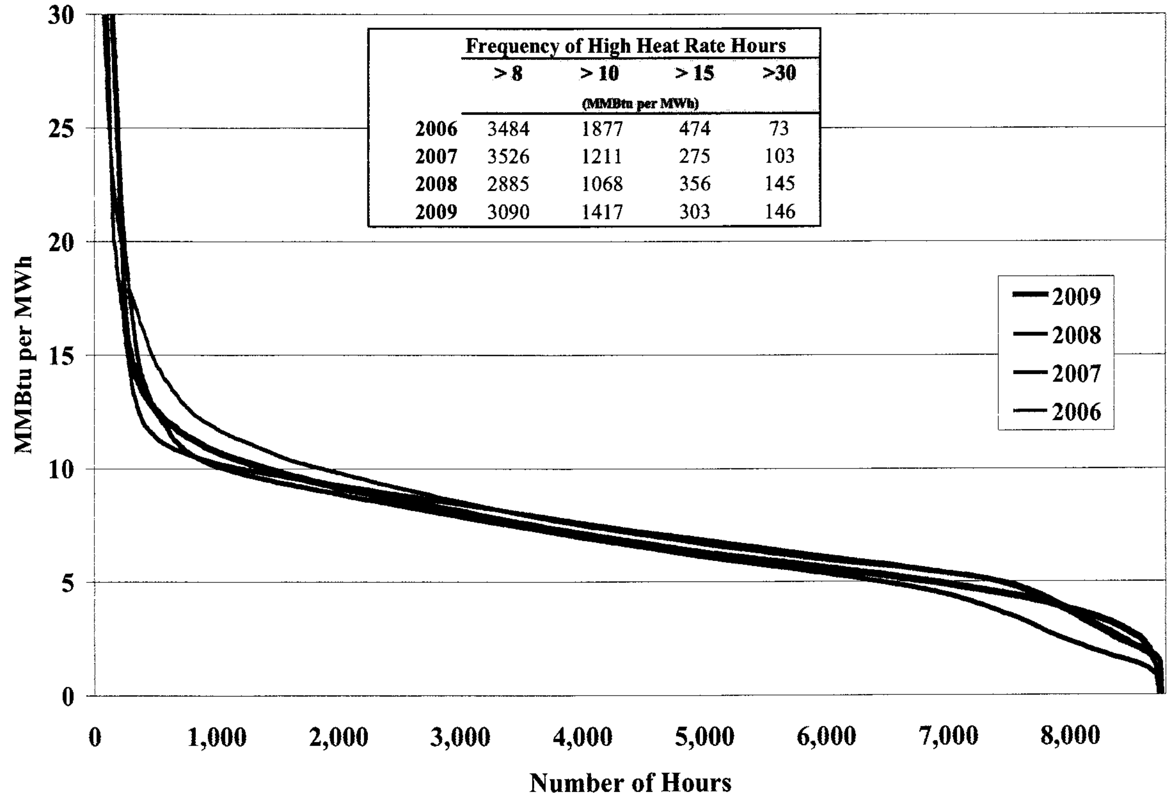
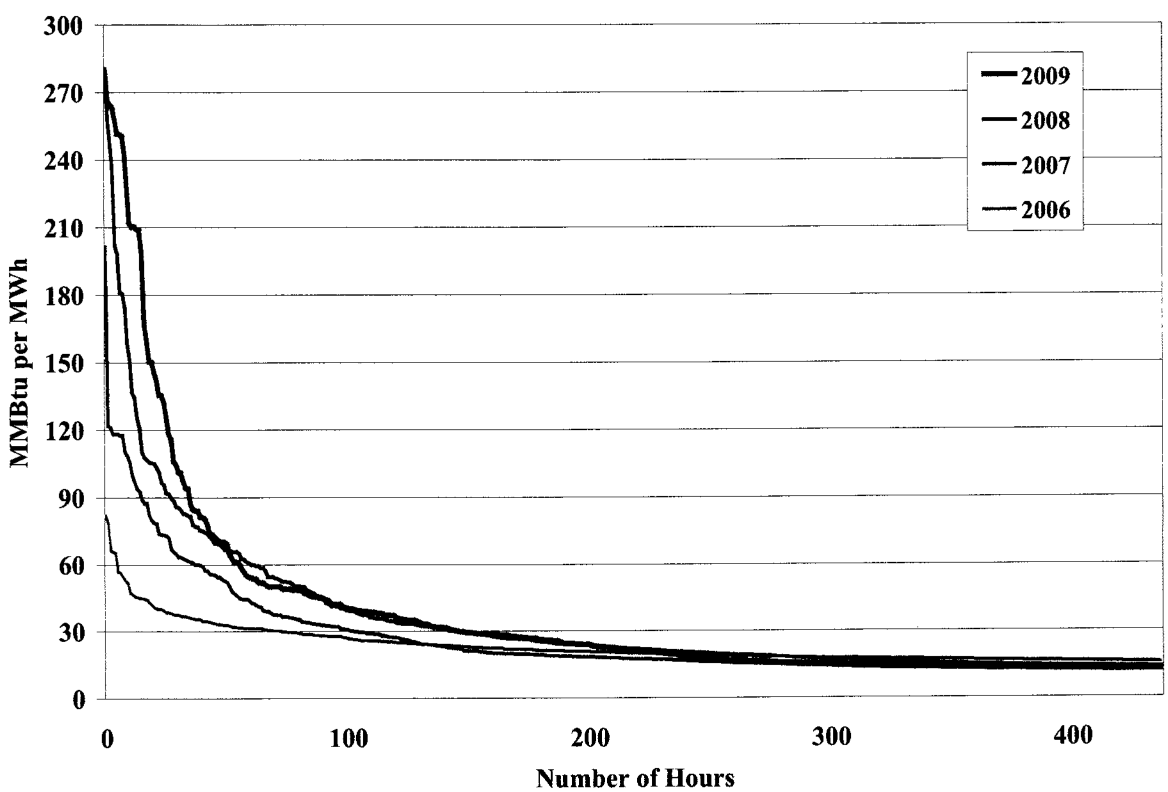
Figure 7: Implied Marginal Heat Rate Duration Curve – All Hours



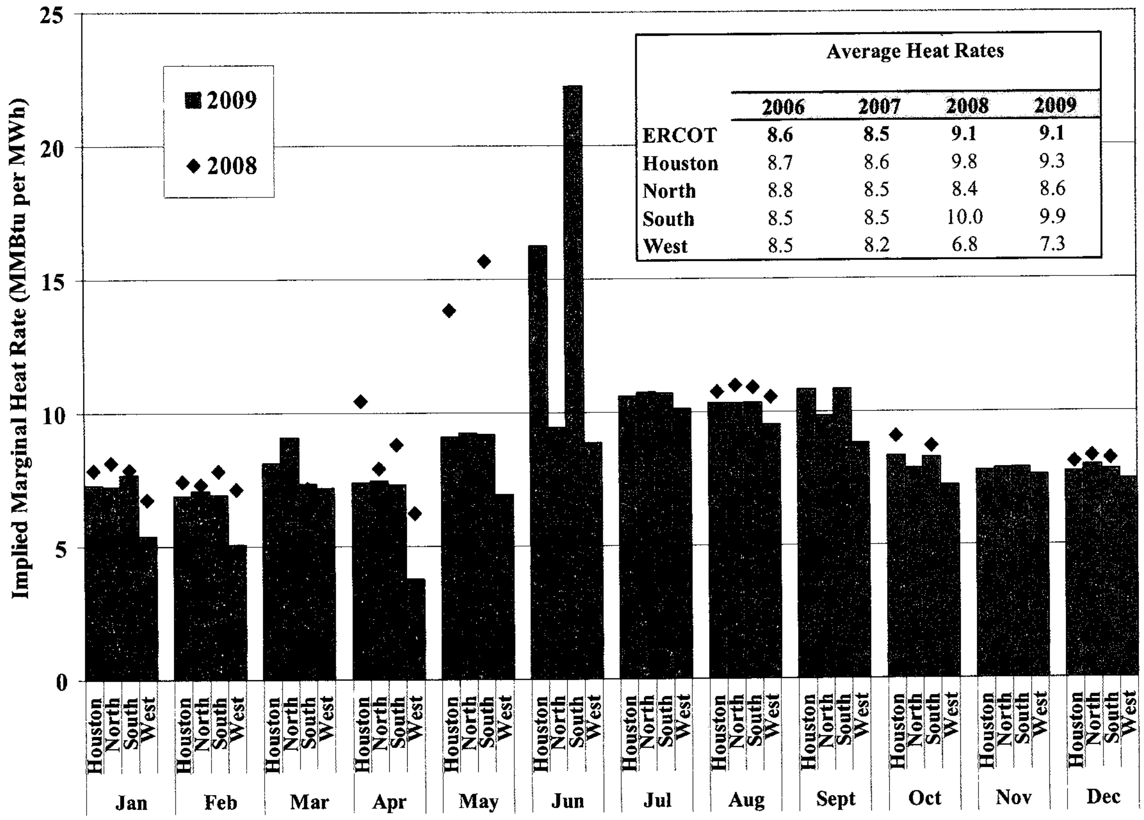
Frequency of High Heat Rate Hours	> 8	> 10	> 15	>30
	(MMBtu per MWh)			
2006	3484	1877	474	73
2007	3526	1211	275	103
2008	2885	1068	356	145
2009	3090	1417	303	146

Figure 8: Implied Marginal Heat Rate Duration Curve – Top 5% of Hours



To better illustrate these differences, the next figure shows the implied marginal heat rates on a monthly basis in each of the ERCOT zones in 2008 and 2009, with annual average heat rate data for 2006 through 2009. This figure is the fuel price-adjusted version of Figure 1 in the prior sub-section. Adjusting for gas price influence, Figure 9 shows that average implied heat rate for all hours of the year was comparable in 2009 to 2008.

Figure 9: Monthly Average Implied Marginal Heat Rates



	2006	2007	2008	2009
ERCOT	**8.6**	**8.5**	9.1	9.1
Houston	8.7	8.6	9.8	9.3
North	8.8	8.5	8.4	8.6
South	8.5	8.5	10.0	9.9
West	8.5	8.2	6.8	7.3

The average implied heat rate was significantly higher in 2008 than in 2009 during the months of April and May due to significant zonal congestion on the North to South and North to Houston interfaces that materialized in these months in 2008. Similarly, the magnitude of zonal congestion on the North to South interface increased significantly in late June 2009, causing the implied heat rate in June to be significantly higher in 2009 than in 2008. The implied heat rate in July was higher in 2009 than in 2008, primarily because of a stretch of extremely high temperatures and load levels, including the setting of a new record peak demand of 63,400 MW on July 13, 2009. Finally, the implied heat rate in September was much lower in 2008 than in

2009 because of the landfall of Hurricane Ike in September 2008 that resulted in widespread and prolonged loss of load in the Houston area.

3. Price Convergence

One indicator of market performance is the extent to which forward and real-time spot prices converge over time. In ERCOT, there is no centralized day-ahead market so prices are formed in the day-ahead bilateral contract market. The real-time spot prices are formed in the balancing energy market. 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.

These two conditions are largely satisfied in the current ERCOT market. Relaxed balanced schedules allow QSEs to increase and decrease their purchases in the balancing energy market. This flexibility should better enable them to arbitrage forward and real-time energy prices. While this should result in better price convergence, it should also reduce QSEs' total energy costs by allowing them to increase their energy purchases in the lower-priced market. However, volatility in balancing energy prices can create risks that affect convergence between forward prices and balancing energy prices. For example, risk-averse buyers are willing to pay a premium to purchase energy in the bilateral market thereby locking in their energy costs and avoiding the more volatile costs of the balancing energy market.

In this section, we measure two aspects of price convergence between forward and real-time markets. The first analysis investigates whether there are significant differences in prices between forward markets and the real-time market. The second tests whether there is a large spread between real-time and forward prices on a daily basis.

To determine whether there are significant differences between forward and real-time prices, we examine the difference between the average forward price and the average balancing energy price in each month between 2006 and 2009.[13] This analysis reveals whether persistent and predictable differences exist between forward and real-time prices, which participants should arbitrage over the long-term.

To measure the short-term deviations between real-time and forward prices, we also calculate the average of the absolute value of the difference between the forward and real-time price on a daily basis during peak hours. It is calculated by taking the absolute value of the difference between a) the average daily peak period price from the balancing energy market (*i.e.*, the average of the 16 peak hours during weekdays) and b) the day-ahead peak hour bilateral price. This measure captures the volatility of the daily price differences, which may be large even if the forward and balancing energy prices are the same on average. For instance, if forward 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 price difference between the forward 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. These two statistics are shown in Figure 10 for each month between 2006 and 2009.

13 Day-ahead bilateral prices as reported by Megawatt Daily are used to represent forward prices. For 2005-2007, we use the ERCOT Seller's Choice product. For 2008 and 2009, we use the average of the North, South and Houston Zone products.

Figure 10: Convergence between Forward and Real-Time Energy Prices

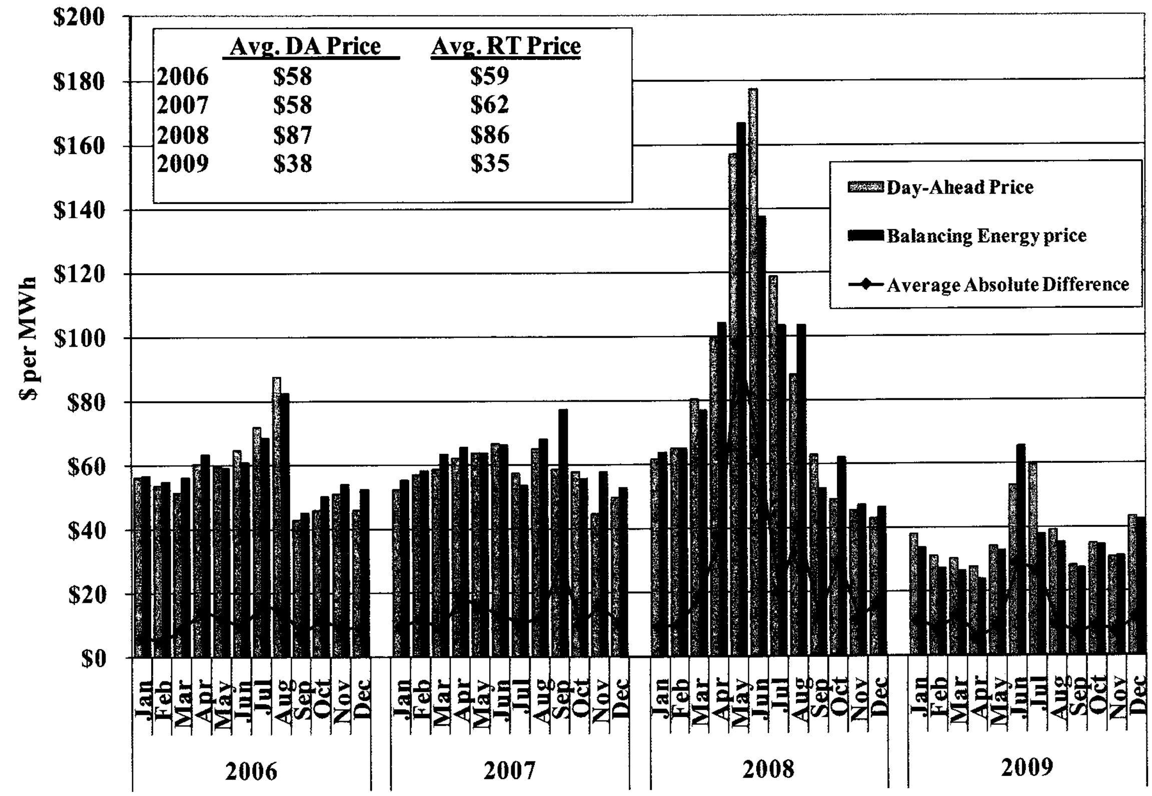


Figure 10 shows price convergence during peak periods (*i.e.*, weekdays between 6 AM and 10 PM). Day-ahead prices averaged $38 per MWh in 2009 compared to an average of $35 per MWh for real-time prices. Although the day-ahead and real-time prices exhibit relatively good average convergence in 2009, Figure 10 also shows that the average absolute price difference increased during the months of June and July 2009.

The average absolute difference was $10 in 2006, $14 in 2007, $31 in 2008 and $12 in 2009. As noted above, the average absolute difference measures the volatility of the price differences. Similar to the months of April, May and June 2008, the price volatility in June 2009 was due in large part to the significant and unpredictable transmission congestion experienced in that timeframe that caused average real-time prices to exceed day-ahead prices in June 2009. In contrast, average day-ahead prices were significantly higher than real-time prices in July 2009, which may be associated with transmission congestion expectations based on the experience in the prior month, as well as real-time pricing expectations associated with the extremely high temperatures and loads experienced during July 2009.

4. Volume of Energy Traded in the Balancing Energy Market

The primary purpose of the balancing energy market is to match supply and demand in real-time and to manage zonal congestion. In addition to fulfilling this purpose, the balancing energy market signals the value of power for market participants entering into forward contracts and plays a role in governing real-time dispatch. This section examines the volume of activity in the balancing energy market.

The average amount of energy traded in ERCOT's balancing energy market is small relative to overall energy consumption, although the balancing energy market can at times represent well over ten percent of total demand. Most energy is purchased and sold through forward contracts that insulate participants from volatile spot prices. Because forward contracting does not precisely match generation with real-time load, there will be residual amounts of energy bought and sold in the balancing energy market. Moreover, the balancing energy market enables market participants to make efficient changes from their forward positions, such as replacing relatively expensive generation with lower-priced energy from the balancing energy market.

Hence, the balancing energy market will improve the economic efficiency of the dispatch of generation to the extent that market participants make their resources available in the balancing energy market. In the limit, if all available resources were offered competitively in the balancing energy market (to balance up or down), prices in ERCOT's current market would be identical to prices obtained by clearing all power through a centralized spot market, even though most of the commodity currently settles bilaterally. It is rational for suppliers to offer resources in the balancing energy market even when they are fully contracted bilaterally because they may be able to increase their profit by reducing the output from their resources and support the bilateral sale with balancing energy purchases. Therefore the balancing energy market should govern the output of all resources, even though only a small portion of the energy is settled through the balancing energy market.

In addition to their role in governing real-time dispatch, balancing energy prices also provide a vital signal of the value of power for market participants entering into forward contracts. As discussed above, the spot prices emerging from the balancing energy market should directly affect forward contract prices, assuming that the market conditions and market rules allow the two markets to converge efficiently.

This section summarizes the volume of activity in the balancing energy market. Figure 11 shows the average quantities of up balancing and down balancing energy sold by suppliers in each month, along with the net purchases or sales (*i.e.*, up balancing energy minus down balancing energy).

