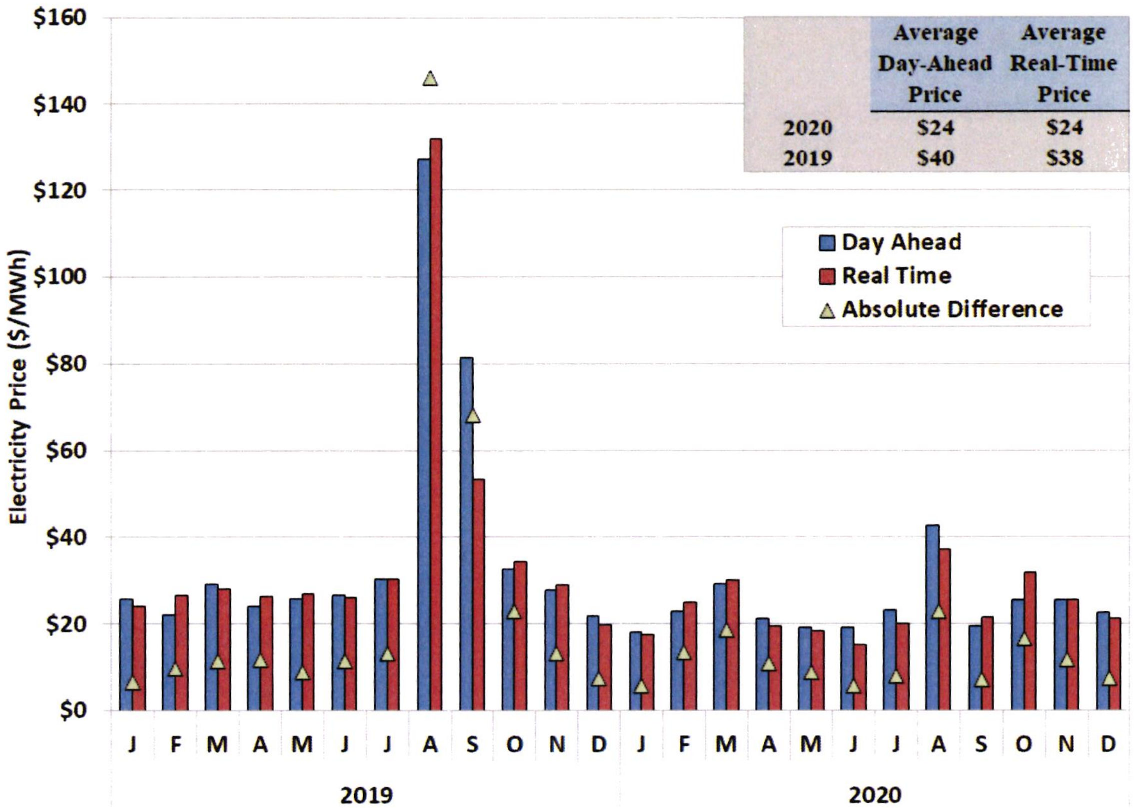
A. Day-Ahead Market Prices

Forward markets provide hedging opportunities for market participants. A primary indicator of the performance of any forward market is the extent to which forward prices converge with real-time prices over time. This price convergence will occur when: (1) there are low barriers to purchases and sales in either market; and (2) sufficient information is available to allow market participants to develop accurate expectations of future real-time prices. These two factors allow participants to arbitrage predictable differences between forward prices and real-time spot prices and bring about price convergence. Price convergence between the day-ahead and real-time markets is important because it leads to improved, i.e., more efficient, commitment of resources needed to satisfy the system's real-time needs. In this subsection, we evaluate the price convergence between the day-ahead and real-time markets.

This average price difference between forward prices and real-time spot prices reveals whether persistent and predictable differences exist between day-ahead and real-time prices that participants should arbitrage over the long term. Figure 22 shows the average day-ahead and real-time prices by month for the past two years. It also shows the average of the absolute value of the difference between the day-ahead and real-time price, calculated on a daily basis. This measure captures the volatility of the daily price differences, which may be large even if the prices converge on average.

