to identify two transmission lines, which when opened greatly reduced congestion around the Odessa North station without causing other reliability concerns. After the lines were opened in mid-September, congestion around Odessa North was almost eliminated for the rest of the year.

The second constraint on the list is the overload of the 138 kV transmission line between China Grove and Bluff Creek. Also located in west Texas, west of Abilene, this constraint was new to the top ten list in 2012. Most of the time this line was constrained in the China Grove to Bluff Creek direction, typically under low wind conditions but also when there were other lines out of service. Depending on the level of wind and outages, the direction of the constraint would be reversed. That is, under high wind conditions flow on the line would be limited in the Bluff Creek to China Grove direction (from the west).

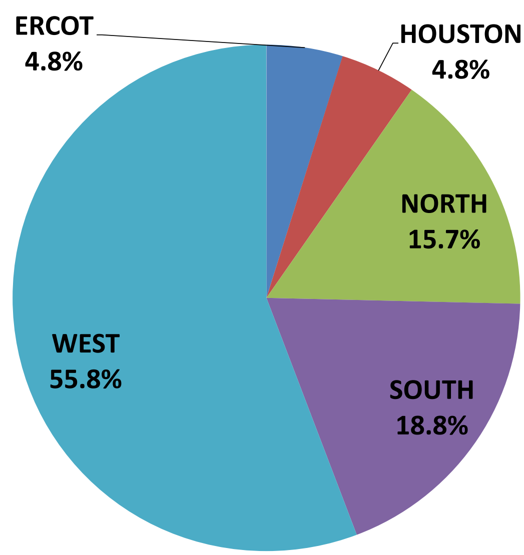
The West to North constraint continued to be active at some point during every month of 2012 but with far less impact than in 2011. Through the years this constraint has been a major impediment to delivering all the wind generation located and produced in the western reaches of ERCOT to the load centers. This constraint had the highest financial impact of all real-time transmission constraints during 2011, but in 2012 its impact dropped to one third that level. The reduction was a result of the combined impact of higher loads in the West and increased transfer capability due to the first CREZ transmission lines being placed in service. We will review the utilization of this constraint in more detail later in this section.

In all, eight of the top ten real-time constraints during 2012 were due to load and wind conditions within the West. In addition to the four constraints previously discussed, the Morgan Creek and San Angelo Red Creek transformer constraints were typically active under high load and low wind conditions. Unlike the previous constraints the Moore to Downie 138 kV line and the Turnersville to Buda 138 kV line are not located in the West, but became constrained during low wind and high load conditions in the West that occurred while other, larger capacity transmission lines were out of service.

The remaining two constraints on the list were of fairly short duration; their high impact reflecting limitations on the ability to import power to a major load center. The Lewisville to Jones Street 138 kV constraint was related to serving load in the DFW area. Belton to Belton Southwest was due to outage conditions that limited the flow to Central Texas.

To further highlight the significance of West zone congestion, Figure 33 depicts that more than half of the costs of real-time congestion that occurred in 2012 were associated with facilities located in the West zone. Further, the cost of congestion on facilities that spanned zonal boundaries, labeled ERCOT in the figure, primarily comprised costs associated with West to North congestion.

Figure 33: Real-Time Congestion Costs



Irresolvable Constraints

When a constraint becomes irresolvable, ERCOT's dispatch software can find no combination of generators to dispatch in a manner such that the flows on the transmission line(s) of concern are below where needed to operate reliably. In these situations, offers from generators are not setting locational prices. Prices are set based on predefined rules, which since there are no supply options for clearing, should reflect the value of reduced reliability for demand. To address the situation more generally, a regional peaker net margin mechanism was introduced such that once local price increases accumulate to a predefined threshold due to an irresolvable constraint; the shadow price of that constraint would drop.

As shown below in Table 3, thirteen contingency and constraint pairs were deemed irresolvable in 2012 and as such, had a maximum shadow price cap imposed according to the Irresolvable Constraint Methodology. Three of the top ten real-time constraints, Odessa North 138/69 kV transformer, China Grove to Bluff Creek 138 kV line, and Morgan Creek #1 345/138 kV Autotransformer were designated as irresolvable in 2012. One of these constraints, the Odessa North 138/69 kV transformer, reached the second level trigger in the procedure and had its maximum shadow price lowered to $2,000 per MW in August 2012.

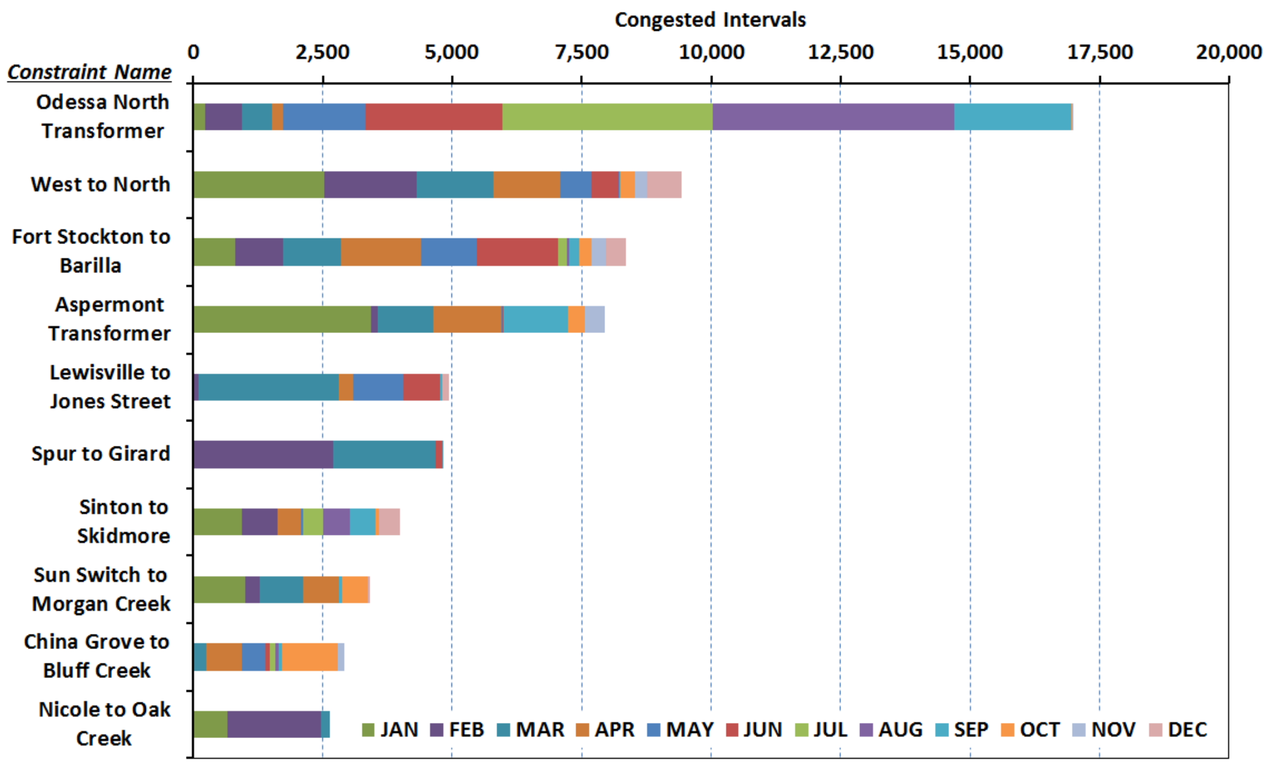
Table 3: Irresolvable Constraints

Loss of:	Overloads:	Maximum Shadow Price	Effective Date
Base case	Valley Import	$2,000.00	Jan 1, 2012
Graham to Long Creek 345 kV line	Bomarton to Seymour 69 kV line	$2,000.00	Jan 1, 2012
Odessa to Morgan Creek / Quail 345 kV lines	Ackerly Vealmoor to Ackerly 69 kV line	$2,000.00	Jan 1, 2012
Denton to Argyle / West Denton 138 kV lines	Jim Crystal to West Denton 69 kV line	$2,000.00	Jan 1, 2012
Cabaniss to Westside 138 kV line	Naval to North Padre 69 kV line	$2,000.00	Jan 1, 2012
Key to Willow Valley 138 kV line	Ackerly Vealmoor to Ackerly 69 kV line	$2,554.65	Jan 1, 2012
Key to Willow Valley 138 kV line	Ackerly to Lyntegar 69 kV line	$2,800.00	Jan 1, 2012
Odessa North to Holt 69 kV line	Odessa Basin to Odessa North 69 kV line	$2,800.00	Jan 1, 2012
Holt to Moss 138 kV line	Odessa North 138/69 kV transformer	$2,274.65 $2000.00	Jan 1, 2012 Aug 6, 2012
Odessa Basin to Odessa North 69 kV line	Holt to Ector Shell Tap 69 kV line	$2,000.00	Jan 1, 2012
Odessa to Morgan Creek / Quail 345 kV lines	China Grove to Bluff Creek 138 kV line	$2,000.00	May 3, 2012
Sun Switch to Morgan Creek 138 kV line	China Grove to Bluff Creek 138 kV line	$2,000.00	Oct 11, 2012
Morgan Creek #4 345 kV/138 kV Autotransformer	Morgan Creek #1 345 kV/138 kV Autotransformer	$2,000.00	Nov 2, 2012

Figure 34 presents a slightly different set of real-time constraints. These are the most frequently occurring. Other than the Lewisville to Jones Street constraint described previously, all constraints in this ranking are related to wind generation. With the exception of the Odessa North 138/69 kV transformer, which typically overloads under low wind and high local load conditions, the other frequently occurring constraints are all related to high West zone wind. Not only was the Odessa North 138/69 kV transformer constraint the most costly real-time constraint, it tops the list as the most frequently occurring constraint in 2012. It was binding more than 15 percent of the time in 2012.

As the second most frequently occurring real-time constraint, West to North was binding less than 10 percent of the time in 2012. During 2011 this constraint was the most frequently occurring real-time constraint; active more than 20 percent of the time.

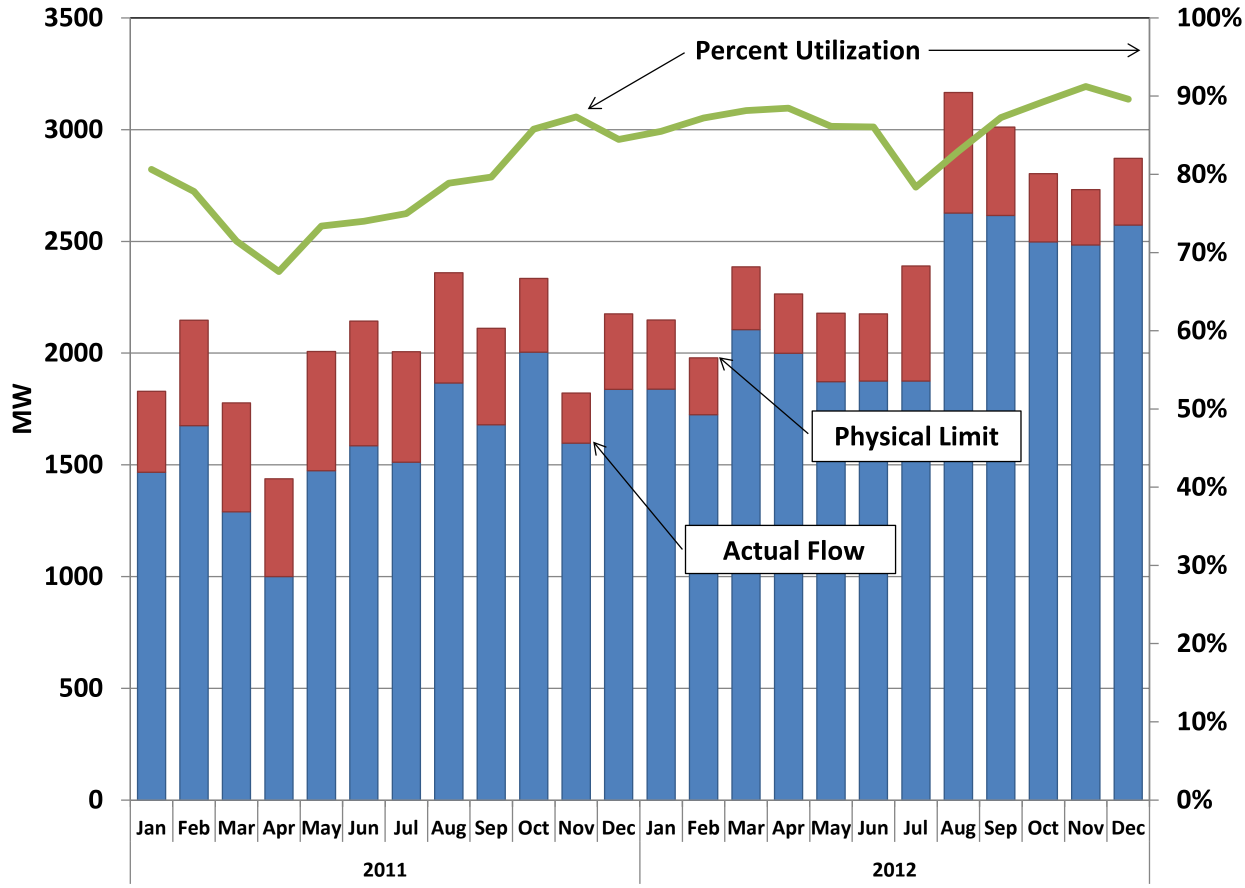
Figure 34: Most Frequent Real-Time Constraints



To maximize the economic use of scarce transmission capacity, the ideal outcome would be for the actual transmission line flows to reach, but to not exceed the physical limits required to maintain reliable operations. Figure 35 presents a summary of the utilization of the West to

North interface transmission constraint during 2011 and 2012. By comparing the actual flow with the physical limit of the constraint for each dispatch interval it was binding, we can compute its average utilization.

Figure 35: Utilization of the West to North Interface Constraint



Through July 2012, the average physical limit was approximately 2,200 MW and the average actual flow during constrained intervals was approximately 1,900 MW. After July 2012, the physical limit increased to an average of 2,900 MW and the actual flow increased to approximately 2,500 MW. In March 2012, a new real-time analysis tool was implemented to better track the dynamic nature of the transient stability limit of the West to North interface. However, there was not a noticeable increase to the transfer limit corresponding to its implementation due to the effects of transmission outages occurring to accommodate maintenance activities and the installation of CREZ lines. Many of these outages were complete by July 2012 which accounts for the increase in the West to North limit. The average annual utilization of the West to North constraint was 87 percent in 2012, which compares favorably to

78 percent utilization experienced in 2011. Over the long term, the physical limit will continue to increase as CREZ transmission projects are completed.

C. Day-Ahead Constraints

In this section we review transmission constraints from the day-ahead market. To the extent the model of the transmission system used for the day-ahead market matches the real-time transmission system, and assuming market participants transact in the DAM similarly to how they transact in real-time, we would expect to see the same transmission constraints appear in the day-ahead market as actually occurred during real-time.

Figure 36: Top Ten Day-Ahead Constraints

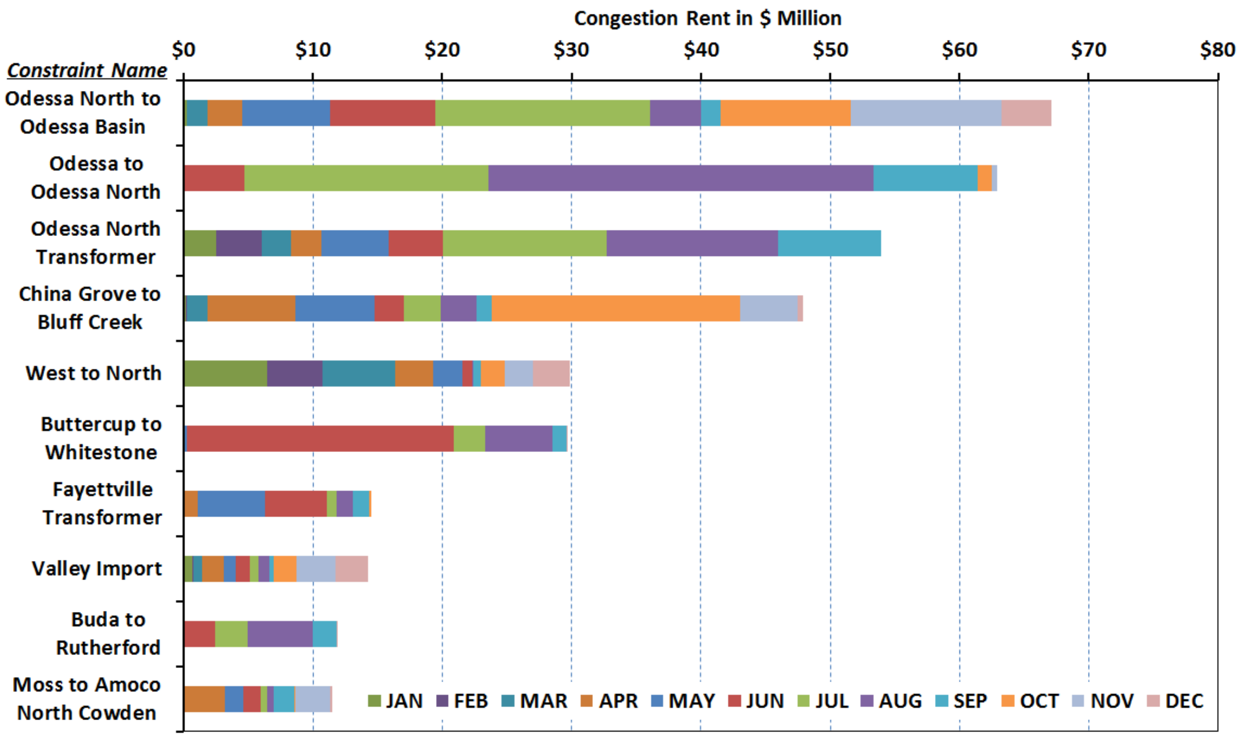


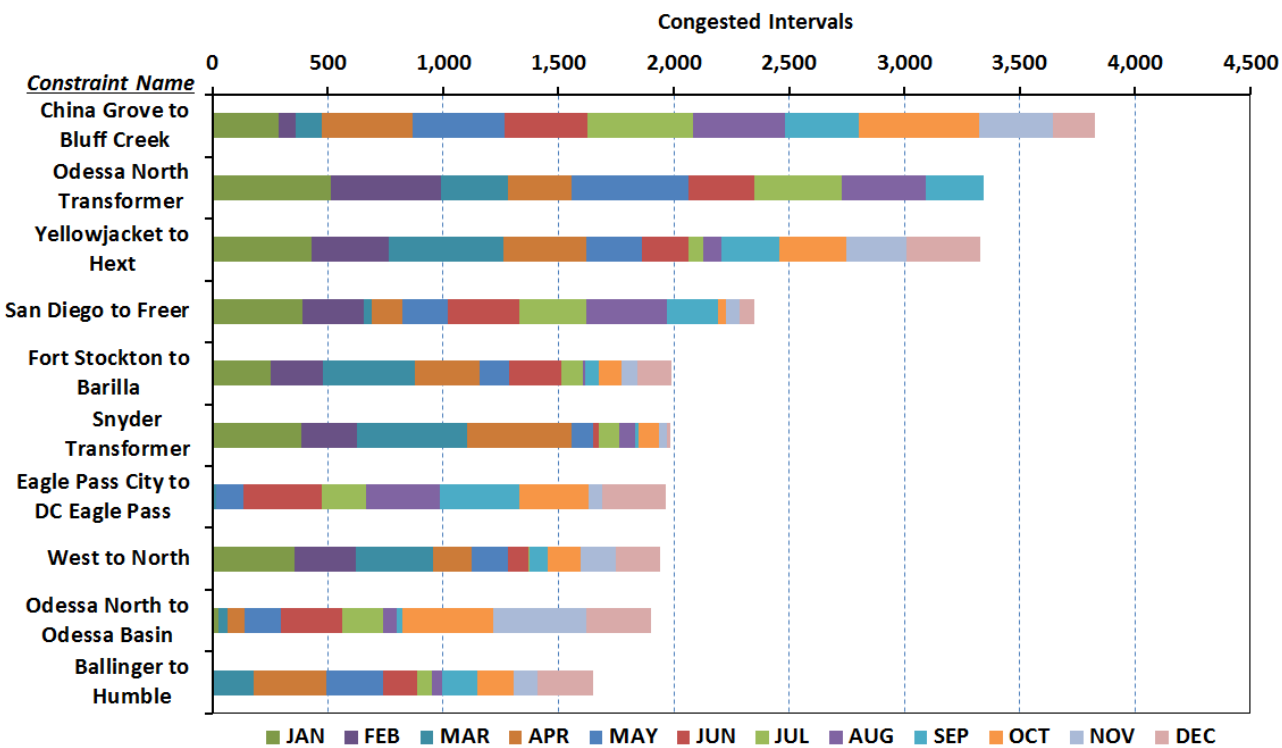
Figure 36 presents the top ten constraints from the day-ahead market, ranked by their financial impact as measured by congestion rent. As was the case with the real-time constraints, the Odessa North transformer constraint had a significant financial impact, although it was not the first on the list. However, the top three constraints are all associated with the Odessa North substation, which further demonstrates the limiting nature of the area. The final constraint on the list, the Moss to Amoco North Cowden 138 kV line, is similar to these Odessa North constraints.

Only the Odessa North transformer, China Grove to Bluff Creek and West to North constraints were previously shown in the real-time list of constraints.

The next two constraints, Buttercup to Whitestone and Buda to Rutherford, are located in Central Texas and were binding in the westward direction. They are both examples of constraints limiting more remote generation reaching the increasing load in west Texas, especially with certain transmission lines out of service. The Fayetteville 345/138 kV transformer constraint limited transfers feeding 138 kV loads in Houston. The Valley import constraint had less of an impact in 2012 than 2011 although it was active during every month.

In our final analysis of this section we review the most frequently occurring day-ahead constraints shown in Figure 37.

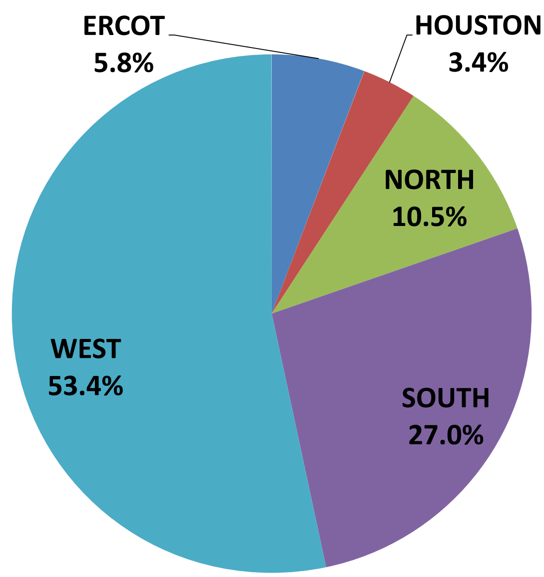
Figure 37: Most Frequent Day-Ahead Constraints



Three of the constraints appearing on the list would not occur in real-time. The Eagle Pass City to Eagle Pass DC Tie constraint appears frequently as a day-ahead constraint, but in real-time operations all transactions with Mexico using this DC Tie are scheduled using a separate process. The process would strictly limit the volume of transactions and not allow a constraint to occur.

The Yellowjacket to Hext and Ballinger to Humble constraints are both affected by nearby phase shifters that depending on the tap setting of the element will have different impedances through the phase shifter. In the day-ahead market, the phase shifters are set at one value throughout the day, typically a mid-setting of the full range. The constraint seen in the day-ahead would likely not bind in real-time due to the fact that the tap settings can be changed to alter the flow over the elements.

Figure 38: Day-Ahead Congestion Costs



With the exception of the San Diego to Freer 69 kV line located in a sparse transmission area of South Texas all of the constraints listed in Figure 37 are associated with West zone conditions.

To further emphasize the effects of West zone congestion in 2012, Figure 38 highlights that, like real-time, day-ahead West zone congestion accounted for more than half the congestion in 2012.

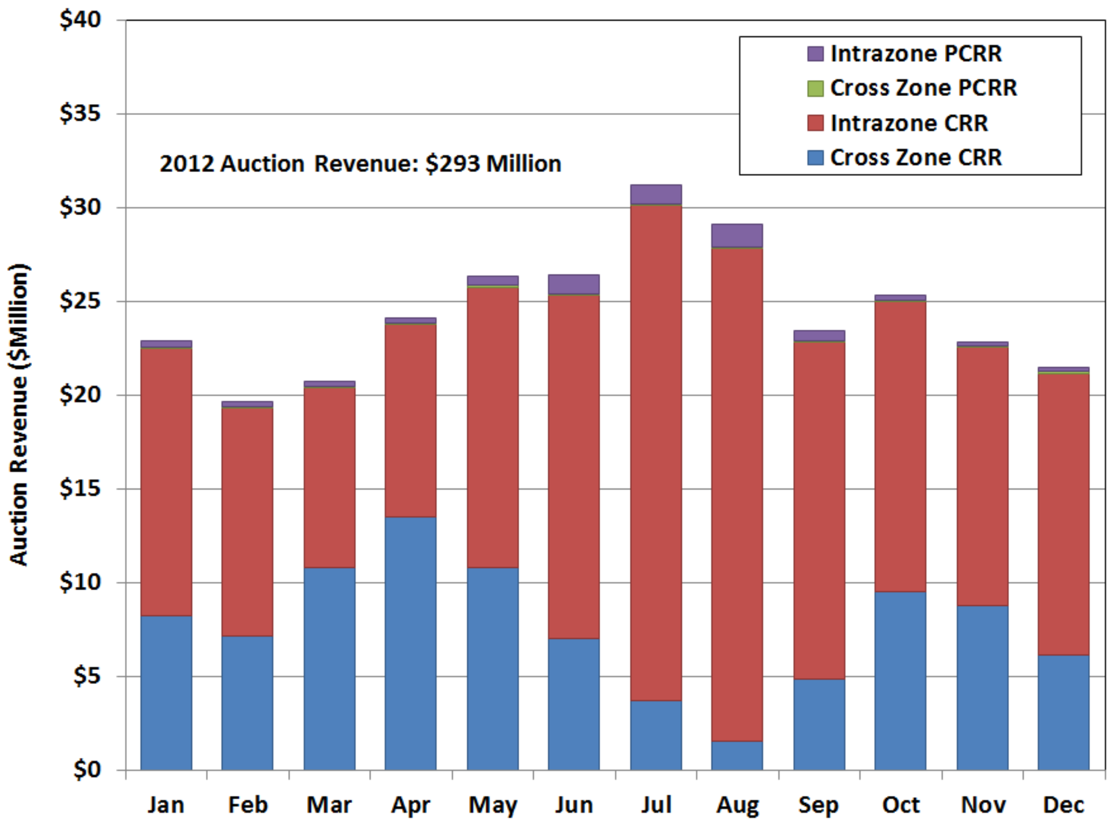
D. Congestion Rights Market

Congestion can be significant from an economic perspective, compelling the dispatch of higher-cost resources because power produced by lower-cost resources cannot be delivered due to transmission constraint(s). Under the nodal market design, one means by which ERCOT market participants can hedge these price differences is by acquiring Congestion Revenue Rights ("CRRs") between any two settlement points.

