

Control Number: 34677



Item Number: 4

Addendum StartPage: 0



34677

2008 STATE OF THE MARKET REPORT FOR THE ERCOT WHOLESALE ELECTRICITY MARKETS

POTOMAC ECONOMICS, LTD.

Independent Market Monitor for the ERCOT Wholesale Market

Hand Clean CED

August 2009

•

TABLE OF CONTENTS

Exec	utive	Summaryi	
	A. B. C. D.	Review of Market Outcomes ii Demand and Resource Adequacy xiii Transmission and Congestion xxv Analysis of Competitive Performance xxxiv	
I.	Revi	ew of Market Outcomes1	
	A. B.	Balancing Energy Market	
II.	Dem	and and Resource Adequacy 43	
	A.	ERCOT Loads in 2008	
	В.	Load Scheduling	
	C.	Generation Capacity in ERCOT	
	D.	Demand Response Capability	
	E.	Net Revenue Analysis	
	F.	Effectiveness of the Scarcity Pricing Mechanism	
III.	Tra	ansmission and Congestion75	
	A.	Electricity Flows between Zones	
	B.	Interzonal Congestion	
	C.	Congestion Rights Market	
	D.	Local Congestion and Local/System Capacity Requirements102	
IV.	Analysis of Competitive Performance10		
	A.	Structural Market Power Indicators107	
	В.	Evaluation of Supplier Conduct111	

LIST OF FIGURES

Figure 1: Average Balancing Energy Market Prices	2
Figure 2: Average All-in Price for Electricity in ERCOT	3
Figure 3: Comparison of All-in Prices across Markets	4
Figure 4: ERCOT Price Duration Curve	5
Figure 5: Zonal Price Duration Curves	6
Figure 6: Average Balancing Energy Prices and Number of Price Spikes	7
Figure 7: Implied Marginal Heat Rate Duration Curve – All Hours	. 10
Figure 8: Implied Marginal Heat Rate Duration Curve - Top 5% of Hours	. 10
Figure 9: Monthly Average Implied Marginal Heat Rates	. 11
Figure 10: Convergence between Forward and Real-Time Energy Prices	. 14
Figure 11: Average Quantities Cleared in the Balancing Energy Market	. 16
Figure 12: 2008 Net Balancing Energy by Load Level	. 17
Figure 13: Magnitude of Net Balancing Energy and Corresponding Price	. 19
Figure 14: Hourly Gas Price-Adjusted Balancing Energy Price vs. Real-Time Load	. 21
Figure 15: Final Energy Schedules during Ramping Up Hours	. 22
Figure 16: Final Energy Schedules during Ramping Down Hours	. 23
Figure 17: Balancing Energy Prices and Volumes	. 24
Figure 18: Balancing Energy Prices and Volumes	. 25
Figure 19: Average Balancing Energy Prices and Load by Time of Day	. 26
Figure 20: Average Balancing Energy Prices and Load by Time of Day	. 27
Figure 21: Monthly Average Ancillary Service Prices	. 30
Figure 22: Ancillary Service Costs per MWh of Load	. 33
Figure 23: Regulation Prices and Requirements by Hour of Day	. 35
Figure 24: Reserves and Regulation Capacity, Offers, and Schedules	. 36
Figure 25: Portion of Reserves and Regulation Procured Through ERCOT	38
Figure 26: Hourly Responsive Reserves Capability vs. Market Clearing Price	39
Figure 27: Hourly Non-Spinning Reserves Capability vs. Market Clearing Price	40
Figure 28: Annual Load Statistics by Zone	44
Figure 29: ERCOT Load Duration Curve – All Hours	45
Figure 30: ERCOT Load Duration Curve – Top 5% of Hours	46
Figure 31: Ratio of Final Load Schedules to Actual Load	47
Figure 32: Average Ratio of Final Load Schedules to Actual Load by Load Level	48
Figure 33: Average Ratio of Final Load Schedules to Actual Load	49
Figure 34: Installed Capacity by Technology for each Zone	51
Figure 35: Short and Long-Term Deratings of Installed Capability*	53
Figure 36: Short-Term Outages and Deratings*	54
Figure 37: Excess On-Line and Quick Start Capacity	56
Figure 38: Provision of Responsive Reserves by LaaRs	58
Figure 39: Estimated Net Revenue	61
Figure 40: Comparison of Net Revenue of Gas-Fired Generation between Markets	64
Figure 41: Peaker Net Margin	66
Figure 42: Average Day Ahead Hourly Load Forecast Error by Season	69
Figure 43: Average Day Ahead Load Forecast Error by Month and Hour Blocks	70
Figure 44: Balancing Energy Market Prices during Shortage Intervals	71

Figure 45:	Average SPD-Modeled Flows on Commercially Significant Constraints	76
Figure 46:	Actual Zonal Net Imports	79
Figure 47:	Average SPD-Modeled Flows on Commercially Significant Constraints	80
Figure 48:	Actual Flows versus Physical Limits during Congestion Intervals	82
Figure 49:	Actual Flows versus Physical Limits during Congestion Intervals	83
Figure 50:	Pricing Contours of Unresolved Congestion in the Zonal Market	85
Figure 51:	Pricing Contours of Unresolved Congestion in the Nodal Market	85
Figure 52:	Average West Zone Wind Production	88
Figure 53:	Actual Flows versus Physical Limits during Congestion Intervals	89
Figure 54:	Actual Flows versus Physical Limits during Congestion Intervals	91
Figure 55:	Actual Flows versus Physical Limits during Congestion Intervals	92
Figure 56:	Transmission Rights vs. Real-Time SPD-Calculated Flows	94
Figure 57:	Quantity of Congestion Rights Sold by Type	96
Figure 58:	TCR Auction Prices versus Balancing Market Congestion Prices	98
Figure 59:	Monthly TCR Auction Price and Average Congestion Value	99
Figure 60:	TCR Auction Revenues, Credit Payments, and Congestion Rent	101
Figure 61:	Expenses for Out-of-Merit Capacity and Energy	104
Figure 62:	Expenses for OOME, OOMC and RMR by Region	105
Figure 63:	Residual Demand Index	108
Figure 64:	Balancing Energy Market RDI vs. Actual Load	109
Figure 65:	Ramp-Constrained Balancing Energy Market RDI vs. Actual Load	110
Figure 66:	Ramp-Constrained Balancing Energy Market RDI Duration Curve	111
Figure 67:	Balancing Energy Offers Compared to Total Available Capacity	113
Figure 68:	Balancing Energy Offers Compared to Total Available Capacity	114
Figure 69:	Short-Term Deratings by Load Level and Participant Size	116
Figure 70:	Incremental Output Gap by Load Level	118
Figure 71:	Output Gap by Load Level and Participant Size	119
Figure 72:	Price Spikes vs. Remaining Up Balancing Capacity	120

LIST OF TABLES

Table 1: Average Hourly Responsive Reserves and Non-Spinning Reserves Prices	32
Table 2: Average Calculated Flows on Commercially Significant Constraints	77
Table 3: Average Calculated Flows on Commercially Significant Constraints during	
Transmission Constrained Intervals	81
Table 4: CSC Average Physical Limits vs. Actual Flows during Constrained Intervals	93

.