Figure 22: Convergence Between Day-Ahead and Real-Time Energy Prices



Day-ahead and real-time prices both averaged $24 per MWh in 2020.[21] This convergence was a change from the day-ahead premium in 2019, which occurred in the summer months and reflected the value of day-ahead energy purchases as a hedge against the volatility of real-time prices under tight conditions. The relative stability of real-time prices and absence of tight conditions reduced the risk premium associated with day-ahead hedges.

Price convergence was evident in all months of 2020 except August, when day-ahead prices were higher, and October, when real-time prices were higher, likely offsetting and creating price convergence on average for the year. Slightly larger quantities of installed reserves for the summer of 2020, coupled with milder temperatures, led to expectations of less frequent shortage conditions and lower associated prices in real-time.

The average absolute difference between day-ahead and real-time prices was $11.60 per MWh in 2020, a sharp decrease from $27.63 MWh in 2019 and $16.21 in 2018, respectively. The largest absolute difference primarily occurred in August as expectations of shortages in the day-ahead market and actual reserve shortages in the real-time market led to relatively larger hourly differences. The largest zonal average absolute price differences occurred in the West zone as transmission congestion led to wide swings in West zone prices. For additional price convergence results in 2020, see Figure A9, Figure A10, and Figure A19 in the Appendix.