CRRs are acquired by annual and monthly auctions while Pre-assigned Congestion Revenue Rights ("PCRRs") are allocated to certain participants based on their historical patterns of transmission usage. Parties receiving PCRRs pay only a fraction of the auction value of a CRR between the same source and sink. Both CRRs and PCRRs entitle the holder to payments corresponding to the difference in locational prices of the source and sink.

Figure 39 summarizes the revenues collected by ERCOT in each month for all CRRs, both auctioned and allocated. These revenues are distributed to loads in one of two ways. Revenues from cross zone CRRs are allocated to loads ERCOT wide. Revenues from CRRs that have their source and sink in the same geographic zone are allocated to loads within that zone. This method of revenue allocation provides a disproportionate share of CRR auction revenues to loads located in the West zone. In 2012, CRRs with both their source and sink in the West zone accounted for 27 percent of CRR Auction revenues. This share of revenue was allocated to West zone loads, which accounted for only 7 percent of the ERCOT total. Allocating CRR Auction revenues in this manner helps reduce the impact of the higher congestion on West zone prices.

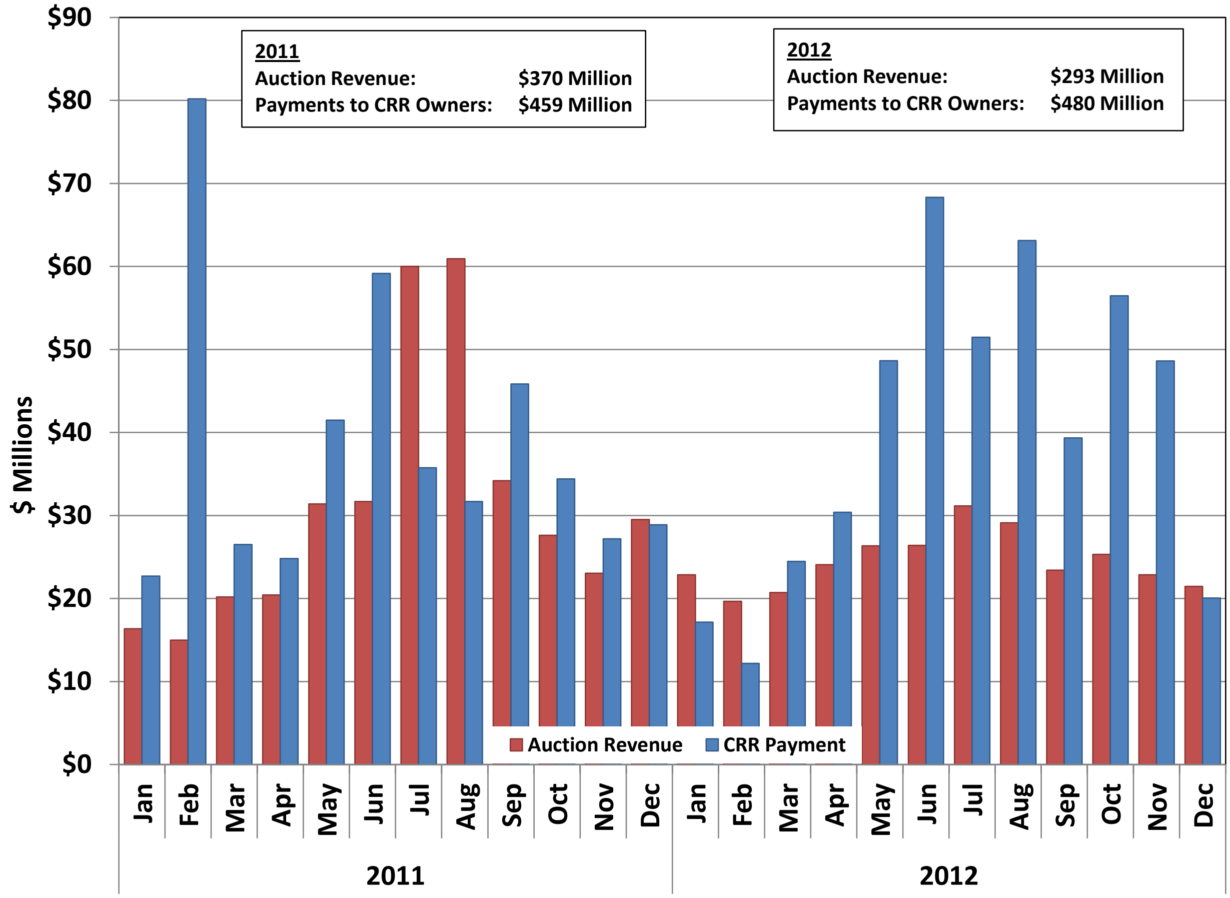
Figure 39: CRR Auction Revenue



As we showed in Section I.A, Real-Time Market Prices, the annual average price for the West zone was $38.24 per MWh, nearly $10 per MWh higher than the ERCOT-wide average. The value of CRR Auction revenues distributed only to the West zone equated to almost $3 per MWh more than the amounts distributed to other zones.

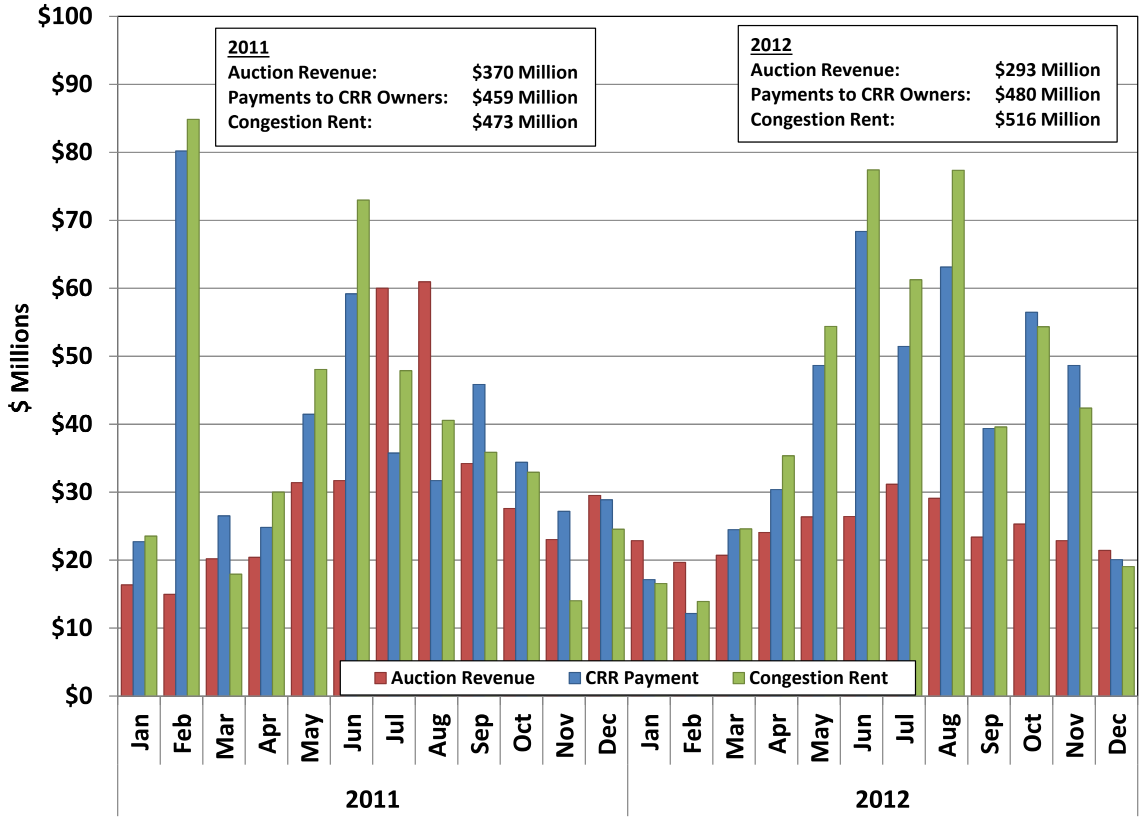
Next, in Figure 40 we examine the value CRR owners (in aggregate) received compared to the price they paid to acquire the CRRs. Although results for individual participants and specific source/sink combinations varied, we find that in most months participants did not over pay in the auction. Across the entire year of 2012, participants spent $293 million to procure CRRs and received $480 million.

Figure 40: CRR Auction Revenue and Payment Received



In our next look at aggregated CRR positions, we add congestion rent to the picture. Simply put, congestion rent is the difference between the total costs that loads pay and the total revenue that generators receive. Congestion rent creates the source of funds used to make payments to CRR owners. Figure 41 presents all three values for each month of 2011 and 2012. Congestion rent for the year 2012, totaled $516 million and payments to CRR owners were $480 million. The only months in 2012 when congestion rent was less than payments to CRR owners were January and October through December.

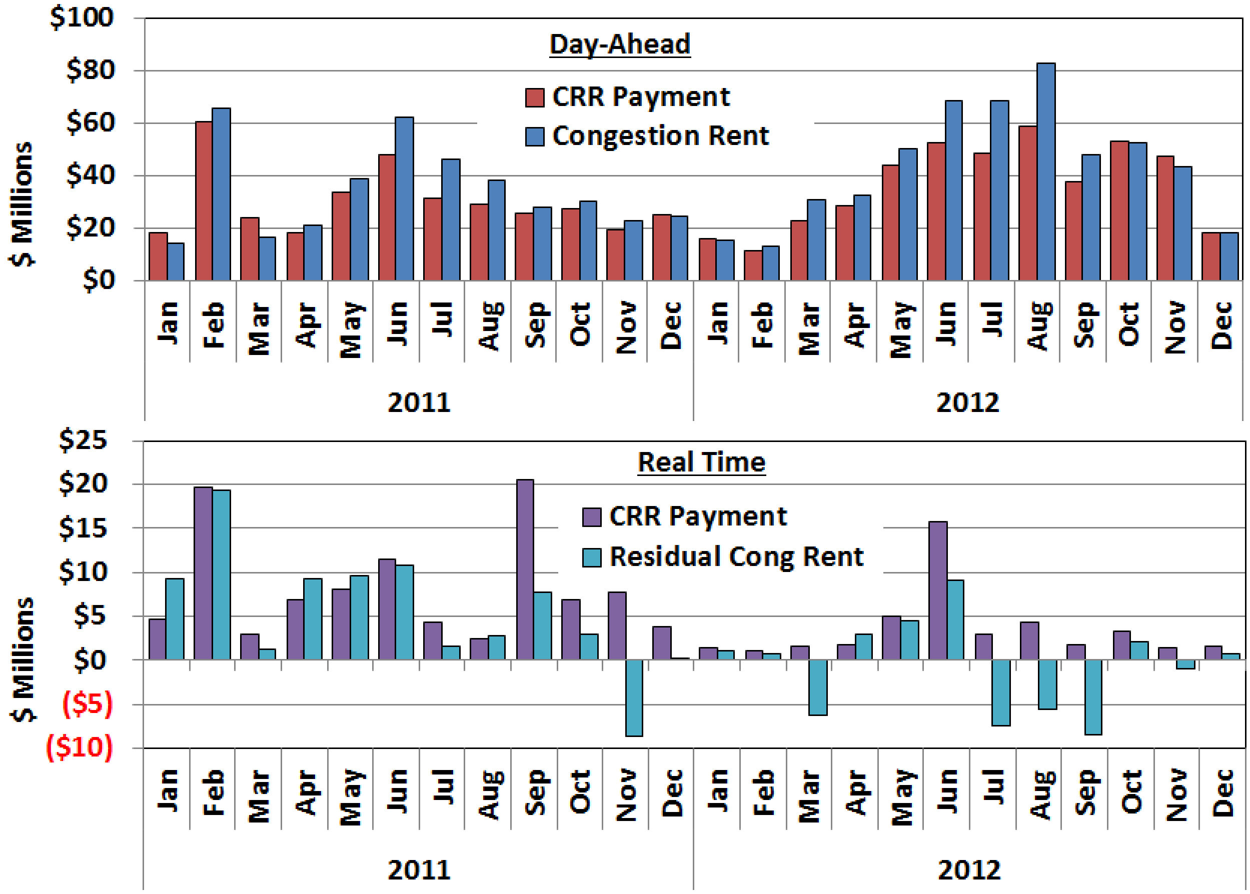
Figure 41: CRR Auction Revenue, Payments and Congestion Rent



We further analyze the relationship between congestion rent and payments to CRR owners by separating the impacts of CRRs that are settled based on day-ahead prices from the subset of CRRs that are paid based on real-time prices.

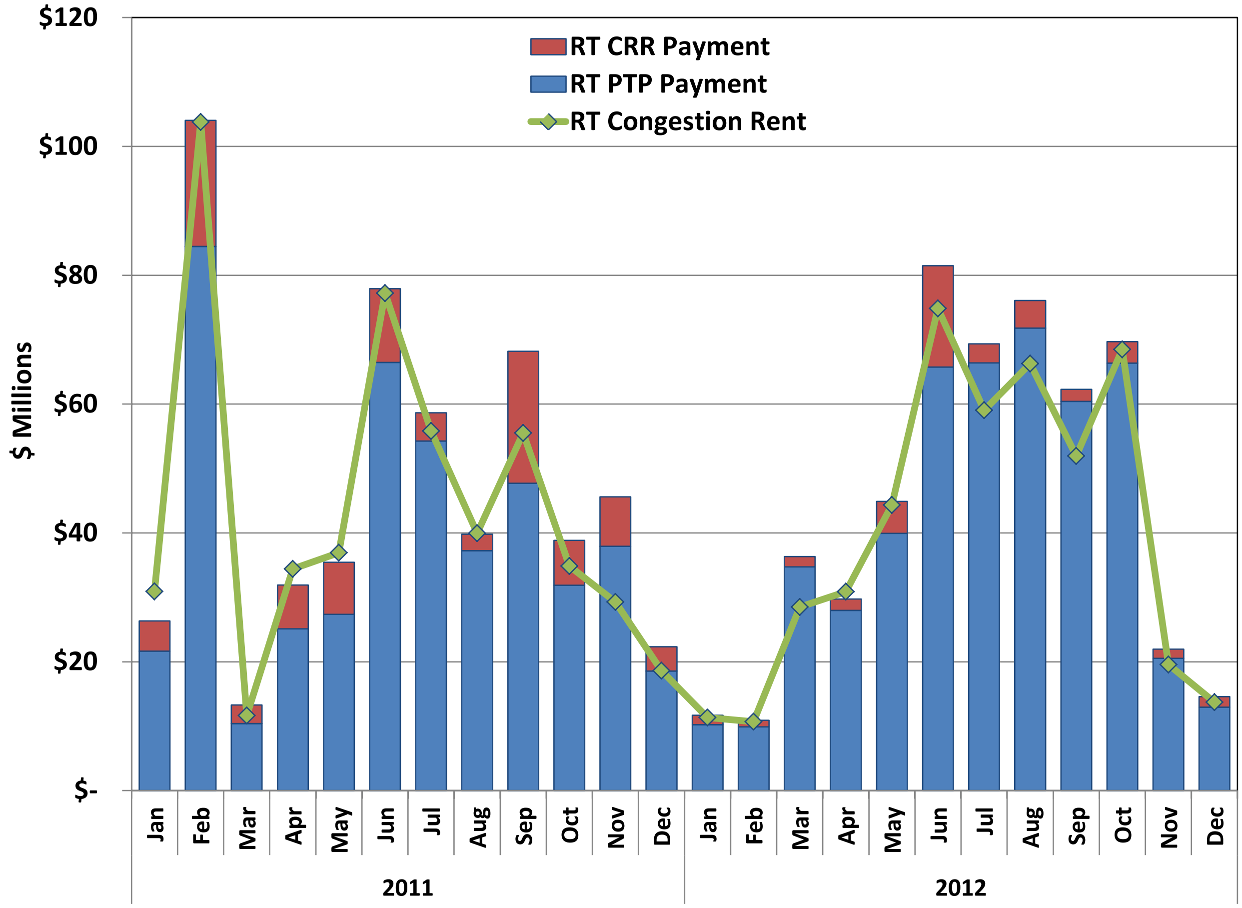
The top portion of Figure 42, shown below, displays the comparison of day-ahead congestion rent to payments received by CRR owners. Congestion rent was larger than payments in most months of 2012 and for the year congestion rent was $523 million compared to $438 million that was paid to CRR owners.

Figure 42: Day-Ahead and Real-Time Congestion Payments and Rent



The bottom portion of Figure 42 presents a different view. For this analysis we have assumed that all PTP Obligations have been fully funded from real-time congestion rent and any residual real-time congestion rent is available to fund payments to the subset of CRR owners that elected to have their CRRs be settled based on real-time prices. In 2012 there was less real-time congestion rent than the payments to holders of PTP Obligations, resulting in a $7 million shortfall. Further, there were real-time CRR payments of $42 Million. Hence, real-time congestion rent was insufficient to fund all PTP Obligations and CRRs being settled in real-time in the amount of $49 million. The next figure shows this explicitly.

Figure 43: Real-Time Congestion Payments



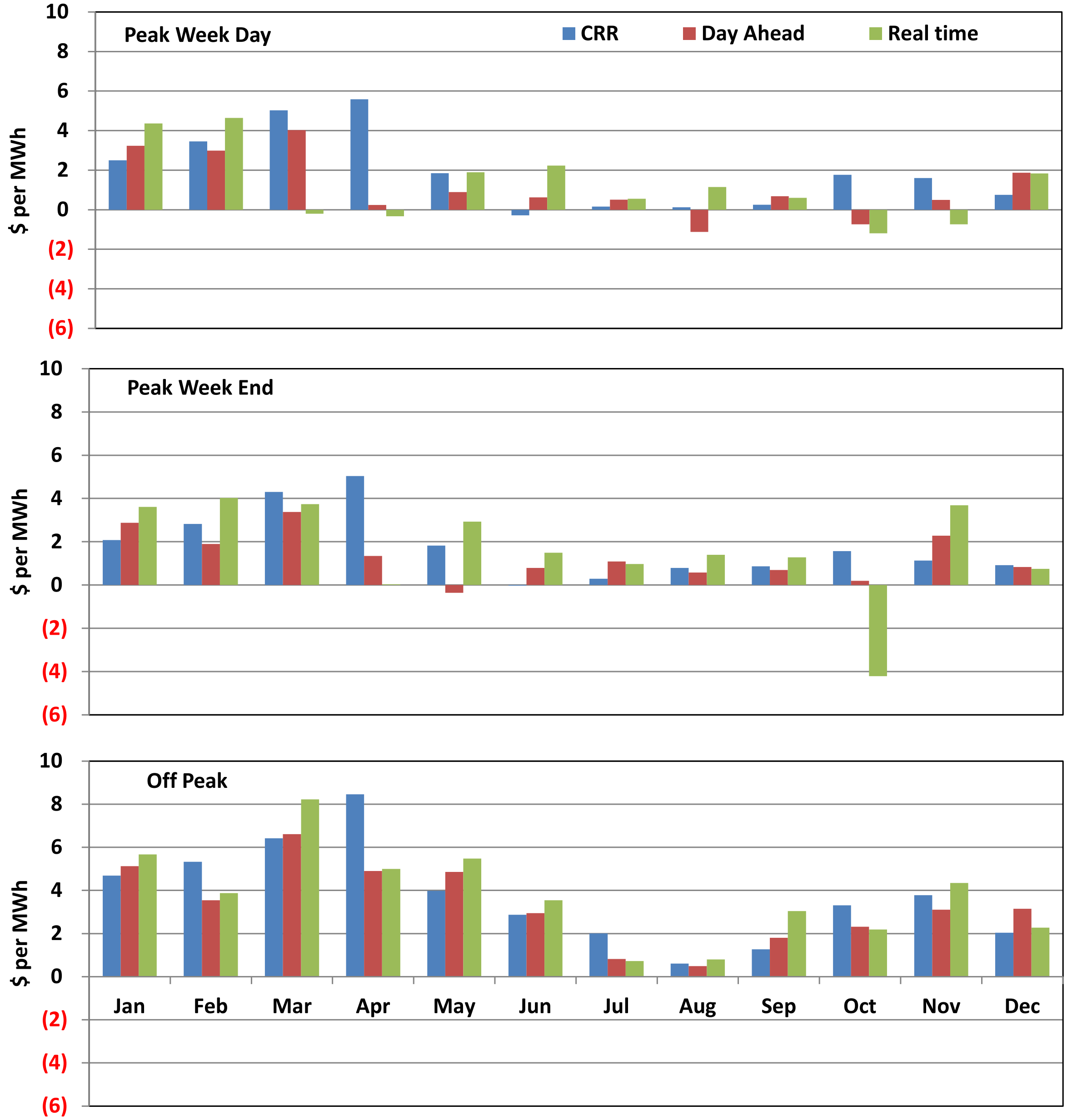
In Figure 43 the combined payments to PTP Obligation owners and CRR owners that have elected to receive real-time payments are compared to the total real-time congestion rent. For the year of 2012, real-time congestion rent was $480 million, payments for PTP Obligations were $487 million and payments for real-time CRRs were $42 million, resulting in a shortfall of approximately $49 million for the year.

This type of shortfall typically results from discrepancies between transmission topology assumptions used when clearing the day-ahead market and the actual transmission topology that exists during real-time. Specifically, if the day-ahead topology assumptions allow too many real-time congestion instruments (PTP Obligations and CRRs settled at real-time prices) with impacts on a certain path, and that path constrains at a lower limit in real-time, there will be insufficient rent generated to pay all of the congestion instruments.

From Figure 43 we can see that March and June through September were the months with the most noticeable deficiencies. A detailed examination of the daily congestion pattern during these months shows that outages of transmission facilities did occur on days with large insufficiency.

For our last look at congestion we examine the impacts of the West to North constraint in more detail. Figure 44 presents the price spreads between the West Hub and North load zone as valued at three separate points in time – at the monthly CRR auction, day-ahead and in real-time.

Figure 44: West Hub to North Load Zone Price Spreads



Since CRRs are sold for one of three defined time periods, weekday on peak, weekend on peak, and off peak, Figure 44 includes a separate comparison for each.

As expected, most real-time congestion, as evidenced by the largest price spread, occurred in the off peak period, for the months of January through June and November. The day-ahead price spreads were very similar for this period, while the prices paid for CRRs in April were more than the value received. Conversely, during the summer months of July and August, there was very little congestion. In October day-ahead and real-time prices were higher at the West Hub at the peak hours, which the results of the CRR auction did not anticipate.

IV. LOAD AND GENERATION

This section reviews and analyzes the load patterns during 2012 and the existing generating capacity available to satisfy the load and operating reserve requirements. We provide specific analysis of the large quantity of installed wind generation and conclude this section with a discussion of the daily generation commitment process.

A. ERCOT Loads in 2012

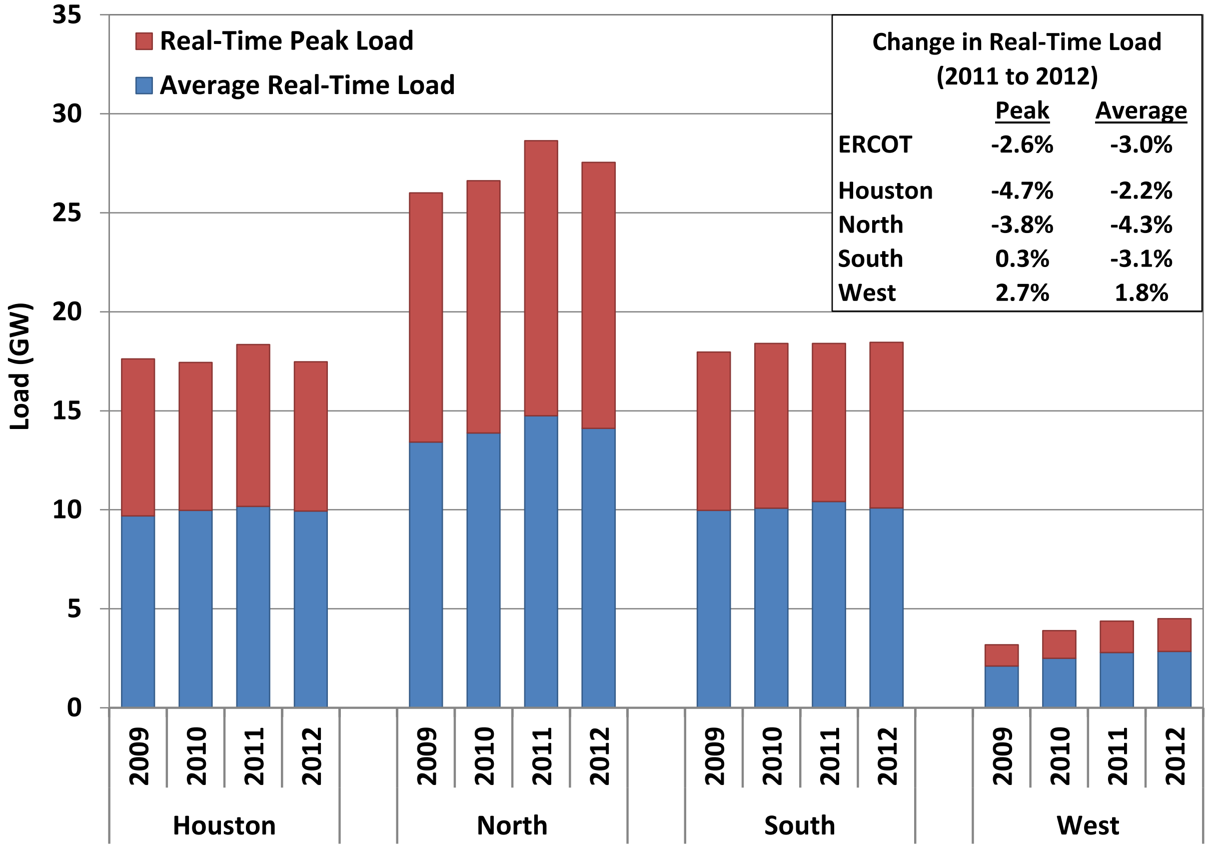
The changes in overall load levels from year to year can be shown by tracking the changes in average load levels. This metric will tend to capture changes in load over a large portion of the hours during the year. It is also important to separately evaluate the changes in the load during the highest-demand hours of the year. Significant changes in peak demand levels play a major role in assessing the need for new resources. They also affect 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 and inconvenience to customers that can result from interruptions to that service. Hence, both of these dimensions of load during 2012 are examined in this subsection and summarized in Figure 45.