Figure 11: Average Quantities Cleared in the Balancing Energy Market

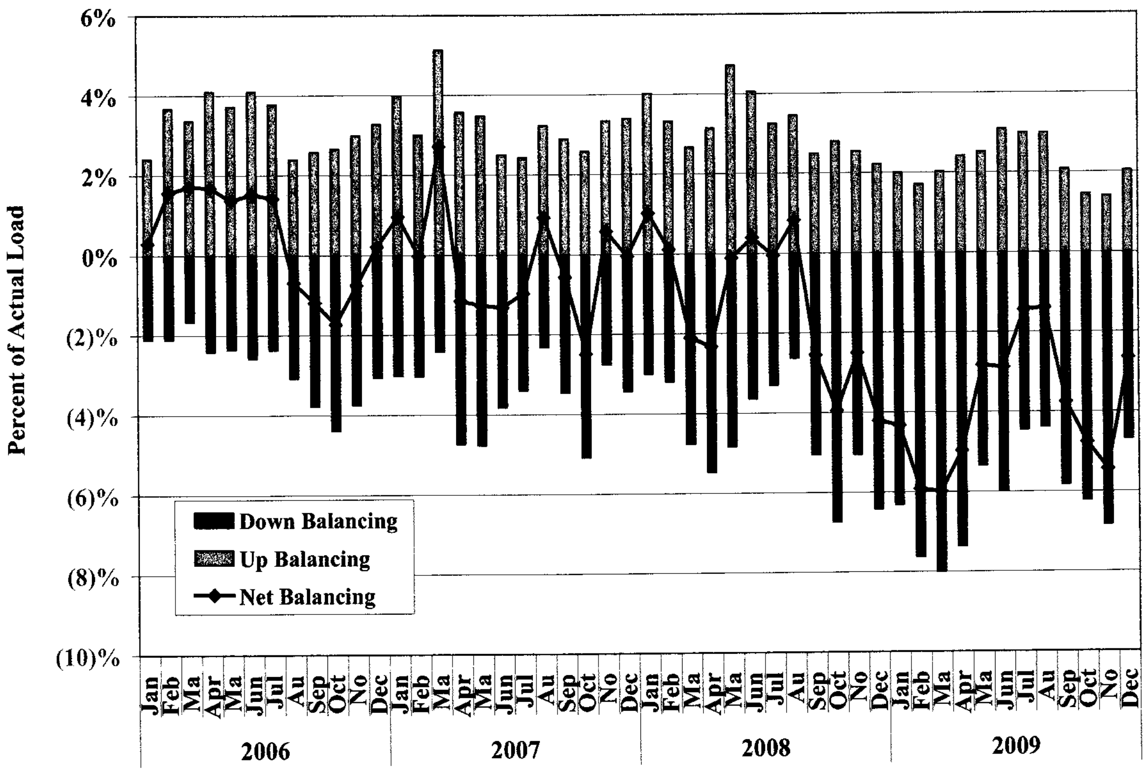
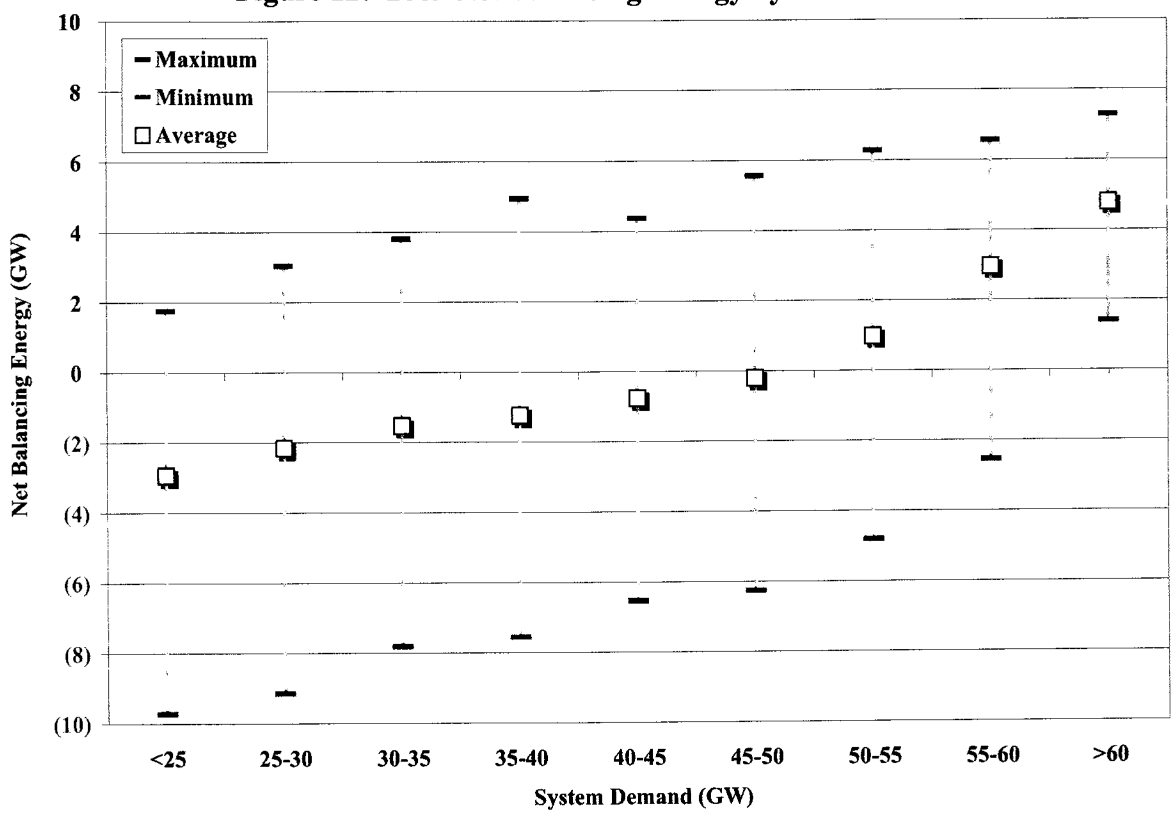


Figure 11 shows that the total volume of up balancing and down balancing energy as a share of actual load increased from an average of 7.7 percent in 2008 to 8.3 percent in 2009. Starting in August 2006, the average volume of down balancing energy began to increase. In 2008, for the first time the average amount of down balancing energy was greater than up balancing energy. This trend continued through 2009. The net quantity of balancing energy for every month in 2009 was negative, meaning that the average quantity of down balancing energy was greater than the quantity of up balancing energy. As discussed in Section II, this trend is related to the large increase in wind generation capacity added to the ERCOT region since the fall of 2008 and the associated scheduling patterns of these resources.

Figure 12 provides additional perspective to the monthly average net balancing energy deployments shown in Figure 11 by showing the net balancing energy deployments by load level for all intervals in 2009.

Figure 12: 2009 Net Balancing Energy by Load Level



While Figure 11 shows average net down balancing energy deployments in 2009, Figure 12 shows that this relationship is quite different when viewed as a function of the ERCOT system demand. Figure 12 shows average net down balancing deployments at load levels less than 50 GW, and average net up balancing deployments for load levels greater than 50 GW. Further, maximum net up balancing deployments exceeded 10 percent of demand at all system load levels in excess of 25 GW, except for levels exceeding 60 GW when net balancing deployments were exclusively in the upward direction.

Relaxed balanced schedules allow market participants to intentionally schedule more or less than their anticipated load, buying or selling in the balancing energy market to satisfy their actual load obligations. This scheduling flexibility allows the balancing energy market to operate as a centralized energy spot market. Although convergence between forward prices and spot prices has not been good on a consistent basis, the centralized nature of the balancing energy market facilitates participation in the spot market and improves the efficiency of the market results.

Aside from the introduction of relaxed balanced schedules, another reason for significant balancing energy quantities is that large quantities of up balancing and down balancing energy are often deployed simultaneously to clear "overlapping" balancing energy offers. Deployment of overlapping offers improves efficiency because it displaces higher-cost energy with lower-cost energy, lowering the overall costs of serving load and allowing the balancing energy price to more accurately reflect the marginal value of energy.

When large quantities of net up balancing or net down balancing energy are scheduled, it indicates that Qualified Scheduling Entities (QSEs) are systematically under-scheduling or over-scheduling load relative to real-time needs. If large hourly under-scheduling or over-scheduling occurs suddenly, the balancing energy market can lack the ramping capability (*i.e.*, how quickly on-line generation can increase or decrease its output) and sometimes the volume of energy offers necessary to achieve an efficient outcome. In these cases, large net balancing energy purchases can lead to transient price spikes when capacity exists to supply the need, but is not available in the 15-minute timeframe of the balancing energy market. The remainder of this sub-section and the next section will examine in detail the patterns of over-scheduling and under-scheduling that has occurred in the ERCOT market, and the effects that these scheduling patterns have had on balancing energy prices.

To provide a better indication of the frequency with which net purchases and sales of varying quantities are made from the balancing energy market, Figure 13 presents a distribution of the hourly net balancing energy. The distribution is shown on an hourly basis rather than by interval to minimize the effect of short-term ramp constraints and to highlight the market impact of persistent under- and over-scheduling. Each of the bars in Figure 13 shows the portion of the hours during the year when balancing energy purchases or sales were in the range shown on the x-axis. For example, the figure shows that the quantity of net balancing energy traded was between zero and positive 0.5 gigawatts (*i.e.*, loads were under-scheduled on average) in approximately 7 percent of the hours in 2009.

Figure 13: Magnitude of Net Balancing Energy and Corresponding Price

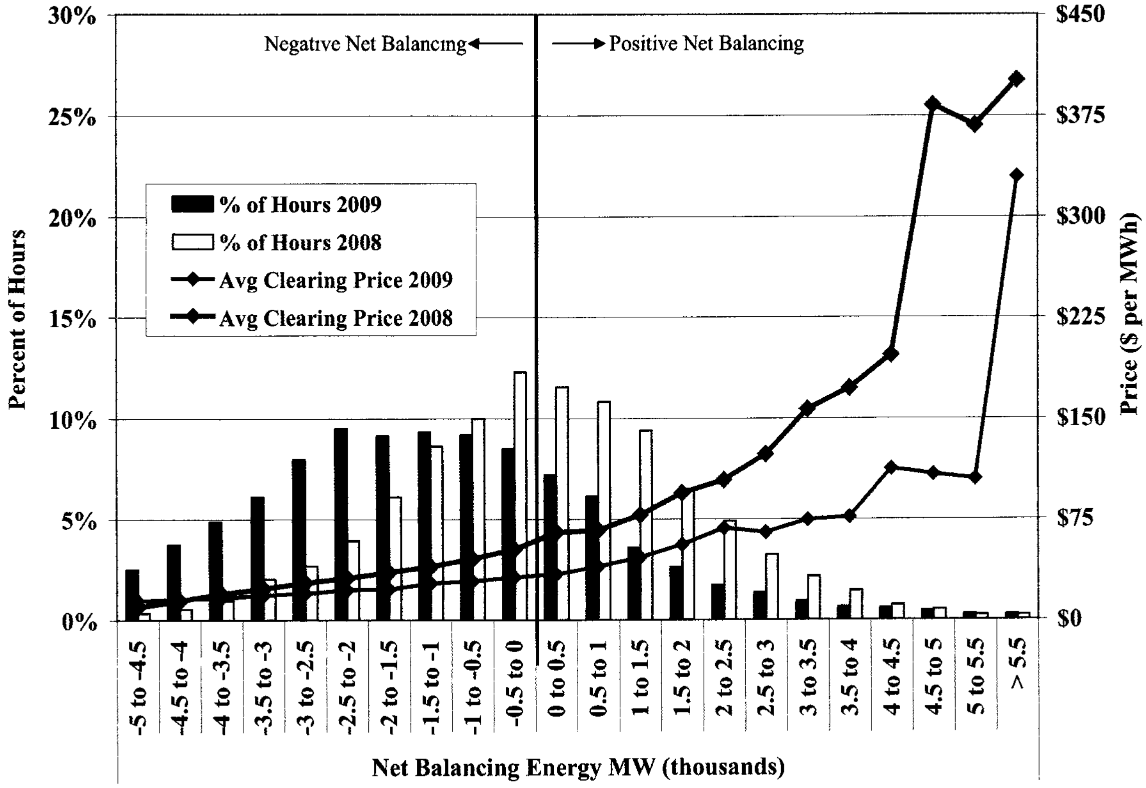


Figure 13 shows that the distribution of net balancing energy deployments in 2009 is shifted well to the left of zero, meaning that more down balancing energy was deployed than up balancing energy. This change in 2009 is consistent with the data shown in Figure 11, and is discussed in more detail in Section II. The lines plotted in Figure 13 show the average balancing energy prices corresponding to each level of balancing energy volumes for 2008 and 2009. In an efficiently functioning spot market, there should be little relationship between the balancing energy prices and the net purchases or sales. Instead, one should expect that prices would be primarily determined by more fundamental factors, such as actual load levels and fuel prices. However, this figure clearly indicates that balancing energy prices increase as net balancing energy volumes increase. This relationship is explained in part by the fact that net balancing energy deployments tend to be positively correlated with the level of demand as shown in Figure 12. However, scheduling practices and ramping issues contribute significantly to the observed pattern. We analyze this relationship more closely in the next subsections.

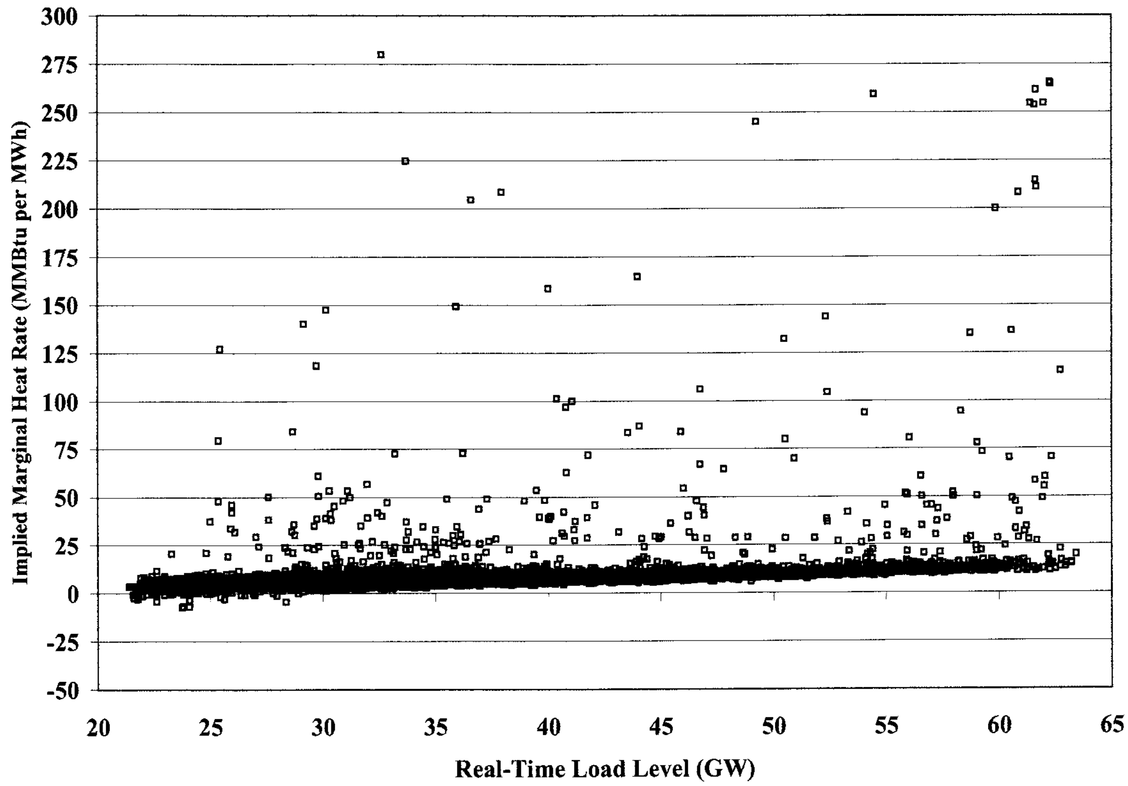
5. Determinants of Balancing Energy Prices

The prior section shows that the level of net sales in the balancing energy market appears to play a significant role in explaining the balancing energy prices. In this section, we examine this relationship in more detail, as well as the role of more fundamental determinants of balancing energy prices, such as the ERCOT load and fuel prices.

In an efficient market, we expect peak prices to occur under extreme demand conditions or as a result of unforeseen conditions that cause brief shortages, such as the loss of a large generator or an unanticipated rise in load. In ERCOT, prices in the balancing market can reach extremely high levels even when demand is not particularly high and absent such unforeseen operating conditions. This is primarily due to structural inefficiencies in the balancing energy market that are inherent to the zonal market model and the lack of a centralized unit commitment.

To further examine the relationship between actual load in ERCOT and balancing energy prices, Figure 14 shows the hourly average gas price-adjusted balancing energy prices versus the hourly average loads in ERCOT irrespective of time. This type of analysis shows more directly the relationship between balancing energy prices adjusted for natural gas prices and actual load. In a well-performing market, one should expect a clear positive relationship between these variables since resources with higher marginal costs must be dispatched to serve rising load.

Figure 14: Hourly Gas Price-Adjusted Balancing Energy Price vs. Real-Time Load



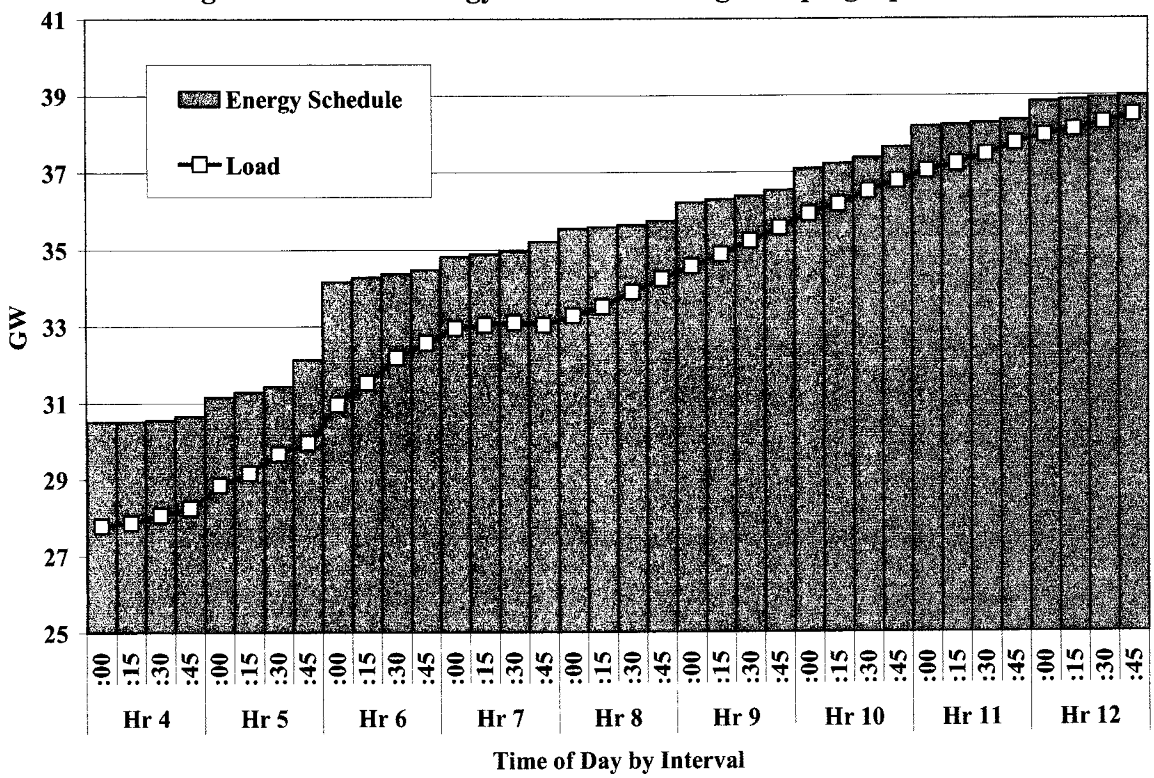
The figure indicates a positive correlation between real-time load and the clearing price in the balancing market. Although prices were generally higher at higher load levels, the data in Figure 13 indicate that the net volume of energy purchased in the balancing energy market is often a stronger determinant of price spikes than the level of demand.