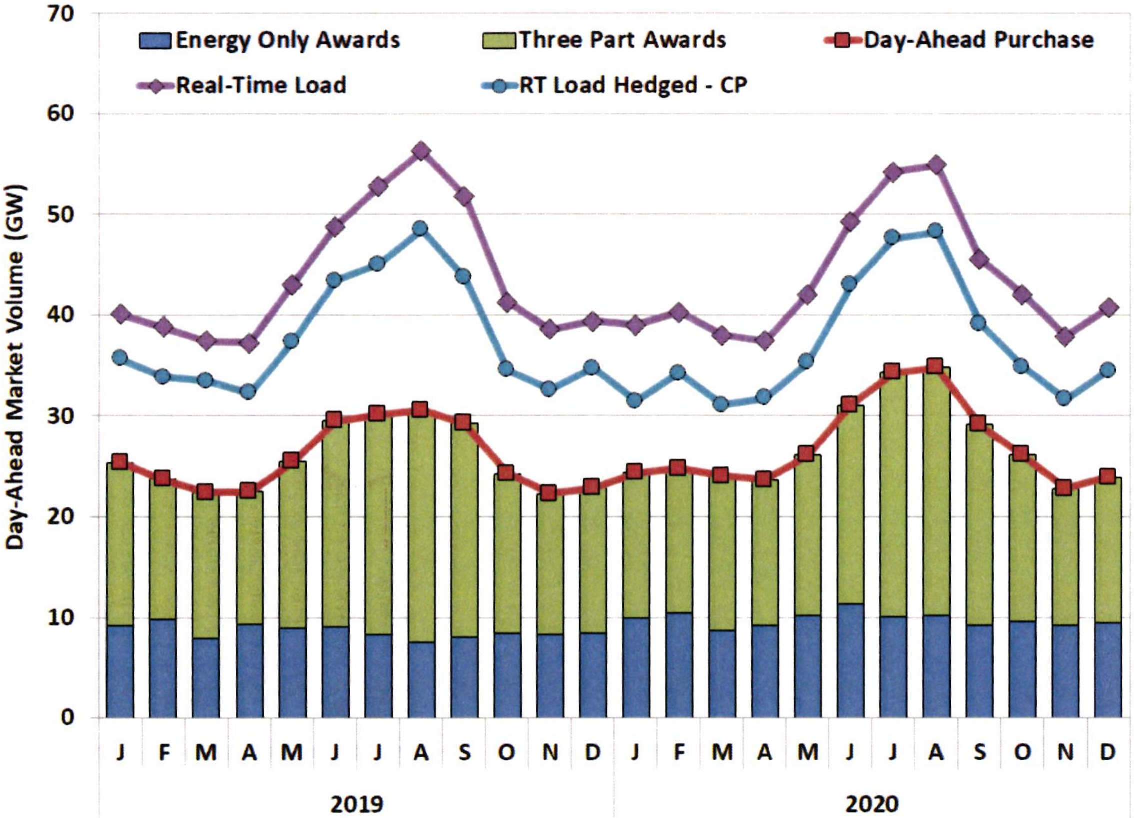
B. Day-Ahead Market Volumes

Figure 23 summarizes the volume of day-ahead market activity by month, which includes both purchases and sales of energy, for the last two years. The additional load shown as hedged in this figure (the difference between the red day-ahead purchases and the blue real-time load hedged) is load served by PTP obligations scheduled to a load zone from other locations.

Figure 23 shows that the volume of day-ahead energy purchases provided through a combination of generator-specific offers (also known as three-part offers) and virtual energy offers was 64% of real-time load in 2020, an increase from 59% in 2019. Although it may appear that many loads are still subjecting themselves to greater risk by not locking in a day-ahead price and instead exposed to real-time volatility, other transactions or arrangements outside the organized market are used to hedge real-time prices. In these cases, often PTP obligations are scheduled to hedge real-time congestion costs to complement those transactions.

21 These values are simple averages, rather than load-weighted averages as presented in Figures 1 and 2.

Figure 23: Volume of Day-Ahead Market Activity by Month



PTP obligations are financial transactions purchased in the day-ahead market. Although PTP obligations do not themselves involve the direct supply of energy, a PTP obligation allows a participant to, in effect, buy the network flow from one location to another.[22] When coupled with a self-committed generating resource, the PTP obligation allows a participant to serve its load while avoiding the associated real-time exposure because the only remaining settlement would correspond to the congestion costs between the locations. PTP obligations are also scheduled by financial participants seeking to arbitrage locational congestion differences between the day-ahead and real-time markets.

PTP volumes have been growing quickly in recent years, with important implications for the day-ahead market performance and ability to publish within the protocol timeline. They have increased four-fold over the last decade. According to ERCOT, the highest correlation to day-ahead market performance issues in unawarded PTP obligations bids, i.e., the volume of bids submitted that are unlikely to be awarded is driving the problem.