EXECUTIVE SUMMARY

This report reviews and evaluates the outcomes of the ERCOT wholesale electricity markets in 2008. 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 pursuant to the provisions of Public Utility Commission of Texas ("PUCT") Substantive Rule 25.505(g).

Our analysis indicates that the market performed competitively in 2008. However, the report generally confirms prior findings that the current market rules and procedures are resulting in systemic inefficiencies. Many of these findings can be found in six previous reports we have issued regarding the ERCOT electricity markets.¹ These reports included a number of recommendations designed to improve the performance of the current ERCOT markets. Many of these recommendations were considered by ERCOT working groups and some were embodied in protocol revision requests ("PRRs"). Most of the remaining recommendations will be addressed by the introduction of the nodal market design in late 2010.

One of the most important functions of any electricity market is to manage the flows of power over the transmission network, limiting additional power flows over transmission facilities when they reach their operating limits. As discussed in previous reports, this is also one of the most significant shortcomings of the current ERCOT zonal market design. The zonal market structure is an inherently inefficient model for managing transmission congestion. The zonal market model also suffers from the need to predict and define ahead of time those constraints that can be reasonably managed by using zonal congestion management techniques. Given the dynamic nature of supply, demand and the topology of the transmission system, such predictions can often be incorrect. This was the case in 2008, resulting in significant price excursions in the South and

¹ "ERCOT State of the Market Report 2003", Potomac Economics, August 2004 ("2003 SOM Report"); "2004 Assessment of the Operation of the ERCOT Wholesale Electricity Markets", Potomac Economics, November 2004; "ERCOT State of the Market Report 2004", Potomac Economics, July 2005 ("2004 SOM Report"); "ERCOT State of the Market Report 2005", Potomac Economics, July 2006 ("2005 SOM Report"); "ERCOT State of the Market Report 2006", Potomac Economics, August 2007 ("2006 SOM Report"); and "ERCOT State of the Market Report 2007", Potomac Economics, August 2008 ("2007 SOM Report"); and "ERCOT State of the Market Report 2007", Potomac Economics, August 2008 ("2007 SOM Report").

Houston Zones during the months of April, May and early June until an expedited PRR that modified ERCOT congestion management procedures was implemented.

The wholesale market should function more efficiently under the nodal market design by providing better incentives to market participants, facilitating more efficient commitment and dispatch of generation, and improving ERCOT's operational control of the system. The congestion on all transmission paths and facilities will be managed through market-based mechanisms in the nodal market. In contrast, under the current zonal market design, transmission congestion is most frequently resolved through non-transparent, non-market-based procedures.

Under the nodal market, unit-specific dispatch will allow ERCOT to more fully utilize generating resources than the current market, which frequently exhibits price spikes even when generating capacity is not fully utilized. The nodal market will also allow ERCOT to increase the economic and reliable utilization of scarce transmission resources well beyond that attainable in the zonal market. Finally, the nodal market will produce price signals that better indicate where new generation is most needed for managing congestion and maintaining reliability. In the long-term, these enhancements to overall market efficiency should translate into substantial savings for consumers.

A. Review of Market Outcomes

1. Balancing Energy Prices

The balancing energy market allows participants to make real-time purchases and sales of energy to supplement their forward bilateral contracts. While on average only a relatively small portion of the electricity produced in ERCOT is cleared through the balancing energy market, its role is critical in the overall wholesale market. The balancing energy market governs real-time dispatch of generation by altering where energy is produced to: a) balance supply and demand; b) manage interzonal congestion, and c) displace higher-cost energy with lower-cost energy given the energy offers of the Qualify Scheduling Entities ("QSEs").

In addition, the balancing energy prices also provide a vital signal of the value of power for market participants entering into forward contracts. Although most power is purchased through

forward contracts of varying duration, the spot prices emerging from the balancing energy market should directly affect forward contract prices.

As shown in the following figure, balancing energy market prices were 37 percent higher in 2008 than in 2007, with May and June 2008 showing the largest increases from the same months in 2007. The average natural gas price in 2008 increased 28 percent over 2007 levels, with monthly changes ranging from a 87 percent increase in July (\$5.91 per MMBtu in July 2007 and \$11.05 per MMBtu in July 2008) to an 20 percent decrease in December (\$6.63 per MMBtu in December 2007 and \$5.29 per MMBtu in December 2008). Natural gas is typically the marginal fuel in the ERCOT market. Hence, the movements in wholesale energy prices from 2007 to 2008 were largely a function of natural gas price levels.



Balancing Energy Market Prices

Although fuel price fluctuations are the dominant factor driving electricity prices in the ERCOT wholesale market, fuel prices alone do not explain all of the price outcomes. At least five other factors provided a meaningful contribution to price outcomes in 2008.

First, as discussed in Section II, ERCOT peak demand and installed capacity were relatively flat in 2008, and energy production increased only slightly in 2008 compared to 2007. These results were similar to 2007 compared to 2006. In contrast to years prior to 2007 that experienced increasing demand and decreasing supply, the static supply and demand characteristics from 2007 to 2008 contributed to comparable wholesale pricing outcomes over the course of these two years, with the exception of the second factor, which is transmission congestion.

As discussed in Section III, chronic and severe transmission congestion from North to South and North to Houston materialized in April, May and June 2008 that had a significant effect on balancing energy pricing outcomes, particularly in the Houston and South zones. In addition, significant increases in installed wind generation in the West Zone led to an increase in West to North congestion, in turn producing a load-weighted average price in the West Zone that was approximately 26 percent below the ERCOT average price in 2008, with wind resources frequently being the marginal generation source in the West Zone.

Third, aside from the effect of wind generation on the West Zone prices, the continued increase in wind production in 2008 served to displace more costly generation resources when the wind was producing. This will tend to lower average prices across the market, but the intermittent nature of wind can also lead to transitory price spikes as other generation resources may be required on short notice to fill the gap left by significantly lower than expected or rapidly declining wind output.

Fourth, the balancing energy offer cap increased to \$2,250 per MWh on March 1, 2008, consistent with Commission rule. Prior to March 1, the rule had set the offer cap at \$1,500 per MWh. The increased offer cap is intended to produce higher prices during system shortage conditions as a part of the PUCT's rules that rely upon energy prices exclusively to ensure generation resource adequacy as opposed to the reliance on both capacity and energy prices used in most other domestic organized electricity markets. As discussed in Section II, this mechanism was not always effective in achieving this intended outcome, and some of ERCOT's reliability-based actions can often disrupt the market-based balance of supply and demand, thereby frustrating the long-term success of the energy-only market.

Finally, as discussed in Section IV, the overall competitive performance of the market exhibited continued improvement in 2008, which will tend to lower prices. The following figure presents ERCOT balancing energy market prices adjusted for natural gas price fluctuations to better highlight variations in electricity prices not related to fuel costs.