6. Balancing Energy Market Scheduling

In the previous subsection, we analyzed balancing energy prices adjusted for fuel and load and found that while balancing energy prices are correlated to real-time load levels, other factors also have substantial effects on balancing energy levels. In this subsection, we investigate whether balancing energy prices are influenced by market participants' scheduling practices that tend to intensify the demand for balancing energy during hours when load is ramping.

We begin our analysis by examining factors that determine the demand for balancing energy during periods when load is ramping up and periods when it is ramping down. Figure 15 shows average energy schedules and actual load for each interval from 4 AM to 1 PM during 2009.

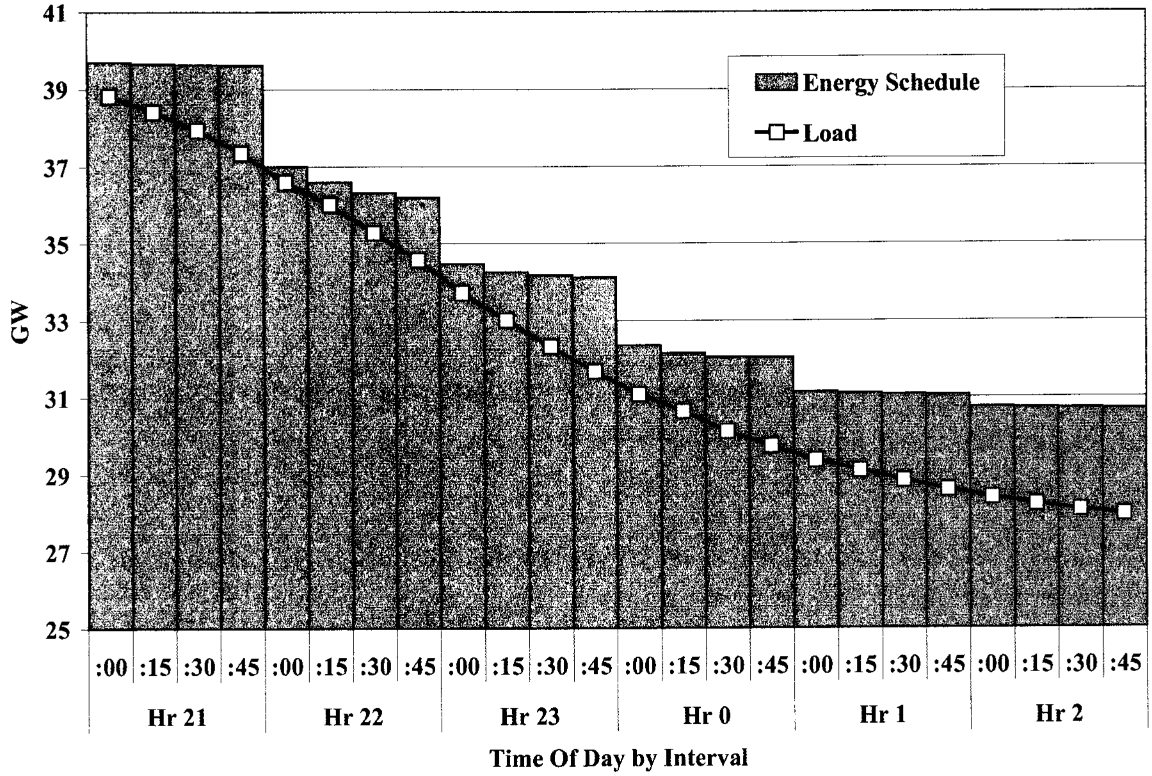
Figure 15: Final Energy Schedules during Ramping Up Hours



For ERCOT as a whole, energy schedules that are less than the actual load result in balancing energy purchases while energy schedules higher than actual load result in balancing energy sales. On average, load increases from approximately 28 GW to almost 39 GW in the nine hours shown in Figure 15, resulting in an average increase per 15-minute interval of approximately 330 MW.

The increase in load during ramping up hours is steady relative to the increase in energy schedules. Energy schedules rise less smoothly, with small increases from the first to fourth interval in each hour and larger increases from the fourth interval to the first interval of the next hour. For instance, the average energy schedule increases by more than 2.7 GW from the last interval of the hour ending 6 AM to the interval beginning at 6 AM, while the average energy schedule increases by only 160 megawatts in the subsequent three intervals. The same scheduling patterns exist in the ramping down hours. Figure 16 shows average energy schedules and load for each interval from 9 PM to 3 AM during 2009.

Figure 16: Final Energy Schedules during Ramping Down Hours



On average, load drops from approximately 39 GW to less than 29 GW in the six hours shown in Figure 16. The average decrease per 15-minute interval is 417 MW, although the rate of decrease is greatest from 9:45 PM to midnight. The progression of load during ramping down hours is steady relative to the progression of energy schedules. As was the case during ramping up hours, energy schedules change (decrease) in relatively large steps at the beginning of each hour. For example, the average energy schedule drops nearly 3.7 GW from the last interval before 10 PM to the interval beginning at 10 PM.

The sudden changes in energy schedules that occur at the beginning of each hour during ramping up hours and at the end of each hour during ramping down hours arise from the fact that much of the generation in ERCOT is scheduled by QSEs that submit energy schedules that change hourly. In addition, as indicated in Figure 15 and Figure 16, a number of schedules are based on bilateral contracts for 16-hour service, beginning as 6 AM and ending at 10 PM. Differences between energy schedules submitted by QSEs and load forecasted by ERCOT will result in purchases or sales in the balancing energy market. Specifically, the amount of net up balancing energy is equal to ERCOT's load forecast minus scheduled energy.

To evaluate the effects of systematic over- and under-scheduling more closely, we analyzed balancing energy prices and deployments in each interval during the ramping up period and ramping down period (consistent with the periods shown in Figure 15 and Figure 16). This analysis is similar to that shown in Figure 11 and Figure 12, except instead of showing balancing energy prices relative to load, we show balancing energy prices relative to net balancing energy deployments. Figure 17 shows the analysis for ramping up hours.

Figure 17: Balancing Energy Prices and Volumes Ramping Up Hours

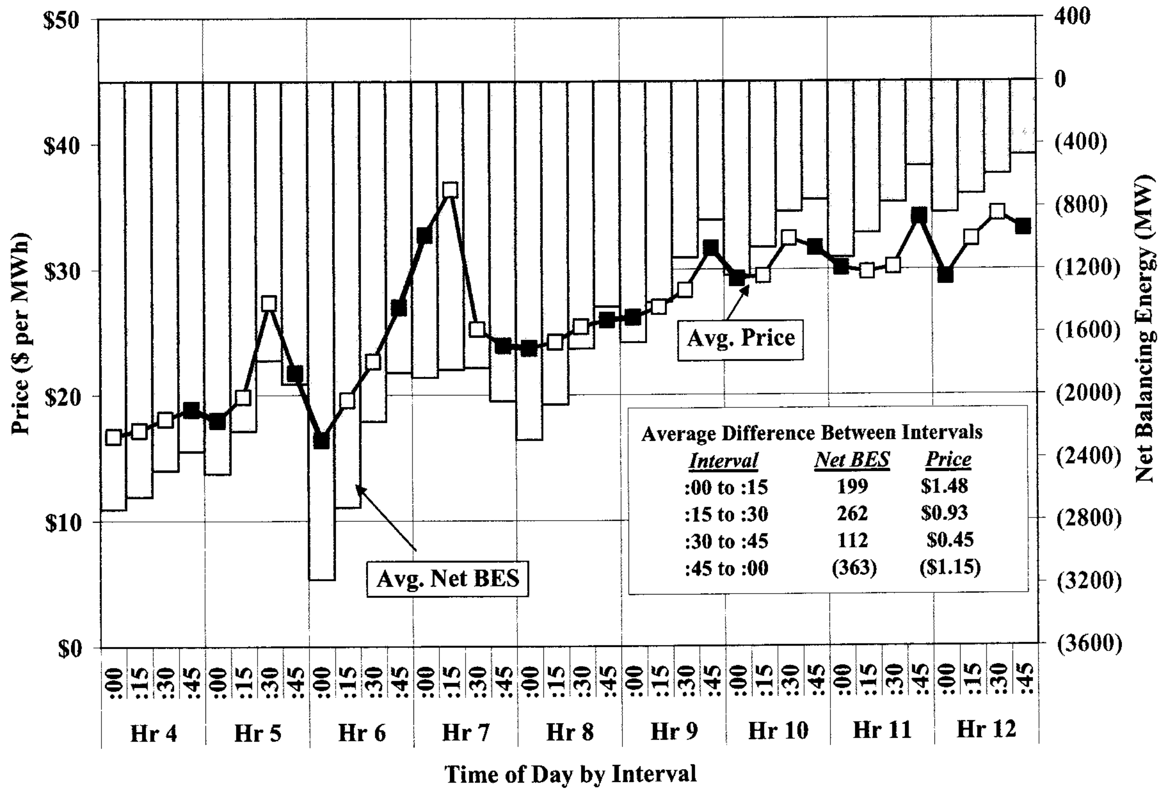


Figure 17 reveals two key aspects of the balancing energy market. First, as discussed above, balancing energy prices are highly correlated with balancing energy deployments. Second, with the exception of hour 7, there is a distinct pattern of increasing net balancing energy deployments during the hour. This is consistent with the notion that hourly schedules are established at a level that corresponds to an average expected load for the hour. The scheduling patterns that create these balancing deployments result in inefficient prices that are relatively volatile and could result in erratic dispatch signals to the generators.

Figure 18: Balancing Energy Prices and Volumes Ramping Down Hours

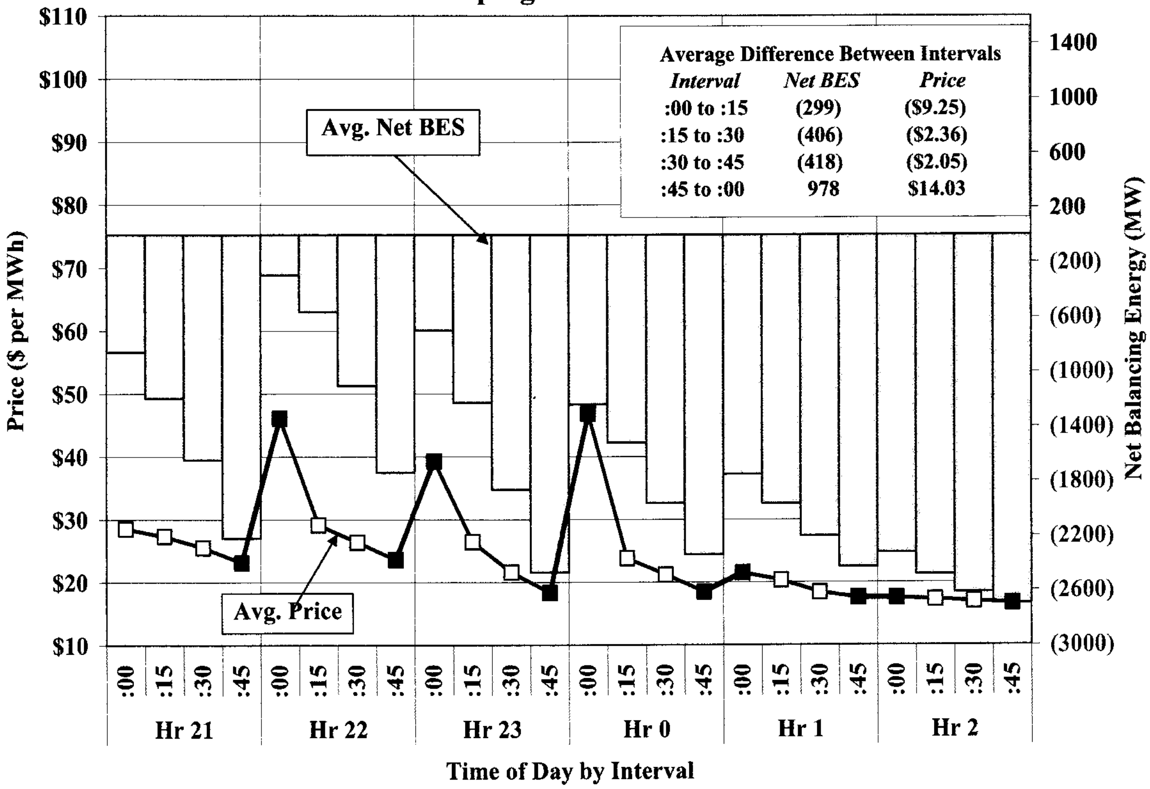


Figure 18 shows the same analysis for the ramping down hours. During ramping down hours, at the beginning of the hour, actual load tends to be higher than energy schedules, resulting in substantial balancing energy purchases. At the end of the hour actual load tends to be lower relative to the energy schedules, resulting in lower balancing energy demand.

To further examine how balancing energy prices relate to actual load levels, the final analysis in this subsection shows the average balancing energy prices by interval during the hours each day when load is increasing or decreasing rapidly (*i.e.*, when load is ramping up and ramping down). ERCOT load increases during the day from an average of almost 28 GW at 4 AM to 39 GW at 1 PM. Thus, the change in load averages 1,290 MW per hour (322 MW per 15-minute interval) during the morning and early afternoon. Figure 19 shows the average load and balancing energy price in each interval from 4 AM through 1 PM during 2009.

Figure 19: Average Balancing Energy Prices and Load by Time of Day Ramping Up Hours

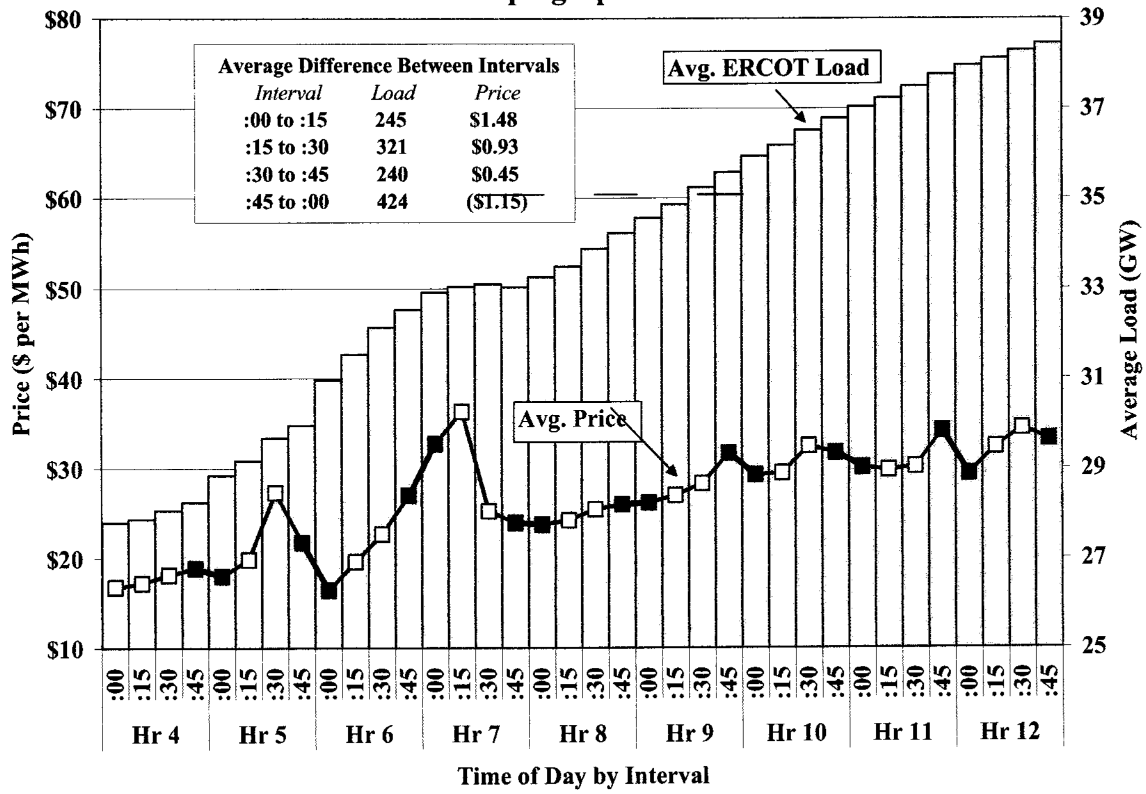


Figure 19 shows that, with the exception of hour 7, load steadily increases in every interval and prices generally move upward from an average of $18 per MWh at 4:00 AM to $32 per MWh at 12:45 PM. If actual load were the primary determinant of energy prices, the balancing energy prices would rise gradually as the actual load rises. However, Figure 19 shows this is not the case. In most hours the balancing energy price rises throughout each hour and drops substantially in the first interval of the next hour. In the figure, the red lines highlight the transition from one hour to the next hour. The average price change from the last interval of one hour to the first interval of the next hour is -$1.15 per MWh. This occurs because participants tend to change their schedules once per hour, bringing on additional substantial quantities of generation at the beginning of the hour which reduces the balancing energy prices.

A similar pattern is observed at the end of the day when load is decreasing. In ERCOT, load tends to decrease in the evening more quickly than it increases early in the day. Most of the decrease occurs over a six hour period, averaging a decrease of 1,891 MW per hour (473 MW per 15-minute interval) during the late evening. Figure 20 shows this decrease in load by interval, together with the average balancing energy prices for the intervals from 9 PM to 3 AM.