Because the large and increasing quantities of PTP transactions are the principal cause of the delays, and the delays are costly to the market at large, cost causation principles dictate that PTP

22 PTP obligations are equivalent to scheduling virtual supply at one location and virtual load at another.

volumes bear some of the costs they are causing to provide incentives to resolve the issue. DAM software capability can be thought of as a scarce resource that must be allocated efficiently. Charging no fee for PTP bids allow participants to submit very large quantities of bids that are unlikely to clear provide very little value to the market. Additionally, they bear no share of ERCOT's administrative expenses even though they are consuming a large portion of the software and supporting resources. Applying a small bid fee to the PTP bids is consistent with cost causation principles and would incent participants to submit smaller quantities of bids that are more valuable and more likely clear. Because even a small fee would likely reduce or eliminate the bids that are very unlikely to clear, this should substantially eliminate the delays in the day-ahead market process. In recommendation No. 2020-4 above, the IMM recommends a PTP bid fee as an economically rational way to manage this volume.

Figure 23 also shows the portion of the real-time load that is hedged either through day-ahead energy purchases or PTP obligations scheduled by Qualified Scheduling Entities (QSEs).[23] Although QSEs are the party financially responsible to ERCOT, their financial obligations are aggregated and held by a counterparty. When measured at this level, the percentage of real-time load hedged dropped slightly to 85% in 2020, similar to the 87% seen in 2019.

The volume of three-part offers comprised less than half of day-ahead market clearing. To determine whether this was due to small volumes of three-part offers being submitted, Figure 24 shows the total capacity from three-part offers submitted in the day-ahead market for 2020.