Monthly Average Implied Marginal Heat Rate

Adjusted for gas price influence, the above figure shows that average implied heat rate for all hours of the year increased by 6.8 percent from 8.50 in 2007 to 9.08 in 2008.² The average implied heat rate was higher in 2008 than in 2007 for the months of April, May, June, August and October. The increases in implied heat rates during April through June compared to 2007 are explained primarily by significant transmission congestion that affected the Houston and South Zones most significantly, and is discussed in more detail in Section III. The increase in the implied heat rate in August and October was due to a greater number of shortage intervals in these two months in 2008 compared to 2007, as well as the effect of *ex post* pricing adjustments

²

The Implied Marginal Heat Rate equals the Balancing Energy Market Price divided by the Natural Gas Price.

during the deployment of non-spinning reserves applied under then-existing ERCOT Protocols, which is discussed in more detail in Section II. In contrast, the implied heat rate in September 2008 was significantly lower than in September 2007. This is explained by two factors. First, September 2007 experienced more shortage intervals than September 2008, which led to an increase in the implied heat rate in September 2007. Second, demand in the ERCOT region was significantly reduced in September 2008 because of the landfall of Hurricane Ike causing widespread and prolonged outages in the Houston area. This suppressed demand and in turn resulted in a significant reduction in the implied heat rate in September 2008.

The report evaluates two other aspects of the balancing energy prices: 1) the correlation of the balancing energy prices with forward electricity prices in Texas, and 2) the primary determinants of balancing energy prices. Natural market forces should push forward market prices to levels consistent with expectations of spot market prices. Day-ahead prices averaged \$87 per MWh in 2008 compared to an average of \$86 per MWh for real-time prices. Although the day-ahead and real-time prices exhibited good average convergence in 2008, the average absolute price difference, which measures the volatility of the price differences, was large for several months in 2008, particularly in April, May and June. The price volatility in April, May and June 2008 was due in large part to the significant and unpredictable transmission congestion experienced in that timeframe. Relatively smaller spikes in the absolute price difference occurred in August and October 2008. These spikes were associated in part with certain *ex post* pricing revisions during the deployment of non-spinning reserves. As discussed in more detail in Sections II and III, the rules and procedures associated with both of these issues have since been revised. The introduction of the nodal market, which will include an integrated day-ahead market, should also improve the convergence between day-ahead and real-time energy prices.

As discussed in prior reports, we continue to observe in 2008 a clear relationship between the net balancing energy deployments and the balancing energy prices. This is not expected in a well-functioning market. This relationship is partly due to the hourly scheduling patterns of most market participants. Energy schedules change by large amounts at the top of each hour while load increases and decreases smoothly over time. This creates extraordinary demands on the balancing energy market and erratic balancing energy prices, particularly in the morning when loads are increasing rapidly and in the evening when loads are decreasing rapidly.

Hr 21

Hr 22



Average Balancing Energy Prices and Load by Time of Day

The previous two figures summarize these erratic price patterns by showing the balancing energy prices and actual load in each 15-minute interval during the morning "ramping up" hours and

Time of Day by Interval

Hr 0

Hr 23

Hr 1

Hr 2

evening "ramping down" hours, with the red lines highlighting the transition from one hour to the next. These pricing patterns raise significant efficiency concerns regarding the operation of the balancing energy market. Moreover, this pattern has been consistently observed for several years and is likely to continue until changes are made to the market rules.³ In prior reports, we have made several recommendations to address the issue under the current zonal design, although many have not been implemented because of the effort to timely implement the nodal market. The nodal market will provide for a comprehensive solution to the operational issues described in this and prior reports.

2. All-In Electricity Prices

In addition to the costs of energy, loads incur costs associated with operating reserves, regulation, and uplift. The uplift costs include payments for out-of-merit capacity ("OOMC"), Replacement Reserve ("RPRS"), out-of-merit energy ("OOME"), and reliability must run agreements ("RMR"), but exclude administrative charges such as the ERCOT fee. These costs, regardless of the location of the congestion, are borne proportionally by all loads within ERCOT.



See 2003 SOM Report, Assessment of Operations, 2004 SOM Report, 2005 SOM Report 2006 SOM Report and 2007 SOM Report.

3

The monthly average all-in energy prices for the past four years are shown in the figure above along with the monthly average price of natural gas. This figure indicates that natural gas prices were the primary driver of the trends in electricity prices from 2005 to 2008. Average natural gas prices increased in 2008 by 28 percent over 2007 levels, although gas prices in 2008 for the higher electricity demand months of May through September were 46% higher than the same months in 2007. The average all-in price for electricity was \$58.47 in 2007 and \$80.97 in 2008, an increase of 38.5 percent.

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.: California ISO, New York ISO, ISO New England, and PJM. 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.



Comparison of All-In Prices across Markets

** 2008 Capacity, Ancillary Services and Uplift data unavailable for PJM

This figure shows that energy prices increased in wholesale electricity markets across the U.S. in 2008, primarily due to increases in fuel costs.

3. Ancillary Services Markets

The primary ancillary services are up regulation, down regulation, and responsive reserves. 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. QSEs 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 2008.

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). The figure below shows the monthly average prices for regulation and responsive reserve services from 2005 to 2008.

This figure shows that after two years of relatively stability, 2008 experienced a significant increase in ancillary service capacity prices. These price movements can be primarily attributed to the variations in energy prices that occurred over the same timeframe.



In addition to the effect of higher energy prices on ancillary service prices, the following factors had a significant effect on ancillary service prices in 2008:

- ERCOT increased its procurement of responsive reserve quantities from January through August 2008 from the historical constant quantity of 2,300 MW to as high as 2,800 MW during peak hours in the summer. Also, the required quantity of non-spinning reserves when procured was increased in most of 2008, and non-spinning reserves were procured more frequently in 2008 than in 2007.
- Significant transmission congestion materialized in April, May and June 2008 leading to significantly higher prices in the Houston and South Zones. These pricing outcomes had the effect of increasing the opportunity costs for providers of responsive reserve in these locations, thereby causing an upward shift in the supply curve for responsive reserve in these months.
- West to North congestion increased significantly in 2008, leading to over 1,100 hours of average negative prices in the West Zone. For providers of responsive reserves in the West Zone, exposure to negative prices significantly increases the cost to provide reserves because the resources must operate uneconomically at minimum load levels. Hence, in periods of expected high wind production, responsive reserve offers from suppliers in the West Zone are expected to be higher to reflect these economics risks.

• The quantity of Loads acting as Resources ("LaaRs") providing responsive reserves was moderately reduced in March through May, and experienced more significant reductions in September, part of October, November and December. The reduction in the provision of responsive reserves by LaaRs in these months resulted in a corresponding increase in the quantity of responsive reserve provided by generation resources, which are typically more expensive, thereby placing an upward pressure on responsive reserve prices.

The current Nodal Protocols specify that energy and ancillary services will be jointly optimized in a centralized day-ahead market. This is likely to improve the overall efficiency of the dayahead unit commitment. Additionally, although it is not possible to implement at the inception in the nodal market, we also recommend the development of real-time markets that co-optimize energy and reserves to further enhance the efficient dispatch of resources and pricing in realtime.

While the previous figure shows the individual ancillary service capacity prices, the following figure shows the monthly total ancillary service costs per MWh of ERCOT load and the average balancing energy price for 2005 through 2008.





This figure 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 increased to \$3.07 per MWh in 2008 compared to \$1.46 per MWh in 2007, an increase of more than 110 percent. However, while the all-in wholesale costs increased by more than 38 percent in 2008 compared to 2007, ancillary service costs accounted for only 2.8 percent of the increase in all-in wholesale power costs in 2008 over 2007.