Figure 20: Average Balancing Energy Prices and Load by Time of Day Ramping Down Hours

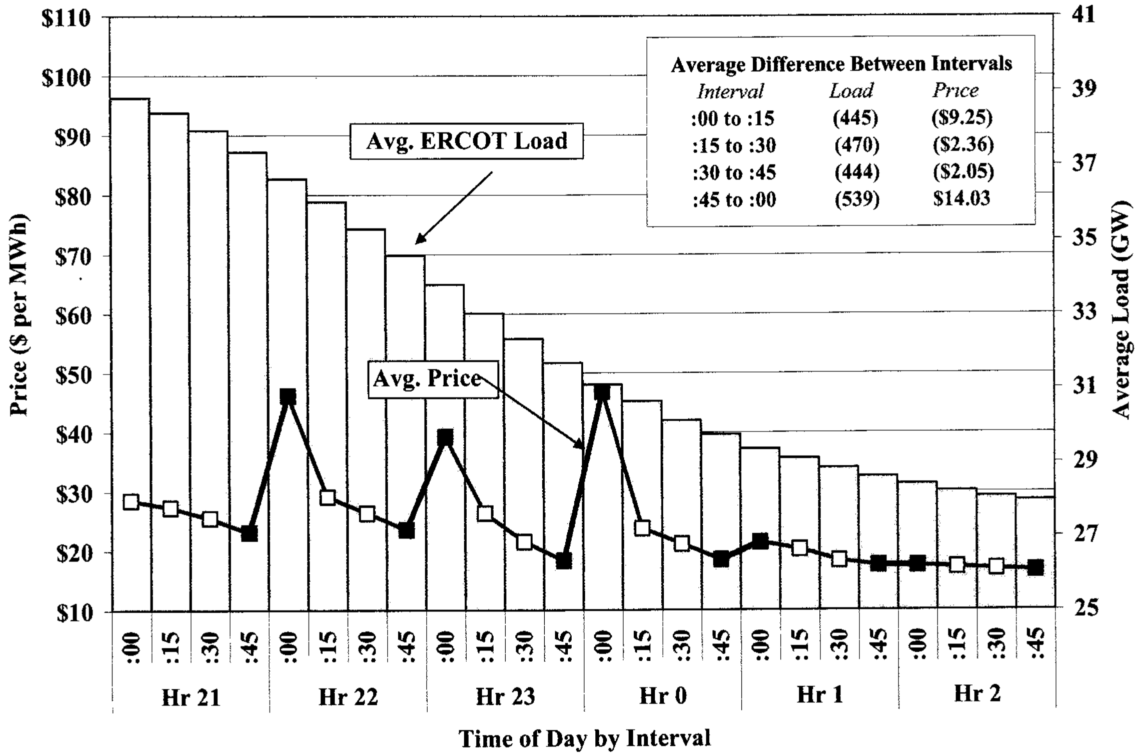


Figure 20 shows that while balancing energy prices decrease over these intervals, the pattern is similar to that exhibited in the ramping up hours. The balancing energy price decreases in each interval of the hour before rising substantially in the first interval of the following hour. The balancing energy price increases by an average of $14.03 per MWh from the last interval of one hour to the first interval of the next hour during this period. This occurs because participants tend to change their schedules once per hour, de-committing generating resources at the beginning of the hour. Because the supply decreases at the beginning of these hours by much more than load decreases, the balancing energy prices generally increase. This is consistent with the patterns of energy schedules and balancing prices in 2006 through 2008.[14]

[14] See 2006, 2007 and 2008 SOM Reports.

Collectively, these figures show that this pattern of balancing energy prices by interval is not explained by changes in actual load. Rather, changes in balancing energy deployments by interval underlie this pricing pattern. Sizable changes in balancing energy deployments occur between intervals, particularly in the first interval of the hour. These changes are associated with large hourly changes in energy schedules.

While QSEs have the option to submit schedules that change for every 15 minute interval, many QSEs schedule only on an hourly basis, making little or no changes on a 15-minute basis. It is primarily the scheduling patterns by the QSEs that schedule on an hourly basis that result in the balancing energy deployments and prices shown in Figure 17 and Figure 18.

The analysis in this section shows that one of the significant issues in the current ERCOT market is the tendency of most QSEs to alter their energy schedules hourly. This tendency may be related to the fact that balancing energy bids and offers are submitted hourly and are made relative to the energy schedule. For example, if a QSE schedules 200 MW from a 300 MW resource, it may offer the remaining 100 MW in the balancing energy market. If it schedules 230 MW, it may offer 70 MW. However, if the energy schedule changes on a 15-minute basis, it may be difficult to reconcile the schedule with the hourly balancing energy offer, leading most QSEs to simply submit hourly schedules. This places a burden on the balancing energy market to reconcile the differences between the hourly schedules and the 15-minute actual load levels, which can result in inefficient price fluctuations. This issue should not continue to be a problem under the nodal market design since resource-specific offers will not be interpreted as a deviation from an energy schedule.

As discussed in this subsection, a significant portion of the volatility of the balancing energy prices in each interval is related to the energy scheduling patterns. This volatility can be exacerbated when portfolio ramp rates are binding. Portfolio ramp rates are constraints QSEs submit with their balancing energy offers to limit the quantity of up balancing or down balancing energy that may be deployed in one interval. These ramp rates are important because they prevent a QSE from receiving deployment instructions that it cannot meet physically. Large changes in balancing energy deployments from interval to interval can cause the ramp rate constraints to bind, preventing the deployment of lower-cost offers and compelling the deployment of higher-cost offers from other QSEs. Ramp rate constraints can also be limiting when resources are instructed to ramp down quickly, although this is less common.

In many cases, the lack of ramp capable resources offered to the balancing energy market results in inefficient price spikes.[15] The efficiency implications associated with these issues continued in 2009 and will likely continue until the current zonal market design is replaced. However, ERCOT implemented 14 minute ramp rates in late October 2009 that are expected to help make more balancing energy ramping capability available, which in turn is expected to reduce the frequency and magnitude of price spikes associated with large schedule changes.[16]

B. Ancillary Services Market

The primary ancillary services are up regulation, down regulation, and responsive reserves. Market participants may self-schedule ancillary services or purchase their required ancillary services through the ERCOT markets. Historically, ERCOT has also procured non-spinning reserves as needed during periods of increased supply and demand uncertainty. However, beginning in November 2008, ERCOT began procuring non-spinning reserves across all hours based on its assessment of "net load" error, where "net load" is equal to demand minus wind production. This section reviews the results of the ancillary services markets in 2009.

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 at least 2,300 MW of responsive reserves to ensure adequate protection against the loss of the two largest units. Non-spinning reserves are procured as a means for ERCOT to implement supplemental generator commitments to increase the supply of energy in the balancing energy market if needed. The balancing energy market deployments that occur in the 15-minute timeframe and regulation deployments that occur in the 4-second timeframe are the primary means for meeting load fluctuations across and within each 15-minute interval.

---

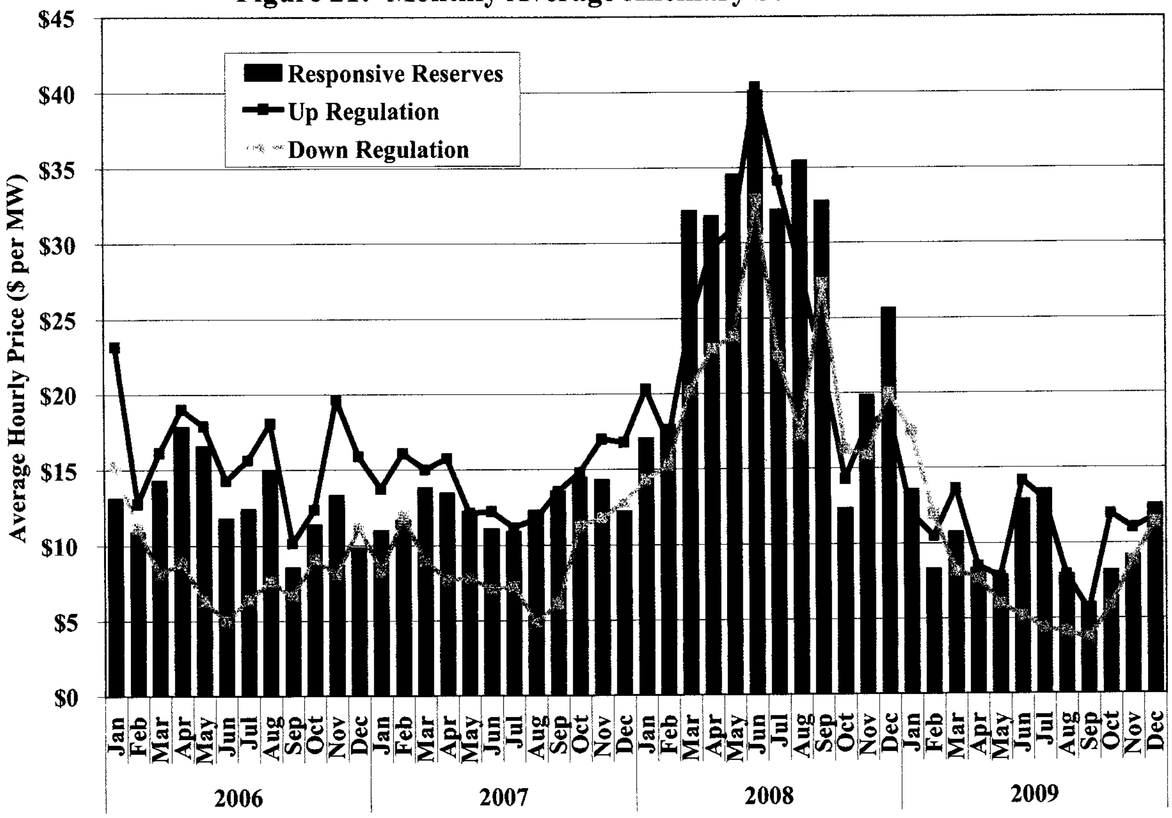
15 2005 SOM Report at 68-76.

16 There are insufficient data to perform an assessment of the effects of the 14-minute ramp implementation in 2009.

1. Reserves and Regulation Prices

Our first analysis in this section provides a summary of the ancillary services prices over the past four years. Figure 21 shows the monthly average ancillary services prices between 2006 and 2009. Average prices for each ancillary service are weighted by the quantities required in each hour.

Figure 21: Monthly Average Ancillary Service Prices



This figure shows that ancillary service capacity prices generally returned to levels seen in 2006 and 2007 after reaching significantly higher levels in 2008. These price movements can be primarily attributed to the variations in energy prices that occurred over the same timeframe.

Because ancillary services markets are conducted prior to the balancing energy market, participants must include their expected value of foregone sales in the balancing energy market in their offers for responsive reserves and regulation. Providers of both responsive reserves and up regulation can incur such opportunity costs if they reduce the output from economic units to make the capability available to provide these services. Likewise, providers of down regulation can incur opportunity costs in real-time if they receive instructions to reduce their output below the most profitable operating level. Further, because generators must be online to provide regulation and responsive reserves, there is an economic risk during low-price periods of operating uneconomically at minimum output levels (or having to operate above minimum output levels if providing down regulation).

Figure 21 shows that average down regulation prices have been lower than prices for up regulation service over the last four years, indicating that the opportunity costs were greater for providers of up regulation, with the exception of September 2008 through February 2009 when the average down regulation price was slightly higher than the average up regulation price.

Figure 21 also shows that, on average, the price of up regulation is slightly higher than the price of responsive reserves from 2006 through 2009. This is consistent with expectations because a supplier incurs opportunity costs to provide either service, while providing up regulation can generate additional costs. These additional costs include (a) the costs of frequently changing resource output levels, and (b) the risk of having to produce output when regulating at balancing energy prices that are less than the unit's variable production costs. However, during periods of persistent high prices, up regulation providers may have lower opportunity costs than responsive reserves providers to the extent that they are dispatched up to provide regulation. This factor explains in part the reversal in the relationship between responsive reserve and up regulation prices in 2008 when average responsive reserve prices were greater than or equal to average up regulation prices in seven out of twelve months.

One way to evaluate the rationality of prices in the ancillary services markets is to compare the prices for different services to determine whether they exhibit a pattern that is reasonable relative to each other. Table 1 compares the average prices for responsive reserves and non-spinning reserves over the past four years in those hours when ERCOT procured non-spinning reserves. Non-spinning reserves were purchased in approximately 20 and 14 percent of hours in 2006 and 2007, respectively, but increased to 51 percent of the hours in 2008. ERCOT began procuring non-spinning reserves in every hour beginning in November 2008, primarily to address the increasing uncertainty in net load associated with increasing levels of intermittent generation resources.

Table 1: Average Hourly Responsive Reserves and Non-Spinning Reserves Prices During Hours When Non-Spinning Reserves Were Procured

	2006	2007	2008	2009
Non-Spin Reserve Price	$21.75	$6.07	$7.97	$3.08
Responsive Reserve Price	$25.55	$16.74	$36.39	$9.68

Table 1 shows that responsive reserves prices are higher on average than non-spinning reserves prices during hours when non-spinning reserves were procured. It is reasonable that responsive reserves prices would generally be higher since responsive reserves are a higher quality product that must be delivered in 10 minutes from on-line resources while non-spinning reserves must be delivered in 30 minutes. Further, the significant reduction in the price of non-spinning reserves relative to responsive reserves beginning in 2007 was associated with the implementation of Protocol Revision Request ("PRR") 650, which significantly reduced the risk of uneconomic deployments for providers of non-spinning reserves, thereby reducing the capacity price for the provision of this service.

In contrast to the previous data that show the individual ancillary service capacity prices, Figure 22 shows the monthly total ancillary service costs per MWh of ERCOT load and the average balancing energy price for 2006 through 2009.

Figure 22: Ancillary Service Costs per MWh of Load

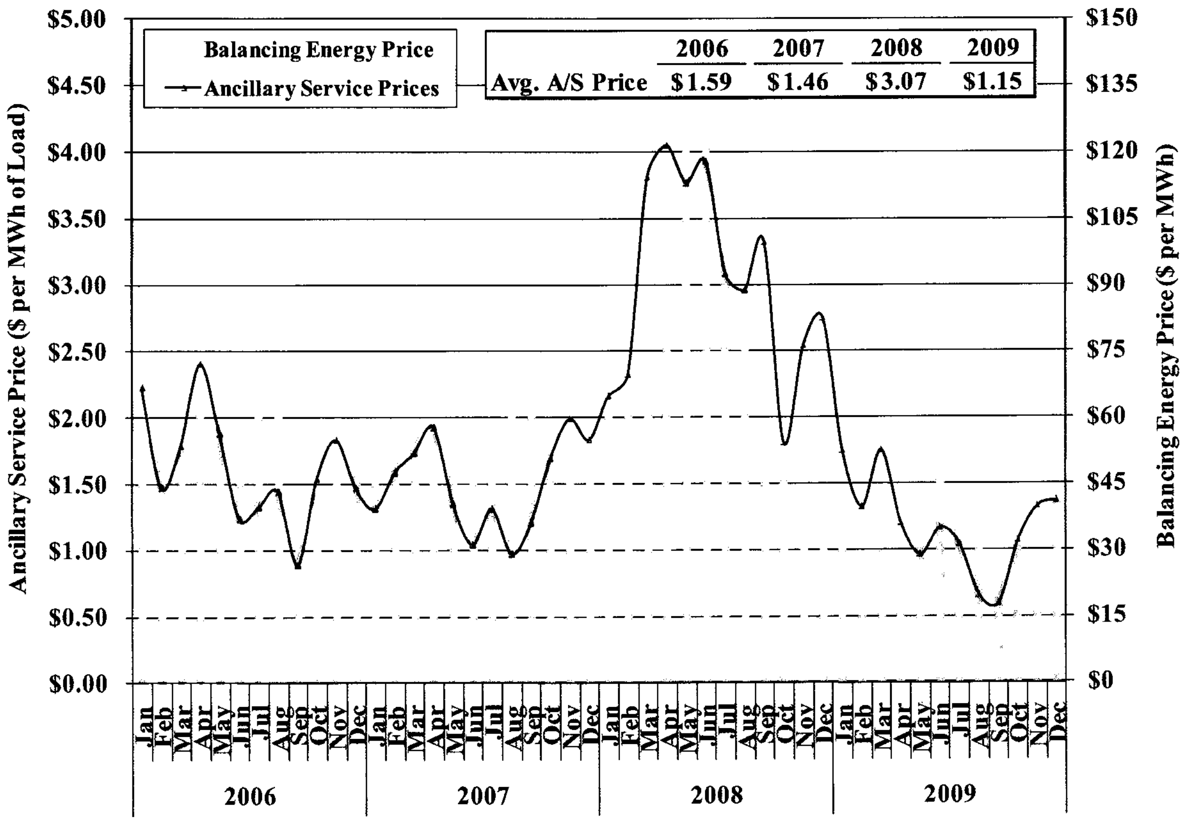


Figure 22 shows that total ancillary service costs are generally correlated with balancing energy price movements which, as previously discussed, are highly correlated with natural gas price movements. The average ancillary service cost per MWh of load decreased to $1.15 per MWh in 2009 compared to $3.07 per MWh in 2008, a decrease of more than 63 percent. Ancillary service costs were equal to 4.0 and 3.5 percent of the load-weighted average energy price in 2008 and 2009, respectively.

Our next analysis evaluates the variations in regulation prices. Regulation providers continuously vary their output levels to keep ERCOT-wide load and generation continually in balance during the time between SPD instructions, which are issued every fifteen minutes. When load and generation fluctuate by larger amounts, additional regulation resources are needed to keep the system in balance. This is particularly important in ERCOT due to the limited interconnections with adjacent areas, which results in much greater variations in frequency when generation does not precisely match load. Movements in load and generation are greatest when the system is ramping, thus ERCOT needs substantially more regulating capacity during ramping hours

Figure 23 shows the relationship between the quantities of regulation required by ERCOT and regulation price levels. This figure compares regulation prices to the average regulation quantity (both up and down regulation) procured, shown for each hour of the day. Regulation prices are weighted by the quantities of each service procured.

The figure shows that ERCOT requires approximately 1,350 MW of regulation capability prior to the initial ramping period (beginning at 6 AM). The requirement then increases to more than 1,900 MW during the steepest ramping hours from 6 AM to 9 AM. The requirement declines to about 1,400 MW during the late morning and afternoon hours when system load is relatively steady. From 6 PM until midnight, the system is ramping down rapidly and demand for regulation averages approximately 1,800 MW.