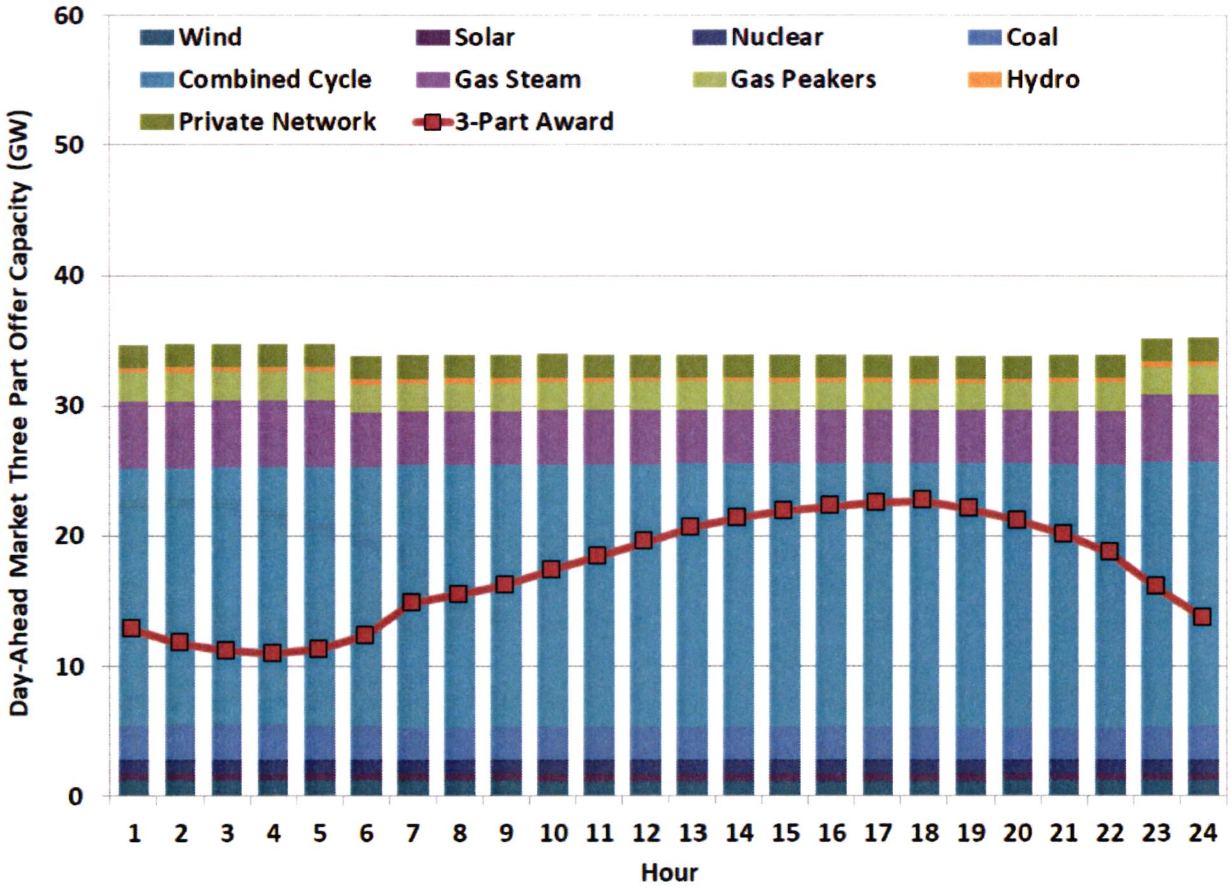
The submitted capacity has been averaged for each month and is shown to be significantly more than the amount of capacity cleared. This is not unusual, given that load in most periods does not require all available generation to be scheduled. The portion of the generation cleared in the peak hours increases as one would expect.

With the largest share of installed capacity, it follows that combined cycle units are the predominant type of generation submitting offers in the day-ahead market. More importantly, because combined cycle units are typically marginal units, offering that capacity into the day-ahead market allows a market participant to determine whether its unit is economic. Conversely, few wind units offer in day-ahead because of uncertainty on whether wind will be available in real-time to cover any award. Further analysis on day-ahead market activity volume can be found in Figure A20 in the Appendix.