B. Demand and Resource Adequacy

1. Installed Capacity and Peak Demand

Because electricity cannot be stored, the electricity market must ensure that generation matches load on a continuous basis. Thus, one critical issue for a wholesale electricity market is whether sufficient supplies exist to satisfy demand under peak conditions. In 2008, the load served by ERCOT reached a peak of over 62.2 GW, which was almost identical to the peak demand in 2007. Changes in the peak demand levels are very important because they are a key determinant of the probability and frequency of shortage conditions, although daily unit commitment practices, load uncertainty and unexpected resource outages are also contributing factors.

More broadly, peak demand levels and the capability of the transmission network are the primary factors that determine whether the existing generating resources are adequate to maintain reliability. The report provides an accounting of the current ERCOT generating capacity, which is dominated by natural gas-fired resources. These resources account for 70 percent of generation capacity in the ERCOT region while providing only 43 percent of total production in 2008.

ERCOT has more than 80 GW of installed capacity. This includes import capability, resources that can be switched to the SPP, and Loads acting as Resources ("LaaRs"). However, significant amounts of this are not kept constantly in service, with about 5 GW of mothballed capacity existing in 2008. Furthermore, ambient temperature restrictions increase during the summer months when demand is highest, leading to substantial deratings. Although ERCOT had sufficient capacity to meet load and ancillary services needs during the 2008 peak, it is important to consider that electricity demand will continue to grow and that a significant number of

generating units in Texas will soon reach or are already exceeding their expected lifetimes. Without significant capacity additions, these factors may cause the resource margins in ERCOT to diminish over the next three to five years. Moreover, although several baseload facilities are currently under construction, the rapidly increasing penetration of intermittent resources such as wind and solar facilities will likely create the reliability need for additional operationally flexible resources, such as modern gas turbines. This reinforces the importance of ensuring that efficient economic signals are provided by the ERCOT market.

2. Generator Outages and Commitments

Despite adequate installed capacity, 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.



Short and Long Torm Doratings of Installed Canability

A derating is the difference between the installed capability of a generating resource and its maximum capability (or "rating") in a given hour. Generators can be fully derated (rating equals 0) due to a forced or planned outage. However, it is very common for a generator 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.*, ambient temperature conditions). The previous figure shows the daily available and derated capability of generation in ERCOT.

The figure shows that long-term outages and other deratings fluctuated between 9 and 18 GW. These outages and deratings reduce the effective resource margins in ERCOT from the levels reported above. With regard to short-term deratings and outages, the patterns of planned outages and forced outages were consistent with expectations in that: (1) forced outages occur randomly over the year and the forced outage rates were relatively low; and (2) planned outages were relatively large in the spring and fall and extremely small during the summer.

In addition to the generation outages and deratings, the report evaluates the results of the generator commitment process in ERCOT, which is decentralized and largely the responsibility of the QSEs. This evaluation includes analysis of the real-time excess capacity in ERCOT. We define excess capacity as the total online capacity plus quick-start units each day minus the daily peak demand for energy, responsive reserves provided by generation, and up regulation. Hence, it measures the total generation available for dispatch in excess of the electricity needs each day.

The report finds that the excess on-line capacity during daily peak hours on weekdays averaged 2,723 MW in 2008, which is approximately 8 percent of the average load in ERCOT. The overall trend in excess on-line capacity also 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 provide a market-wide optimal solution. Also contributing to the suboptimal results of the current unit commitment process is that the decentralized unit commitment is reported to ERCOT through 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. Consequently, ERCOT frequently takes additional actions to ensure reliability that may be more costly and less efficient than necessary. Under the nodal market design, the introduction of a day-ahead energy market with centralized Security Constrained Unit

Commitment ("SCUC") that is financially binding promises substantial efficiency improvements in the commitment of generating resources.

3. Load Participation in the ERCOT Markets

The ERCOT Protocols allow for loads to participate in the ERCOT-administered markets as either Load acting as Resources ("LaaRs") or Balancing Up Loads ("BULs"). LaaRs are loads that are qualified by ERCOT to offer responsive reserves, non-spinning reserves, or regulation into the day-ahead ancillary services markets and can also offer blocks of energy in the balancing energy market.

As of December 2008, 2,158 MW of capability were qualified as LaaRs. In 2008, LaaRs were permitted to supply up to 1,150 MW of the responsive reserves requirement. Although the participants with LaaR resources are qualified to provide non-spinning reserves and up balancing energy in real-time, LaaR participation in the non-spinning reserve and balancing energy market was negligible in 2008.⁴ This is not surprising because the value of curtailed load tends to be relatively high, and providing responsive reserves offers substantial revenue with very little probability of being deployed. In contrast, resources providing non-spinning reserves are 70 times more likely to be deployed. Hence, most LaaRs will have a strong preference to provide responsive reserves over non-spinning reserves or balancing energy. The following figure shows the daily average provision of responsive reserves by LaaRs in the ERCOT market from 2005 through 2008.

⁴ Although there was no active participation in the balancing energy market, loads can and do respond to market prices without actively submitting a bid to ERCOT. This is often referred to as passive load response.



Provision of Responsive Reserves by LaaRs

The high level of participation by demand response participating in the ancillary service markets sets ERCOT apart from other operating electricity markets. The figure above shows that the amount of responsive reserves provided by LaaRs has held fairly constant at 1,150 MW since the beginning of 2005. Notable exceptions were a period in September/October 2005 corresponding to Hurricane Rita, and a more prolonged decrease in September/October of 2008 corresponding to the Texas landfall of Hurricane Ike. Of interest in late 2008 is the post-hurricane recovery of the quantity of LaaRs providing Responsive Reserve followed by a steady reduction for the remainder of the year, which was likely a product of the economic downturn and its effect on industrial operations.

4. **Net Revenue Analysis**

The next analysis of the outcomes in the ERCOT markets in 2008 is the analysis of "net revenue". Net revenue is defined as the total revenue that can be earned by a new generating unit less its variable production costs. It represents the revenue that is available to recover a unit's fixed and capital costs. Hence, this metric shows the economic signals provided by the market for investors to build new generation or for existing owners to retire generation. In longrun equilibrium, the markets should provide sufficient net revenue to allow an investor to breakeven on an investment in a new generating unit, including a return of and on the investment.

In the short-run, if the net revenues produced by the market are not sufficient to justify entry, then one of three conditions likely exists:

- (i) New capacity is not currently needed because there is sufficient generation already available;
- (ii) Load levels, and thus energy prices, are temporarily low due to mild weather or economic conditions; or
- (iii) Market rules are causing revenues to be reduced inefficiently.

Likewise, the opposite would be true if the markets provide excessive net revenue in the shortrun. Excessive net revenue that persists for an extended period in the presence of a capacity surplus is an indication of competitive issues or market design flaws.

The report estimates the net revenue that would have been received in 2007 and 2008 for four types of units: a natural gas combined-cycle generator, a simple-cycle gas turbine, a coal-fired steam turbine with scrubbers, and a nuclear unit.