Figure 23: Regulation Prices and Requirements by Hour of Day

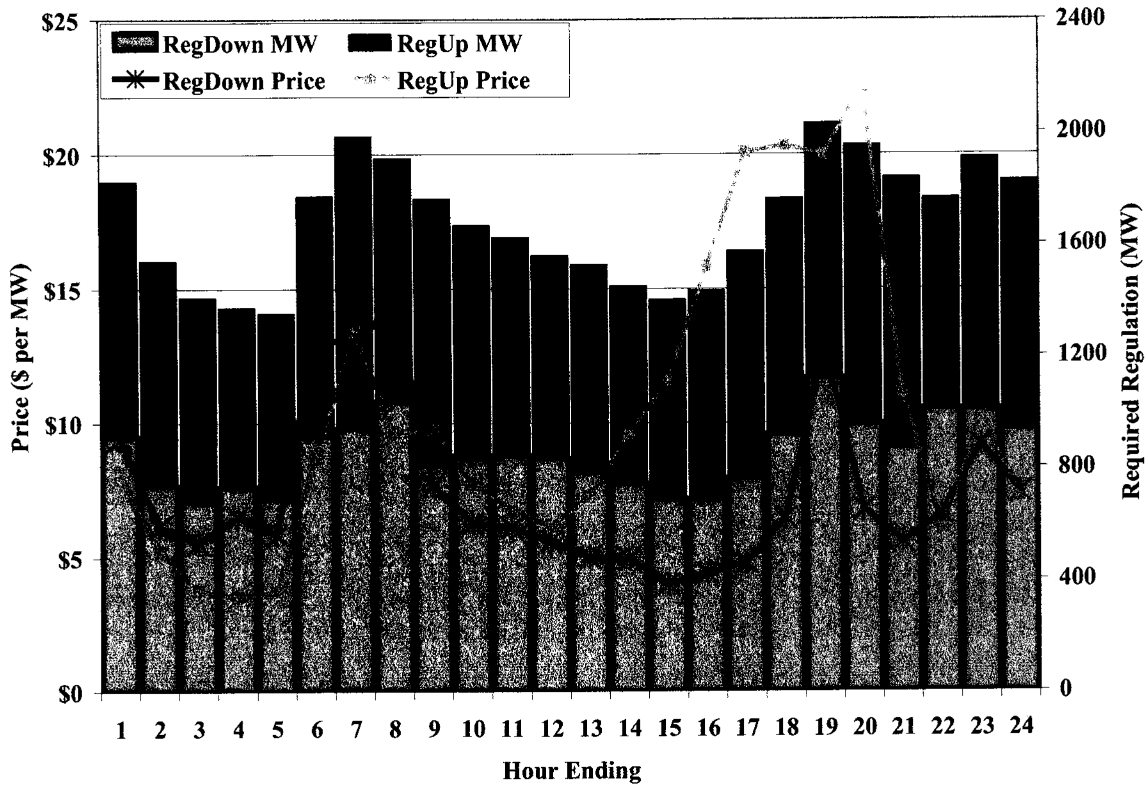


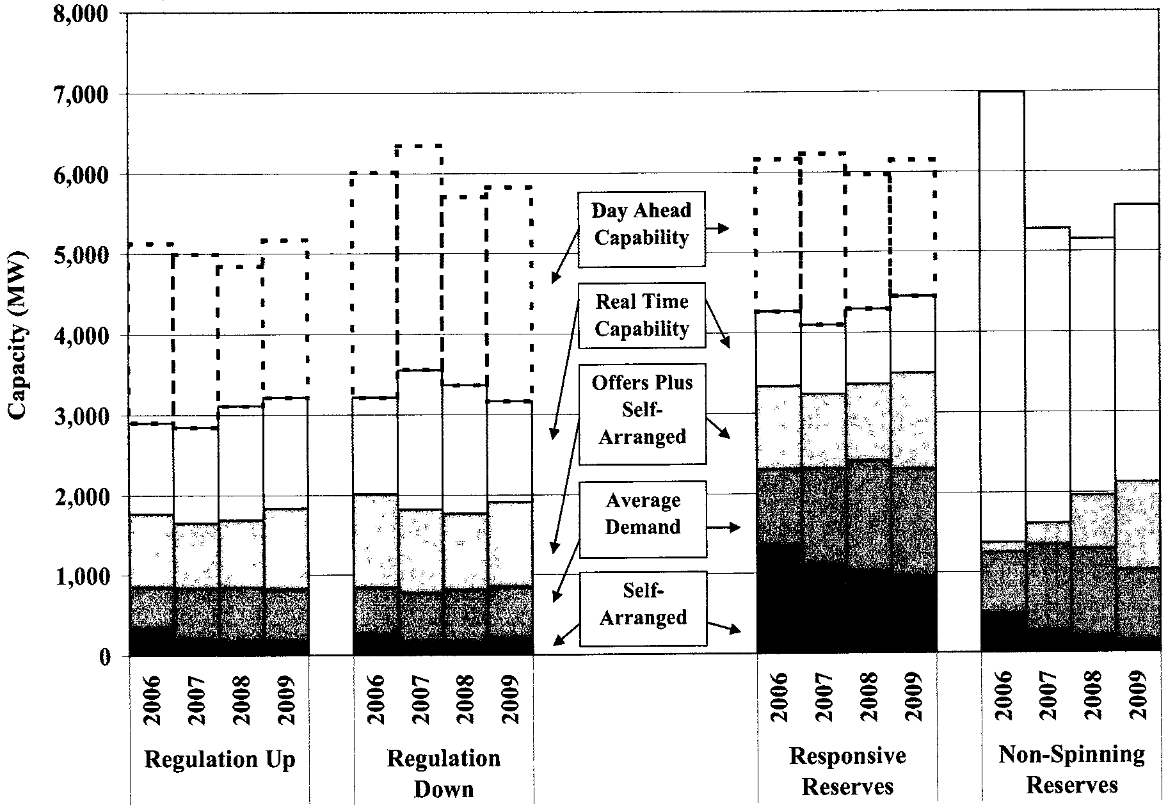
Figure 23 indicates that average regulation prices are generally correlated with the regulation quantity purchased and the typical load pattern in ERCOT. During non-ramping hours, such as overnight and late morning, up and down regulation prices are at their lowest levels. During the ramping hours in early morning average up and down regulation prices reached approximately $11 per MW. During evening ramping hours, down regulation prices also reached $7 per MW, while up regulation prices topped out at almost $14 per MW. Up regulation prices are higher on average in the late afternoon hours because load levels and balancing energy prices are typically higher in these hours and the amount of capacity available to supply up regulation is lower than in other hours.

2. Provision of Ancillary Services

To better understand the reserve prices and evaluate the performance of the ancillary services markets, we analyze the capability and offers of ancillary services in this section. The analysis is shown in Figure 24. This figure summarizes the quantities of ancillary services offered and self-arranged relative to the total capability and the typical demand for each service. The bottom segment of each bar in Figure 24 is the average quantity of ancillary services self-arranged by owners of resources or through bilateral contracts. The second segment of each bar is the average amount offered and cleared in the ancillary services market. Hence, the sum of the first two segments is the average demand for the service.

The third segment of each bar is the quantity offered into the auction market that is not cleared. Therefore, the sum of the second and third segments is the total quantities offered in each ancillary services auction on average, including the quantities cleared and not-cleared. The empty segments correspond to the ancillary services capability that is not scheduled or offered in the ERCOT markets. The lower part of the empty segments correspond to the amount of real-time capability that is not offered while the top part of the empty segments correspond to the additional quantity available in the day-ahead that was not offered. Capabilities are generally lower in the real-time because offline units that require significant advance notice to start-up will not be capable of providing responsive reserves or regulation in real time (only capability held on online resources is counted).

Figure 24: Reserves and Regulation Capacity, Offers, and Schedules



The capability shown in Figure 24 incorporates ERCOT's requirements and restrictions for each type of service. For regulation, the capability is calculated based on the amount a unit can ramp in five minutes for those units that have the necessary equipment to receive automatic generation control signals on a continuous basis. For responsive reserves, the capability is calculated based on the amount a unit can ramp in ten minutes. This is limited by an ERCOT requirement that no more than 20 percent of the capacity of a particular resource is allowed to provide responsive reserves. However, the responsive reserve capability shown in Figure 24 is not reduced to account for energy produced from each unit, which causes the capability on some resources to be overstated in some hours.

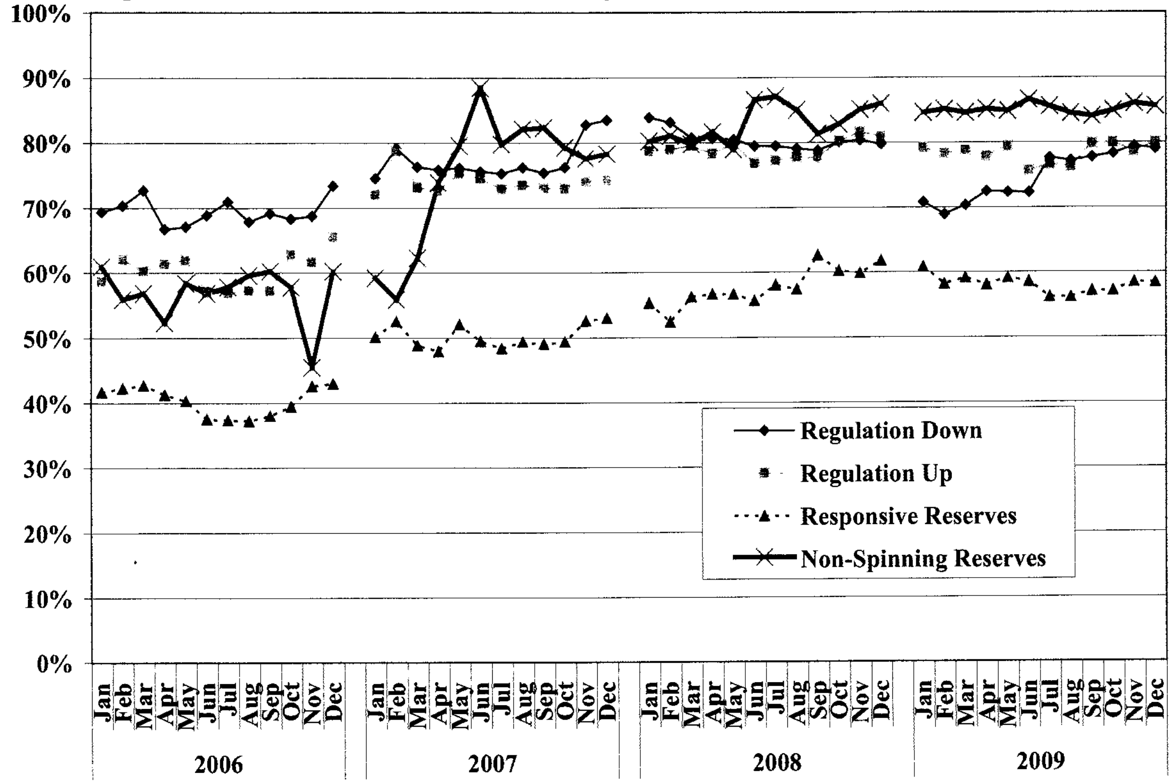
For non-spinning reserves, Figure 24 includes the capability of units that QSEs indicate are able to ramp-up in thirty minutes and able to start-up on short notice. The total capability shown in this figure does not account for capacity of online resources. However, it should be noted that any on-line resource with available capacity can provide non-spinning reserves, so the actual capability is larger than shown in the figure.

Figure 24 shows that except for responsive reserves, for which approximately 55 percent of available responsive reserve capacity was offered, less than one-half of each type of ancillary services capability was offered during the year from 2006 to 2009. One explanation for these levels of offers is that the ancillary services markets are conducted ahead of real time so participants may not offer resources that they expect to dispatch to serve their load or to support sales in the balancing energy market. In other words, some of the available reserves and regulation capability becomes unavailable in real time because the resources are dispatched to provide energy. The current market design creates risk and uncertainty for suppliers who must predict one day in advance whether their resources will be more valuable as energy or as ancillary services.

In addition, participants may not offer the capability of resources they do not expect to commit for the following day. Suppliers could submit offer prices high enough to ensure that their costs of committing additional resources to support the ancillary services offers are covered. However, under the current market design, ancillary services are procured independently for each hour and not optimized over the entire day (e.g., including minimum run times and minimum quantities), which greatly increases the risk for generators. The nodal market will include co-optimized procurement of energy and reserves over the entire operating day, which should enhance the efficiency of the procurement of reserves.

These services can be self-supplied from owned resources or from resources purchased bilaterally. To evaluate the quantities of ancillary services that are not self-supplied more closely, Figure 25 shows the share of each type of ancillary service that is purchased through the ERCOT market.

Figure 25: Portion of Reserves and Regulation Procured Through ERCOT

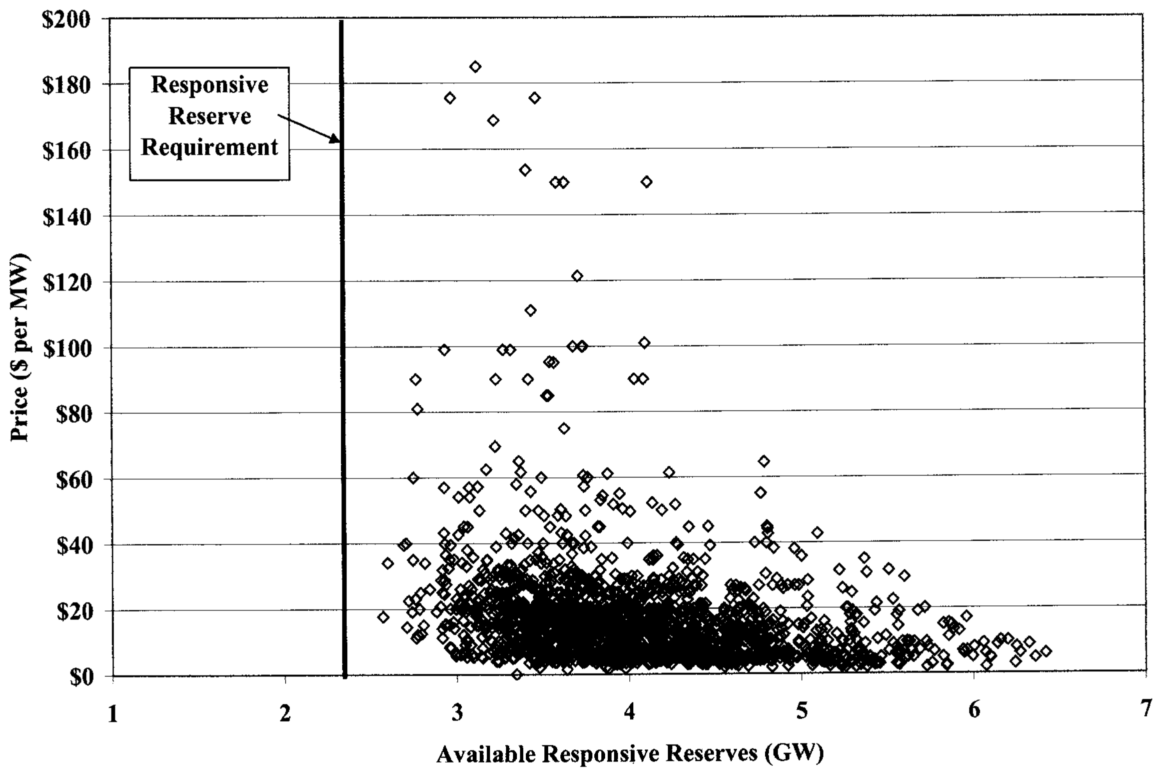


As market participants have gained more experience with the ERCOT markets, larger portions of the available reserves and regulation capability have been offered into the market, thereby increasing the market's liquidity. Nevertheless, Figure 25 shows that a fair share of these services is still self-supplied, particularly responsive reserves.

Prices in the ERCOT responsive reserve market tend to be somewhat higher than in other markets that co-optimize the procurement and dispatch of energy and responsive reserves. Responsive reserve prices in the ERCOT market are also affected by relatively higher requirements than other markets, as well as reliability restrictions that limit the quantity of responsive reserves that can be provided by each generating unit. Lower prices occur in co-optimized markets because the procurement is optimized with energy over the entire operating day and in most hours there is substantial excess online capacity that can provide responsive reserves at very low incremental costs. For example, a steam unit that is not economic to operate at its full output in all hours will have output segments that can provide responsive reserves at very low incremental costs. If the surplus responsive reserves capability from online resources is relatively large in some hours, one can gauge the efficiency of the ERCOT reserves market by evaluating the prices in these hours.

Figure 26 plots the hourly real-time responsive reserves capability against the responsive reserves prices during the peak afternoon hours of 2 PM to 6 PM. The capability calculated for this analysis reflects the actual energy output of each generating unit and the actual dispatch point for LaaRs. Hence, units producing energy at their maximum capability will have no available responsive reserves capability and, consistent with ERCOT rules, the responsive reserve that can be provided by each generating unit is limited to 20 percent of the unit's maximum capability. The figure also shows the responsive reserves requirement of 2,300 MW in 2009 to show the amount of the surplus in each hour.

Figure 26: Hourly Responsive Reserves Capability vs. Market Clearing Price Afternoon Peak Hours



In a well functioning-market for responsive reserves, we would expect excess capacity to be negatively correlated with the clearing prices. The data in this figure indicate only a weak negative correlation. Particularly surprising is the frequency with which price exceeds $20 per MW when the responsive reserve capability is more than 2,000 MW higher than the requirement.

In these hours the marginal costs of supplying responsive reserves should be very low. These results reinforce the potential benefits which should result from jointly optimizing the operating reserves and energy markets. The upcoming nodal market implementation will include day ahead co-optimization, but not real-time.

II. Demand and Resource Adequacy

The first section of this report reviewed the market outcomes and provided analyses of a variety of factors that have influenced the market outcomes. This section reviews and analyzes the load patterns during 2009 and the existing generating capacity available to satisfy the load and operating reserve requirements.

A. ERCOT Loads in 2009

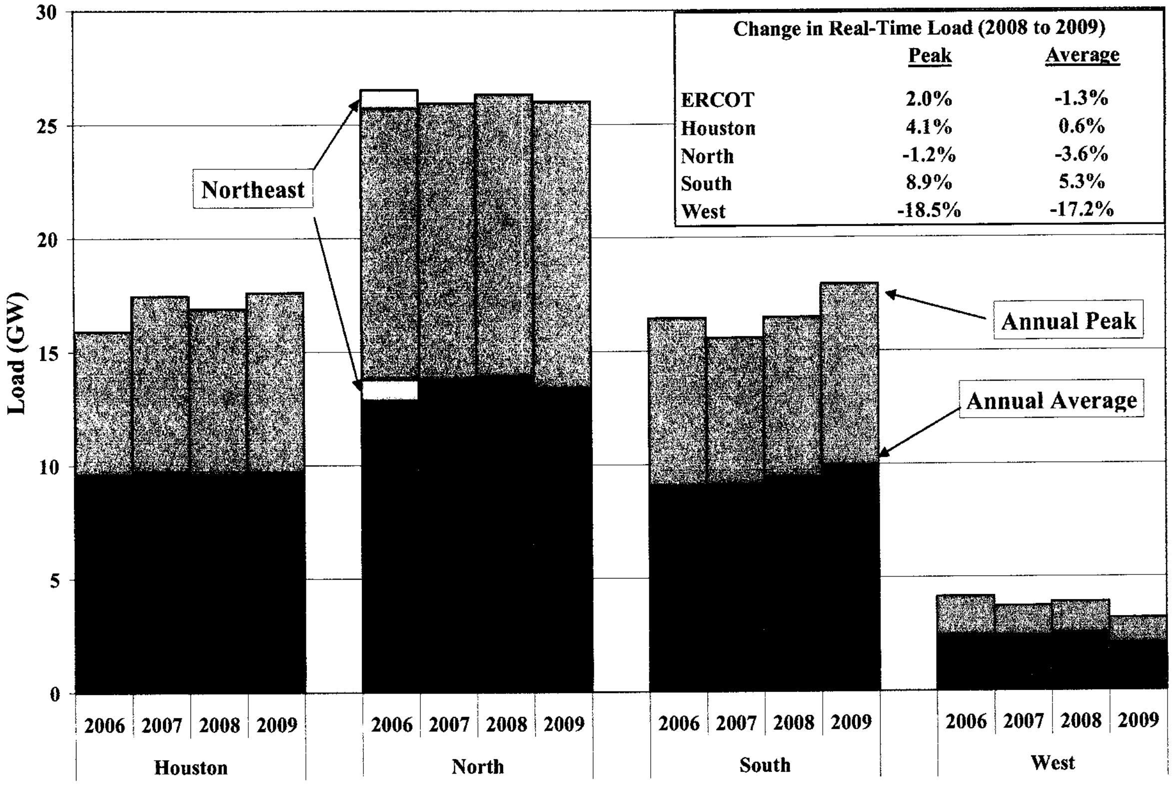
There are two important dimensions of load that should be evaluated separately. First, the changes in overall load levels from year to year can be shown by tracking the changes in average load levels. This metric will tend to capture changes in load over a large portion of the hours during the year. Second, it is important to separately evaluate the changes in the load during the highest-demand hours of the year. Significant changes in these peak demand levels have historically been very important and played a major role in assessing the need for new resources. The expectation in a regulated environment was that adequate resources would be acquired to serve all firm load, and this expectation remains in the competitive market. The expectation of resource adequacy is based on the value of electric service to customers and the damage and inconvenience to customers that can result from interruptions to that service. Additionally, significant changes in peak demand levels affect the probability and frequency of shortage conditions (*i.e.*, conditions where firm load is served but required operating reserves are not maintained). Hence, both of these dimensions of load during 2009 are examined in this subsection and summarized in Figure 27.