This figure shows peak load and average load in each of the ERCOT zones from 2009 to 2012.[13] In each zone, as in most electrical systems, peak demand significantly exceeds average demand. The North zone is the largest zone (with about 38 percent of the total ERCOT load); the South and Houston zones are comparable (27 percent) while the West zone is the smallest (8 percent of the total ERCOT load).

Figure 45 also shows the annual non-coincident peak load for each zone. This is the highest load that occurred in a particular zone for one hour during the year; however, the peak can occur in different hours for different zones. As a result, the sum of the non-coincident peaks for the zones was greater than the annual ERCOT peak load.

[13] For purposes of this analysis NOIE Load Zones have been included with the proximate geographic Load Zone.

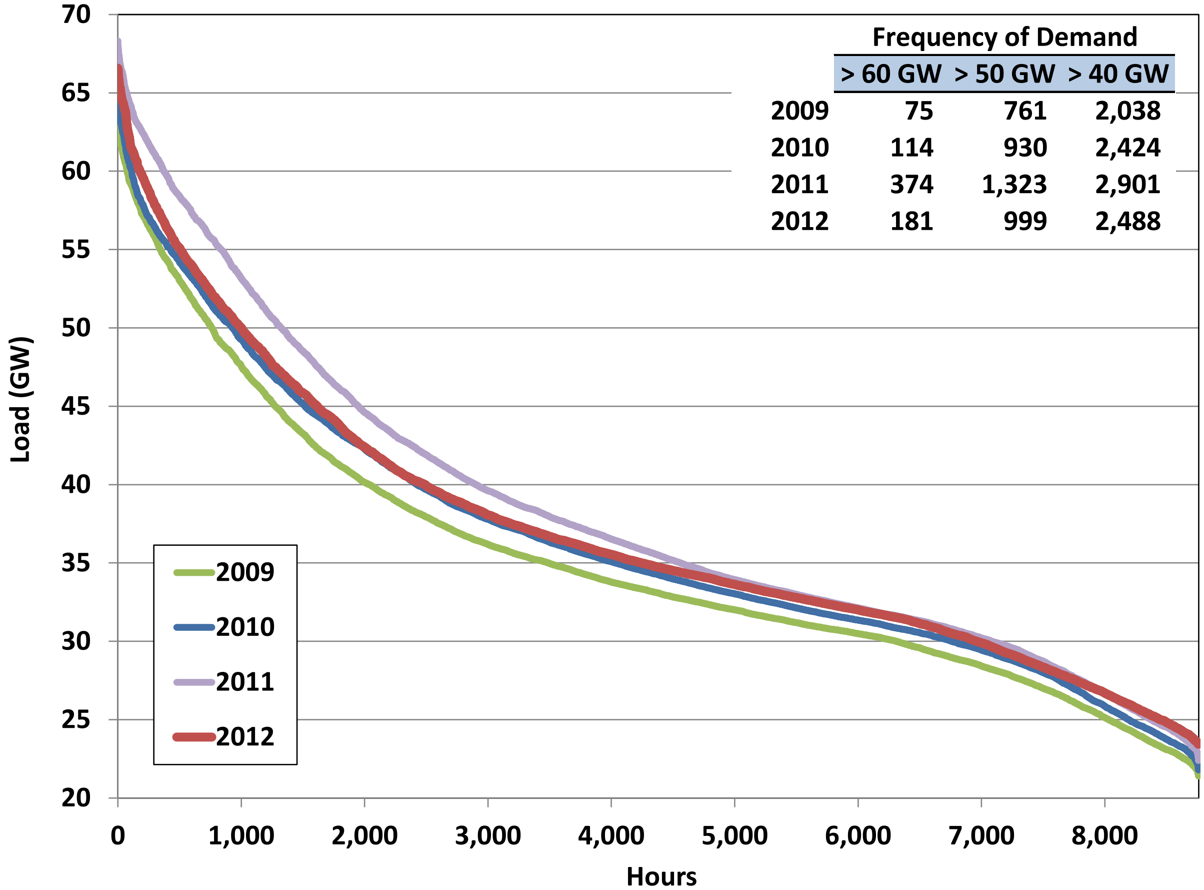
Figure 45: Annual Load Statistics by Zone



Total ERCOT load decreased from 334 TWh in 2011 to 325 TWh in 2012, a decrease of 2.7 percent or an average of 1,130 MW every hour. Similarly, the ERCOT coincident peak hourly demand decreased from 68,311 MW in 2011 to 66,559 MW, a decrease of 1,752 MW, or 2.6 percent. The results at the zonal level are not consistent. Average load decreased in three of the four zones, but grew by 1.8 percent in the West zone.

To provide a more detailed analysis of load at the hourly level, Figure 46 compares load duration curves for each year from 2009 to 2012. A load duration curve shows 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.

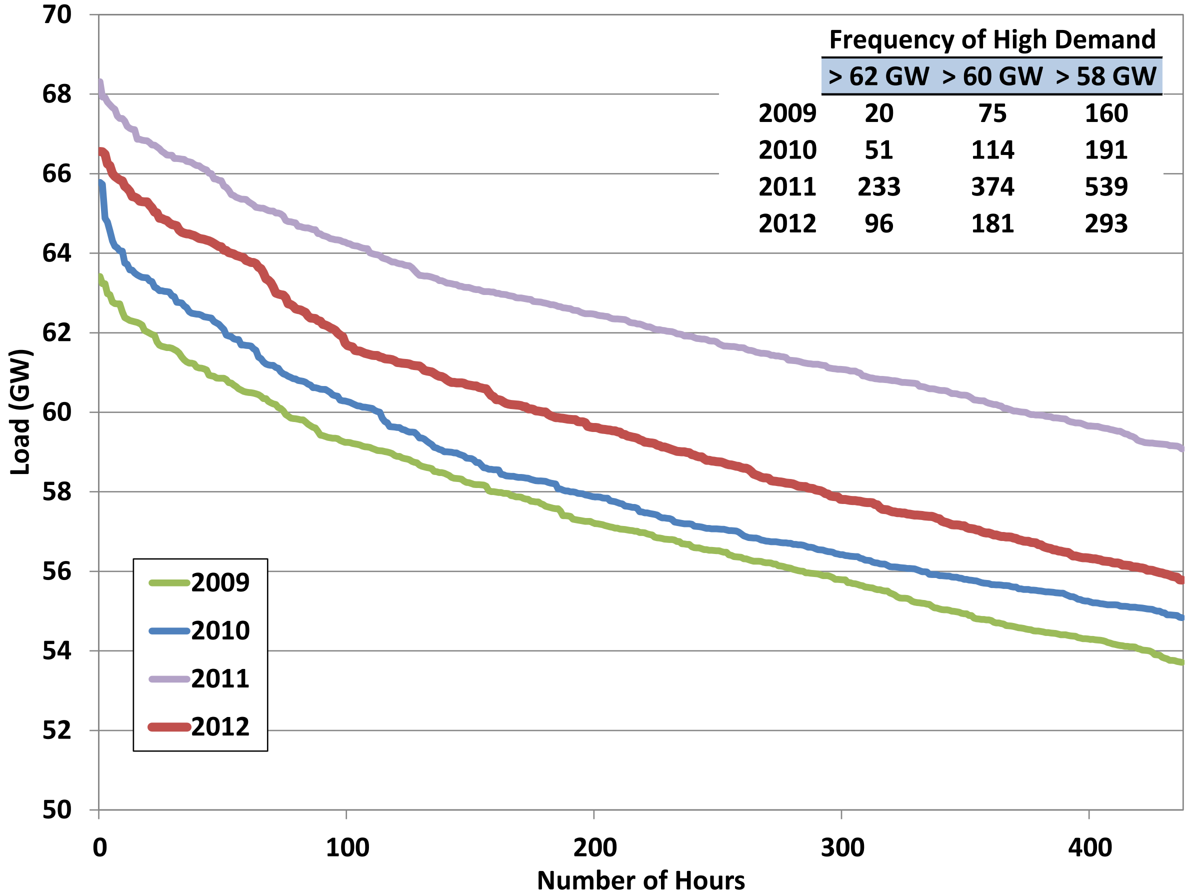
Figure 46: Load Duration Curve – All hours



As shown in Figure 46, the load duration curve for 2012 is significantly lower than in 2011 for approximately half of the hours in the year. This is consistent with the aforementioned 2.7 percent load decrease from 2011 to 2012.

To better illustrate the differences in the highest-demand periods between years, Figure 47 shows the load duration curve for the five percent of hours with the highest loads. This figure also shows that the peak load in each year is significantly greater than the load at the 95th percentile of hourly load. From 2009 to 2012, 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 – almost 10 GW – is needed to supply energy in less than 5 percent of the hours.

Figure 47: Load Duration Curve – Top five percent of hours

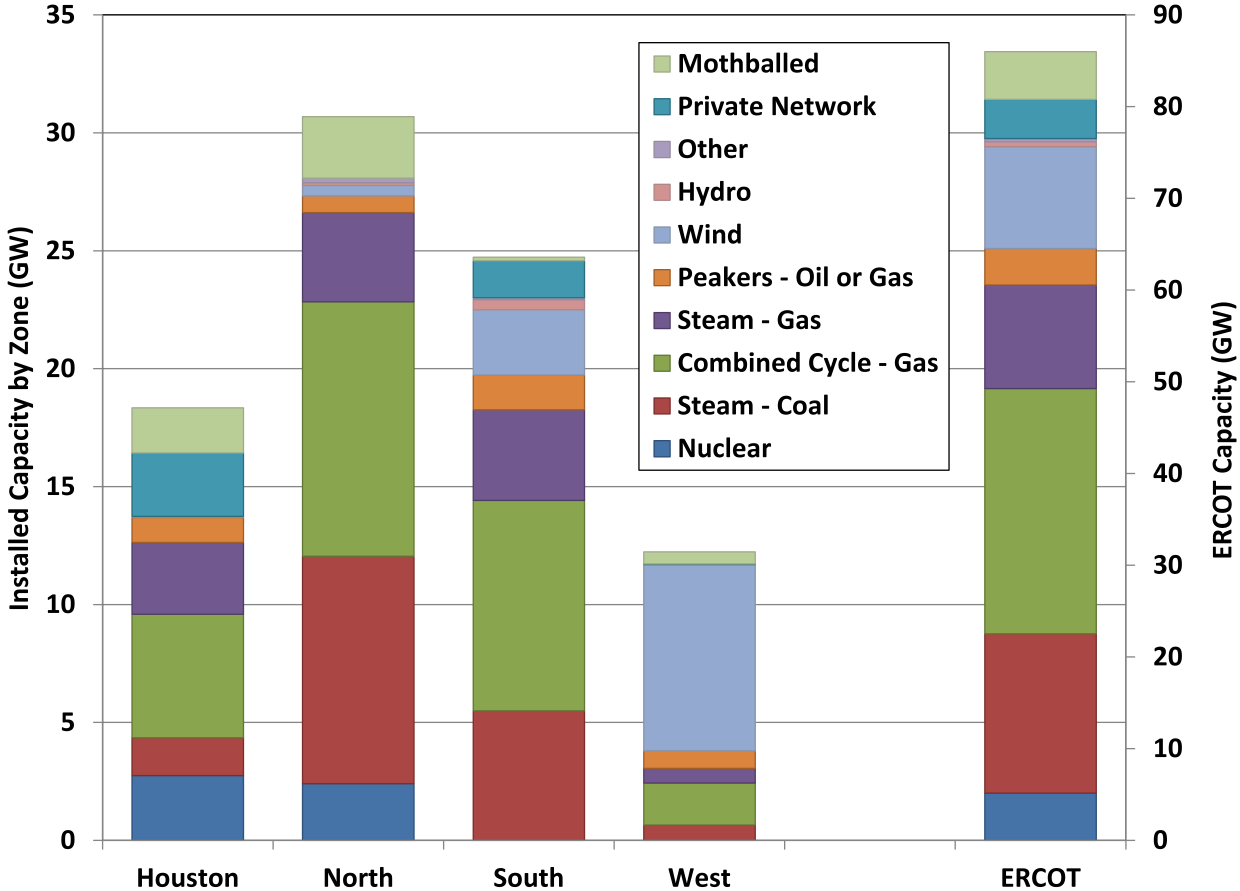


B. Generation Capacity in ERCOT

In this section we evaluate the generation mix in ERCOT. 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 36 percent of capacity, the South zone 29 percent, the Houston zone 21 percent, and the West zone 14 percent. The Houston zone typically imports power, while the West zone typically exports power. Excluding mothballed resources and including only 8.7 percent of wind capacity as capacity available to reliably meet peak demand, the North zone accounts for approximately 39 percent of capacity, the South zone 31 percent, the Houston zone 23 percent, and the West zone 6 percent. Figure 48 shows the installed generating capacity by type in each of the ERCOT zones.[14]

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

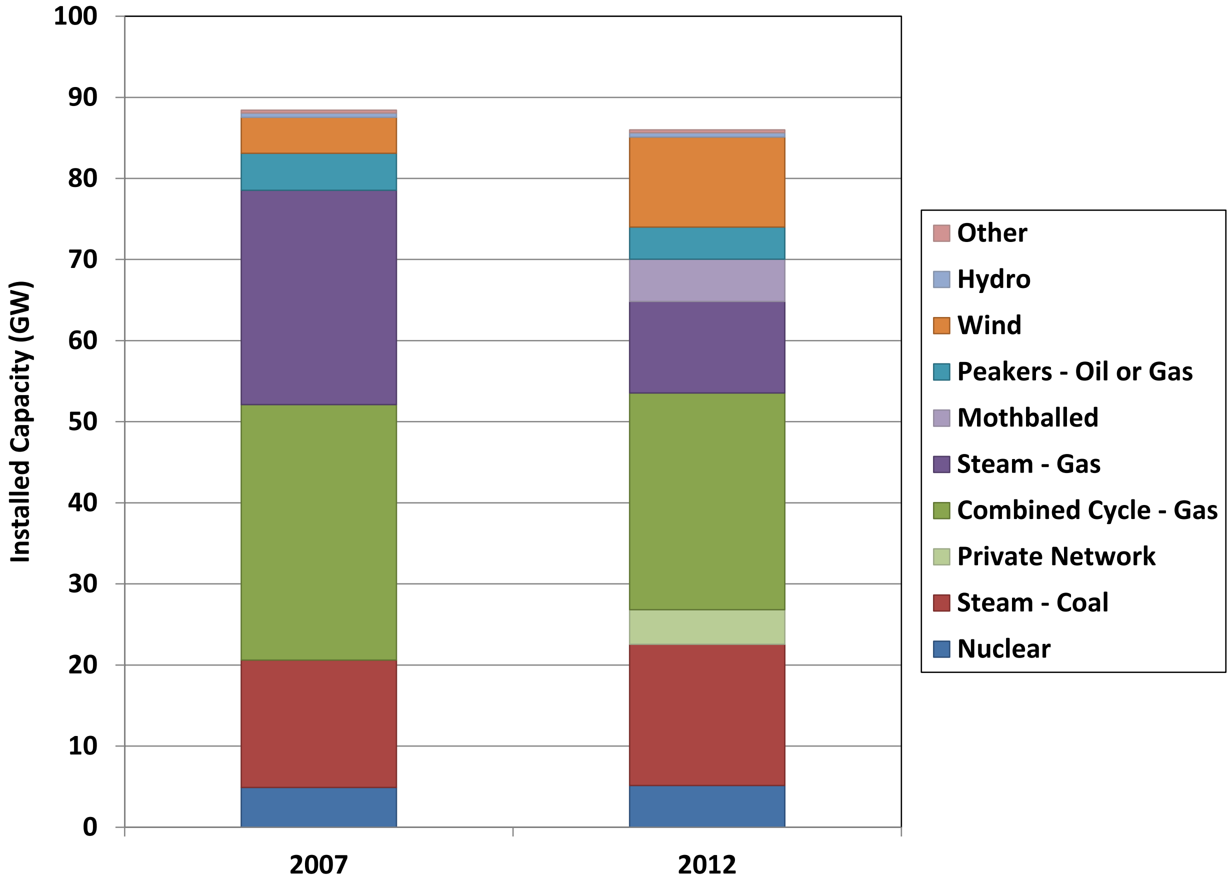
Figure 48: Installed Capacity by Technology for each Zone



Approximately 1 GW of generation resources came online in 2012; most of which were wind units and the remainder were solar and biomass. Even with no new natural gas units added during 2012, natural gas generation accounts for approximately 49 percent of total ERCOT installed capacity, and coal generation accounts for approximate 20 percent.

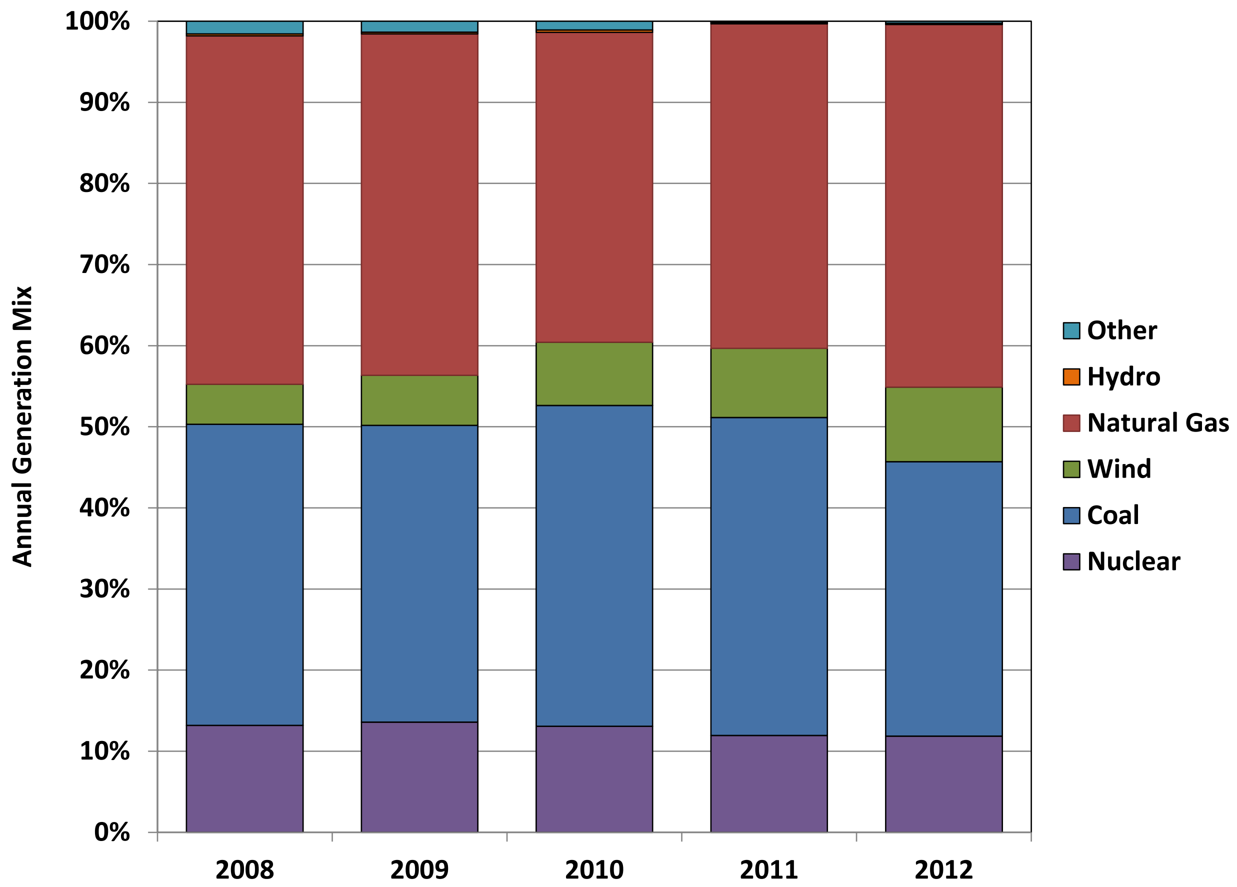
By comparing the current mix of installed generation capacity to that in 2007, as shown in Figure 49, we can see the effects of longer term trends. Over these six years, wind and coal generation are the two categories with the most increased capacity. The sizable additions in these two categories have been more than offset by retirements of natural gas-fired steam units, resulting in less installed capacity in 2012 than there was in 2007.

Figure 49: Installed Capacity by Type: 2007 to 2012



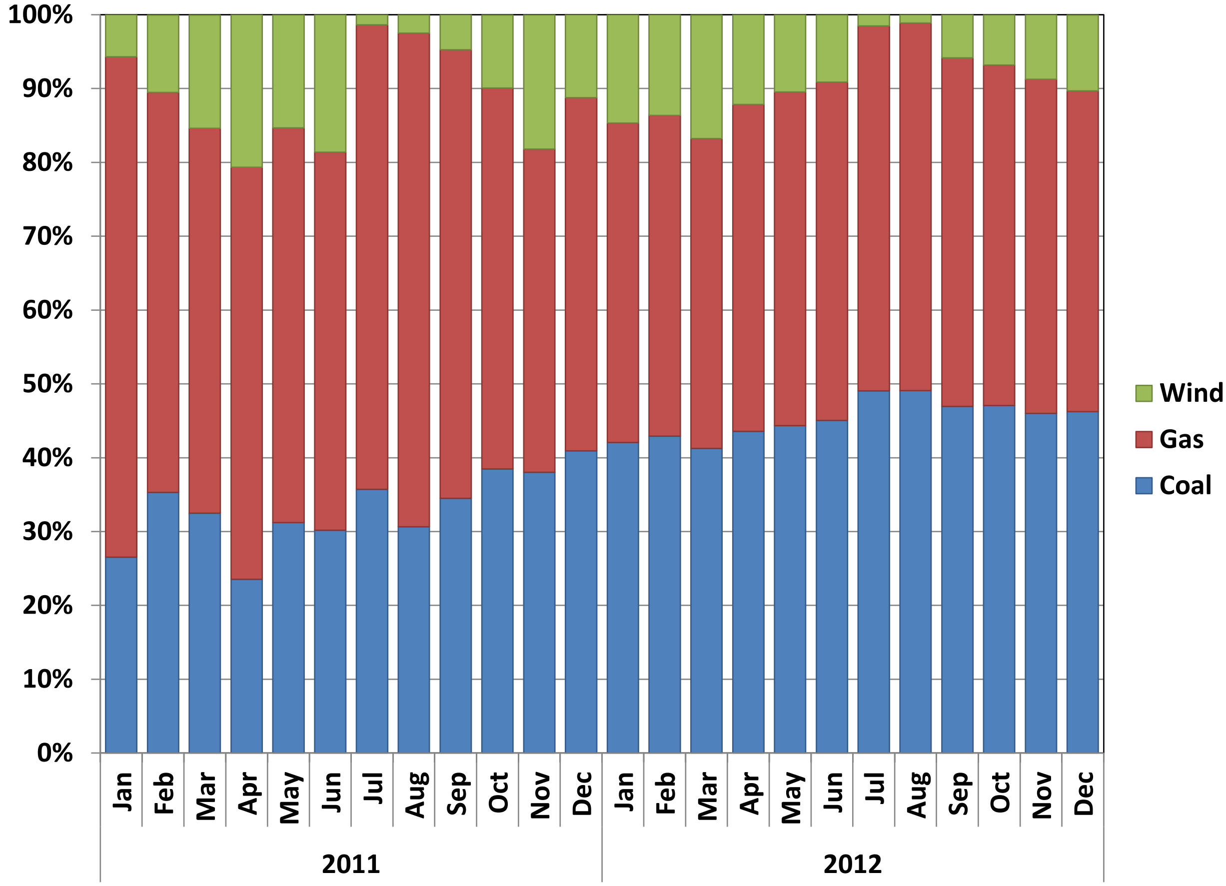
The shifting contribution of coal and wind generation is evident in Figure 50, which shows the percent of annual generation from each fuel type for the years 2008 through 2012. The generation share from wind has increased every year, reaching 9 percent of the annual generation requirement in 2012, up from 5 percent in 2008. During the same period the percentage of generation provided by natural gas decreased from 43 percent in 2008 to 38 percent in 2010, before increasing to 45 percent in 2012. Correspondingly, the percentage of generation produced by coal units increased from 37 percent to 40 percent in 2010 before decreasing to 34 percent in 2012. The increase in the share of generation produced by natural gas, and corresponding reduction in coal generation is due to historically low price of natural gas in 2012.

Figure 50: Annual Generation Mix



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 23 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. However, due to the low price of natural gas in 2012, we observe that the share of generation produced from coal-fired and nuclear units decreased to less than half of the energy in ERCOT, with the reduction coming from decreased coal generation. This reduction in the share of coal generation results in an increase in the occurrences when coal units were setting the price. This happens because the decrease in natural gas price results in those units becoming infra-marginal; that is, less costly than the last unit needed to satisfy total demand. As natural gas units are marginal less frequently, coal units increasingly become marginal. We can see the results of this tradeoff in Figure 51 which shows that the frequency with which coal was the marginal fuel was greater than 40 percent in all months during 2012, a noticeable increase from 2011.