Estimated Net Revenue

The figure above shows that the net revenue increased substantially in 2008 in each zone compared to 2007. 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 \$70 to \$95 per kW-year. The estimated net revenue in 2008 for a new gas turbine was approximately \$120, \$113 and \$61 per kW-year in the South, Houston and North Zones, respectively. For a new combined cycle unit, the estimated net revenue requirement is approximately \$105 to \$135 per kW-year. The estimated net revenue in 2008 for a new combined cycle unit was approximately \$191, \$185 and \$124 per kW-year in the South, Houston and North Zones, respectively. These values indicate that the estimated net revenue in 2008 was sufficient to support new entry for a new gas turbine in the South and Houston zones and for a combined cycle unit in the South, Houston and North zones. However, as discussed later in this subsection, significant portions of the net revenue results for gas turbine and combined cycle units in 2008 can be attributed to anomalous market design related inefficiencies rather than fundamentals that would support an investment decision for new gas turbines and combined cycle units.

Prior to 2005, net revenues were well below the levels necessary to justify new investment in coal and nuclear generation. However, high natural gas prices have 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. For a new coal unit, the estimated net revenue requirement is approximately \$190 to \$245 per kW-year. The estimated net revenue in 2008 for a new coal unit was approximately \$427, \$421 and \$351 per kW-year in the South, Houston and North Zones, respectively. For a new nuclear unit, the estimated net revenue requirement is approximately \$280 to \$390 per kW-year. The estimated net revenue in 2008 for a new nuclear unit was approximately \$567, \$562 and \$492 per kW-year in the South, Houston and North Zones, respectively. These values indicate that the estimated net revenue for a new coal and nuclear unit in the South, Houston and North Zones, respectively. These values indicate that the estimated net revenue for a new coal and nuclear unit in the South, Houston and North Zones was sufficient to support new entry in 2008, as was the case in 2005, 2006 and 2007. Thus, it is not surprising that some market participants are building new baseload facilities and that several others have initiated activities that may lead to the construction of additional baseload facilities in the ERCOT region.

Although estimated net revenue grew considerably in 2008 compared to prior years, there are other factors that determine incentives for new investment. First, market participants must anticipate how prices will be affected by the new capacity investment, future load growth, and increasing participation in demand response. Second, net revenues can be inflated when prices clear above competitive levels as a result of market power being exercised. Thus, a market participant may be deterred from investing in new capacity if it believes that prevailing net revenues are largely due to an exercise of market power that would not be sustainable after the entry of the new generation. Third, the nodal market design will have an effect on the profitability of new resources. In a particular location, nodal prices could be higher or lower than the prices in the current market depending on the pattern of congestion. Finally, and most importantly in 2008, net revenues can be inflated when prices clear at high levels due to inefficiencies in the market design. Similar to the case of market power, a market participant may be deterred from investing in new capacity if it believes that prevailing net revenues are largely due to market design.

Such market design inefficiencies were apparent in 2008. As discussed in Section III, the vast majority of price excursions in 2008 – particularly in the South and Houston Zones – were not a function of market fundamentals; rather, the price excursions were driven by inefficient congestion management techniques that have since been corrected and are not expected to materialize in the future, especially upon implementation of the nodal market in 2010. In addition to these transmission congestion issues, in 2008 the ERCOT Protocols provided for *ex post* re-pricing provisions in intervals in which non-spinning reserve prices were deployed that frequently resulted in scarcity-level prices at times when ERCOT's operating reserve levels were not deficient. These rules were recently changed, thereby reducing the probability of scarcity-level prices during non-scarcity conditions going forward. Hence, a significant portion of the net revenue produced in 2008 is not reflective of fundamentals that would support an investment decision for new gas turbines and combined-cycle units.

5. Effectiveness of the Scarcity Pricing Mechanism

The 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 gradually increasing it to \$1,500 per MWh on March 1, 2007, \$2,250 per MWh on March

1, 2008, and to \$3,000 per MWh shortly after the implementation of the nodal market. Additionally, market participants controlling less than five percent of the capacity in ERCOT by definition do not possess 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. However, because of the competition faced by the smaller market participants, the quantity offered at such high prices – if any – is very small.

Unlike markets with a long-term capacity market where fixed capacity payments are made to resources across the entire year regardless of the relationship of supply and demand, the objective of the energy-only market design is to allow energy prices to rise significantly higher during legitimate shortage conditions (*i.e.*, when the available supply is insufficient to simultaneously meet both energy and operating reserve requirements) such that the appropriate price signal is provided for demand response and new investment when required. During non-shortage conditions (*i.e.*, most of the time), the expectation of competitive energy market outcomes is no different in energy-only than in capacity markets.

As noted in the net revenue analysis, 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. Both of these issues have been addressed in the zonal market and will be further improved with the implementation of the nodal market in 2010. Absent these inefficiencies, net revenues would not have been sufficient to support new peaker entry in 2008. Beyond these anomalies, there were three other factors that significantly influenced the effectiveness of the SPM in 2008:

- A substantial decrease in out-of-merit deployments by ERCOT during declared shortsupply conditions;
- A continued strong positive bias in ERCOT's day-ahead load forecast that tended to regularly commit online resources in excess of the quantity required to meet expected demand and operating reserve requirements; and
- The dependence on market participants to submit offers at or near the offer cap to produce scarcity level prices during legitimate shortage conditions.

The first item relates to issues that occurred in 2007 and were successfully resolved through changes in ERCOT procedures and the implementation of PRR 750 during 2008. However, the

other two items represent ongoing concerns with respect to the successful operation of the ERCOT energy-only market.

The second issue that can adversely affect the successful operation of the ERCOT energy-only market is ERCOT's day-ahead load forecast. The following figure shows the ERCOT day-ahead load forecast error by hour in 2007 and 2008, with the summer and non-summer months presented separately. In this figure, positive values indicate that the day-ahead load forecast was greater than the actual load in real-time.



The existence of such a strong and persistent positive bias in the day-ahead load forecast will tend to lead to an inefficient over-commitment of resources and to the depression of real-time prices relative to a more optimal unit commitment. To the extent load uncertainty is driving the bias in the day-ahead load forecast, such uncertainty is more efficiently managed through the procurement of ancillary services such as non-spinning reserve, or through supplemental commitments of short-lead time resources at a time sufficiently prior to, but closer to real-time as uncertainty regarding real-time conditions diminishes. As a general principle, competitive and efficient market prices should be consistent with the cost of the marginal action taken to satisfy the market's demand. In the vast majority of hours, the marginal action is the dispatch of the most expensive online generator. 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 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 emergency 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 the market is in shortage, the marginal action 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.

Under the PUCT rules governing the energy-only market, the mechanism that allows for such pricing during shortage conditions relies upon the submission of high-priced offers by small market participants. The following figure shows the balancing market clearing prices during the 103 15-minute shortage intervals in 2008.



Balancing Energy Market Prices during Shortage Intervals

The figure above shows that the prices during these 103 shortage intervals in 2008 ranged from \$105 per MWh to the offer cap of \$2,250 per MWh (prior to March 1, 2008, the offer cap was \$1,500 per MWh), with an average price of \$534 per MWh and a median price of \$293 per MWh. The results in 2008 are similar to those in 2007 when there were 108 shortage intervals with an average price of \$476 per MWh and a median price of \$299 per MWh.