This figure shows peak load and average load in each of the ERCOT zones from 2006 to 2009. It indicates that in each zone, as in most electrical systems, peak demand significantly exceeds average demand. The North Zone is the largest zone (about 38 percent of the total ERCOT load);[17] the South and Houston Zones are comparable (with about 28 percent) while the West Zone is the smallest (with about 6 percent of the total ERCOT load). Figure 27 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

[17] The Northeast Zone was integrated into the North Zone in 2007.

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

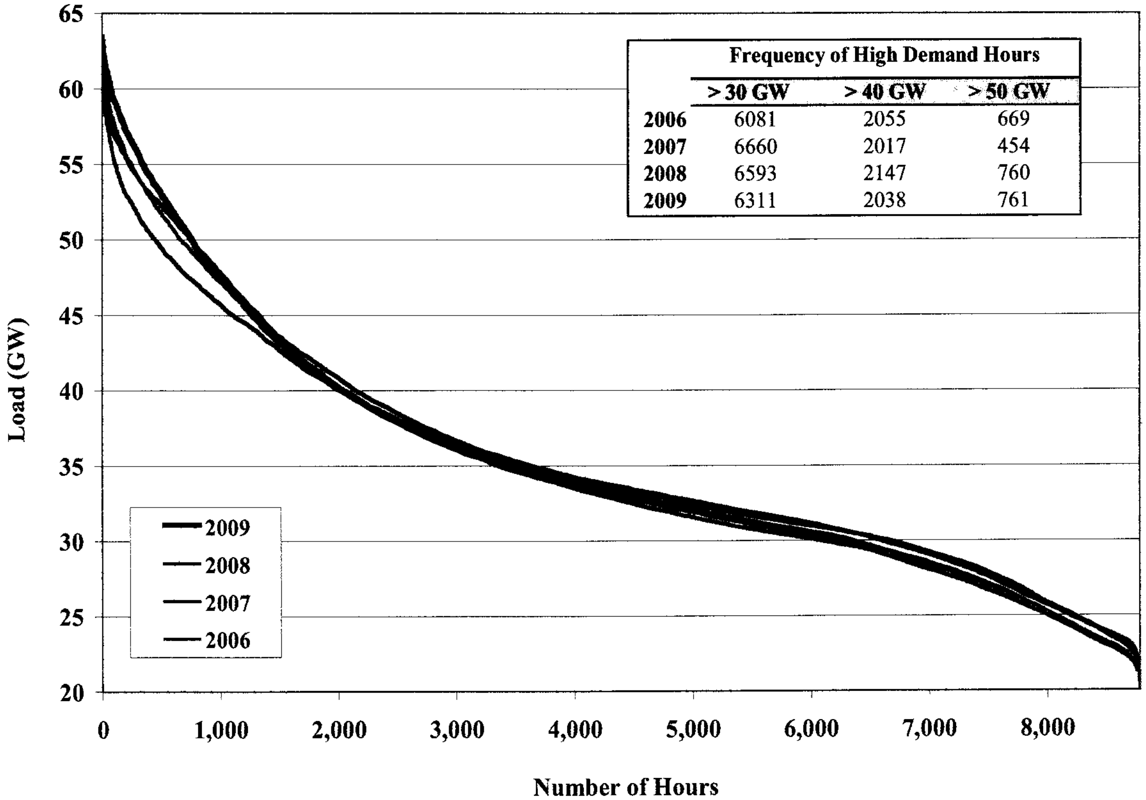
Figure 27: Annual Load Statistics by Zone



Some of the changes in zonal peak and average loads from 2008 to 2009 can be attributed to changes to the zonal definitions that resulted in some loads moving to a different zone in 2009. Overall, the ERCOT average load decreased from 312,401 GWh in 2008 to 308,278 GWh in 2009, a decrease of 1.3 percent. In contrast, the ERCOT coincident peak demand increased from 62,174 MW in 2008 to 63,400 MW in 2009, an increase of 2.0 percent.

To provide a more detailed analysis of load at the hourly level, Figure 28 compares load duration curves for each year from 2006 to 2009. 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, as most hours exhibit low to moderate electricity demand, with peak demand usually occurring during the afternoon and early evening hours of days with exceptionally high temperatures.

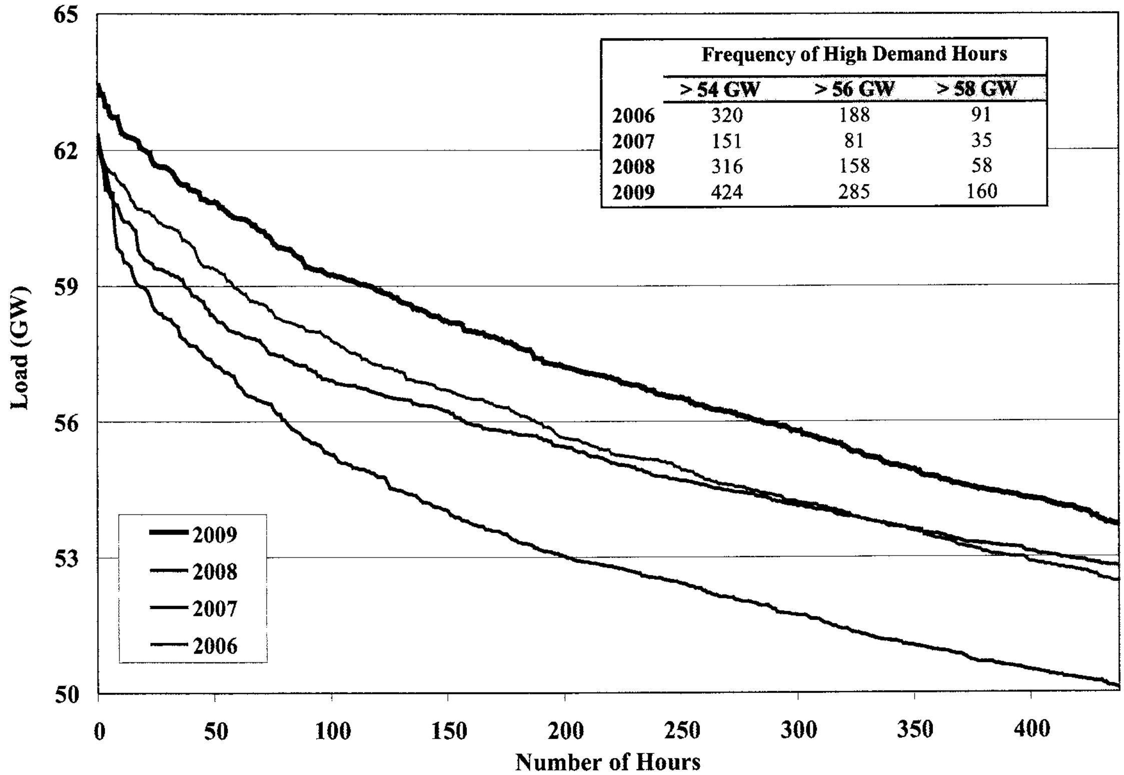
Figure 28: ERCOT Load Duration Curve – All Hours



As shown in Figure 28, the load duration curve for 2009 is slightly lower than in 2008 at load levels less than 45 GW, which accounts for approximately 85 percent of the hours in 2009 and is consistent with the load reduction of 1.3 percent from 2008 to 2009. However, the number of high demand hours (more than 50 GW) in 2008 and 2009 are at comparable levels (760 and 761 hours respectively).

To better show the differences in the highest-demand periods between years, Figure 29 shows the load duration curve for the five percent of hours with the highest loads. This figure shows that while average load increased in each year from 2006 to 2008 and decreased in 2009, the frequency of high-demand hours in 2009 increased compared with year 2008. Load exceeded 58 GW in 160 hours in 2009, more than double the hours in 2008.

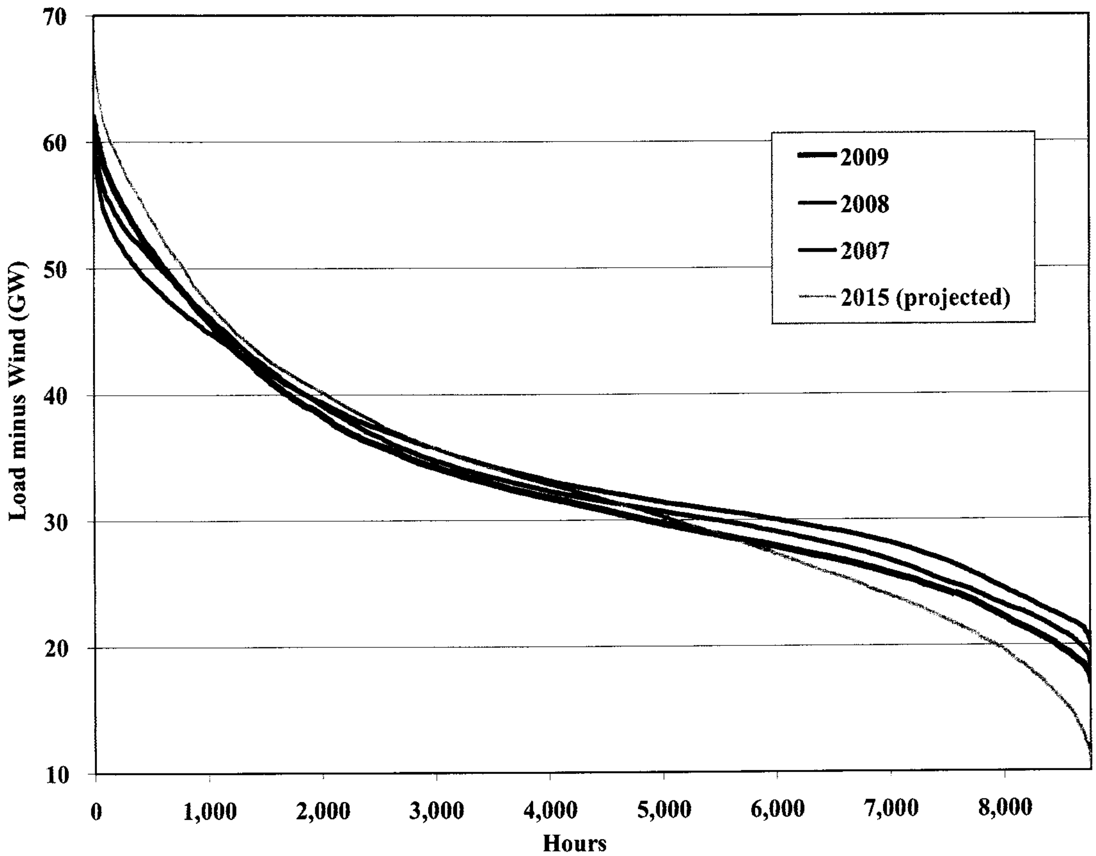
Figure 29: ERCOT Load Duration Curve – Top 5% of Hours



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 2006 to 2009, the peak load value averaged 19.7 percent greater than the load at the 95th percentile. These load characteristics imply that a substantial amount of capacity – over 10 GW – is needed to supply energy in less than 5 percent of the hours. Additionally, another 8 GW of capacity is required to meet the ERCOT planning reserve requirement of expected peak demand plus 12.5 percent. These factors serve to emphasize the importance of efficient energy pricing during peak demand conditions and other times of system stress that send accurate economic signals for the investment in and retention of the resources required to meet these real-time system demands as well as achieving long-term resource adequacy requirements.

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 30 shows the net load duration curves for 2007 through 2009, with projected values for 2015 based on ERCOT data from its Competitive Renewable Energy Zones assessment.

Figure 30: Net Load Duration Curves



The data in Figure 30 show that while the peak net load has increased from 2007 to 2009, the remainder of the net load duration curve has been reduced. This is due in part in to the 1.3 percent decline in energy consumption in 2009, but is largely associated with the increase in wind production in the ERCOT region over this time period. Over 90 percent of the wind resources in the ERCOT region are located in West Texas, and 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 projection for 2015 indicates that the trend shown from 2007 to 2009 is expected to continue and amplify with the addition of significant new wind resources and the reduction in the curtailment of existing wind resources. Focusing on the left side of the net load duration curve, the average difference between peak net load and the 95th percentile of net load was 10.7 GW in 2007 to 2009, but this differential is projected to increase to over 15 GW by 2015. With an additional capacity requirement of approximately 9 GW to meet the 12.5 percent reserve margin requirement, this means that over 24 GW of non-wind capacity will be required to exist on the system with an expectation of operating five percent of the hours in a year or less. On the right side of the net load duration curve, the minimum net load was 17 GW in 2007 to 2009, but the minimum is projected to decrease to less than 11 GW by 2015.

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, particulary within the context of the ERCOT energy-only market design.

B. Load Scheduling

In this subsection, we evaluate load scheduling patterns by comparing load schedules to actual real-time load. Under the ERCOT Protocols, scheduled load must be balanced with scheduled resources for each QSE for each settlement interval; however, there is no requirement that the scheduled load be consistent with the actual load of a QSE. Additionally, a QSE may balance its scheduled load with resources scheduled from ERCOT. Because the financial effect of scheduling resources from ERCOT to balance a load schedule is the same as if the load were unscheduled, in this section, we adjust the load schedules by subtracting the amount that consists of resources scheduled from ERCOT.

To provide an overview of the scheduling patterns, Figure 31 shows a scatter diagram that plots the ratio of the final load schedules to the actual load level during 2009. The ratio shown in the figure will be greater than 100 percent when the final load schedule is greater than the actual load.

Figure 31: Ratio of Final Load Schedules to Actual Load
All ERCOT

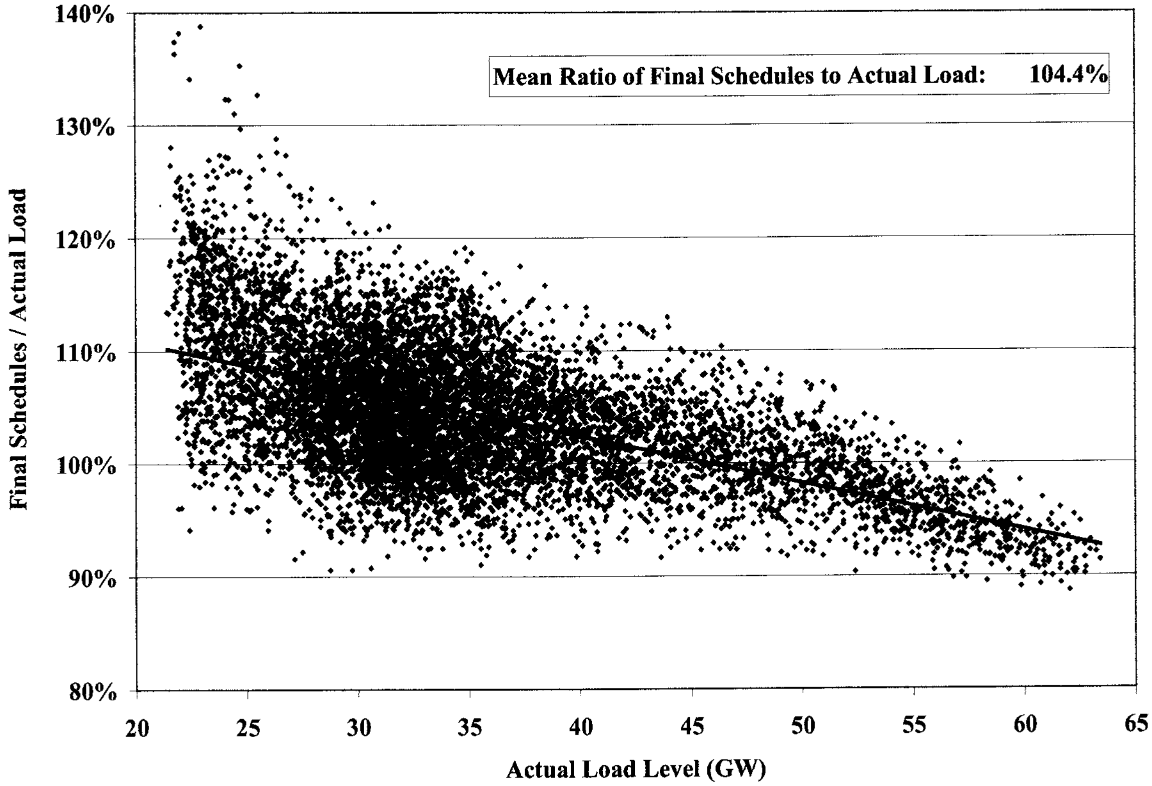


Figure 31 shows that final load schedules on average was higher than the actual load in aggregate, as indicated by an average ratio of the final load schedules to actual load of 104.4 percent. However, the figure also includes a trend line indicating that the ratio of final load schedules to actual load tends to decrease as load rises. In particular, the ratio given by the trend line is above 100 percent for loads under 45 GW and declines to 92 percent at higher load levels. The overall pattern shown in the figure above is similar to previous years, which exhibited the same downward trend in final load schedules relative to actual load.

On average, balancing energy prices are higher and more volatile at high load levels, although the previous subsection showed that spikes can occur under all load conditions. Market participants that are risk averse might be expected to schedule forward to cover a significant portion of their load during high load periods rather than reducing their forward scheduling levels during those periods. There are several explanations for the apparent under-scheduling during high load conditions. First, while the data suggest that QSEs rely more on the balancing energy market at higher load levels, doing so does not necessarily subject them to greater price risk. Financial contracts or derivatives may be in place to protect market participants from price risk in the balancing energy market, such as a contract for differences. Second, market participants who own generation can offer their expensive generation into the market to cover their load needs if balancing energy market prices are high but otherwise allow their load obligations to be met with lower-priced balancing energy. Third, some market participants may not have contracted for sufficient resources to cover their peak load and may, therefore, not be able to fully schedule their load.