Figure 51: Marginal Unit Frequency by Fuel Type



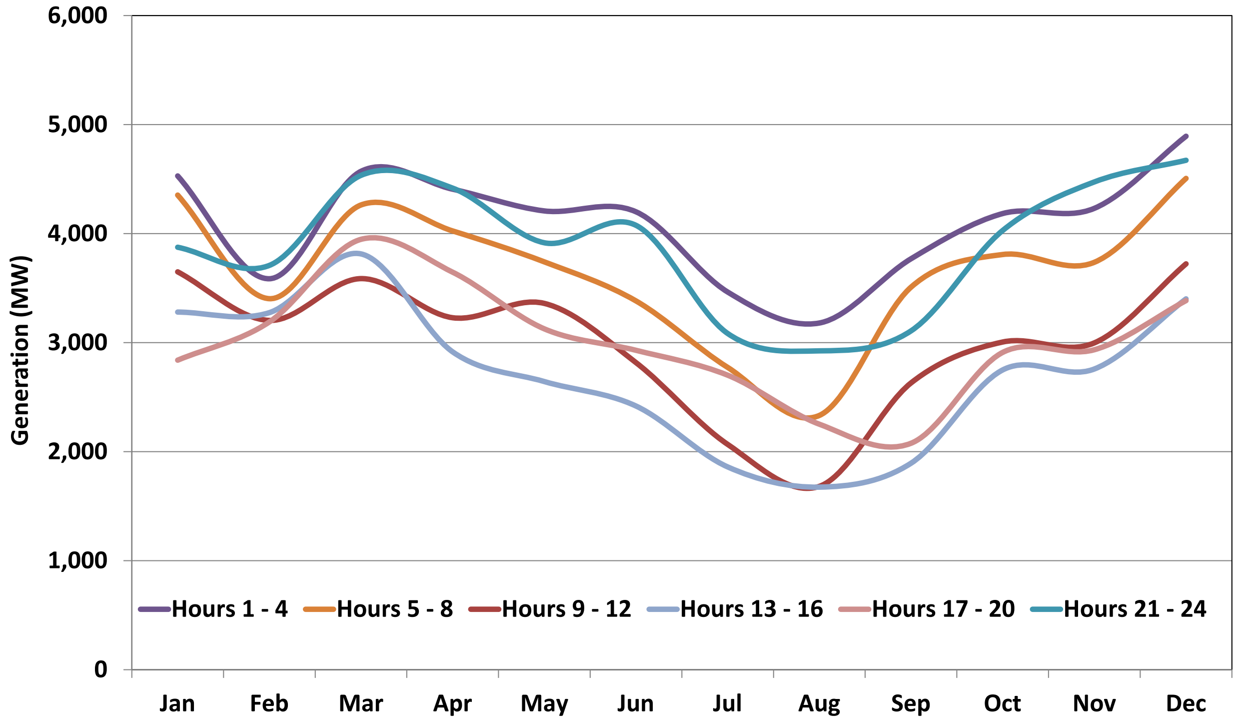
The methodology used in this analysis reflects the details of the unit specific dispatch that are available under the nodal market design. For every five-minute interval we determine which units are marginal, that is they are being dispatched and their offer price is contributing to the locational marginal price. When there is congestion, units with different prices can be marginal at the same time. With all the marginal units identified, we aggregate by their fuel type to compute monthly percentages. This aggregation ignores all locational price differences.

Although natural gas units continue to be marginal most of the time, the frequency at which coal units were marginal has steadily increased since late 2011. As previously discussed this can be attributed to the decline in natural gas prices. The frequency of wind units being marginal also declined in 2012. This can be attributed to the reduced necessity for wind curtailments due to transmission constraints.

1. Wind Generation

The amount of wind generation installed in ERCOT exceeded 10 GW by the end of 2012. Although the large majority of wind generation is located in the West zone, there has been more than 2 GW of wind generation recently installed in the South zone. Additionally, a private transmission line went into service in late 2010 allowing nearly another 1 GW of West zone wind to be delivered directly to the South zone. This section will more fully describe the characteristics of wind generation in ERCOT.

Figure 52: Average Wind Production

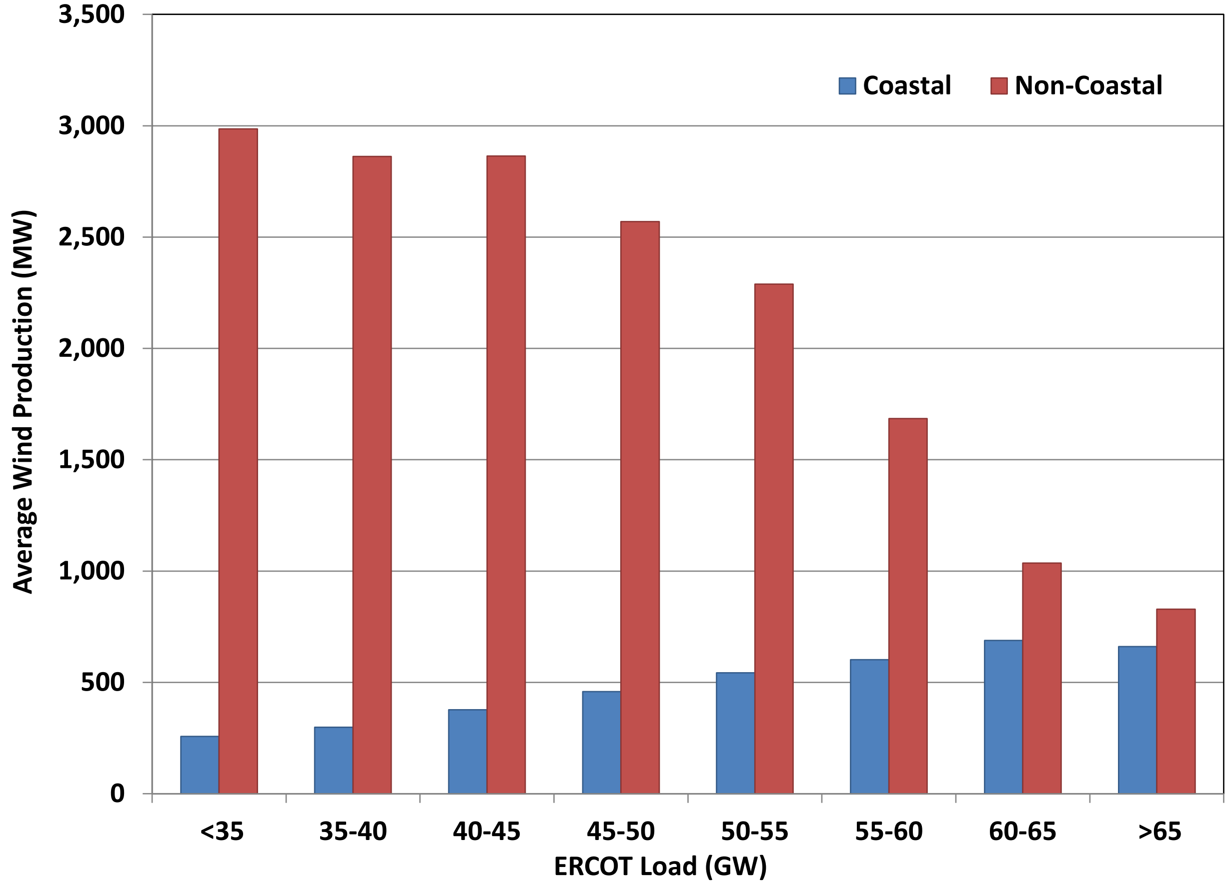


The average profile of wind production is negatively correlated with the load profile, with the highest wind production occurring during non-summer months, and predominately during off-peak hours. Figure 52 shows average wind production for each month in 2012, with the average production in each month shown separately in four hour blocks.[15]

15 Figure 52 shows actual wind production, which was affected by curtailments at the higher production levels. Thus, the higher levels of actual wind production in Figure 52 are lower than the production levels that would have materialized absent transmission constraints.

Examining wind generation in total masks the different wind profiles that exist for locations across ERCOT. Wind developers have more recently been attracted to the site facilities along the Gulf coast of Texas due to the higher correlation of winds with electricity demands. Next we compare the differences in output for wind units located in the coastal area of the South zone and those located elsewhere in ERCOT.

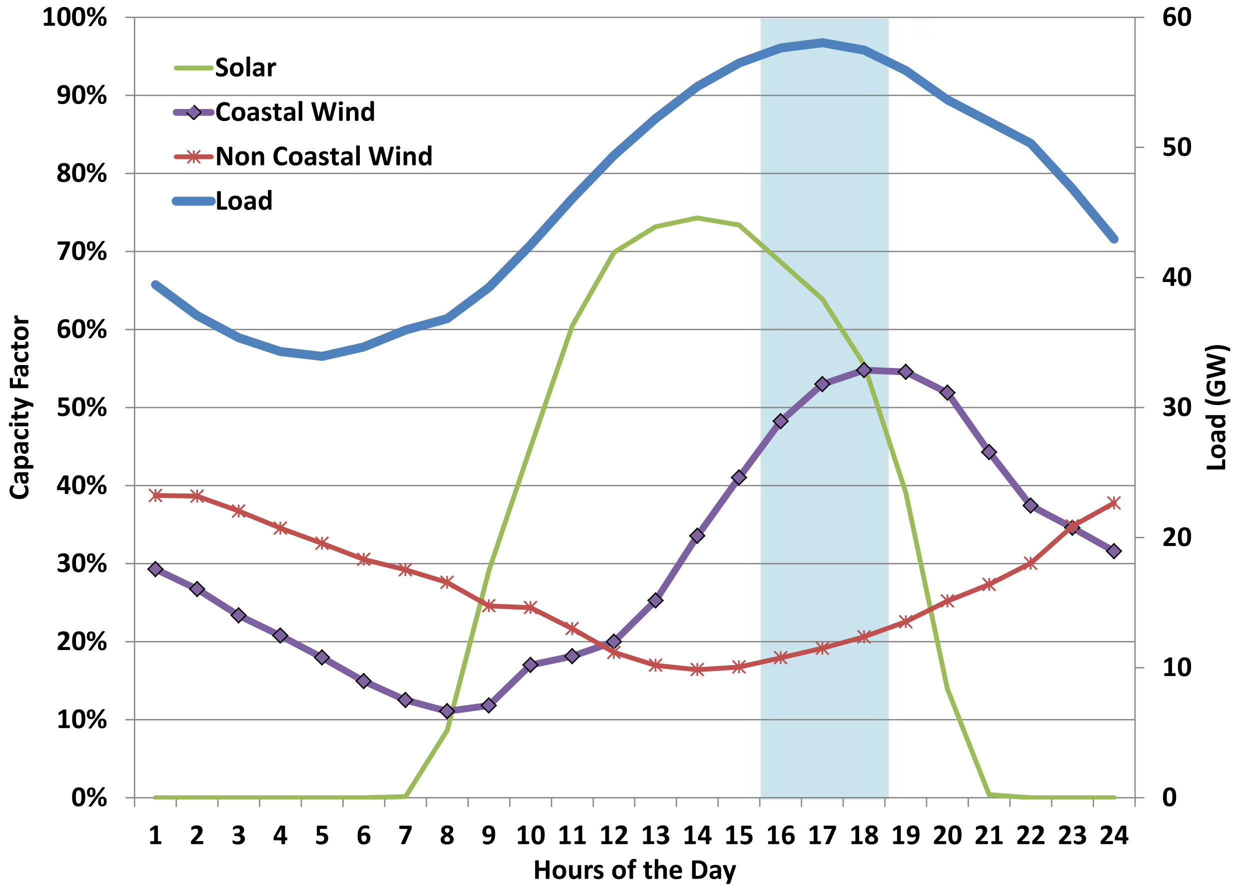
Figure 53: Summer Wind Production vs. Load



In Figure 53 data is presented for the summer months of June through August, comparing the average output for wind generators located in coastal and non-coastal areas in ERCOT across various load levels. It shows a strong negative relationship between non-coastal wind output 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.

The growing numbers of solar generation facilities in ERCOT also have an expected generation profile highly correlated with peak summer loads. Figure 54 below 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. The solar output shown is from relatively small central station photovoltaic facilities totaling approximately 50 MW. However, its production as a percentage of installed capacity is the highest, exceeding 70 percent in the early afternoon, and producing more than 50 percent of its installed capacity during peak.

Figure 54: Summer Renewable Production



The contrast between coastal wind and non-coastal wind is also clearly displayed in Figure 54. Coastal wind produced greater than 50 percent of its installed capacity during summer peak hours while output from non-coastal wind was approximately 20 percent.

Figure 55: Wind Production and Curtailment

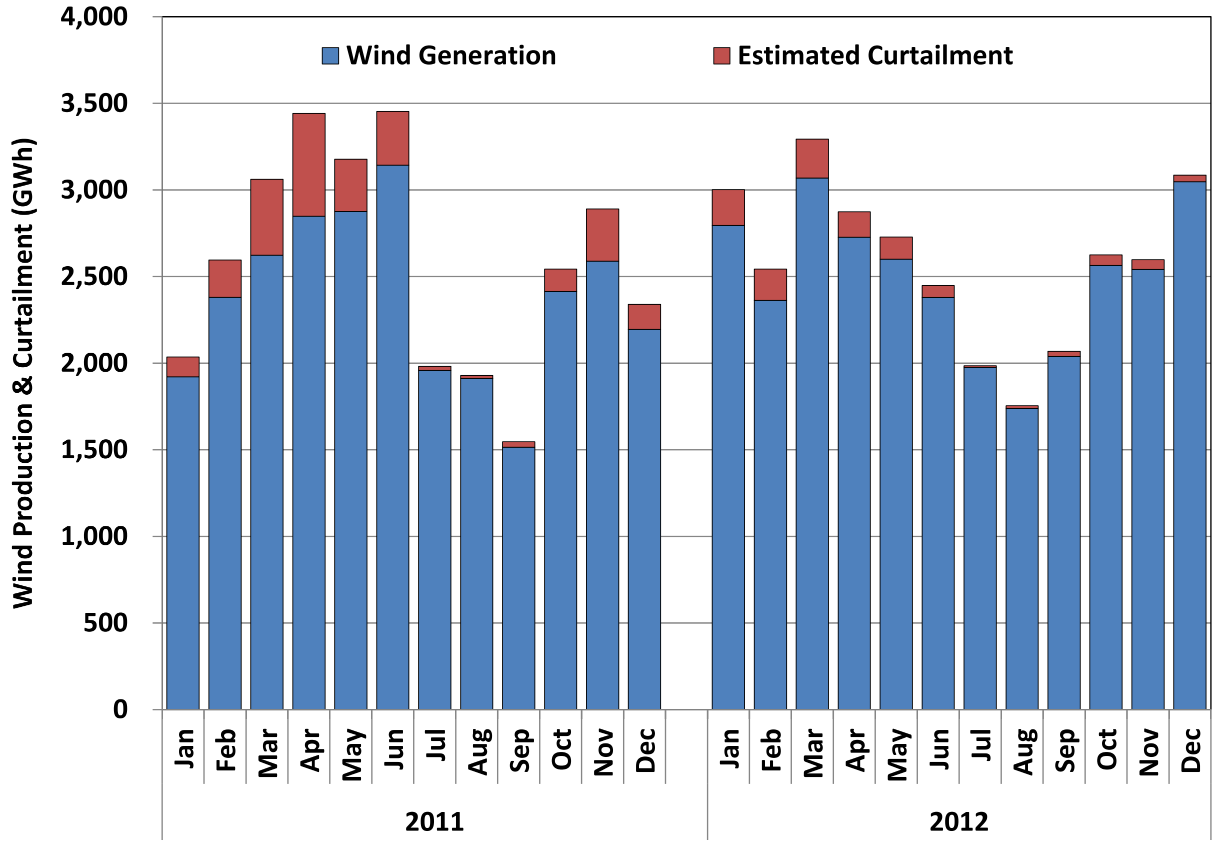
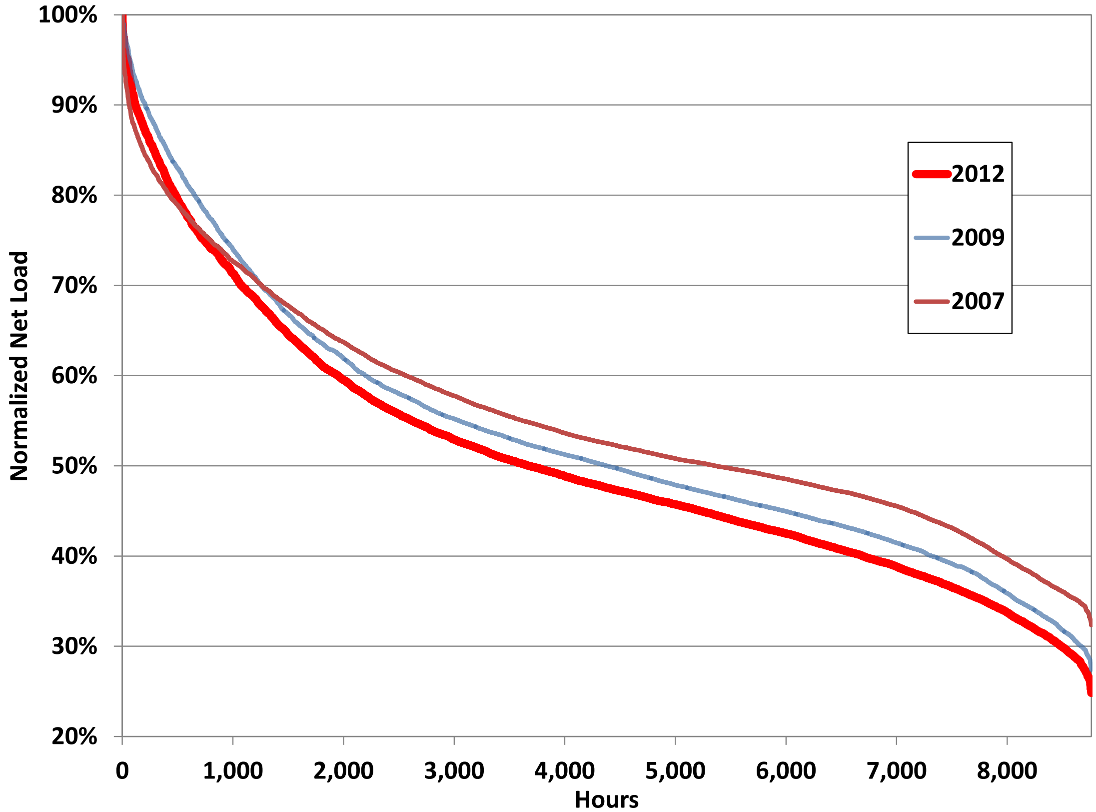


Figure 55 shows the wind production and estimated curtailment quantities for each month of 2011 and 2012. This figure reveals that the total production from wind resources increased again in 2012. More importantly, the quantity of curtailments was lower in 2012 when compared to 2011. The volume of wind actually produced was approximately 96 percent of the total available wind in 2012, up from approximately 92 percent in 2011.

Increasing levels of wind resource in ERCOT also has important implications for the net load duration curve faced by the non-wind fleet of resources. Net load is defined as the system load minus wind production. Figure 56 shows the net load duration curves for selected years since 2007, normalized as a percent of peak load.

Figure 56: Net Load Duration Curves

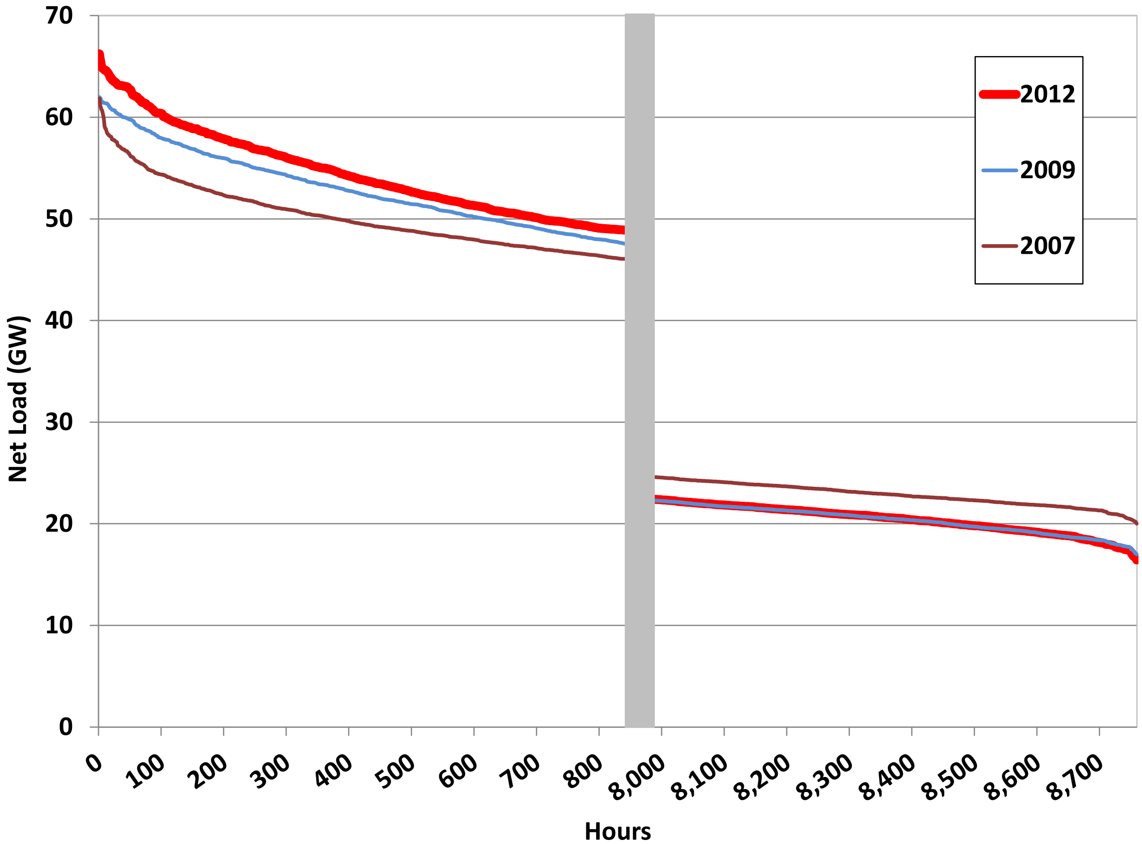


This figure shows the continued erosion of remaining energy available for non-wind units to serve during most hours of the year, with much less impact during the highest loads.

Even with the increased development activity in the coastal area of the South zone, more than 80 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 relatively 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 more significant reductions in the net load relative to the actual load in the other hours of the year. The trend shown from 2007 in Figure 56 may continue with the addition of new wind resources and the reduction in the curtailment of existing wind resources after completion of new transmission facilities.

Focusing on the left side of the net load duration curve shown in Figure 57, the difference between peak net load and the 95th percentile of net load has been between 9.5 and 12.5 GW for the previous six years.

Figure 57: Top and Bottom Ten Percent of Net Load



On the right side of the net load duration curve, the minimum net load has remained approximately 17 GW for the past four years, even with sizable growth in total annual load. This continues to put operational pressure on the nearly 25 GW of nuclear and coal-fired generation currently installed in ERCOT.

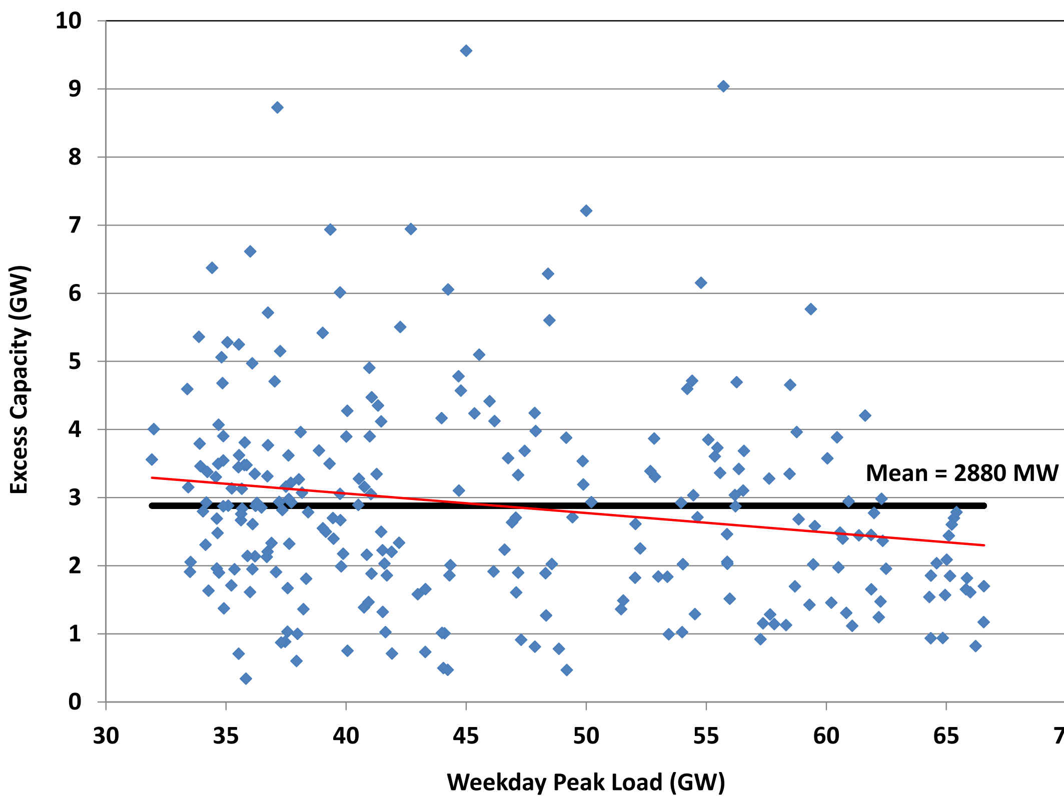
Thus, although the peak net load and reserve margin requirements are projected to continue to increase and create an increasing need for non-wind capacity to meet net load and reliability requirements, the non-wind fleet is expected to operate for fewer hours as wind penetration continues to increase. 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.

2. Daily Generator 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 real-time and inefficiently high energy prices while over-commitment can result in excessive start-up costs, uplift charges, and inefficiently low energy prices.