In this figure, the data are separated into solid blue and red outlined points. The blue points (79) represent true shortage conditions, whereas the red points (24) represent artificial shortage prices occurring as a result of large generation schedule reductions at the top of the hours from 10 PM to 1 AM. As discussed in more detail in Section I, the production of such artificial shortage prices under these conditions is the result of inefficiencies inherent to the current market design that will be significantly improved with the implementation of the nodal market.

Although each of the blue data points represents system conditions in which the market was in shortage, the pricing outcomes are widely varied, with the majority of prices reflecting the marginal cost of the most expensive generation resource dispatched as opposed to the value of foregone operating reserves. These results indicate that relying exclusively upon the submission of high priced offers by market participants was generally not a reliable means of producing efficient scarcity prices during shortage conditions in 2007 and 2008.

Despite the mixed and widely varied results of the SPM, private investment in generation capacity in ERCOT has continued, although such investment has been dominated by baseload (non-natural gas fueled) and wind generation. As indicated in the net revenue analyses, these investments are largely driven by significant increases in natural gas prices in recent years. In contrast, private investment in mid-merit and peaking resources in ERCOT has been relatively thin. In an energy-only market, net revenue expectations for mid-merit and peaking resources are much more sensitive the effectiveness of the shortage pricing mechanism than to factors such as the magnitude of natural gas prices.

More reliable and efficient shortage pricing could be achieved by establishing pricing rules that automatically produce scarcity level prices when operating reserve shortages exist. Such an approach would be more reliable because it would not be dependent upon the submission of high-priced offers by small market participants to be effective. It would also be more efficient during the greater than 99 percent of time in which shortage conditions do not exist because it would not be necessary for small market participants to effectively withhold lower cost resources by offering at prices dramatically higher than their marginal cost.

At least for the pendency of the zonal market, shortage pricing will continue to remain dependent upon the existence of high-priced offers by market participants, and results such as those experienced in 2007 and 2008 will continue to frustrate the objectives of the energy-only market design. Further, although presenting some improvements, the nodal market design does not have a complete set of mechanisms to ensure the production of efficient prices during shortage conditions. While important even in markets with a capacity market, efficient shortage pricing is a particularly critical element in the ERCOT energy-only market to ensure that the long-term resource adequacy requirements are achieved.

C. Transmission and Congestion

One of the most important functions of any electricity market is to manage the flows of power over the transmission network, limiting additional power flows over transmission facilities when they reach their operating limits. In ERCOT, constraints on the transmission network are managed in two ways. First, ERCOT is made up of zones with the constraints between the zones managed through the balancing energy market. The balancing energy market increases energy production in one zone and reduces it in another zone to manage the flows between the two zones when the interface constraint is binding (*i.e.*, when there is interzonal congestion). Second, constraints within each zone (*i.e.*, local congestion) are managed through the redispatch of individual generating resources. The report evaluates the ERCOT transmission system usage and analyzes the costs and frequency of transmission congestion.

1. Electricity Flows between Zones and Interzonal Congestion

The balancing energy market uses the Scheduling, Pricing, and Dispatch ("SPD") software that dispatches energy in each zone to serve load and manage congestion between zones. The SPD model embodies the market rules and requirements documented in the ERCOT protocols. To manage interzonal congestion, SPD uses a simplified network model with four zone-based locations and five transmission interfaces. The transmission interfaces are referred to as Commercially Significant Constraints ("CSCs"). The following figure shows the average flows modeled in SPD during 2008 over each of these CSCs.



Average Modeled Flows on Commercially Significant Constraints

When interzonal congestion exists, higher-cost energy must be produced within the constrained zone because lower-cost energy cannot be delivered over the constrained interfaces. When this occurs, participants must compete to use the available transfer capability between zones. To allocate this capability in the most efficient manner possible, ERCOT establishes a clearing price for each zone and the price difference between zones is charged for any interzonal transactions.

The analysis of these CSC flows in this report indicates that:

- The simplifying assumptions made in the SPD model can result in modeled flows that are considerably different from actual flows.
- A considerable quantity of flows between zones occurs over transmission facilities that are not defined as part of a CSC. When these flows cause congestion, it is beneficial to create a new CSC to better manage congestion over that path.
- The differences between SPD-modeled flows and actual flows on CSCs create operational challenges for ERCOT that result in the inefficient use of scarce transmission resources.

The levels of interzonal congestion increased considerably in 2008, particularly for the North to Houston and North to South CSCs (which was new in 2008), and for the West to North CSC.

Beginning in April and continuing into May 2008, the frequency of congestion on the North to Houston and North to South CSCs began to increase, at times becoming so significant that the constraint was unable to be resolved with available balancing energy in the zones where it was required. When congestion on a CSC cannot be resolved, maximum shadow prices are produced for the CSC that, in turn, produce balancing energy market prices in the deficient zones that can approach or even exceed the system-wide offer caps (the system-wide offer cap was \$2,250 per MWh beginning March 1, 2008).⁵ Historically, the inability to resolve a zonal constraint has been a relatively rare occurrence. In fact, excluding the North to Houston and North to South CSCs, the other CSCs together averaged only 15 intervals in 2008 with shadow prices that were greater than or equal to the current maximum CSC shadow price of \$5,000 per MW. In contrast, the North to South and North to Houston CSCs experienced shadow prices greater than or equal to \$5,000 per MW in 92 and 87 intervals, respectively.

The sharp increase in the frequency of occurrence of unresolved congestion on the North to Houston and North to South CSCs prompted the IMM, in consultation with ERCOT and the PUCT, to initiate in early May 2008 a detailed examination of ERCOT's congestion management procedures. This investigation quickly revealed that ERCOT rules permitted certain transmission elements to be managed with zonal balancing energy deployments when, in actuality, the congestion on these elements was neither effectively nor efficiently resolvable with zonal balancing energy deployments (the transmission elements that can be designated to be managed with zonal balancing energy deployments in the same manner as the CSC are referred to as "Closely Related Elements (CREs)" in the ERCOT Protocols).

Under the current zonal market model, transmission congestion is resolve through a bifurcated process that consists of either (1) zonal balancing energy deployments, or (2) local, unit-specific deployments. Because the pricing and incentives for both load and resources are better aligned with zonal congestion management techniques under the zonal market model, it is preferable to

5

A shadow price represents the marginal generation cost savings that would be achieved if one additional MW of capacity were available on a constraint.

manage transmission congestion by using zonal balancing energy, to the extent it is effective and efficient.

However, for the CREs in question related to the North to Houston and North to South CSCs, zonal balancing energy deployments were neither effective nor efficient in resolving the transmission congestion. The result was an increasing frequency of the deployment of substantial quantities of energy in both the Houston and South Zones up to the point of exhaustion, thereby triggering the maximum shadow prices for these CSCs and associated high balancing energy prices in the South and Houston Zones that approached or even exceeded the system-wide offer cap of \$2,250 per MWh.

To address these market and reliability issues, in late May 2008, again in consultation with ERCOT and the PUCT, the IMM submitted PRR 764, which improved the definition of those transmission elements eligible to be designated as CREs. PRR 764 was processed through the ERCOT committees on an expedited basis and was implemented on June 9, 2008.