---

23 To estimate the volume of hedging activity, energy purchases are added to the volume of PTPs scheduled by QSEs with load that source or sink in load zones, then aggregated to the counterparty (CP) level.

Figure 24: Day-Ahead Market Three-Part Offer Capacity



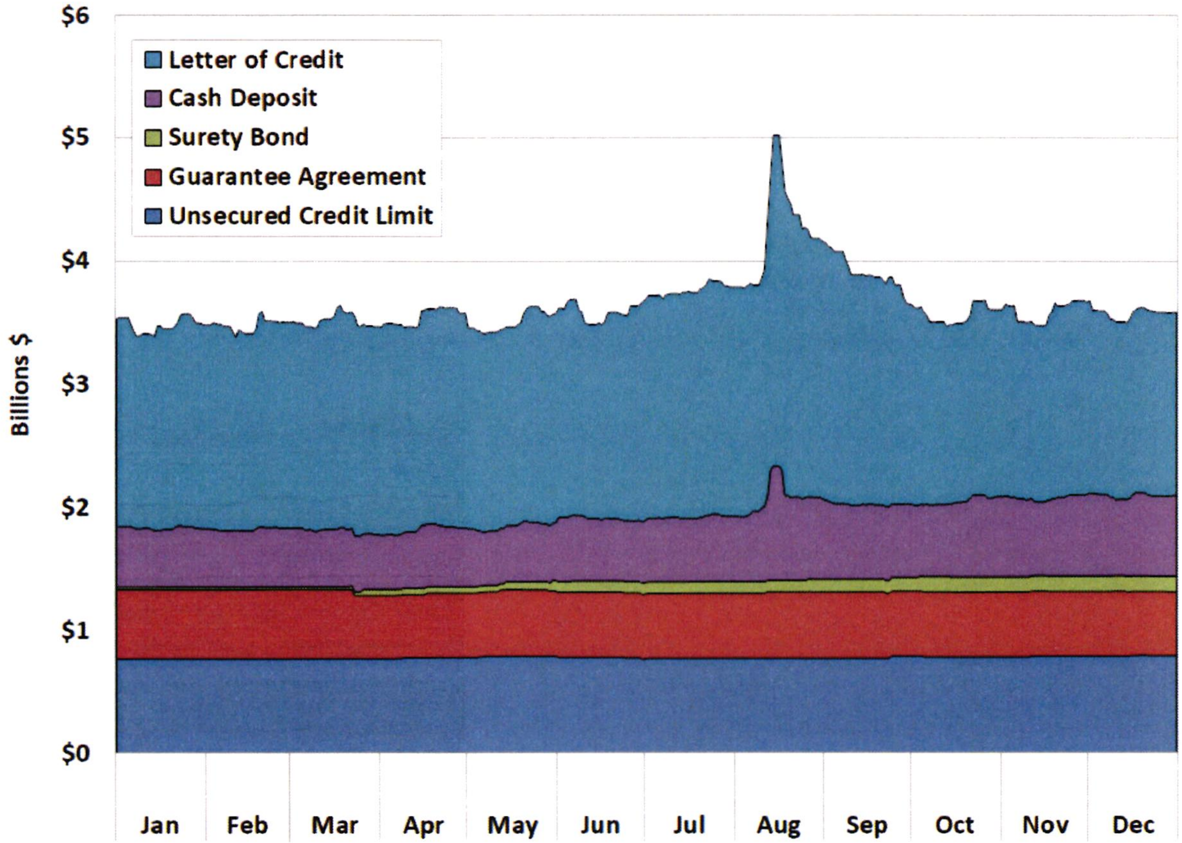
To participate in ERCOT's day-ahead market, a market participant must have sufficient collateral with ERCOT. In 2018, ERCOT introduced forward prices as a determinant in calculating collateral requirements.[24] With even smaller installed reserves in 2019, forward prices were especially high for the summer months of 2019. The effect that forward prices had on the total collateral held by ERCOT throughout the year was quite significant. That trend was reversed in 2020 as installed reserves were higher and forward prices for the summer month of 2020 were down from 2019 levels. The total collateral requirements for 2020, significantly lower than in 2019, are shown below in Figure 25.