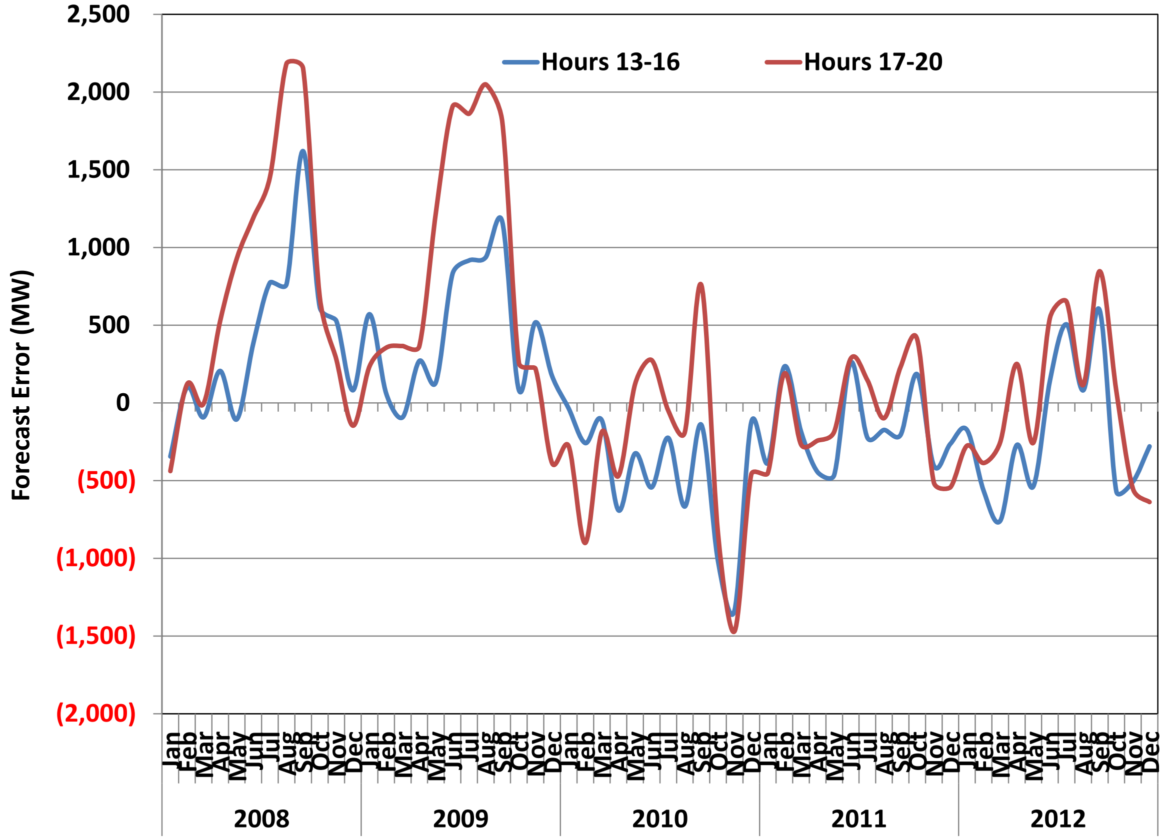
This subsection evaluates the commitment patterns in ERCOT by examining the levels of excess capacity. Excess capacity is defined as the total capacity of online plus quick-start generators minus the demand for energy, responsive reserve, up regulation and non-spinning reserve provided from online capacity or quick-start units. To evaluate the commitment of resources in ERCOT, Figure 58 plots the excess capacity compared to peak load during 2012.

Figure 58: Excess On-Line and Quick Start Capacity



The figure shows the excess capacity in only the peak hour of each weekday because the largest generation commitment usually occurs at the peak hour. Hence, one would expect larger quantities of excess capacity in other hours. Figure 58 shows that the excess on-line capacity during daily peak hours on weekdays averaged 2,880 MW in 2012 which is approximately 7.8 percent of the average load in ERCOT. These values did not change significantly from 2011, when the average excess on-line capacity was 2,901 MW, or 7.6 percent of the average load. Even with the expected improvements in unit commitment coming from having a day-ahead market, if ERCOT's day-ahead load forecast continued to show significant bias toward over-forecasting peak load hours,[16] we would expect to see over commitment of generation using non-market means.

Figure 59: Load Forecast Error



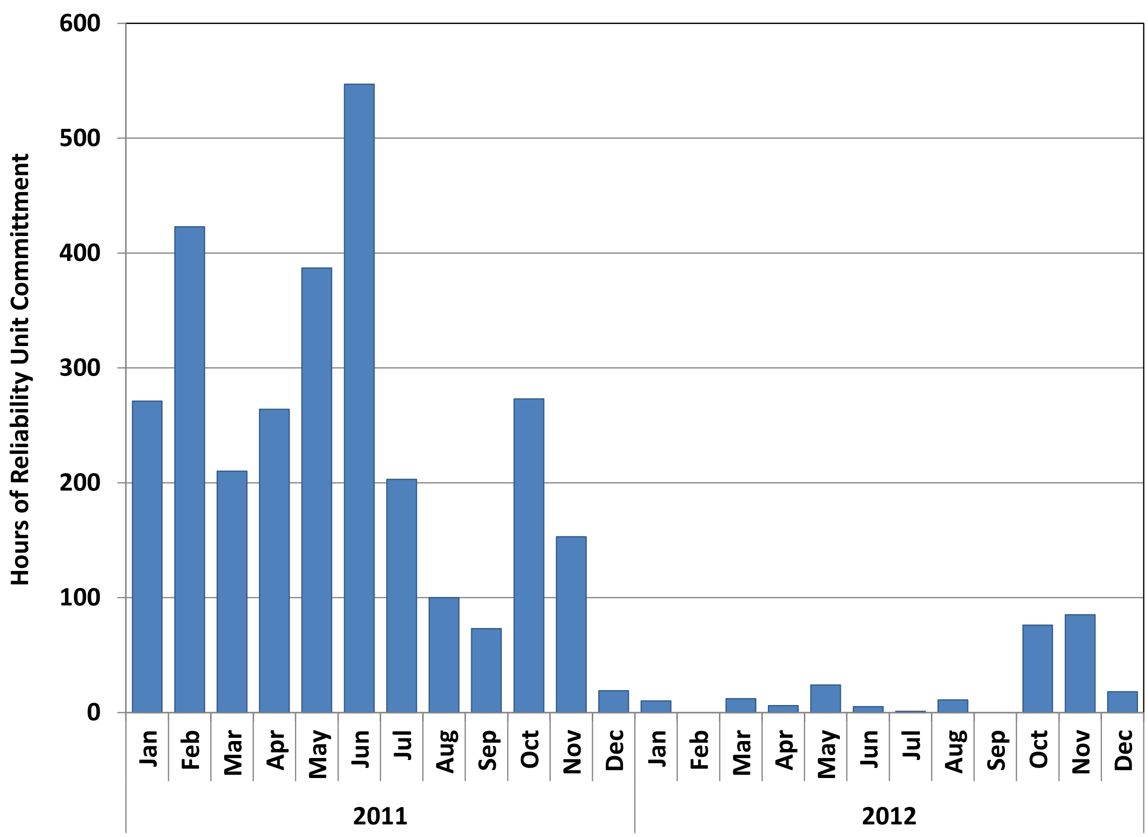
From Figure 59 we can see the noticeable reduction in ERCOT's load forecast bias since 2009. This was due to a procedure change implemented three years ago. ERCOT now specifically identifies and subtracts out the forecast bias and procures additional non-spin capacity in an equal amount.

Once ERCOT assesses the unit commitments resulting from the day-ahead market, additional capacity commitments are made, if needed, using a reliability unit commitment process that executes both on a day-ahead and hour ahead basis. These additional unit commitments may be

16 See 2010 ERCOT SOM report at pages 49-51 and 2009 ERCOT SOM report at pages 68-70.

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 transmission congestion. Figure 60 summarizes, by month, the number of hours with units committed via the reliability unit commitment process. We observe a significant reduction in the reliance upon the reliability unit commitment process in 2012 as compared to 2011. Approximately one third of the hours during 2011 had at least one unit committed by ERCOT through the reliability unit commitment process. The amount of time during 2012 reduced to three percent. The primary reason for the reduction is likely the less extreme weather and resulting lower load levels experienced during 2012. Lower loads resulted in reduced congestion and therefore reduced need for specific units to be brought online to resolve.

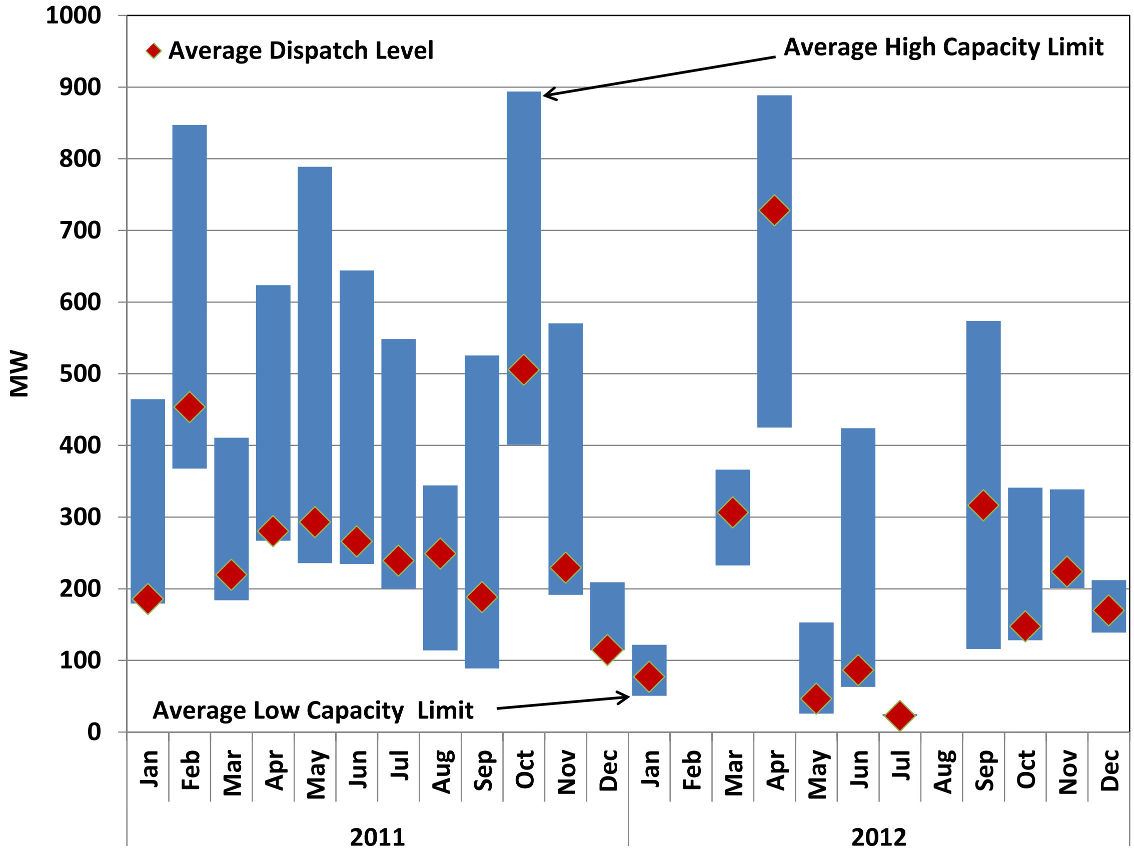
Figure 60: Frequency of Reliability Unit Commitments



There was an operational change midway through 2011 which also contributed to the reduced frequency of reliability unit commitments. During the initial months of operating the nodal market it was common for ERCOT to commit units that were providing non-spin reserves if they were needed to resolve congestion. This practice was greatly reduced starting in July 2011.

The next analysis compares the average dispatched output of the reliability committed units with their operational limits. Figure 61 shows that for most months of 2012 the amount of capacity actually dispatched from units brought on line via the reliability unit commitment process was less than the 200 to 300 MW that was typical for 2011.

Figure 61: Reliability Unit Commitment Capacity



The notable exception was in April 2012, when the large amounts of reliability unit committed capacity were related to brief generator outages resulting in reactive power deficiencies in the Dallas-Fort Worth area. This was similar to the situation that existed during October 2011. The larger quantity of committed capacity in February 2011 was a result of ERCOT operator action taken to attempt to ensure overall capacity adequacy during both the extreme cold weather event that occurred early in the month, and a subsequent bout of cold weather that occurred one week later.

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 the system's demands and reliability needs. We begin this section with an evaluation of these economic signals by estimating the "net revenue" new resources would receive from the markets. Next, our review of the effectiveness of the Public Utility Commission's Scarcity Pricing Mechanism includes two recommendations for market design improvements. We conclude this section with a review of the contributions from demand response toward meeting resource adequacy objectives in ERCOT and our third recommended improvement.

A. Net Revenue Analysis

Net revenue is the total revenue that can be earned by a new 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 energy and ancillary services markets together provide the economic signals that inform suppliers' decisions to invest in new generation or retire existing generation. In long-run equilibrium, markets should provide sufficient net revenue to allow an investor to receive a return of, and on an investment in a new generating unit that is needed. In the short-run, if the net short-run revenues produced by the market are not sufficient to justify entry, then one or more of these conditions exist:

- New capacity is not needed because there is sufficient generation already available;
- Load levels, and thus energy prices, are temporarily low due to mild weather or economic conditions;
- Market rules or operational practices are causing revenues to be reduced inefficiently; or
- Market rules are not sufficiently linked in short-term operations to ensure long-term resource adequacy objectives are met.

Likewise, the opposite would be true if the markets provide excessive net revenues in the short-run. The persistence of excessive net revenues in the presence of a capacity surplus is an indication of competitive issues or market design flaws. In this section, we analyze the net revenues that would have been received by various types of generators in each zone.

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 the bilateral energy prices over time and are appropriate to use for this evaluation. For purposes of this analysis, heat rates of 7 MMBtu per MWh for a combined cycle unit, 10.5 MMBtu per MWh for a combustion turbine, and 9.5 MMBtu per MWh for a new coal unit were assumed. Variable operating and maintenance costs of $4 per MWh for the natural gas units and $5 per MWh for the coal unit and fuel and variable operating and maintenance costs of $8 per MWh for the nuclear unit were assumed. For purposes of this analysis, a total outage rate (planned and forced) of 10 percent was assumed for each technology.

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 masks what may be very high locational values for a specific generator location. Some generators may also receive uplift payments because of their specific reliability contributions, either as a reliability must run, or through the reliability unit commitment. This source of revenue is not considered in this analysis. The analysis also includes simplifying assumptions that can lead to over-estimates of the profitability of operating in the wholesale market. The following factors are not explicitly accounted for in the net revenue analysis: (i) start-up costs, which can be significant; and (ii) minimum running times and ramp restriction, which can prevent the natural gas generators from profiting during brief price spikes. Despite these limitations, the net revenue analysis provides a useful summary of signals for investment in the wholesale market.

Figure 62 shows the results of the net revenue analysis for four types of hypothetical new units in 2011 and 2012. These are: (a) natural gas-fired combined-cycle, (b) natural gas-fired combustion turbine, (c) coal-fired generator, and (d) a nuclear unit. For the natural gas-fired technologies, 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 in other hours that it is available (*i.e.*, when it is not experiencing a planned or forced outage). For coal and nuclear technologies, net revenue is calculated by assuming that the unit will produce at full output.

Figure 62: Estimated Net Revenue by Zone and Unit Type

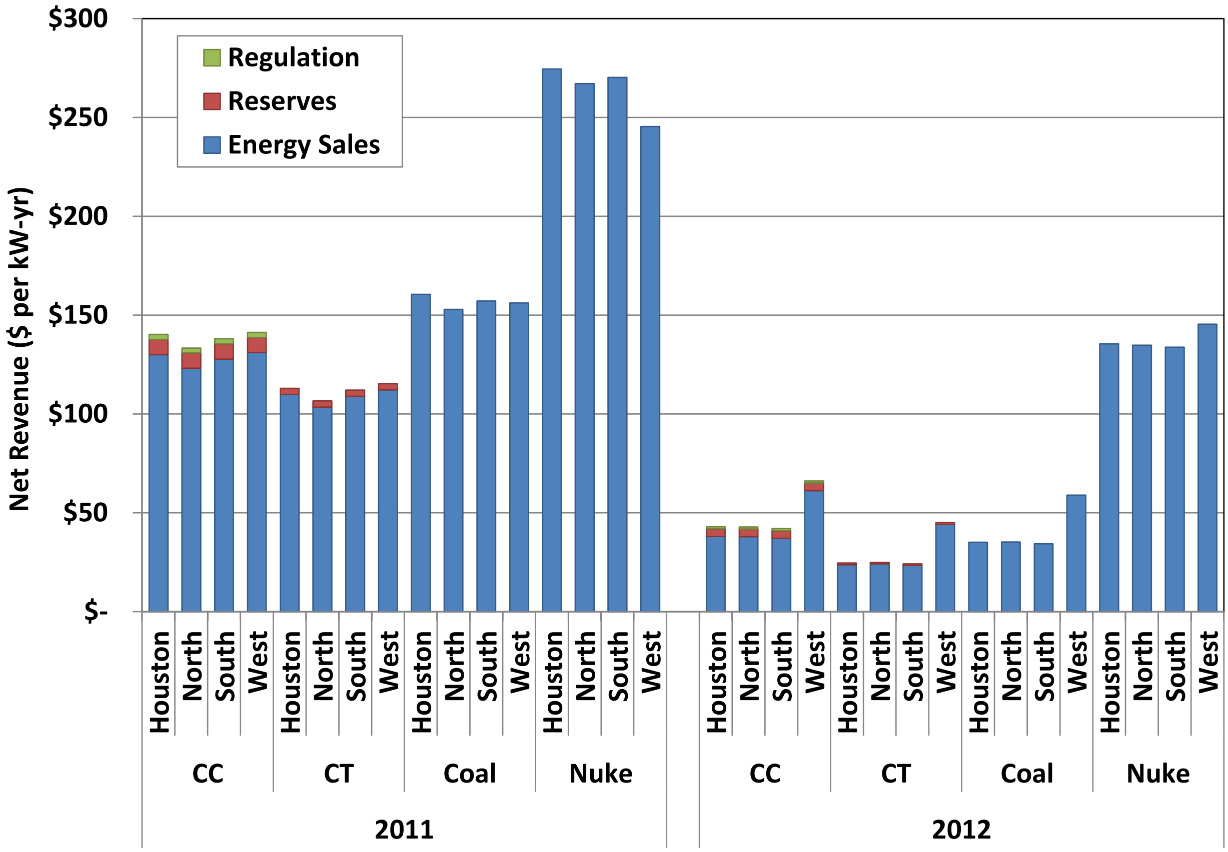


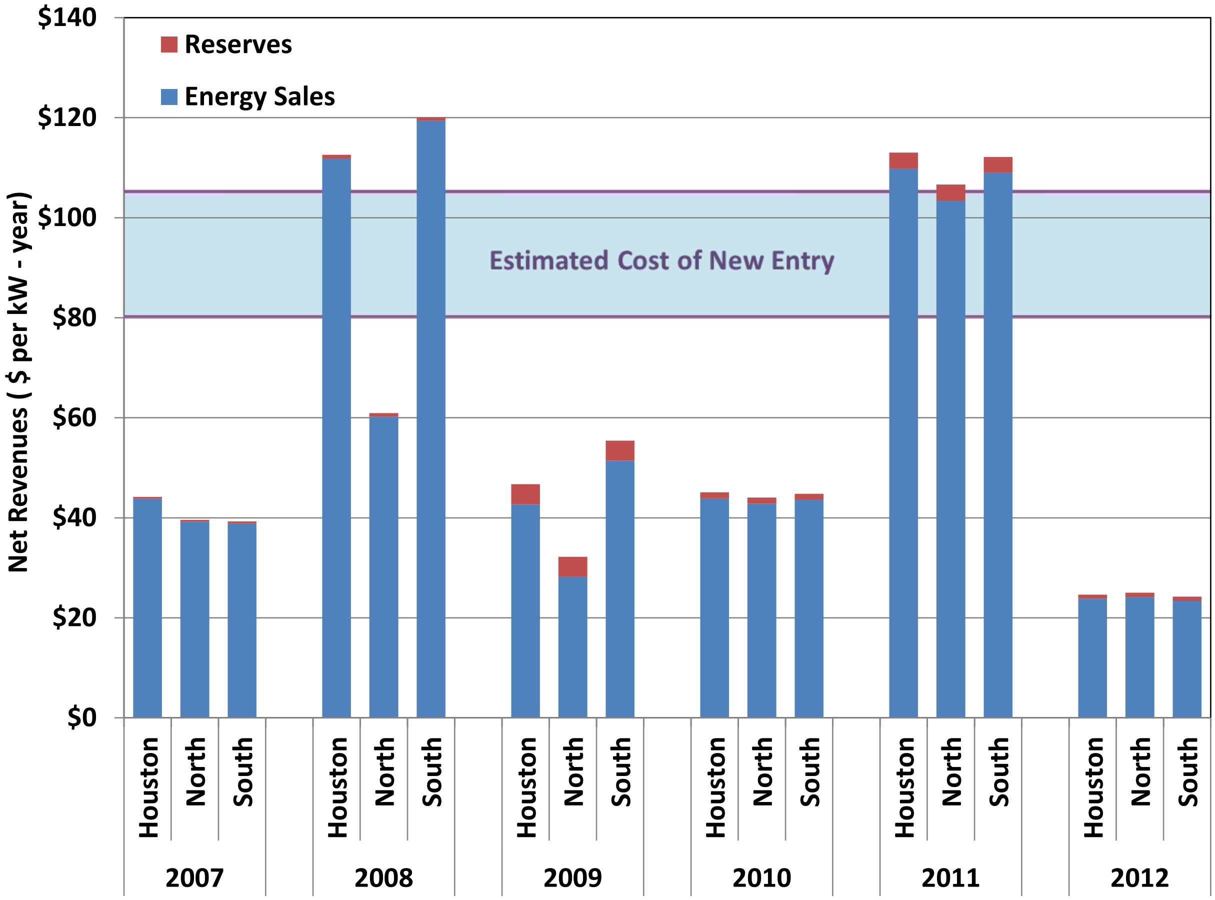
Figure 62 shows that the net revenue for every generation technology type decreased substantially in 2012 compared to each zone in 2011.

- For a new coal unit, the estimated net revenue requirement is approximately $210 to $270 per kW-year. The estimated net revenue in 2012 for a new coal unit was approximately $35 per kW-year.
- For a new nuclear unit, the estimated net revenue requirement is approximately $280 to $390 per kW-year. The estimated net revenue in 2012 for a new nuclear unit was approximately $134 per kW-year.

Prior to 2005, net revenues were well below the levels necessary to justify new investment in coal and nuclear generation. Higher natural gas prices through 2008 allowed energy prices to remain at levels high enough to support new entry for these technologies. The production costs of coal and nuclear units did not change significantly over this period, leading to a dramatic rise in net revenues. However, since 2008 natural gas prices have been on the decline, resulting in reduced net revenues for coal and nuclear technologies. Even with the higher energy prices experienced in 2011, net revenues for these technologies were insufficient to support new entry. With the further decline in natural gas prices and few occurrences of shortage pricing, the estimated net revenue for either a new coal or a nuclear unit in ERCOT was well below the levels required to support new entry in 2012.

The next two figures provide an historical perspective of the net revenues available to support new gas turbine (Figure 63) and combined cycle generation (Figure 64).

Figure 63: Gas Turbine Net Revenues

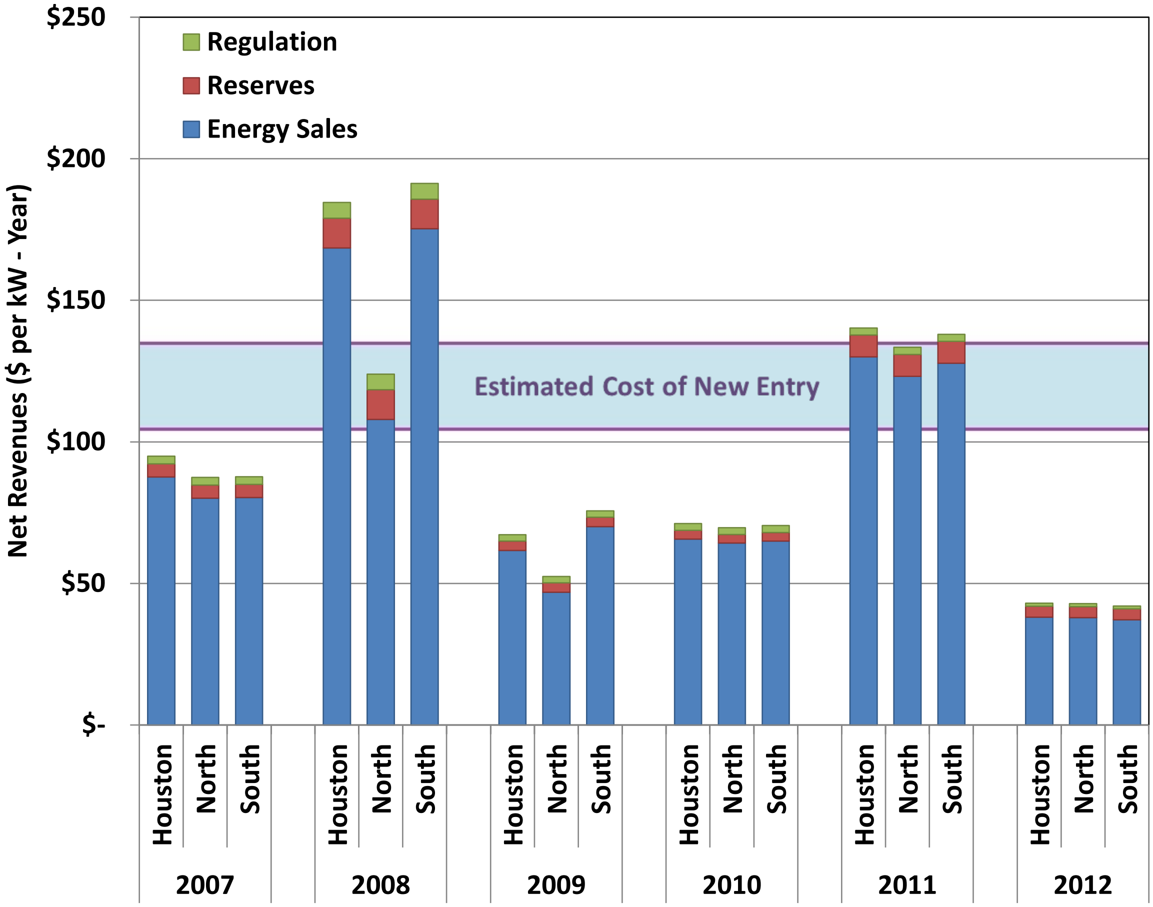


Based on our estimates of investment costs for new units, the net revenue required to satisfy the annual fixed costs (including capital carrying costs) of a new gas turbine unit ranges from $80 to

$105 per kW-year. The estimated net revenue in 2012 for a new gas turbine was approximately $25 per kW-year, far below the levels required to support new gas turbine generation.