While PRR 764 effectively resolved the issues encountered during April through June 2008 within the context of the zonal market model framework, the implementation of the nodal market will eliminate the current bifurcated congestion management process by providing simultaneous, unit-specific solutions that will always presents the most effective and efficient congestion management alternatives to the system operator.

As a point of comparison, the following two figures show the pricing implications of unresolved North to South congestion under the zonal model and under a nodal model. These graphics show that an unresolved constraint produces extremely high market clearing prices are very widespread under the zonal model. In contrast, for the same unresolved constraint, the resulting high prices remain much more localized under the nodal model.



Pricing Contours of Unresolved Congestion in the Zonal Market

Pricing Contours of Unresolved Congestion in the Nodal Market



These differences in pricing outcomes are due to the use of zonal average shift factors under the zonal model compared to the use of location-specific shift factors under the nodal model.

Further, because the nodal market employs unit-specific offers and dispatch, the control of power flows on the system is much more flexible and precise than under the zonal model. Hence, it is much less likely under nodal dispatch to even encounter unresolvable constraints such as those experienced on the North to South and North to Houston interfaces in 2008.

In consideration of these differences, we have estimated the benefits that the nodal market would have produced by allowing more efficient resolution of the congestion on the North to South and North to Houston interfaces in 2008. This analysis indicates that the annual average balancing energy market price in the Houston and South Zones would have been reduced by approximately \$10.42 per MWh. Assuming that only 5 to 10 percent of customers in the South and Houston Zones were directly affected by the significant price increases in the balancing energy market and short-term bilateral markets associated with the North to Houston and North to South congestion, this analysis indicates that the efficiencies of the nodal market, had it been in place, could have reduced the annual costs for customers by \$87 to \$175 million in 2008. This analysis estimates only the savings that could have occurred through more efficient congestion management during the periods of acute North to Houston and North to South congestion, and does not include the benefits that the nodal market will provide more generally with respect to congestion management and other dispatch efficiency improvements.

The other CSC experiencing a significant increase in the level of congestion in 2008 was the West to North CSC, which was binding in 5,320 intervals (15 percent). This was more frequent than any other CSC in 2008 and more frequent than any other CSC since the inception of single control area operations in 2001. The primary reason for the high frequency of congestion on the West to North CSC in 2008 is due to the significant increase in installed wind generation relative to the load in the West Zone and limited transmission export capability to the broader market. The installed wind capacity in ERCOT grew from approximately 4.5 GW at the beginning of the year to approximately 8.1 GW by December 2008, with more than 90 percent of wind capacity located in the West Zone. Depending on load levels, transmission system topology and other system operating conditions, up to 4 GW of wind generation in the West to North CSC.

Although the marginal production cost of wind generators is near zero, the operating economics are affected by federal production tax credits and state renewable energy credits, which lead to negative-priced offers from most wind generators. Thus, when transmission congestion occurs that requires wind generators to curtail their output, negative balancing energy market prices will result in the West Zone. The hourly average balancing energy market price in the West Zone was less than zero in over 1,100 hours during 2008.

Although plans exist through the Competitive Renewable Energy Zone ("CREZ") project to significantly increase the transmission export capability from the West Zone, it is likely given the current transmission infrastructure and the level of existing wind facilities in the West Zone that the quantity of wind production that can be reliably accommodated in the West Zone will continue to be significantly limited for several years until such transmission improvements can be completed.

2. Transmission Congestion Rights and Payments

Participants in Texas can hedge against congestion in the balancing energy market by acquiring Transmission Congestion Rights ("TCRs") between zones, which entitle the holder to payments equal to the difference in zonal balancing energy prices. Because the modeled limits for the CSC interfaces vary substantially, the quantity of TCRs defined over a congested CSC frequently exceeds the modeled limits for the CSC. When this occurs, the congestion revenue collected by ERCOT will be insufficient to satisfy the financial obligation to the holders of the TCRs and the revenue shortfall is collected from loads through uplift charges. The aggregate shortfall increased to \$99 million in 2008, up from \$61 million in 2007. This increase was primarily due to increased interzonal congestion in 2008 and decreased accuracy in the quantity of TCRs sold in the monthly auction, especially for the West-to-North CSC.

In a perfectly efficient system with no uncertainty, the average congestion cost in real-time should equal the auction price of the congestion rights. In the real world, however, we would expect reasonably close convergence with some fluctuations from year to year due to uncertainties. In 2006, market participants over-estimated the value of congestion on the South to North, South to Houston, and North to Houston CSCs. In 2007, market participants still over-estimated the value of congestion on the South to North and South to Houston CSCs, but

significantly under-estimated the value of congestion on the North to Houston, North to West and West to North CSCs. However, market participants generally under-estimated the value of congestion by a wide margin in 2008, particularly during the first half of the year. These outcomes were likely influenced by the congestion management procedures that were applied during the first half of the year and modified by the implementation of PRR 764 in June 2008.

3. Local Congestion and Local Capacity Requirements

ERCOT manages local (intrazonal) congestion using out-of-merit dispatch ("OOME up" and "OOME down"), which causes units to depart from their scheduled output levels. When insufficient capacity is committed to meet local or system reliability requirements, ERCOT commits additional resources to provide the necessary capacity in either the day-ahead market or in the adjustment period (the adjustment period includes the hours after the close of the day-ahead market up to one hour prior to real-time). Capacity required for local reliability constraints is procured through either the Replacement Reserve Service market ("Local RPRS") or as out-of-merit capacity ("OOMC"). Capacity required for system reliability requirements (*i.e.*, the requirement that the total system-wide online capacity be greater than or equal to the sum of the ERCOT load forecast plus operating reserves in each hour) is procured through either the RPRS market ("Zonal RPRS") or as OOMC. ERCOT also enters into RMR agreements with certain generators needed for local reliability that may otherwise be mothballed or retired. When these units are called out-of-merit order, they receive revenues specified in the agreements rather than standard OOME or OOMC payments. The following figure shows the out-of-merit energy and capacity costs, including RMR costs, from 2005 to 2008.



Expenses for Out-of-Merit Capacity and Energy

The results in the figure above show that overall uplift costs for RMR units, OOME units, OOMC/Local RPRS and Zonal RPRS units were \$217 million in 2008, which is a \$20 million increase over the \$197 million in 2007. OOME Down costs accounted for the most significant portion of the change in 2008, increasing from \$49 million in 2007 to \$78 million in 2008. OOMC/Local RPRS costs increased from \$50 million in 2007 to \$60 million in 2008, and RMR costs decreased from \$33 million in 2007 to \$20 million in 2008. This figure also shows that the highest Zonal RPRS costs occur in July and August when electricity demand in the ERCOT region is at its highest levels.

D. Analysis of Competitive Performance

The report evaluates two aspects of market power, structural indicators of market power and behavioral indicators that would signal attempts to exercise market power. The structural analysis in this report focuses on identifying circumstances when a supplier is "pivotal," *i.e.*, when its generation is needed to serve the ERCOT load and satisfy the ancillary services requirements.