Credit requirements are a constraint on submitting bids in the day-ahead market. When the available credit of a QSE is limited, its participation in day-ahead market will necessarily be limited as well. We see no indication that credit represented a barrier to participating in the day-ahead market in 2020.

---

24 NPRR800: *Revisions to Credit Exposure Calculations to Use Electricity Futures Market Prices*

Figure 25: Daily Collateral Held by ERCOT

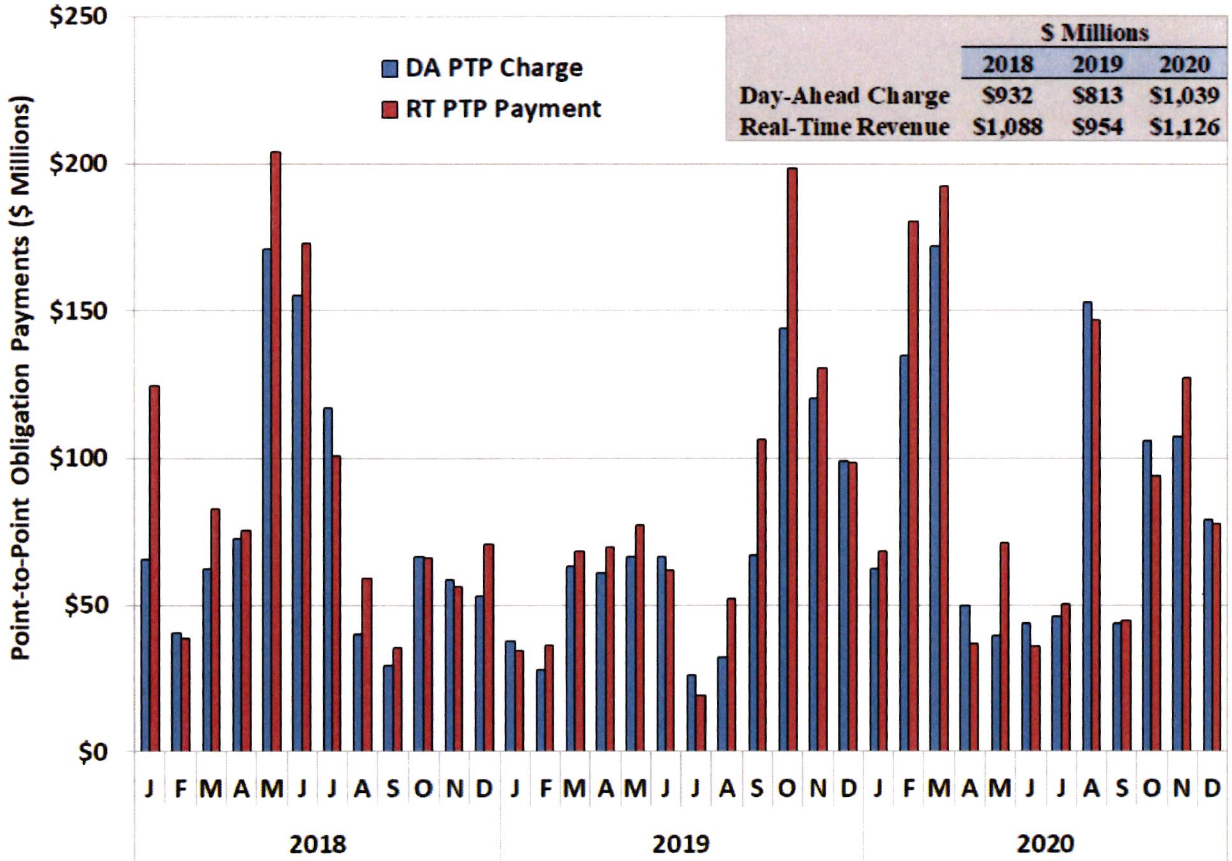


C. Point-to-Point Obligations

Purchases of PTP obligations comprise a significant portion of day-ahead market activity. They are both similar to and can be used to complement Congestion Revenue Rights (CRRs). CRRs, as more fully described in Section IV: Transmission Congestion and Congestion Revenue Rights, are acquired via monthly and annual auctions and allocations. CRRs accrue value to their owner based on locational price differences as determined by the day-ahead market.

Participants buy PTP obligations by paying the difference in prices between two locations in the day-ahead market. The holder of the PTP obligation then receives the difference in prices between the same two locations in the real-time market. Hence, a participant that owns a CRR can use its CRR proceeds from the day-ahead market to buy a PTP obligation between the same two points to transfer its hedge to real-time. Because PTP obligations represent such a substantial portion of the transactions in the day-ahead market, additional details about the volume and profitability of these PTP obligations are provided in this subsection. The first analysis of this subsection, shown in Figure 26, compares the total day-ahead payments made to acquire these products, with the total amount of revenue received by the owners of PTP obligations in the real-time market.

Figure 26: Point-to-Point Obligation Charges and Revenues



As prices and total congestion costs have increased substantially in recent years, so have the costs and revenues associated with PTP obligations. This trend was reinforced again in 2020 after a slight dip in 2019. The average volume of PTP obligations has been stable for the past three years from a quantity standpoint, although the numbers of individual transaction submissions have risen.

Figure 26 shows that the aggregated total revenue received by PTP obligation owners in 2020 was greater than the amount charged to the owners to acquire them, as in prior years. This indicates that, in aggregate, buyers of PTP obligations profited from the transactions, and occurs when real-time congestion costs are greater than day-ahead market congestion costs. Profits were spread throughout 2020 (January, February, March, May, July, September and November), accruing when congestion priced in the day-ahead market was lower than the congestion that occurred in real time.

To provide additional insight on the profits that have accrued to PTP obligations, Figure 27 shows the profitability of PTP obligation holdings for all physical parties and financial parties (those with no real-time load or generation), as well as the profitability of "PTP obligations with links to options" in 2020. These are instruments available only to Non-Opt-In Entities and allow them to receive congestion revenue but not have congestion charges. As such, we show them below as "PTP Options," because they are settled as options, not obligations.