For a new combined cycle unit, the estimated net revenue requirement is approximately $105 to $135 per kW-year. The estimated net revenue in 2012 for a new combined cycle unit was approximately $42 per kW-year, also far below the levels to support new combined cycle generation in ERCOT.

Figure 64: Combined Cycle Net Revenues



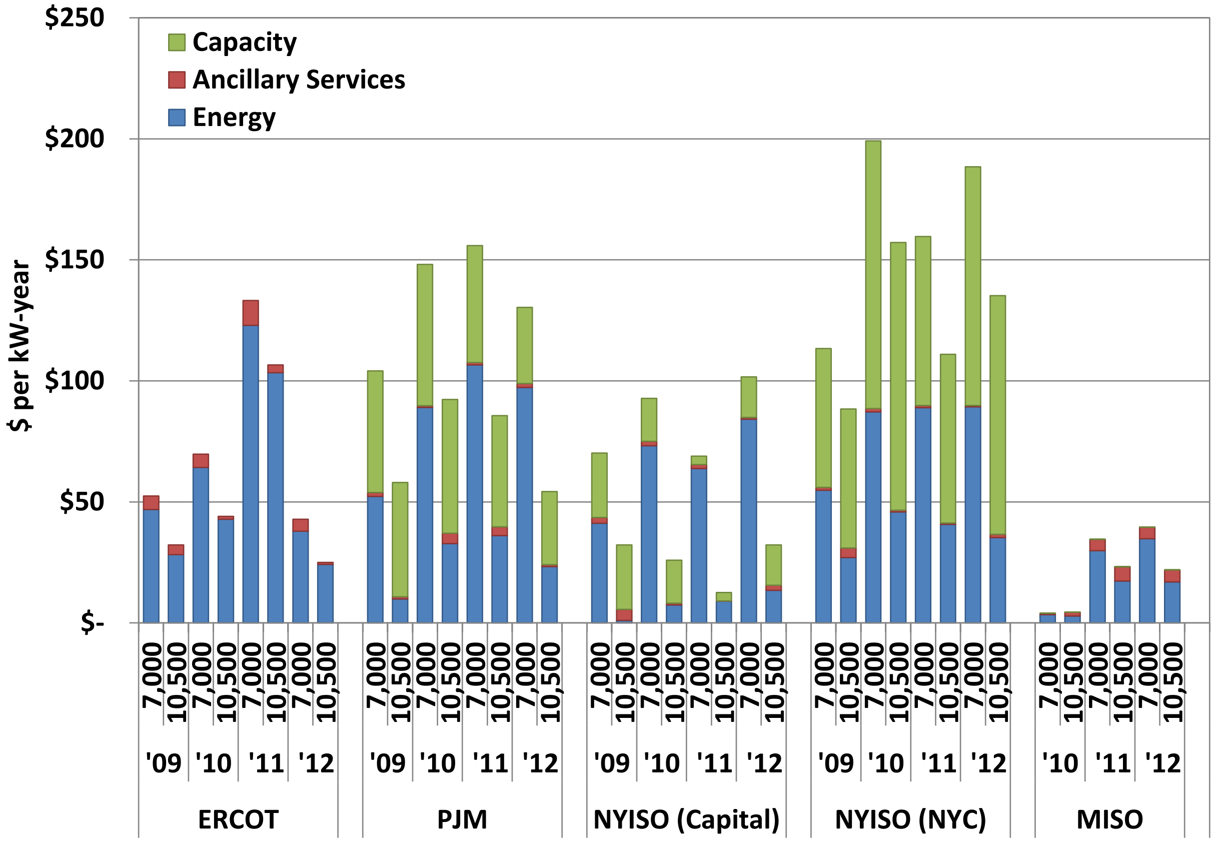
Even though net revenues for the Houston and South zones in 2008 may have appeared to be sufficient to support new natural gas-fired generation, it was actually extremely inefficient transmission congestion management and inefficient pricing mechanisms associated with the deployment of non-spinning reserves which led to high prices and resulting higher than warranted net revenues. Discounting the effect that the 2008 results would have had on forward price signals, we find that 2011 was the first time in five years that net revenues have been sufficient to support either new gas turbine or combined cycle generation.

These results indicate that the ERCOT markets would not have provided sufficient revenues to support profitable investment in any of the types of generation technology evaluated. Higher energy prices in the West zone during 2012 resulted in higher net revenues in that zone, but they were still not high enough to support new entry there. The net revenues in 2012 were much lower than in 2011. However, it is important to recognize that 2011 was highly anomalous, with some of the hottest summer temperatures on record. Net revenues may have been sufficient to cover the costs of a new combined cycle or new combustion turbine in 2011, however, we would not expect this to be consistently true in years with comparable reserve margins absent the extreme weather conditions, as evidenced by the 2012 net revenue results.

To provide additional context for the net revenue results presented in this section, we also compared the net revenue in the ERCOT market for two types of natural gas-fired technologies with the net revenue that those technologies could expect in other wholesale markets with centrally cleared capacity markets. The technologies are differentiated by their assumed heat rate; 7,000 MMBtu per MWh for combined cycle and 10,500 MMBtu per MWh for simple-cycle combustion turbine.

Figure 65 compares estimates of net revenue for the ERCOT North zone, PJM, two locations within the New York ISO, and the Midwest ISO. The figure includes estimates of net revenue from energy, reserves and regulation, and capacity. ERCOT does not have a capacity market, and thus, does not have any net revenue from capacity sales. Figure 65 shows that net revenues increased from 2011 to 2012 for both technologies in NY ISO. In PJM net revenue decreased and in Midwest ISO net revenues remained flat. In the figure below, net revenues are calculated for central locations. However, there are load pockets within each market where net revenue and the cost of new investment may be higher. Thus, even if new investment is not generally profitable in a market, it may be economic in certain areas. Finally, resource investments are driven primarily by forward price expectations, so historical net revenue analyses do not provide a complete picture of the future pricing expectations that will spur new investment.

Figure 65: Comparison of Net Revenue of Gas-Fired Generation between Markets



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. These rules include a Scarcity Pricing Mechanism ("SPM") that relaxed the existing system-wide offer cap by increasing it in multiple steps until it reached $3,000 per MWh shortly after the implementation of the nodal market. PUCT SUBST. R. 25.505 provides that the IMM may conduct an annual review of the effectiveness of the SPM. This subsection provides an assessment of the effectiveness of the SPM in 2012 under ERCOT's energy-only market structure.

Experiencing reduced levels of generation development activity coupled with higher than expected loads resulting in diminishing planning reserve margins, the PUCT has devoted considerable effort over the past two years deliberating issues related to resource adequacy. These deliberations have included the question of whether that planning reserve margin is a target or a minimum requirement. Further, if it is a minimum requirement, whether the energy-only market design can ensure the desired reliability level or whether an alternate market design mechanism may be required. To date, the PUCT has taken no action to change the fundamental energy-only nature of the ERCOT market or the designation of the planning reserve margin as a target. However, there have been changes to the rules governing the system-wide offer cap and peaker net margin mechanism.

Approved during 2012, new PUCT SUBST. R. 25.508 increased the system-wide offer cap to $4,500 per MWh effective August 1, 2012. As shown in Figure 15 on page 15, there was only a brief period when energy prices rose to the cap after this change was implemented. Revisions were also adopted to PUCT SUBST. R. 25.505 which specified future increases to the system-wide offer cap as follows:

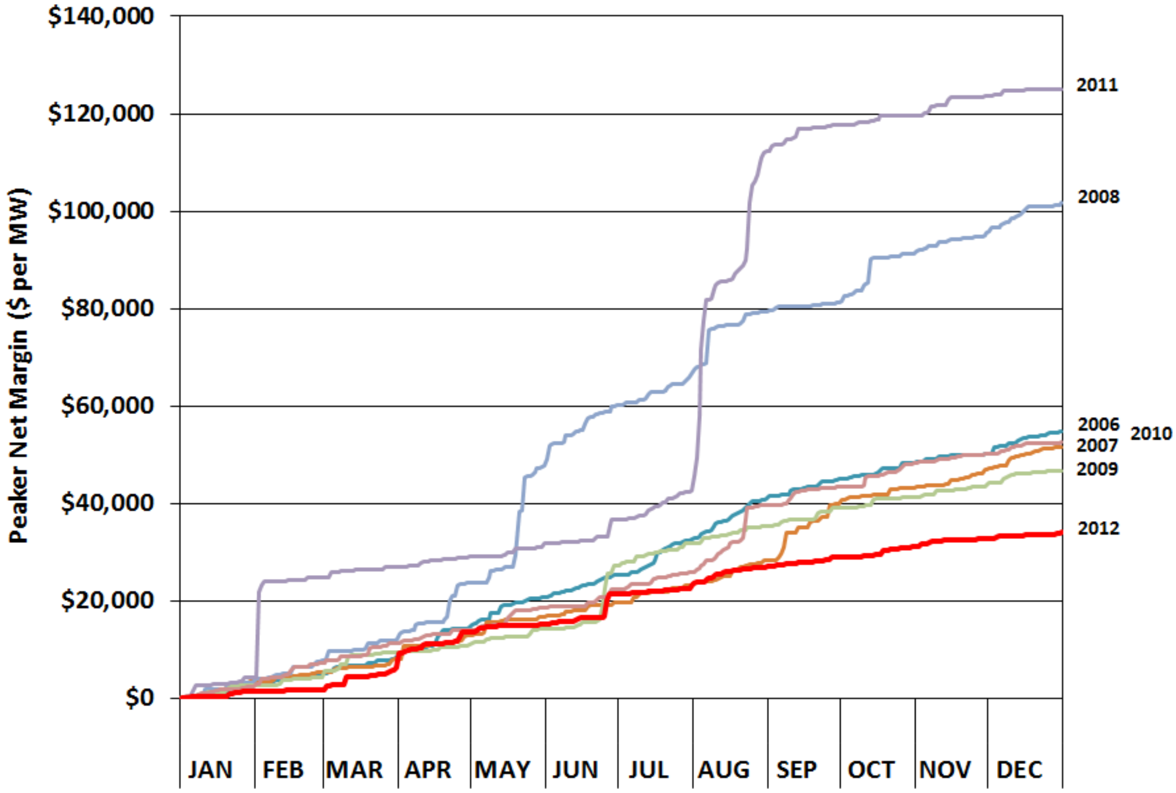
- $5,000 per MWh beginning on June 1, 2013,
- $7,000 per MWh beginning on June 1, 2014, and
- $9,000 per MWh beginning on June 1, 2015.

The SPM includes a provision termed the Peaker Net Margin ("PNM") that is designed to measure the annual net revenue of a hypothetical peaking unit. This aspect of the rule was also amended in 2012. Under the current rule, if the PNM for a year reaches a cumulative total of $300,000 per MW,[17] 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.[18] Figure 66 shows the cumulative PNM results for each year from 2006 through 2012 and shows that PNM in 2012 was the lowest it has been since its implementation.

[17] 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.

[18] These values were increased from a previous threshold of $175,000 per MW and an LCAP of $500 per MWh.

Figure 66: Peaker Net Margin



As previously described, the net revenue required to satisfy the annual fixed costs (including capital carrying costs) of a new gas turbine unit ranges from $80,000 to $105,000 per MW-year. Thus, as shown in Figure 66 and consistent with the previous findings in this section relating to net revenue, the PNM was nowhere near sufficient to support new entry in 2012. Only in two of the seven years since the rule was implemented has the PNM been sufficient – 2008 and 2011. A significant portion of the net revenue increase in 2008 was associated with extremely inefficient transmission congestion management and inefficient pricing mechanisms associated with the deployment of non-spinning reserves.[19] With these issues addressed in the zonal market, the PNM dropped substantially in 2009 and 2010. The extreme weather experienced in 2011 was highly anomalous. Hence, although the PNM may have been sufficient to cover the costs of a new combustion turbine in 2011, we would not expect this to be true on a continuous basis into the future.

19 See 2008 ERCOT SOM Report at pages 81-87.

Shortage Pricing and Resource Adequacy

Efficient electricity markets allow energy prices to rise substantially at 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, which is discussed in the next subsection.

The expectation of competitive energy market outcomes is no different in energy-only than in markets that include a capacity market. However, capacity markets are designed to ensure a specified planning reserve margin, which may be higher than what an energy-only market would achieve. Under this condition the higher planning reserve margin will serve to reduce the frequency of shortages in the energy market.

As a general principle, competitive and efficient market prices should be consistent with the marginal cost of the marginal action taken to satisfy the market’s demand. In the vast majority of hours, the marginal cost of the marginal action is associated with the dispatch of the last generator required to meet demand. It is appropriate and efficient in these hours for this generator to “set the price”. However, this is not true under shortage conditions. When the system is in shortage, the demand for energy and minimum operating reserves cannot be satisfied with the available resources, which will cause the system operator to take one or more of the following actions:

- Sacrifice a portion of the operating reserves by dispatching them for energy;
- Voluntarily curtail load through demand response programs;
- Curtail exports or make emergency imports; or
- Involuntarily curtail load.

A market design that adheres to the pricing principles stated above will set prices that reflect each of these actions. When there is a shortage of supply in the market, the marginal action first taken by the system operator is generally to not satisfy operating reserves requirements (*i.e.*, dispatching reserves for energy). Diminished operating reserves results in diminished reliability, which has a real cost to electricity consumers. In this case, the value of the foregone reserves – which is much higher than the marginal cost of the most expensive online generator – should be reflected in energy prices to achieve efficient economic signals governing investment in generation, demand response and transmission.

With the implementation of the nodal market, more reliable and efficient shortage pricing has been achieved by establishing pricing rules that recognize when operating reserve shortages exist and allowing energy prices to rise automatically. Figure 16 on page 16 clearly shows this relationship between increasing prices as operating reserve levels decline. This approach is more reliable than what existed in the previous zonal market because it is not dependent upon the submission of high-priced offers by small market participants to be effective. It is also more efficient during the vast majority of time in which shortage conditions do not exist because it is not necessary for market participants to effectively withhold lower cost resources by offering relatively small quantities at prices dramatically higher than their marginal cost. At times when there is insufficient capacity available to meet both energy and minimum operating reserve requirements, all available capacity will be dispatched and the clearing price will rise in a predetermined manner to a maximum of the system-wide offer cap.

Although the nodal market implementation brought about more reliable and efficient shortage pricing, there remain aspects of the ERCOT real-time energy pricing that can be improved. As discussed later in this section, prices during the deployment of load resources do not reflect the value of reduced reliability which occurs when responsive reserves have been converted to energy.

Similarly, during the first year of nodal market operation when non-spinning reserves were deployed (converted to energy), prices rarely reflected the marginal cost of the action being taken. Non-spinning reserves are provided primarily by off-line natural gas-fired combustion turbines capable of starting in 30 minutes or less. The implementation of the nodal market significantly increased market efficiencies in a number of areas, including the move to a five minute rather than 15 minute energy dispatch. However, it lacked an efficient economic commitment mechanism for resources such as offline gas turbines and other resources that are not immediately dispatchable in the five minute energy dispatch. This led to prices that were inefficiently low because they did not represent the costs associated with starting and running the gas turbines that were being deployed to meet demand.

The changes described in NPRRs 426 and 428 were implemented in early 2012, and added requirements for providers of non-spinning reserve to make that capacity available to ERCOT's dispatch software, subject to certain offer floors.[20] Providers are now able to specify the price at which they are willing to convert their non-spinning reserve capacity to energy. Further, ERCOT uses this price information to determine which non-spin units to deploy. Real-time energy price formation has been improved, but the current mechanism is sub-optimal from a reliability and efficiency perspective. We continue to recommend that ERCOT develop a mechanism that will rationally commit generation that can start or load resources that can curtail within 30 minutes.

This deficiency in ERCOT's nodal market design should be addressed by implementing "look ahead" dispatch functionality for the real-time energy market to produce energy and ancillary services commitment and dispatch results that are co-optimized and recognize anticipated changes in system demands.[21] This additional functionality represents a major change to ERCOT systems; one we recommend together with improved pricing provisions that will allow offers from load resources to set prices if they are required to meet system demand.

Effective look ahead dispatch functionality should also reduce the price dampening effects of energy produced by units operating below their low sustainable operating limit. Although alternatives have been suggested to address this issue in a standalone manner, we believe the better approach will be to develop a comprehensive look ahead dispatch solution.

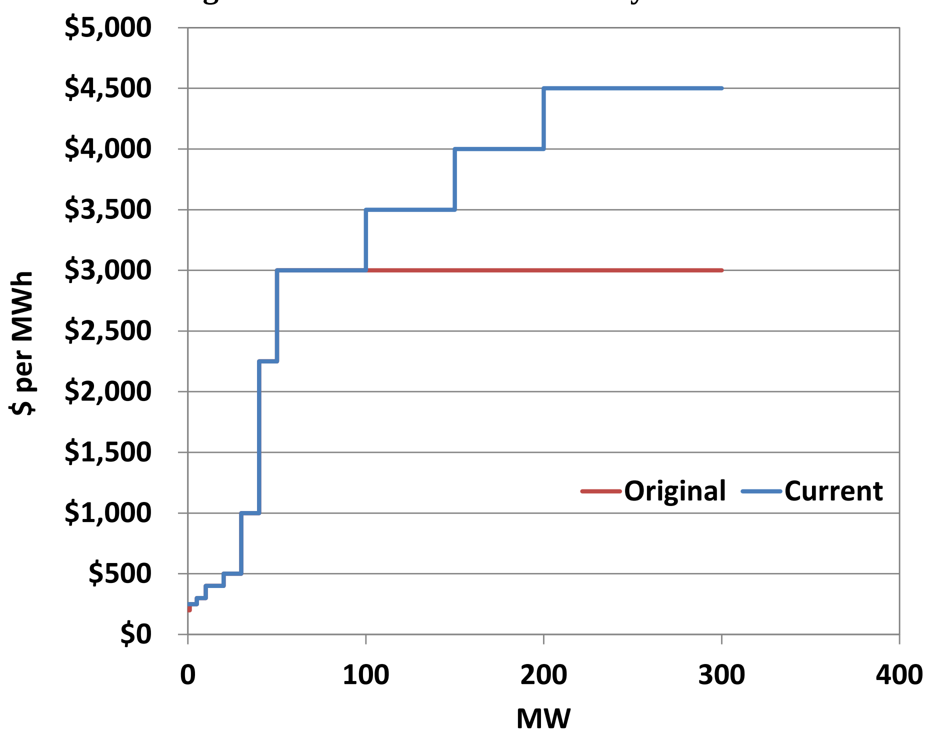
---

[20] The offer floors for online and offline non-spinning reserves are $120 and $180 per MWh, respectively. NPRR427 requires that energy offers from generation resources providing responsive reserve and up regulation reserves be priced at the system-wide offer cap.

[21] See Direct Testimony of David B. Patton, PUCT Docket No. 31540, (Nov. 10, 2005), at pages 35-41.

Aside from the offer floors for non-spinning reserves, the Power Balance Penalty Curve ("PBPC") and the offer floors for up regulation and responsive reserve provided from generation resources defines the relationship between the quantity of operating reserve deficiency and the resulting energy price. The PBPC was modified during 2012 in conjunction with the increase in the system-wide offer cap to $4,500 per MWh. Figure 67 compares the original PBPC in place at the start of the nodal market and the current curve, as modified in 2012.

Figure 67: Power Balance Penalty Curves



Under the current curve, if operating reserves are deficient by 5 MW or less, the energy price will be $250 per MWh. If the deficiency is greater than 150 MW but less than 200 MW, the energy price would be set at $4,000 per MWh. Once the 200 MW from the PBPC is exhausted, the only remaining energy available is from generator provided responsive reserves and up regulation reserves. Since energy provided by these services is required to be offered at the system-wide offer cap, real-time energy prices will be set at that level.

Table 4: Power Balance Penalty Curve

Maximum Operating Reserve Deficiency (MW)	Energy Price ($ per MWh) Original Curve	Energy Price ($ per MWh) Current Curve
1	$200	
5	$250	$250
10	$300	$300
20	$400	$400
30	$500	$500
40	$1,000	$1,000
50	$2,250	$2,250
>50	$3,001	
100		$3,000
150		$3,500
200		$4,000
>200		$4,501

The current relationship between operating reserve deficiency and energy prices defined by the PBPC and the operating reserve offer floors has no real analytic basis other than having its end anchored by the system-wide offer cap. The intermediate values are set at values acceptable to the collective agreement of stakeholders. A more analytically rigorous approach would be to introduce an operating reserve pricing mechanism that reflects the operational loss of load probability ("LOLP") at varying levels of operating reserves multiplied by the value of lost load ("VOLL"). The LOLP would be equal to 1.0 when operating reserves fall to the level where involuntary load shedding is directed by ERCOT. The LOLP would decline exponentially from 1.0 as the level of operating reserves increased.

The implementation of such a curve is currently being evaluated under different approaches. The most complex approach would be to implement real-time co-optimization of energy and reserves with an operating reserve demand curve. The approach named "Solution B+" is likely to be easier to implement and would introduce the operating reserve demand curve but not include real-time co-optimization.[22] And yet another approach would be to adjust the current

[22] See ERCOT Presentation Regarding Potential Implementation of Scarcity Pricing Proposal Offered by Professor Hogan, PUCT Docket No. 40000, (Jan. 22, 2013).

operating reserve offer floors to better reflect LOLP * VOLL at various levels of operating reserves. Each of these approaches could result in more efficient pricing of operating reserve shortages. However, for any of these approaches to result in planning reserve margins over the long-term that meet the historical standard of one loss of load event in ten years would likely require an increase to the level of minimum operating reserves significantly beyond what is required to maintain reliable system operations, or by establishing shortage pricing that is substantially higher than the system-wide offer cap rising to $9,000 per MWh or most reasonable estimates of VOLL. As discussed below, each of these changes can introduce costly operational inefficiencies into the ERCOT energy markets.

As the system-wide offer cap increases to $5,000 per MWh and beyond, it is likely under the current market mechanisms that prices will rise to levels approaching the VOLL when the available reserves are at levels where the LOLP is much less than 1.0 and involuntary load curtailment is not imminent. We have two recommendations to address this concern. The first is to modify the slope of the existing PBPC and the offer floor for responsive reserve service to provide a more gradual slope up to the system-wide offer cap. This could be accomplished by any of three approaches discussed above. We also recommend modifying the automatic pricing of unoffered capacity such that it is not all priced at the system-wide offer cap to avoid the inefficiencies associated with the automated economic withholding of such capacity.

Shortage Pricing, Capacity Markets, and Resource Adequacy

Regardless of the means by which revenues are produced in a wholesale electricity market, it is fundamental that investment will only occur when the total net revenues expected by the investor are greater than its entry costs (including profit on its investment). Additionally, these sources of revenue must be available to all resources, both new and existing, in order to facilitate efficient investment, maintenance, and retirement decisions by all suppliers.

In an energy only market, the primary source of such revenue is the net revenues received during periods of shortage. Expectations about both the magnitude of the energy price during shortage conditions and the frequency of shortage conditions are the primary means to attract new investment in an energy-only market. If the expected revenues are not high enough to facilitate enough investment to satisfy the planning reserve target, one option is to increase the shortage pricing levels to levels that substantially exceed the expected value of lost load. As the planning reserve levels grow, however, the frequency of shortages will tend to drop sharply, which can make it difficult to use this means to meet planning reserve targets.[23] Additionally, as discussed below, such approaches introduce costly operational inefficiencies in to the ERCOT energy markets.

Most other competitive electricity markets do not rely solely on shortage pricing to generate sufficient revenue to support the capacity additions necessary to satisfy their planning reserve requirements. They employ capacity markets to competitively generate capacity payments over the year that are made to suppliers in return for meeting defined capacity obligations. Capacity prices and associated payments vary monthly or annually based on long-term planning reserve levels, independent of the real-time supply and demand conditions. These capacity markets are designed to ensure that a specified planning reserve margin is achieved.

In 2012, ERCOT engaged The Brattle Group to assess its resource adequacy outlook by evaluating a number of market design scenarios.[24] Brattle also supplemented its report with the following table that presents a comparison of costs and reliability for the energy-only market and two capacity market scenarios with 10% and 14% reserve margin requirements.[25] The results of this analysis are summarized in the table below.

Brattle estimates that even with $9,000 per MWh system-wide offer caps, economic equilibrium for the ERCOT energy-only market is achieved at an 8 percent planning reserve margin, although the actual reserve margin outcomes will be uncertain. Brattle further estimates that in the energy-only market at annual equilibrium, wholesale generation costs will be $18.3 billion. In contrast, Brattle's assessment of a capacity market with a more certain 14 percent reserve margin expectation, estimates generation costs at annual equilibrium to be $18.7 billion.

[23] The difficulty of relying primarily on shortage pricing will depend on how high the planning reserve target is relative to the planning reserve levels any energy-only market priced at the expected value of lost load would provide. See the discussion of the Brattle Report below.

[24] *ERCOT Investment Incentives and Resource Adequacy*, The Brattle Group, PUCT Docket No. 37987 (June 1, 2012).

[25] *Customer Cost Comparison*, The Brattle Group, PUCT Docket No. 40000 (Sept. 4, 2012).

It is important to recognize that this increase in cost is not due to the introduction of the capacity market, it is due to the requirement to sustain a planning reserve margin greater than 8 percent. In fact, the Brattle analysis indicates that a capacity market would deliver the higher planning reserve margin at a relatively low incremental cost with much more certainty.