Figure 32: Average Ratio of Final Load Schedules to Actual Load by Load Level

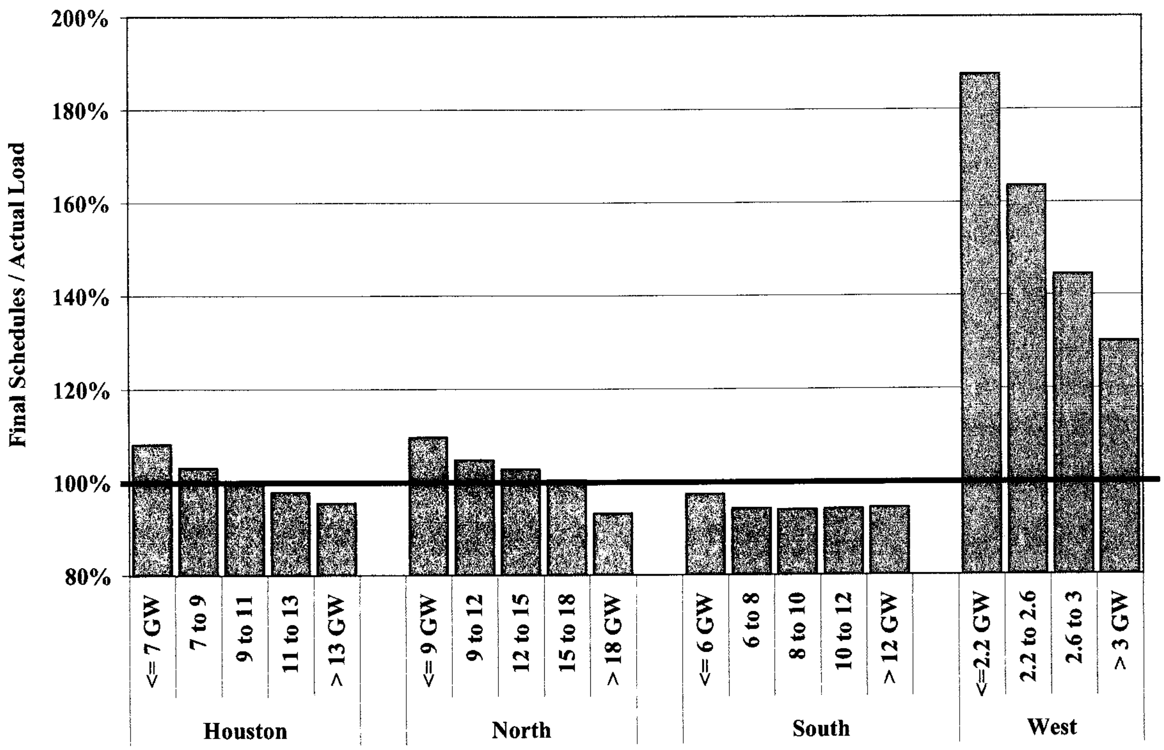


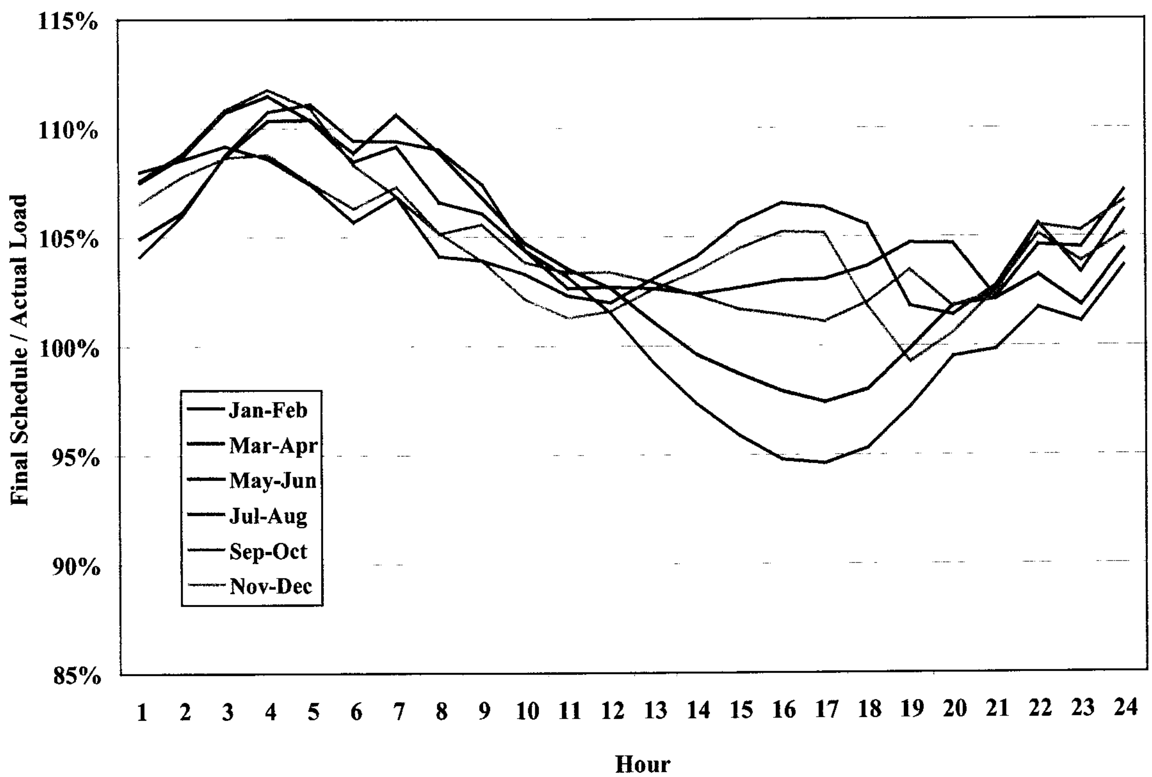
Figure 32 shows the ratio of final load schedules to actual load evaluated at five different load levels for each of the ERCOT zones. Figure 32 shows that:

- The West Zone is significantly over-scheduled, although the ratio declines as load increases.
- The Houston and North Zones are under-scheduled at the highest load levels.
- The South Zone is under-scheduled at all load levels.

It should be noted that regardless of the relationship between the aggregate scheduled load and actual load, individual QSEs may be significant net sellers or purchasers in the balancing energy market. Persistent load imbalances are not necessarily a problem. Imbalances can reflect the fact that some suppliers schedule energy from resources they expect to be economic in the balancing energy market when they have not already sold the power in a bilateral contract. Rather than selling power to the balancing energy market through deployments in the balancing energy market, they sell through load imbalances. Additionally, some load-serving entities may choose to purchase a portion of their load obligations in the balancing energy market. These approaches reflect economic decisions of wholesale buyers and sellers and generally do not present operational concerns.

To further analyze load scheduling, Figure 33 shows the ratio of final load schedules to actual load by hour in two month blocks.

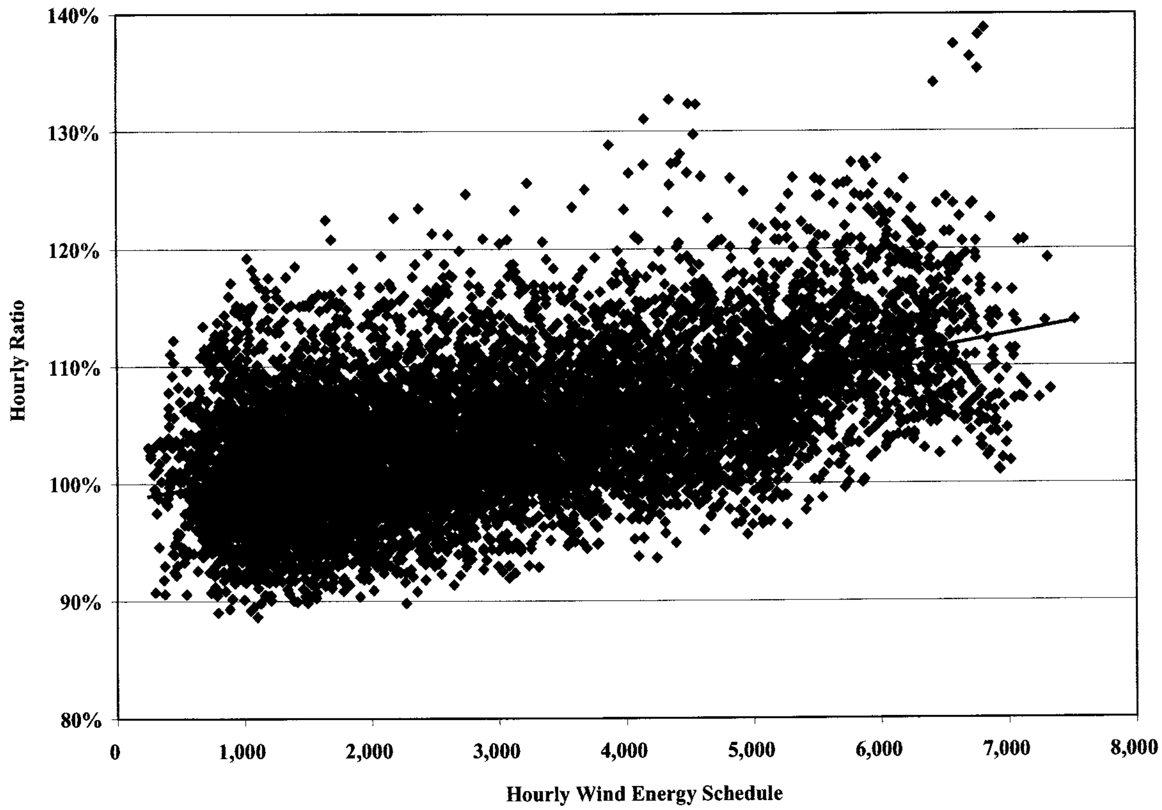
Figure 33: Average Ratio of Final Load Schedules to Actual Load by Hour



This figure shows that the final schedules exceed actual load in all months for the hours 1-12 and 21-24. Final schedules are significantly less than actual load only in the summer months of May through August during the peak demand hours in the afternoon.

A significant factor influencing the relationship between final load schedules and actual load in 2009 was the increased wind generation capacity. Figure 34 shows the load schedule as a percentage of actual load versus wind energy schedules in 2009.

Figure 34: Load Schedule/Actual Load vs. Wind Energy Schedule

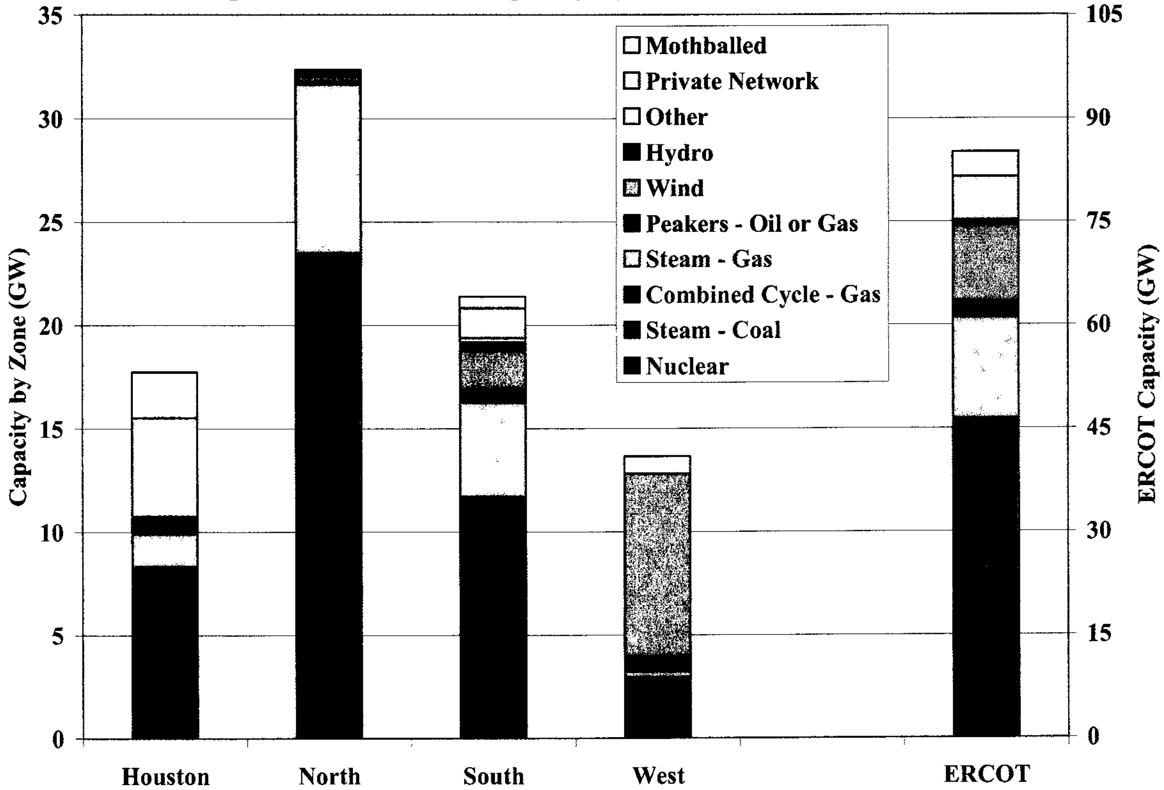


This figure shows a positive correlation between the load schedule as a percentage of actual load and the wind energy schedules. Typically, the production profile for wind resources in the ERCOT market is such that most output occurs during off-peak hours and in the non-summer months. Thus, the data in Figure 34 provide further explanation of the results in Figure 33 that shows that final load schedules exceed actual load most significantly during the off-peak hours and in the non-summer months.

C. Generation Capacity in ERCOT

In this section we evaluate the generation mix in ERCOT. With the exception of the wind resources in the West Zone and the nuclear resources in the North and Houston Zones, the mix of generating capacity is relatively uniform in ERCOT. Figure 35 shows the installed generating capacity by type in each of the ERCOT zones.

Figure 35: Installed Capacity by Technology for each Zone



The nuclear capacity is located in both the North and Houston Zones. Lignite and coal generation is also a significant contributor in ERCOT. However, the primary fuel in ERCOT is natural gas, accounting for nearly 58 percent of generation capacity in ERCOT as a whole and almost 60 percent in the South Zone. Approximately 60 percent of this natural gas-fired capacity represents relatively new combined-cycle units that have been installed throughout ERCOT over the past decade. These new installations have resulted in a small increase in the gas-fired share of installed capacity but have not changed the overall mix significantly, since the generators that have gone out of service during this period were primarily gas-fired steam turbines.

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 38 percent of capacity, the South Zone 25 percent, the Houston Zone 21 percent, and the West Zone 16 percent. The Houston Zone typically imports power, while the West and North Zones typically export 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 45 percent of capacity, the South Zone 27 percent, the Houston Zone 22 percent, and the West Zone 7 percent.

While ERCOT has coal/lignite and nuclear plants that operate primarily as base load units, its reliance on natural gas resources makes it vulnerable to natural gas price spikes. There is approximately 22.6 GW of coal and nuclear generation in ERCOT. Because there are very few hours when ERCOT load drops as low as 20 GW, natural gas resources will be dispatched and set the balancing energy spot price in most hours. Hence, although coal-fired and nuclear units combined produce approximately half of the energy in ERCOT, they have historically played a much less significant role in setting spot electricity prices. However, with the significant increases in wind capacity that has a lower marginal production cost than coal and lignite, the frequency at which coal and lignite are the marginal units in ERCOT is expected to increase in the future, particularly during the off-peak hours in the spring and fall, and even more as additional transmission capacity is added that will accommodate increased levels of wind production in the West Zone.

Figure 36 and Figure 37 show the marginal fuel frequency for the Houston and West Zones, respectively, for each month from 2007 through 2009.[18] The marginal fuel frequency is the percentage of hours that a generation fuel type is marginal and setting the price at a particular location.

As shown in Figure 36, the frequency at which coal was the price setting fuel for the Houston Zone experienced a significant and sustained increase beginning in September 2008. This increase can be attributed to (1) increased wind resource production; (2) a slight reduction in demand in 2009 due to the economic downturn; and (3) periods when natural gas prices were very low thereby makinn coal and combined-cycle natural gas resources competitive from an economic dispatch standpoint. As significant additional wind, coal and potentially nuclear resources are added to the ERCOT region and transmission constraints that serve to limit existing wind production are alleviated, it is likely that the frequency of coal as the marginal fuel will increase in coming years.