The pivotal supplier analysis indicates that the frequency with which a supplier was pivotal in the balancing energy market remained relatively constant in 2008 compared to 2007. The following figure shows the ramp-constrained balancing energy market Residual Demand Index ("RDI") duration curves for 2005 through 2008. When the RDI is greater than zero, the largest supplier's balancing energy offers are necessary to prevent a shortage of offers in the balancing energy market.



Ramp-Constrained Balancing Energy Market RDI Duration Curve

The frequency with which at least one supplier was pivotal (*i.e.*, an RDI greater than zero) has fallen over the last four years from 29 percent of hours in 2005 to 21 percent of the hours in 2006 and less than 11 percent of hours in 2007 and 2008. These results indicate that the structural competitiveness of the balancing energy market in 2008 maintained the improvement exhibited in 2007 compared to prior years.

A behavioral indicator that evaluates potential economic withholding is measured 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 balancing energy price. A participant can economically withhold resources, as measured by the output gap, by raising its balancing energy offers so as not to be dispatched or by not offering unscheduled energy in the balancing energy market.

The figure below compares the real-time load to the average incremental output gap for all market participants as a percentage of the real-time system demand from 2005 through 2008.



Incremental Output Gap by Load Level

The figure above shows that the competitiveness of supplier offers improved considerably in 2006 compared to 2005, followed by even more substantial improvement in 2007 and 2008. Although 2008 exhibited a higher average incremental output gap at the highest load levels, the overall magnitude remains small and does not raise significant economic withholding concerns.

A final measure used to evaluate the competitiveness of the market outcomes in 2008 analyzes the number of balancing energy market price spikes compared to the quantity of remaining Up Balancing capacity. If the market is operating competitively, price spikes should occur during shortage and near shortage conditions, and the number of price spikes should reduce significantly as the amount of available surplus capacity increases. For the purpose of this analysis, a price spike is measured as an interval in which the balancing energy market price exceeded an implied heat rate of 20 MMBtu per MWh, which is greater than the marginal costs of most online generating units. However, the marginal cost of offline quick start units is often greater than this threshold. Thus, some of the price spikes in this figure are indicative of the deployment of quick start gas turbines, particularly in 2007 and 2008 when several market participants had well over 1,000 MW of quick start capability qualified to provide balancing energy. In contrast, in 2005 only one market participant had quick start unit qualified to provide balancing energy (Austin Energy; 7 units and approximately 330 MW), and in 2006 one additional market participant had qualified quick start gas turbines (CPS Energy; 4 units and approximately 200 MW).



The results in the figure above indicate very competitive market outcomes in 2008, with over 95 percent of the price spikes occurring during intervals with less than 500 MW of Up Balancing capacity remaining. These results show significant improvement over 2005 and 2006 when only 74 and 84 percent, respectively, of the price spikes occurred during intervals with less than 500 MW of Up Balancing capacity remaining.

Price Spikes vs. Remaining Up Balancing Capacity

The changes in the market outcomes from 2005 through 2008 shown in the figure above are consistent with expectations given the improvements in structural and supplier conduct competitiveness over this timeframe that are highlighted in the previous two figures.

Overall, based upon the analyses in this section, we find that the ERCOT wholesale market performed competitively in 2008.

I. REVIEW OF MARKET OUTCOMES

A. Balancing Energy Market

1. Balancing Energy Prices During 2008

The balancing energy market is the spot market for electricity in ERCOT. As is typical in other wholesale markets, only a small share of the power produced in ERCOT is transacted in the spot market, although at times such transactions can exceed 10 percent of total demand. Although most power is purchased through bilateral forward contracts, outcomes in the balancing energy market are very important because of the expected pricing relationship between spot and forward markets (including bilateral markets).

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 balancing energy market will translate to artificially-low forward prices. Likewise, price spikes in the balancing energy market will increase prices in the forward markets. This section evaluates and summarizes balancing energy market prices during 2008.

To summarize the price levels during the past four years, Figure 1 shows the monthly loadweighted average balancing energy market prices in each of the ERCOT zones during 2007 and 2008, with annual summary data for 2005 and 2006.⁶

.

⁶

The load-weighted average prices are calculated by weighting the balancing energy price for each interval and each zone by the total zonal load in that interval. This is not consistent with average prices reported elsewhere in this report that are weighted by the balancing energy procured in the interval. For this evaluation, balancing energy prices are load-weighted since this is the most representative of what loads are likely to pay (assuming that balancing energy prices are generally consistent with bilateral contract prices).



Figure 1: Average Balancing Energy Market Prices

ERCOT average balancing energy market prices were 37 percent higher in 2008 than in 2007. May, June and July experienced the highest balancing energy market price increases in 2008 at 222, 206 and 187 percent, respectively, of the prices in the same months in 2007. In contrast, average balancing energy market prices in September, November and December 2008 decreased to 69, 80 and 82 percent, respectively, of the prices in the same months in 2007.

The average natural gas price increased 28 percent in 2008 over 2007. May, June and July experienced the highest natural gas price increases in 2008 at 147, 170 and 187 percent, respectively, of the prices in the same months in 2007. In contrast, average natural gas prices in October, November and December 2008 decreased to 95, 99 and 80 percent, respectively, of the prices in the same months in 2007. Natural gas is typically the marginal fuel in the ERCOT market. Hence, the changes in energy prices from 2007 to 2008 were largely a function of natural gas price movements, although significant transmission congestion during April, May and June was also responsible for approximately one-quarter of the average balancing energy market price increase in 2008.

The next analysis evaluates the total cost of serving load in the ERCOT wholesale market. In addition to the costs of energy, loads incur costs associated with ancillary services and "uplift".⁷ 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 2 shows the monthly average all-in price for all of ERCOT from 2005 to 2008 and the associated natural gas price.





The components of the all-in price of electricity include:

7

- Energy costs: Balancing energy market prices are used to estimate energy costs, under the assumption that the price of bilateral energy purchases converges with balancing energy market prices over the long-term, as discussed above.
- Ancillary services costs: These are estimated based on the demand and prices in the ERCOT markets for regulation, responsive reserves, and non-spinning reserves.

As discussed in more detail in Section III, uplift costs are costs that are allocated to load that pay for out-ofmerit dispatch, out-of-merit commitment, and Reliability Must Run contracts.

Uplift costs: Uplift costs are assigned market-wide on a load-ratio share basis to pay for . out-of-merit energy dispatch, out-of-merit commitment, and Reliability Must Run contracts.

Figure 2 indicates that natural gas prices were a primary driver of the trends in electricity prices from 2005 to 2008. Again, this is not surprising given that natural gas is the predominant fuel in ERCOT, especially among generating units that most frequently set the balancing energy market prices.

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.: California ISO, New York ISO, ISO New England, and PJM. 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 energy prices increased in wholesale electricity markets across the U.S. in 2008, primarily due to increases in fuel costs.

Figure 4 presents price duration curves for the ERCOT balancing energy market in each year from 2005 to 2008. 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 hourly load-weighted average prices for the ERCOT balancing energy market.





Balancing energy prices exceeded \$50 in more than 4,900 hours in 2008 compared to more than 4,100 hours in 2007. These year-to-year changes reflect the effects of higher fuel prices in 2008 which impact electricity prices in a broad range of hours.

Figure 5 shows the hourly average price duration curve for each of the four ERCOT zones in 2008.