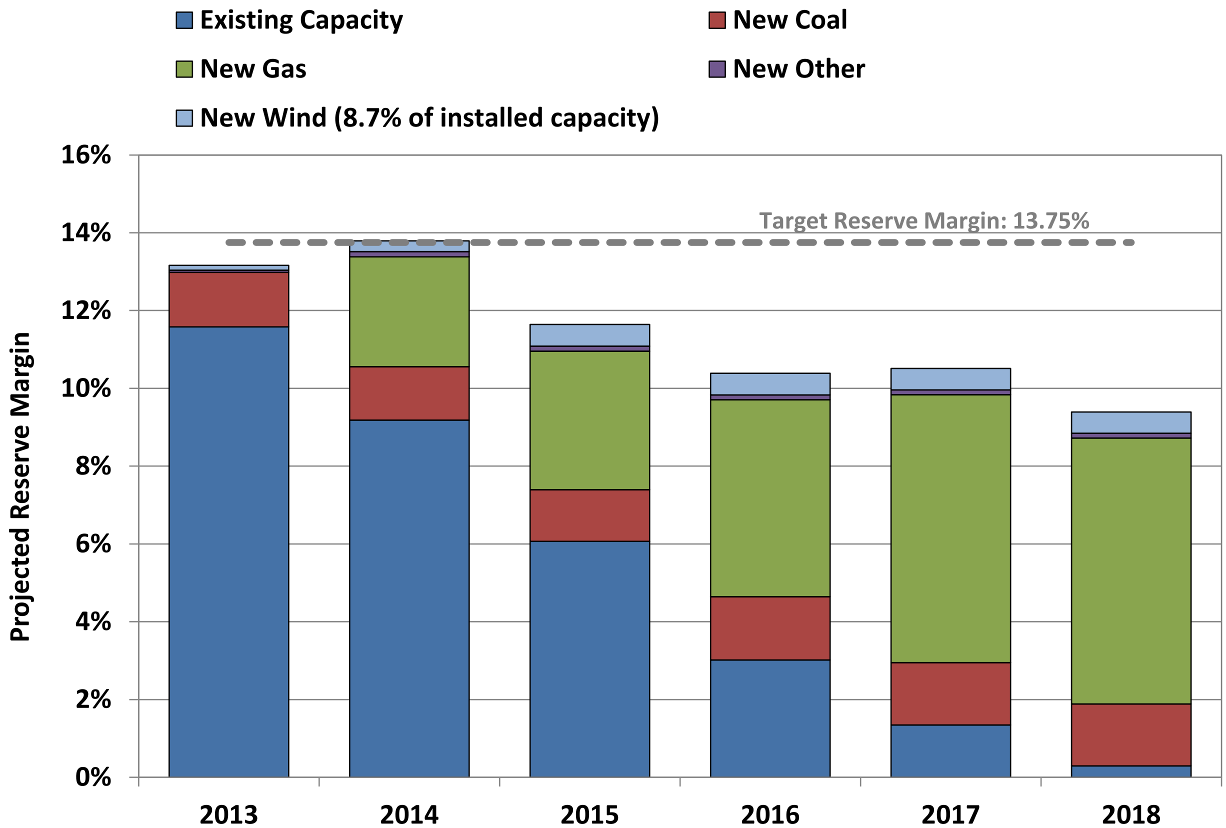
COMPARISON OF COSTS AND RELIABILITY

	Energy-Only Equilibrium	10% Reserve Margin Requirement	14% Reserve Margin Requirement
Reliability			
Reserve Margin	8%	10%	14%
Reserve Margin Certainty	Uncertain	More Certain	More Certain
Annual Avg. Loss of Load Hours	4.1	2.2	0.3
Customer Costs			
Energy Costs ($billions)	$18.3	$16.3	$14.0
Capacity Costs ($billions)	$0	$2.1	$4.7
Total Costs ($billions)	$18.3	$18.4	$18.7
Cost Increase over Energy-Only Equilibrium (%)	NA	0.7%	2.4%
Rate Increase over Energy-Only Equilibrium (%)	NA	0.4%	1.4%
Combustion Turbine Energy Margins and Capacity Revenues			
Energy Margins ($/kW-y)	$105	$75	$41
Capacity Revenues ($/kW-y)	$0	$30	$64
Total Margins ($/kW-y)	$105	$105	$105

Notes: 8% energy-only equilibrium reserve margin based on The Brattle Group's simulations with a $9,000 price cap and gradually sloping scarcity pricing function. Rate impacts assume generation costs comprise 60% of total retail rates.

Recent studies have indicated that to maintain the same small level of risk of having an involuntary curtailment of firm load, the planning reserve target should be increased from 13.75 percent to approximately 16 percent. Hence, the difficulty of satisfying ERCOT's planning needs with shortage pricing alone will grow if this recommendation is adopted. Shown below in Figure 68 is ERCOT's most current projection of reserve margins. It indicates that the region will have a 13.2 percent reserve margin heading into the summer of 2013. With the addition of recently announced generation additions, in 2014 the reserve margin is expected to reach 13.8 percent -- just barely above the current target. The bulk of the new capacity being added is natural gas-fired generation, approximately a quarter of which is expansions at existing facilities.

Figure 68: Projected Reserve Margins



Source: ERCOT Capacity Demand Reserve Reports / 2013 data from Winter 2012, 2014 - 2018 from May 2013

In response to these observations, proposals have been put forth that would introduce significant operational inefficiencies into the ERCOT energy markets, such as a requirement to substantially increase the quantity of operating reserves ERCOT procures and to, by rule, economically withhold these surplus reserves from the market. Such approaches would introduce significant inefficiencies into ERCOT day ahead and real time operations in an effort to manufacture more frequent shortage pricing and a higher planning reserve margin than would be achieved in a pure energy-only market framework. However, such approaches will not guarantee that the planning reserve targets will be satisfied and, because of the resulting inefficiencies, will be more costly for ERCOT's consumers. Hence, consistent with Brattle's findings, it is our view that, if the planning reserve margin is viewed as a minimum requirement, implementation of a capacity market is the most efficient mechanism to achieve this objective. As observed by Brattle, a well-designed capacity market can efficiently meet a planning reserve requirement without impairing the efficiency of energy market operations. However, there are many determinations required in the design, implementation and maintenance of a capacity market construct.[26]

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. Additionally, loads may participate passively in the market by simply adjusting consumption in response to observed prices. Unlike active participation in ERCOT administered markets, passive demand response is not directly tracked by ERCOT.

ERCOT allows qualified load resources to offer responsive reserves and non-spinning reserves into the day-ahead ancillary services markets. Those providing responsive reserves must have high set under-frequency relay equipment, which enables the load to be automatically tripped when the frequency falls below 59.7 Hz, which will typically occur only a few times in each year. Deployments of non-spinning reserves occur much more frequently. To date, load resources have shown a clear preference for providing responsive reserve service.

As of December 2012, approximately 2,500 MW of capability were qualified as Load Resources. Figure 69 shows the amount of responsive reserves provided from load resources on a daily basis in 2012. The high level of participation by demand response in the ancillary service markets sets ERCOT apart from other operating electricity markets. For reliability reasons the maximum amount of responsive reserves that can be provided by load resources was limited to 1,150 MW until April 2012. At that time, the limitation on load resources providing responsive reserve increased to 1,400 MW, corresponding with the increase in total responsive reserve requirements.

---

[26] *ERCOT Investment Incentives and Resource Adequacy,* The Brattle Group, at 115-119, PUCT Docket No. 37987 (June 1, 2012).

Figure 69: Daily Average of Responsive Reserves provided by Load Resources

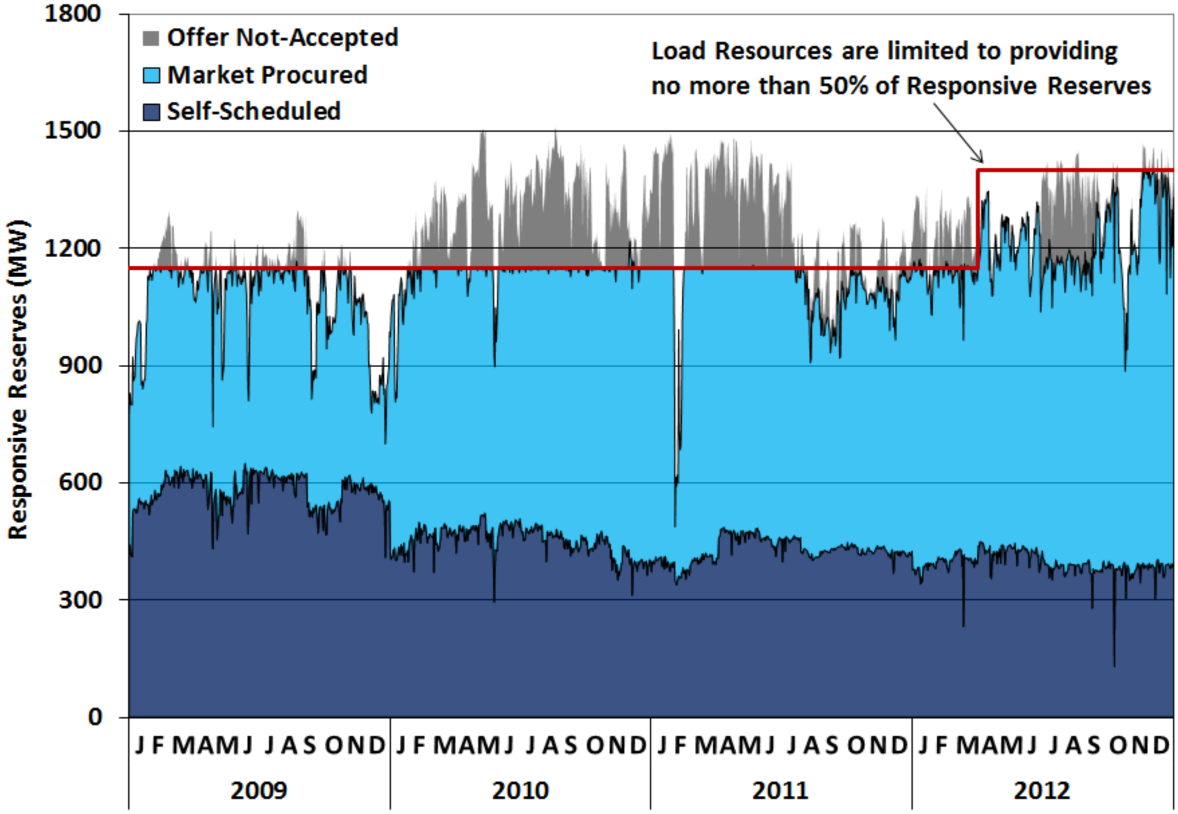


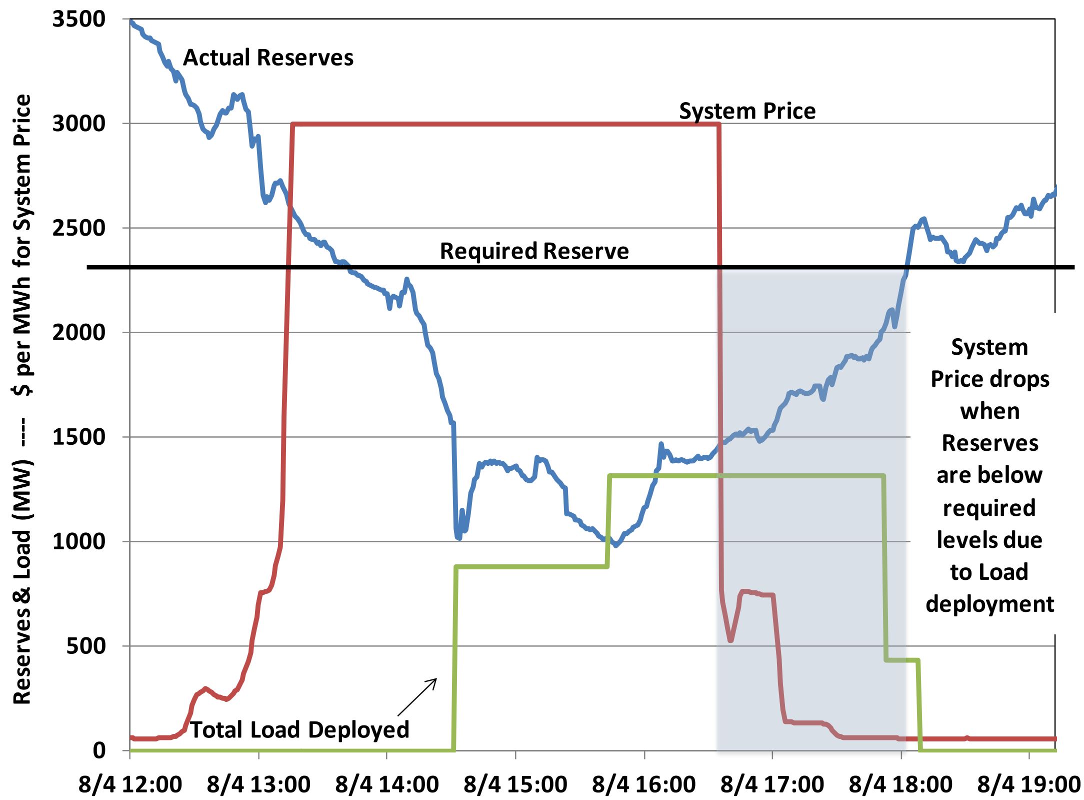
Figure 69 shows that it took a few months after implementing the increased requirement for the amount of offers by load resources to routinely reach this level. Notable exceptions include a prolonged reduction from November 2008 through January 2009 that was likely a product of the economic downturn and its effect on industrial operations. Seasonal reductions were also observed during late 2009 and 2012.

During 2011 there was a significant reduction in loads offering to provide responsive reserve during early February and again starting in mid-July. Both of these times corresponded with expected high real-time prices. Since load resources provide capacity by reducing their consumption, they have to actually be consuming energy to be eligible to provide the capacity service. During periods of expected high prices the price paid for the energy can exceed the value received from providing responsive reserves.

Pricing During Load Deployments

During times when there are shortages of supply offers available for dispatch and Responsive Reserves are deployed, that is, converted to energy as one of the last steps taken before shedding firm load, the value of the foregone reserves – which is much higher than the marginal cost of the most expensive online generator – should be reflected in energy prices to achieve efficient economic signals governing investment in generation, demand response and transmission. Unfortunately, ERCOT's dispatch software does not recognize that load has been curtailed, and computes prices based on supplying only the remaining load. A good example of this situation occurred on August 4, 2011. Figure 70 displays available reserves and the system price for that afternoon and shows that even though reserves were below required levels, system price dropped to $60 per MWh. At this level prices are being set based on supply offers and do not reflect the value of the load that is being curtailed to reliably serve the remaining system demand.

Figure 70: Pricing During Load Deployments



We recommend that ERCOT implement system changes that will ensure that ***all*** demand response that is actively deployed by ERCOT be incorporated into the dispatch software so that such deployments will be able to set the price at the value of load at times when such deployments are necessary to reliably serve the remaining system demand. This includes load resources and Emergency Response Service (ERS) providers being deployed for the services they contracted to provide or when firm load is involuntarily curtailed.

VI. Analysis of Competitive Performance

In this section we evaluate market power from two perspectives, structural (does market power exist) and behavioral (have attempts been made to exercise it). We examine market structure by using a pivotal supplier analysis that indicates the frequency with which a supplier was pivotal increased at higher levels of demand. This is consistent with observations in prior years. To evaluate participant conduct we estimate measures of physical and economic withholding. We examine withholding patterns 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 2012.

A. Structural Market Power Indicators

We analyze market structure by using the Residual Demand Index ("RDI"), a statistic that measures the percentage of load that could not be satisfied without the resources of the largest supplier. 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.[27] 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,

[27] 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 71 shows the RDI relative to load for all hours in 2012. The trend line indicates a strong positive relationship between load and the RDI. This 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. It is possible that they also control other capacity through bilateral arrangements, although we do not know whether this is the case. To the extent that the resources scheduled by the largest QSEs are not controlled or providing revenue to the QSE, the RDIs will tend to be slightly overstated.

Figure 71: Residual Demand Index

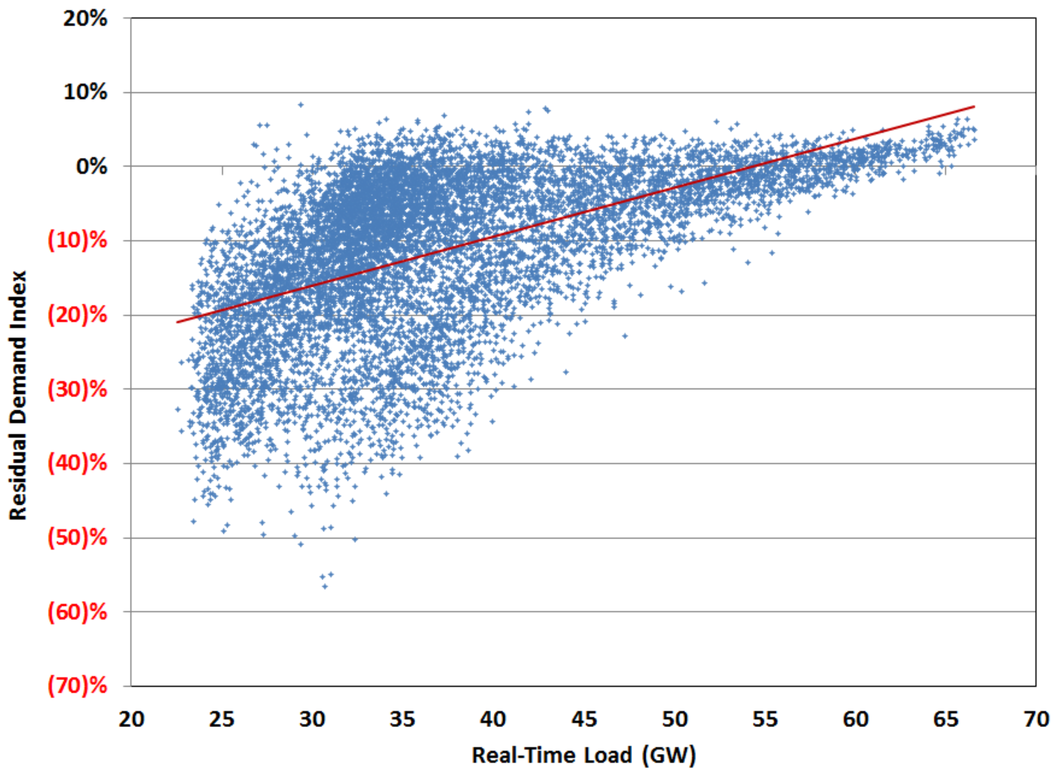
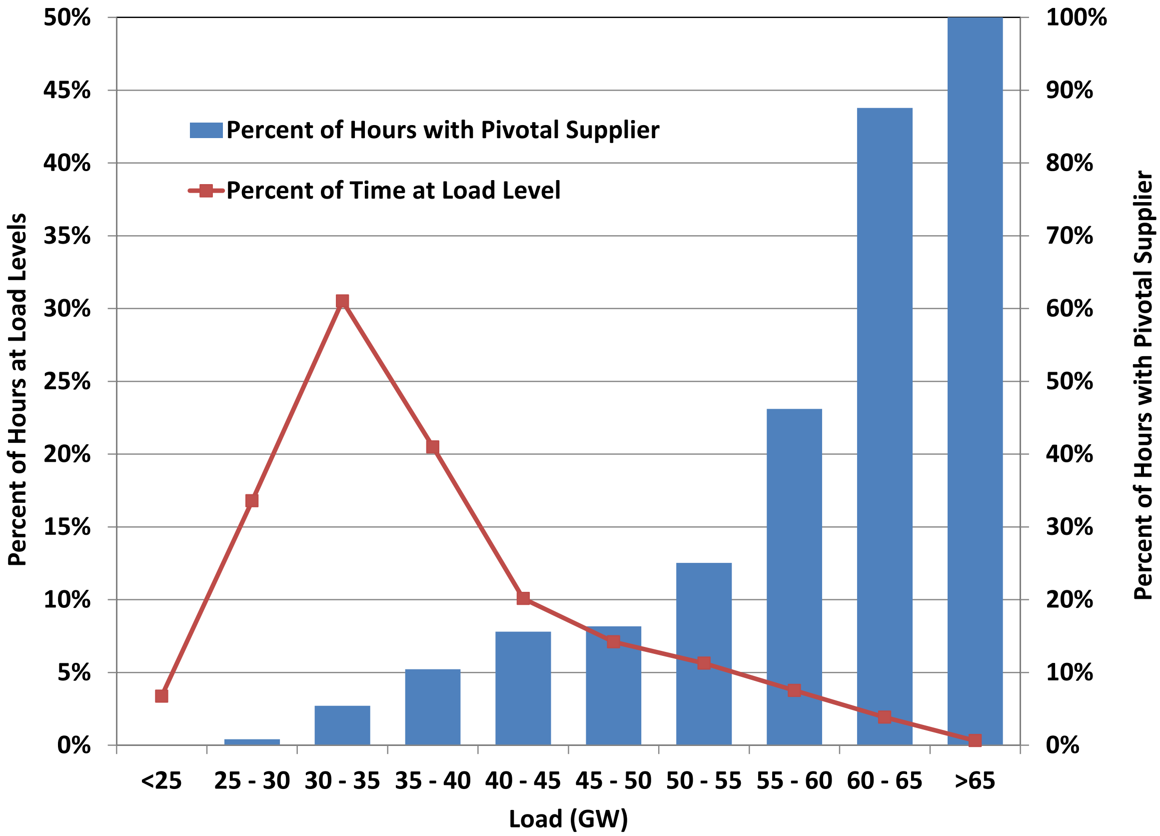


Figure 72 below summarizes the results of our RDI analysis by displaying the percent of time at each load level there was a pivotal supplier. At loads greater than 65 GW there was a pivotal supplier 100 percent of the time. The figure also displays the percent of time each load level occurs. Combining these values we find that there was a pivotal supplier in approximately

12 percent of all hours of 2012. As a comparison, the same system-wide measure for the Midwest ISO resulted in less than 1 percent of all hours with a pivotal supplier.

Figure 72: Pivotal Supplier Frequency by Load Level



It is important to recognize that inferences regarding market power cannot be made solely from this data. Bilateral contract obligations can affect a supplier's potential market power. For example, a smaller supplier selling energy in the real-time energy market and through short-term bilateral contracts may have a much greater incentive to exercise market power than a larger supplier with substantial long-term sales contracts. The RDI measure shown in the previous figure does not consider the contractual position of the supplier, which can increase a supplier's incentive to exercise market power compared to the load-adjusted capacity assumption made in this analysis.

In the next analysis of RDI, we impose ramp rate limitations on the capacity available to meet load. As shown in Figure 73, the ramp constrained RDI shows the same pattern of becoming increasingly positive at higher load levels, but is much more likely to be positive as the total capacity available to the market is smaller than in the previous analysis. We observe that the ramp rate constrained RDI was usually positive, indicating the presence of a pivotal supplier, except when load was below 25 GW.