18 The marginal fuel frequency for the North and South Zones are very similar to the Houston Zone.

Figure 36: Marginal Fuel Frequency (Houston Zone)

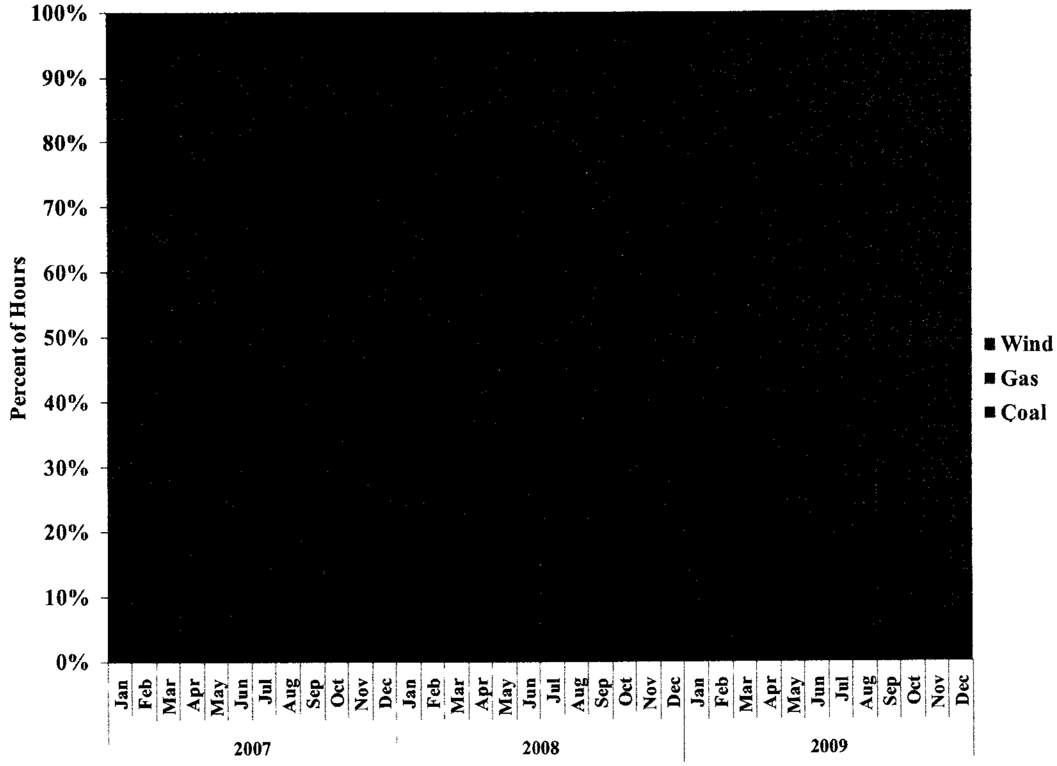


Figure 37: Marginal Fuel Frequency (West Zone)

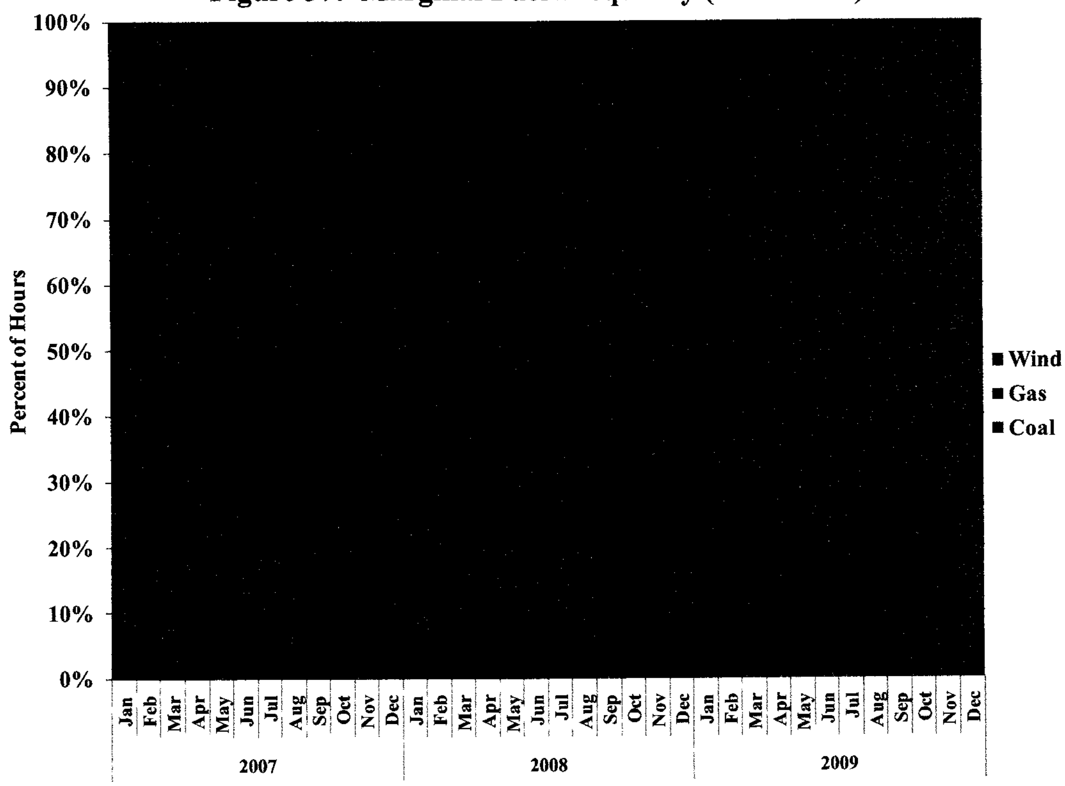


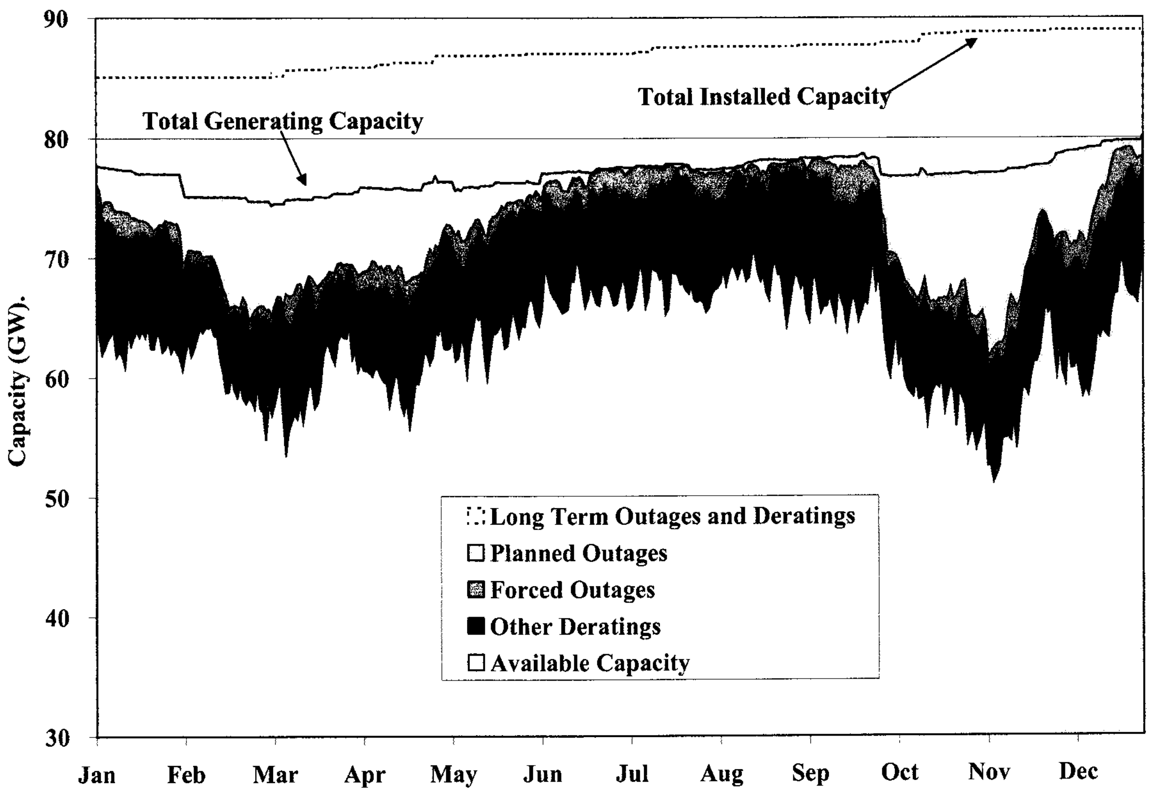
Figure 37 shows that the frequency at which coal was the price setting fuel for the West Zone also experienced a significant and sustained increase beginning in September 2008. This figure also shows that beginning in late 2007 the frequency at which wind was the price setting fuel for the West Zone increased dramatically. This increase is attributable to the growth in installed wind capacity that far exceed the load in the West Zone combined with existing transmission capability that limits the export capability from the West Zone, as discussed in more detail in Section III.

1. Generation Outages and Deratings

Figure 35 in the prior subsection shows that installed capacity is approximately 85 GW including mothballed units and all wind capacity, and approximately 71 GW excluding mothballed capacity and including only 8.7 percent of wind capacity. Hence, the installed capacity exceeds the capacity required to meet annual peak load plus ancillary services requirements of 67 GW. This might suggest that the adequacy of resources is not a concern for ERCOT in the near-term. However, resource adequacy must be evaluated in light of the resources that are actually available on a daily basis to satisfy the energy and operating reserve requirements in ERCOT. A substantial portion of the installed capability is frequently unavailable due to generator deratings. A derating is the difference between the maximum installed capability of a generating resource and its actual capability (or "rating") in a given hour. Generators may be fully derated (rating equals 0) due to a forced or planned outage. It is also 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).

In this subsection, we evaluate long-term and short-term deratings to inform our evaluation of ERCOT capacity levels. Figure 38 shows a breakdown of total installed capability for ERCOT on a daily basis during 2009. This analysis includes all in-service and switchable capacity. The capacity in this analysis is separated into five categories: (a) long-term outages and deratings, (b) short-term planned outages, (c) short-term forced outages, (d) other short-term deratings, and (e) available and in-service capability.

Figure 38: Short and Long-Term Deratings of Installed Capability*



* Includes all outages and deratings lasting greater than 60 days and all mothballed units.

* Switchable capacity is included under installed capacity in this figure.

Figure 38 shows that long-term outages and other deratings fluctuated between 14 and 22 GW. These outages and deratings reduce the effective resource margins in ERCOT from the levels reported above. A large component of the "other deratings" is associated with limited wind resources resulting in generating resources that are not capable of producing up to the full installed capability. Other causes of these deratings reflect:

- Cogeneration resources unavailable to serve market load because they are being used to serve self-serve load;
- Resources out-of-service for economic reasons (*e.g.*, mothballed units); or
- Resources out-of-service for extended periods due to maintenance requirements.

With regard to short-term deratings and outages, the patterns of planned outages and forced outages were consistent with expectations:

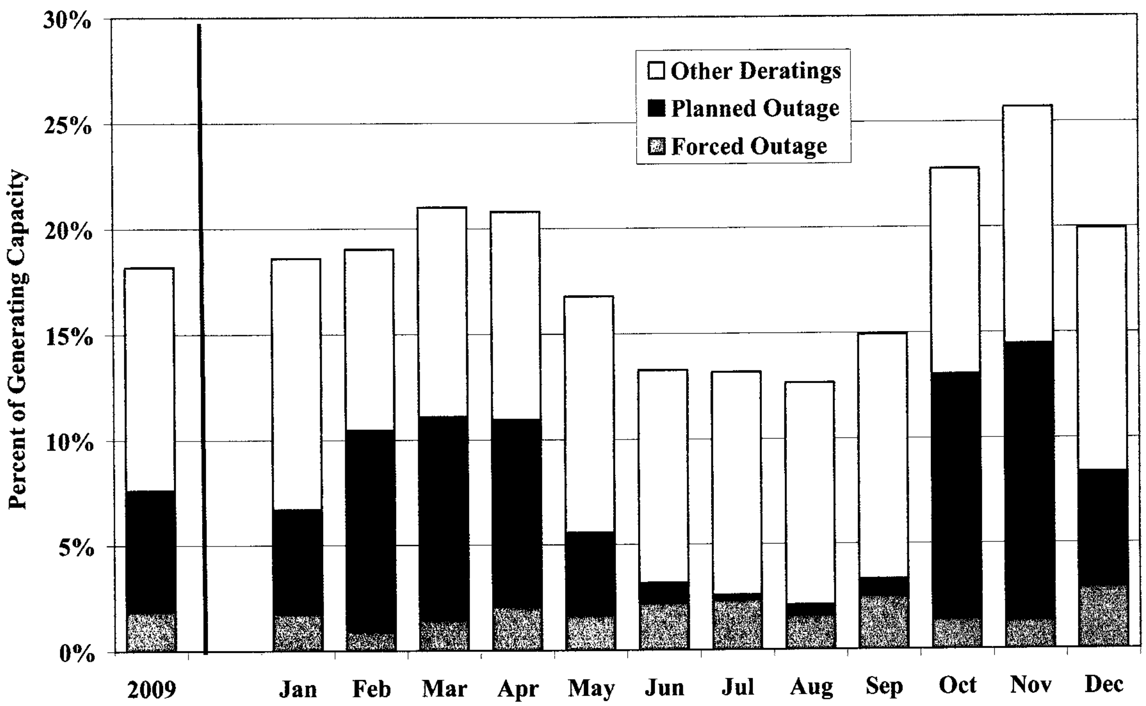
- Forced outages occurred randomly over the year and the forced outage rates were relatively low (although all forced outages may not be reported to ERCOT).

- Planned outages were relatively large in the spring and fall and extremely small during the summer.

Although the total installed capacity was higher in 2009 than in 2008, the annual average of daily available capacity was unchanged. Further, the average of daily available capacity during the summer months (May through September) decreased 1,180 MW from 2008, which can be primarily attributed to higher quantities of derating due to wind resource availability.

The next analysis focuses specifically on the short-term forced outages and other short-term deratings. Figure 39 shows the average magnitude of the outages and deratings lasting less than 60 days for the year and for each month during 2009.

Figure 39: Short-Term Outages and Deratings*



* Excludes all outages and deratings lasting greater than 60 days and all mothballed units.

Figure 39 shows that total short-term deratings and outages were as large as 25 percent of installed capacity in the spring and fall, and dropping to as low as 12 percent for the summer. Most of this fluctuation was due to anticipated planned outages, which ranged as high as 8 to 13 percent of installed capacity during February through April, and October through November. Short-term forced outages occurred more randomly, as would be expected, ranging between one and three percent of total capacity on a monthly average basis during 2009. These rates are relatively low in comparison to other operating markets for two reasons. First, these outages include only full outages (*i.e.*, where the resource's rating equals zero). In contrast, an equivalent forced outage rate is frequently reported for other markets, which includes both full and partial outages. Hence, the forced outage rate shown in Figure 39 can be expected to be lower than equivalent forced outage rates of other markets. Second, because forced outage information is self-reported by generators, we are not confident that the available data includes all forced outages that actually occurred.

The largest category of short-term deratings was the "other deratings" that occur for a variety of reasons. The other deratings would include any short-term forced or planned outage that was not reported or correctly logged by ERCOT. This category also includes deratings due to ambient temperature conditions, cogeneration uses, wind deratings due to variable wind conditions and other factors described above. Furthermore, suppliers may delay maintenance on components such as boiler tubes, resulting in reduced capability. Because these deratings can fluctuate day to day or seasonally, some of the deratings are included in the "long-term outages and deratings" category while the others are included in this category. The other deratings were approximately 10 percent on average during the summer in 2009 and as high as 11 percent in other months. In conclusion, the patterns of outages do not indicate patterns of physical withholding or raise other competitive concerns. However, this issue is analyzed in more detail in Section IV of this report.

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 online capacity plus quick-start[19] units 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,

19 For the purposes of this analysis, "quick-start" includes simple cycle gas turbines that are qualified to provide balancing energy.

Figure 40 plots the excess capacity in ERCOT during 2009. 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 40: Excess On-Line and Quick Start Capacity During Weekday Daily Peaks

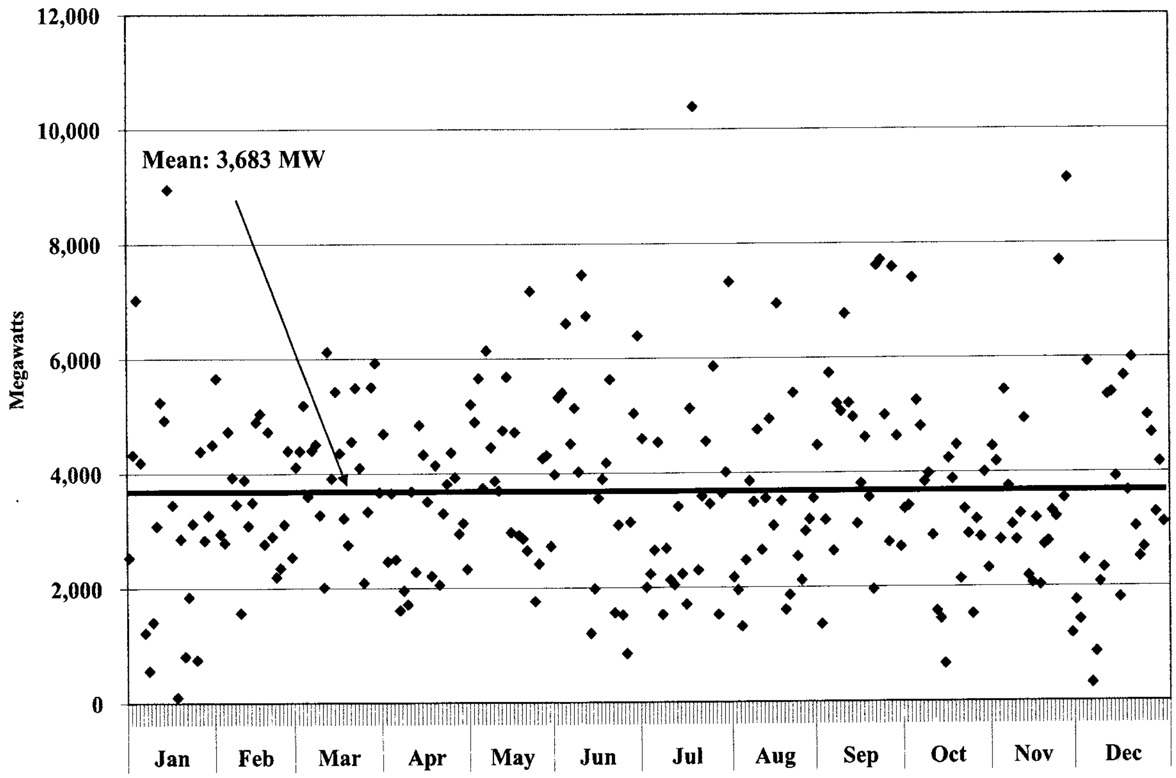


Figure 40 shows that the excess on-line capacity during daily peak hours on weekdays averaged 3,683 MW in 2009, which is approximately 11.9 percent of the average load in ERCOT. This is an increase of more than 600 MW from prior years. One explanation for the increase in excess on-line capacity in 2009 is the increase in the number of quick-start resources that are qualified to provide balancing energy service. Quick-start resources are actually off-line until dispatched; however, these resources are included in the on-line capacity calculation. The use of quick-start resources for balancing energy service results in a more efficient commitment of resources to managed uncertainties that materialize near real-time than does a process of making firm commitment decisions in the day ahead. For this reason, increases in the excess on-line capacity that are associated with the existence of additional quick -tart resources are not an efficiency concern.

The overall trend in excess on-line capacity in recent years indicates a movement toward more efficient unit commitment across the ERCOT market; however, the current market structure is still based primarily upon a decentralized unit commitment process whereby each participant makes independent generator commitment decisions that are not likely to be optimal. Further contributing to the suboptimal results of the current unit commitment process is that the decentralized unit commitment is comprised of non-binding resource plans that form the basis for ERCOT's day-ahead planning decisions. However, these non-binding plans can be modified by market participants after ERCOT's day-ahead planning process has concluded causing ERCOT to take additional actions that may be more costly and less efficient. Hence, the introduction of a day-ahead energy market with centralized Security Constrained Unit Commitment ("SCUC") that is financially binding under the nodal market design promises substantial efficiency improvements in the commitment of generating resources.

D. Demand Response Capability

Demand response is a term that broadly refers to actions that can be taken by end users of electricity to reduce load in response to instructions from ERCOT or in response to certain market or system conditions. The ERCOT market allows participants with demand-response capability to provide energy and reserves in a manner similar to a generating resource. The ERCOT Protocols allow for loads to actively participate in the ERCOT administered markets as either Loads acting as Resources ("LaaRs") or Balancing Up Loads ("BULs"). 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 LaaRs to offer responsive reserves and non-spinning reserves into the day-ahead ancillary services markets. Qualified LaaRs can also offer blocks of energy in the balancing energy market. LaaRs providing up balancing energy must have telemetry and must be capable of responding to ERCOT energy dispatch instructions in a manner comparable to generation resources. Those providing responsive reserves must have high set under-frequency