Figure 73: Ramp-Constrained Residual Demand Index

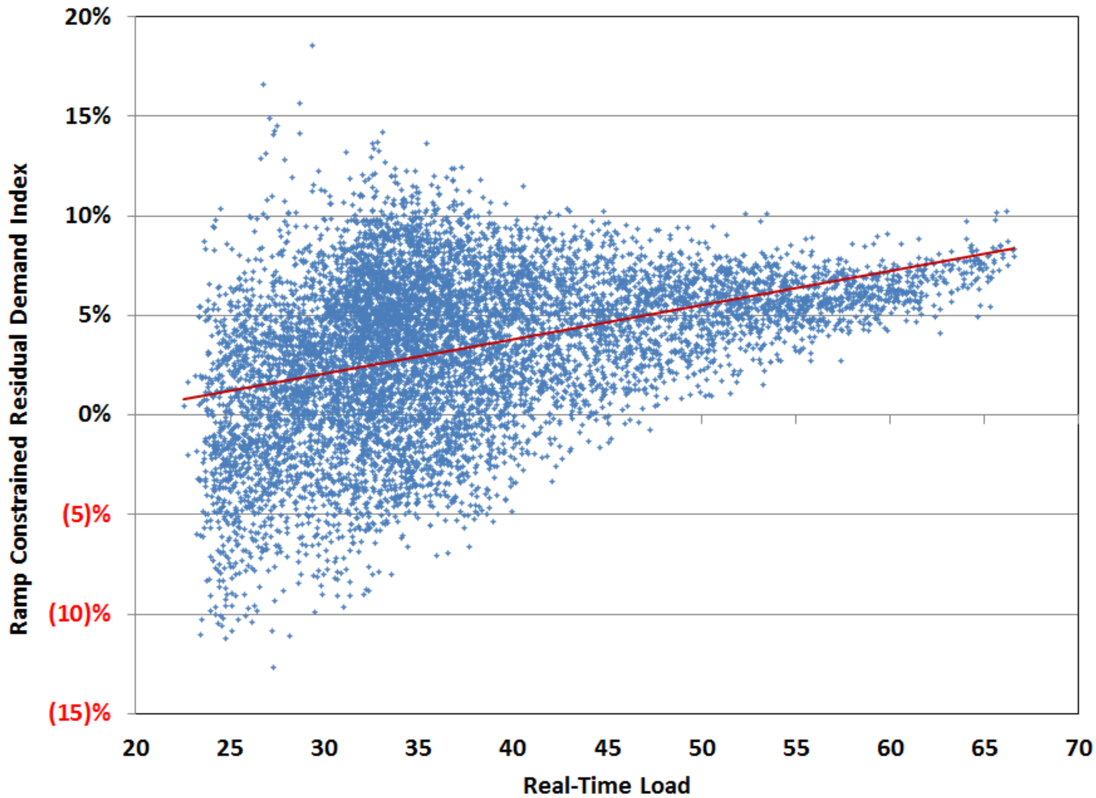
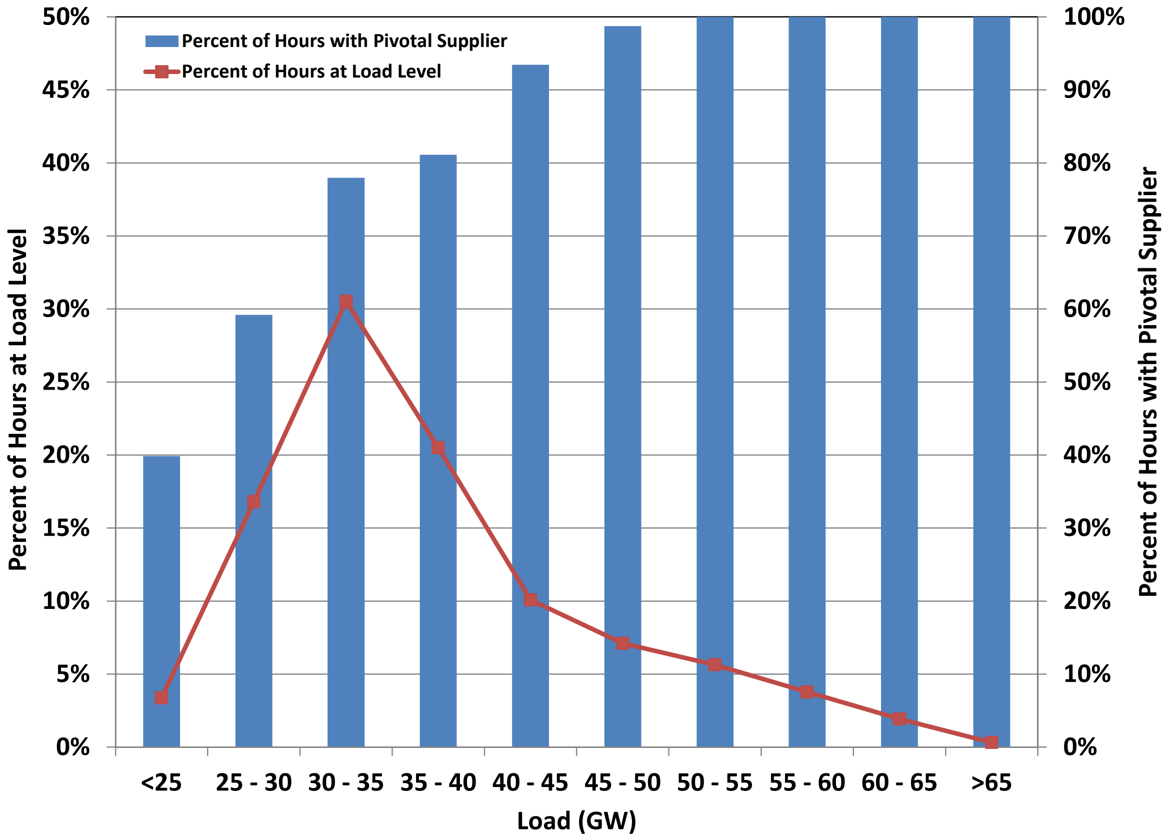


Figure 74 displays the percent of time at each load level there was a pivotal supplier when ramp rate constraints are considered. At loads greater than approximately 50 GW there is a pivotal supplier 100 percent of the time. Ramp rate constrained RDI indicates that there was a pivotal supplier in approximately 80 percent of all hours in 2012. It is important to note that this ramp rate constraint is being imposed for every dispatch interval, or approximately every 5 minutes.

Figure 74: Frequency of Ramp Constrained Pivotal Supplier by Load Level



Voluntary Mitigation Plans

The PUCT approved Voluntary Mitigation Plans ("VMP") for two market participants – NRG and GDF SUEZ – during 2012. Action on the request to approve a VMP for a third participant, Calpine, was pending at the end of the year. Generation owners are motivated to enter into VMPs because adherence to a plan approved by the commission constitutes an absolute defense against an allegation of market power abuse through economic withholding with respect to behaviors addressed by the plan. This increased regulatory certainty afforded to a generation owner regarding its energy offers in the ERCOT real-time market, must be balanced by appropriate protections against a potential abuse of market power in violation of PURA §39.157(a) and PUCT SUBST. R. 25.503(g)(7).

It is our position that 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 the 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.

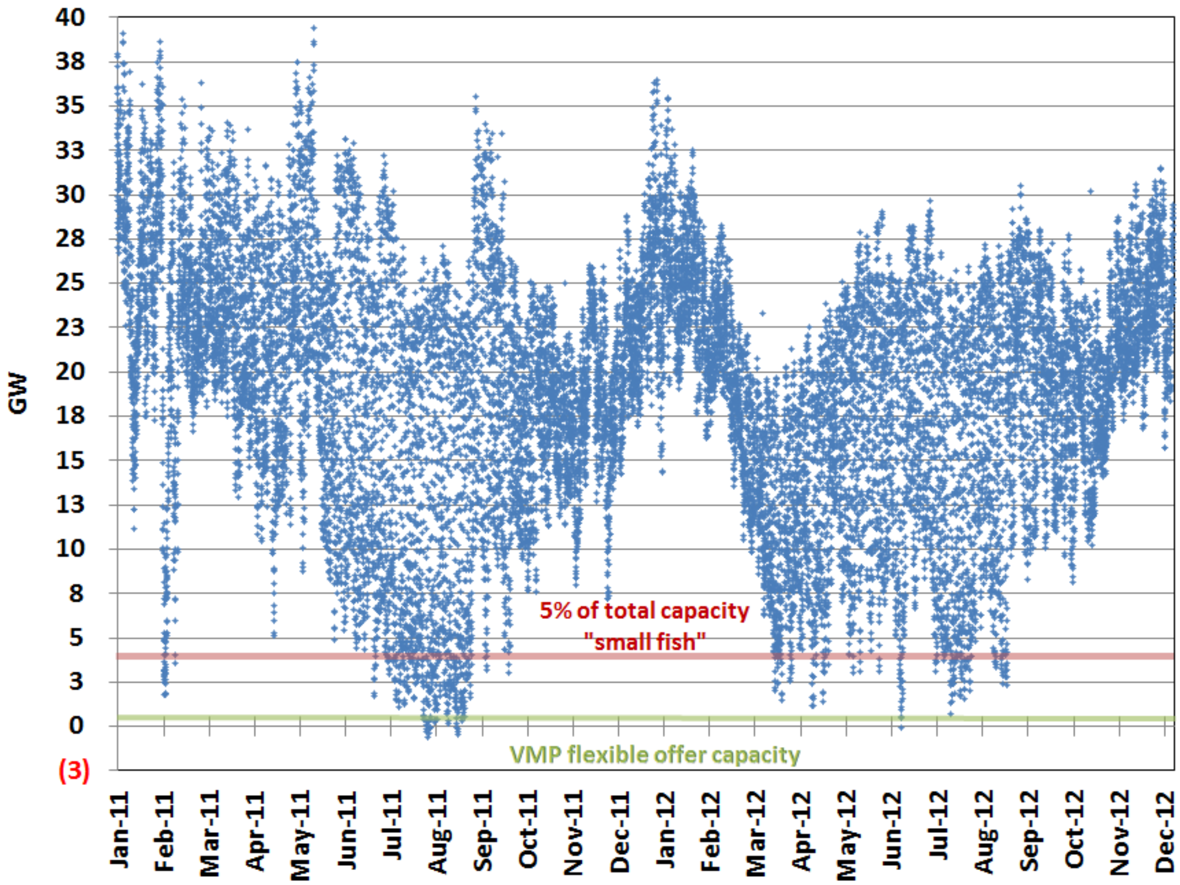
The plan approved for NRG allows the company to offer some of its capacity at prices up to the system-wide offer cap. Specifically, up to 12 percent of the difference between the high sustained limit and the low sustained limit for each natural gas-fired unit (5 percent for each coal/lignite unit) may 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 difference between the high sustained limit and the low sustained limit for each natural gas-fired unit may be offered no higher than the system-wide offer cap. The amount of capacity covered by these provisions could be as much as 400 MW.

Allowing offers up to these high levels 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, NRG's VMP contains 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 for tuning these offers in response to particular market conditions and significantly increases the likelihood that such offers, if offered, are based on legitimate marginal cost considerations.

Under P.U.C. Subst. R. §25.505(d), market participants controlling less than five percent of the capacity in ERCOT by definition do not possess ERCOT-wide market power under the PUCT rules. Hence, these participants can submit very high-priced offers that, per the PUCT rule, will not be deemed to be an exercise of market power. 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 could be large.

The figure below shows the amount of surplus capacity available in each hour of every day during 2011 and 2012. 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 and regulation up capacity, minus load. Over the past two years there were 13 hours with no surplus capacity. 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 75. There were 450 hours over the past two years with less than 4,000 MW of surplus capacity. During these times a large “small fish” would be pivotal and able through their offers to increase the market clearing price, potentially as high as the system-wide offer cap.

Figure 75: Surplus Capacity



To date, the over-mitigation issue discussed in Section I.E, Mitigation at page 20 has meant that mitigation measures have been applied much more broadly than intended or necessary in the ERCOT real-time energy market. Market system changes to narrow the scope of mitigation are scheduled to be implemented in June 2013 to address this issue. Although “small fish” market participants have always been allowed to offer up to all their 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 will become more noticeable after the scope of mitigation is narrowed.

The approved NRG VMP affords the company offer flexibility for up to approximately 400 MW.[28] As indicated by the green line in Figure 75, the ability for NRG to raise the clearing price as a result of its offers would have occurred in less than 30 hours over the past two years.

The final key element in a VMP is the timing of termination. The approved VMP for NRG may be terminated after three business days' notice. PURA §39.157(a) 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..." 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 involve profitably raising prices significantly above the competitive level for a significant period of time. Thus, although the offer thresholds provided in the VMP are designed based on experience to promote competitive market outcomes, the short termination provision provides additional assurance that any unintended consequences associated with the potential exercise of market power can be addressed in a timely manner rather than persisting and rising to the level of an abuse of market power.

B. Evaluation of Supplier Conduct

The previous subsection presented a structural analysis that supports inferences about potential market power. In this section we evaluate actual participant conduct to assess whether market participants have attempted to exercise market power through physical or economic withholding. First, we examine unit deratings and forced outages to detect physical withholding and then we evaluate the "output gap" to detect economic withholding.

[28] Under the terms of their VMP, NRG may offer a certain portion of their dispatchable capacity from online units at prices up to $500 per MWH – 5 percent of coal units and 12 percent of gas units. Additionally, NRG may offer up to 3 percent of their dispatchable capacity from online gas units at prices up to the system-wide offer cap. Any capacity offered under either of these terms must be offered in the same price/quantity pairs for all hours of the operating day.

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. The 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.

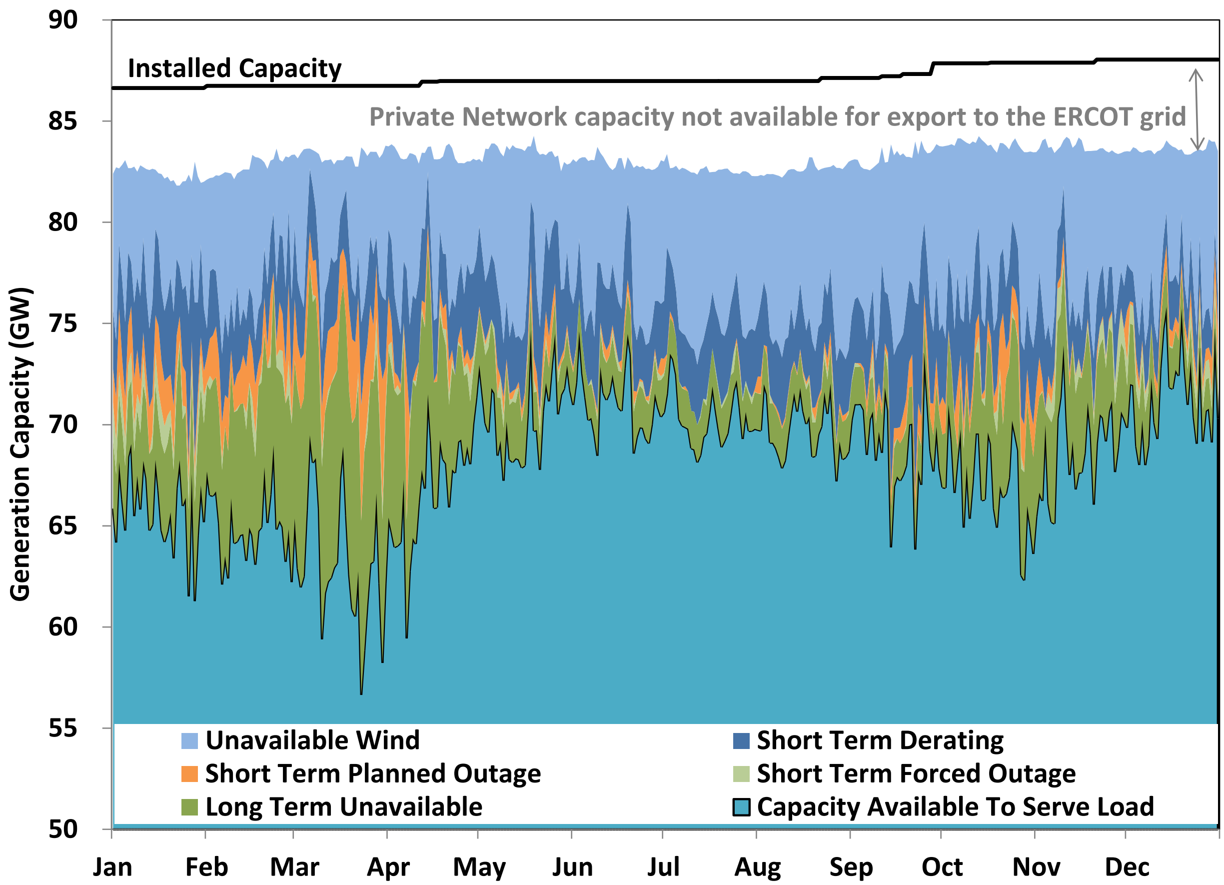
1. Generation Outages and Deratings

A substantial portion of the installed capability is frequently unavailable due to generator outages and deratings. For this analysis we start with the unit status information communicated to ERCOT on a continuous basis. For those units with a status of OUT, meaning they are unavailable, we then cross check to see if an outage had been scheduled. If there is a corresponding scheduled outage we consider the unit on planned outage. If not, it is considered to be a forced outage. We further define derated capacity as the difference between the summertime maximum capability 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 capability level due to technical or environmental factors (e.g., component equipment failures or ambient temperature conditions). It is rare for wind generators to produce at their installed capacity rating due to variations in available wind input. Because such a large portion of derated capacity is related to wind generation we show it separately. In this subsection, we evaluate long-term and short-term deratings to inform our evaluation of ERCOT capacity levels.

Figure 76 shows a breakdown of total installed capability for ERCOT on a daily basis during 2012. This analysis includes all in-service and switchable capacity. From the total installed capacity we subtract away (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 – greater than 30 day -- outages and deratings. What remains is the capacity available to serve load.

Outages and deratings of non-wind generators fluctuated between 3 and 18 GW, as shown in Figure 76, while wind unavailability varied between 1 and 10 GW. Short term planned outages were largest in March, April and October and small during the summer, which are consistent with expectations. Short term forced outages also declined during the summer. Short term deratings peaked during September.

Figure 76: Reductions in Installed Capability



The quantity of long term (greater than 30 days) unavailable capacity, peaked in March at nearly 10GW, reduced to 2 GW during the summer months and increased to almost 6GW in October. This pattern reflects the choice by some owners to mothball certain generators 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. Figure 77 shows the average magnitude of the outages and deratings lasting less than 30 days for the year and for each month during 2012.

Figure 77: Short-Term Outages and Deratings

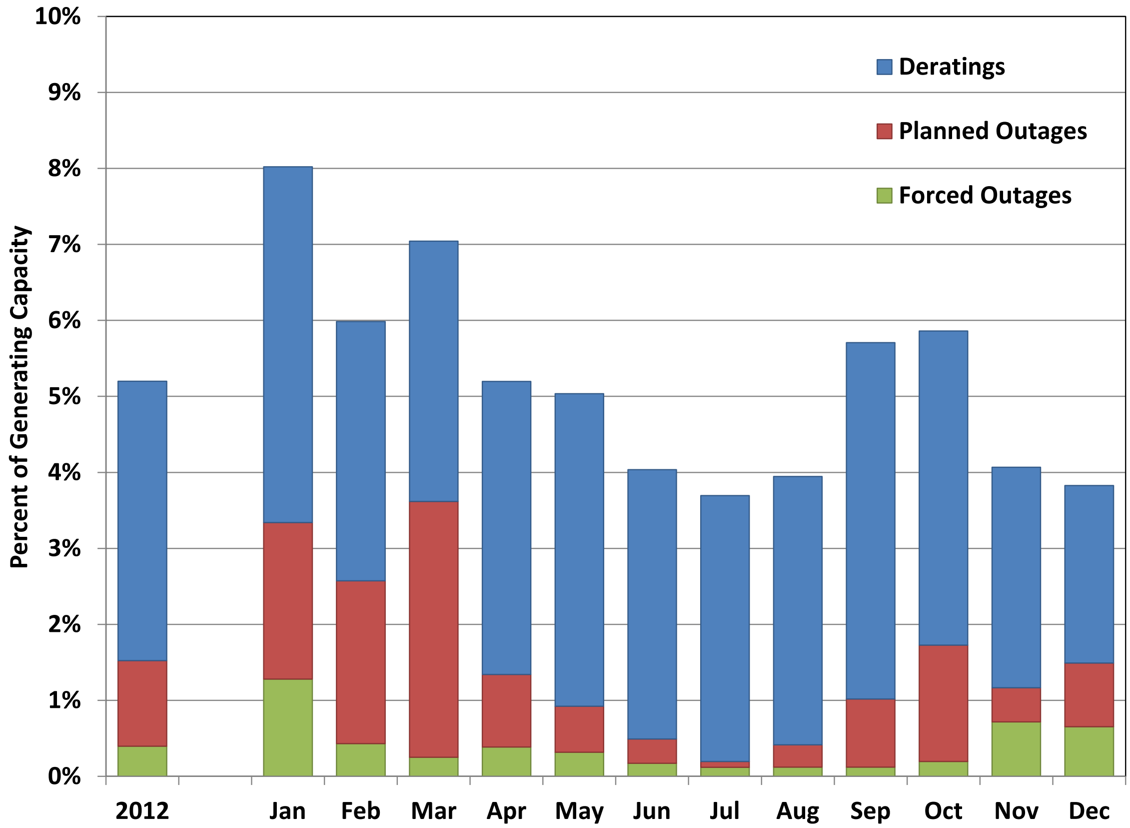


Figure 77 shows that total short-term deratings and outages were as large as 8 percent of installed capacity in January, dropping to below 4 percent for the summer. Most of this fluctuation was due to anticipated planned outages. The amount of capacity unavailable during 2012 averaged slightly above 5 percent of installed capacity. This is a decrease from 2011, when the amount was greater than 6 percent. Similar metrics from the zonal market were consistently above 15 percent. The large disparity between values from the zonal and nodal markets is likely due to combined effects of improved incentives in the nodal market and the lack of unit specific data available from zonal market systems.

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 section is to differentiate justifiable deratings and outages from physical withholding. We test for physical withholding 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 71 through Figure 74 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 a problem in ERCOT, we 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 these peak periods.

Figure 78: Outages and Deratings by Load Level and Participant Size
June to August, 2012

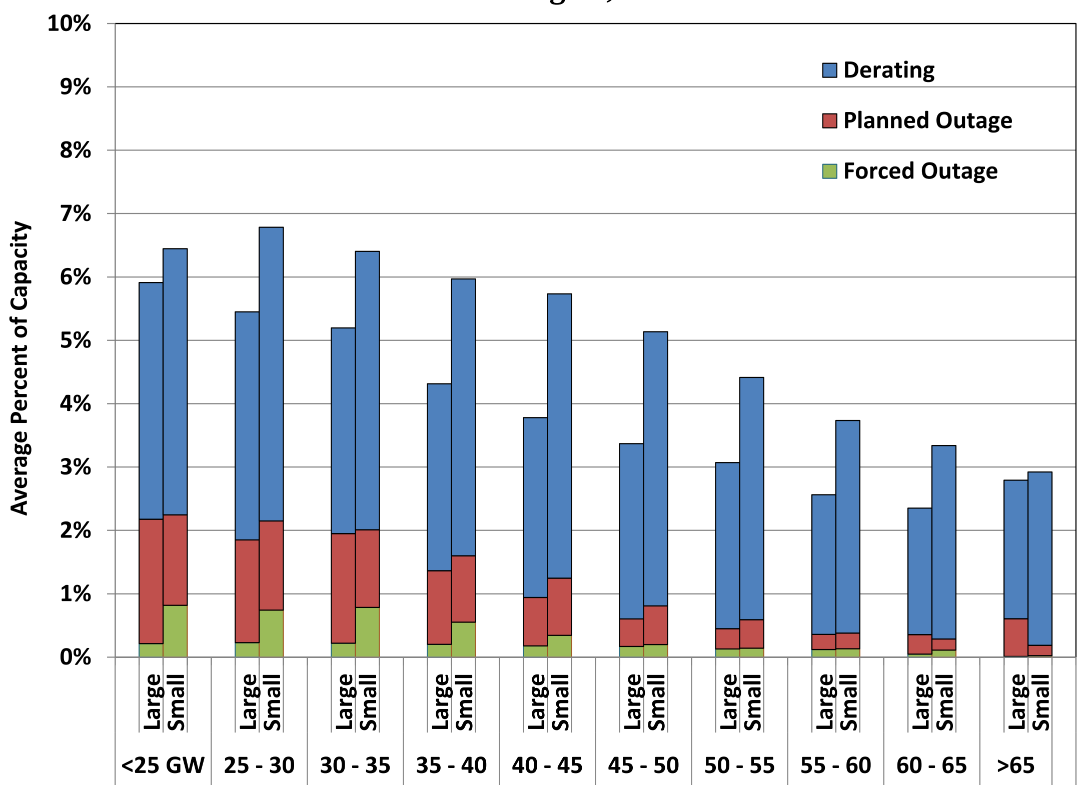


Figure 78 shows the average relationship of short-term deratings and forced outages as a percentage of total installed capacity to real-time load level for large and small suppliers.

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 four largest suppliers in ERCOT. The small supplier category includes the remaining suppliers.

Figure 78 suggests that as demand for electricity increases, all market participants tend to make more capacity available to the market. For both small and large suppliers, the combined short-term derating and forced outage rates decreased from 6 to 7 percent at low demand levels to less than 3 percent at load levels above 65 GW. We observe that at all load levels the percent of unavailable capacity from large suppliers is less than that from small suppliers.

3. Evaluation of Potential Economic Withholding

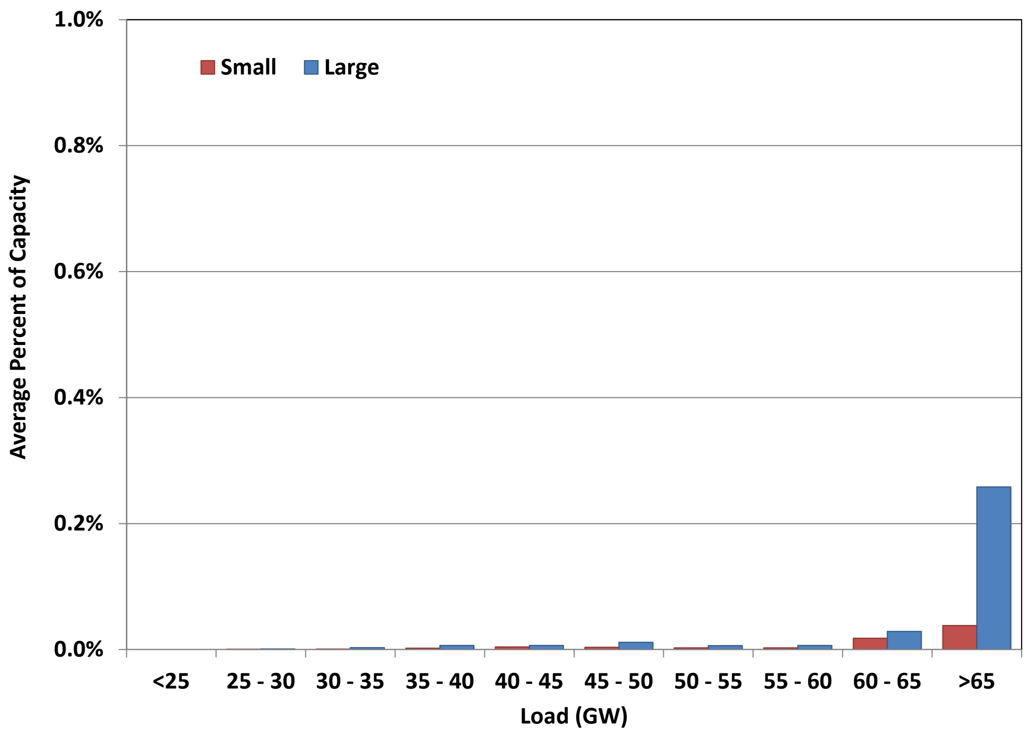
To complement the prior analysis of physical withholding, this subsection evaluates potential economic withholding by calculating an "output gap". The output gap is defined as the quantity of energy that is not being produced by in-service capacity even though the in-service capacity is economic by a 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.

Resources are considered for inclusion in the output gap when they are committed and producing at less than full output. Energy not produced from committed resources is included in the output gap if the real-time energy price exceeds by at least $50 per MWh that unit's mitigated offer cap, which serves as an estimate of the marginal production cost of energy from that resource.

Before presenting the results of the Output Gap analysis, a description of the two-step aspect of ERCOT's 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 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.

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

Figure 79: Incremental Output Gap by Load Level and Participant Size – Step 1



From the results of this analysis, shown in Figure 79, we observe only very small amounts of capacity at only the very highest loads that would be considered part of this output gap. These small quantities raise no competitive concerns.

Figure 80 shows the ultimate output gap, 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 due to a market participant exerting market power.

Similar to the previous analysis, Figure 80 shows the magnitude of the output gap to be very small, even at the highest load levels. These small quantities raise no competitive concerns.

Figure 80: Incremental Output Gap by Load Level and Participant Size – Step 2