Figure 27: Average Profitability of Point-to-Point Obligations

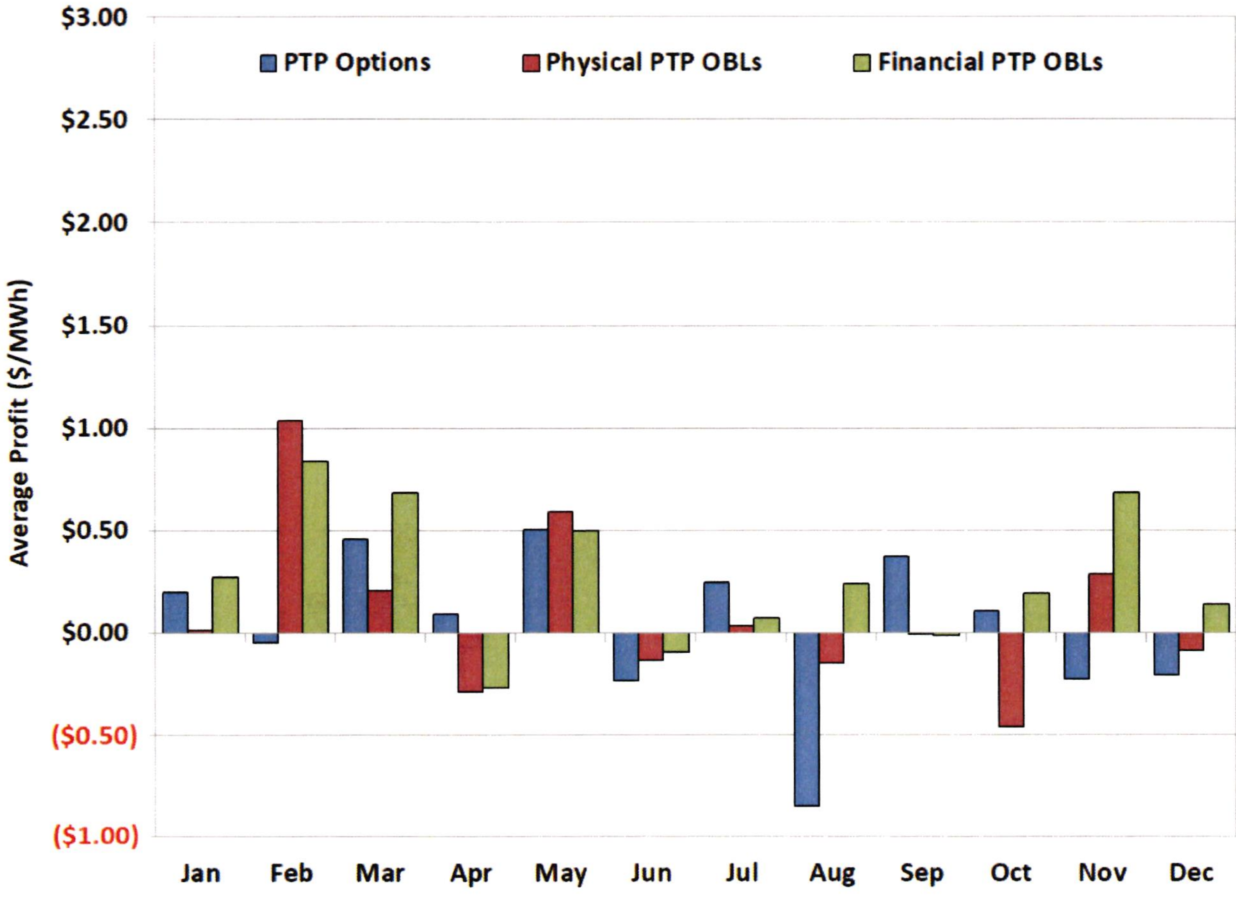


Figure 27 shows that in aggregate, PTP obligation transactions in 2020 were profitable for the year, yielding an average profit of $0.13 per MWh. This is however much less than the average profit of $0.22 per MWh in 2019. PTP obligations were profitable during 2020 for all types of parties, with average profits of $0.07 per MWh for physical parties, $0.27 per MWh for financial parties, and $0.02 per MWh for PTP obligations settled as options. For analysis of the total volume of PTP obligation purchases in 2020, see Figure A21 in the Appendix.

D. Ancillary Services Market

The primary ancillary services are regulation up, regulation down, responsive reserves, and non-spinning reserves. Market participants may self-schedule ancillary services or have them purchased on their behalf by ERCOT. In general, the purpose of responsive and non-spinning reserves is to protect the system against unforeseen contingencies (*e.g.*, unplanned generator outages, load or wind forecast errors), rather than for meeting normal load fluctuations. ERCOT procures responsive reserves to ensure that the system frequency can quickly be restored to appropriate levels after a sudden, unplanned outage of generation capacity. Non-spinning reserves are provided from offline resources that can start quickly to respond to contingencies and to restore responsive reserve capacity.

Regulation reserves are capacity that responds every four seconds, either increasing or decreasing as necessary to keep output and load in balance from moment to moment. The quantity of regulation needed is affected by the accuracy of the supply and demand reflected in the 5-minute dispatch. ERCOT increased this accuracy in 2018 by including a new factor in the determination of generation to be dispatched in the 5-minute dispatch based on the wind forecasts. ERCOT tuned the new parameters multiple times in 2019 to improve the dispatch of other generators and the efficiency of regulation deployments. These improvements continued with the implementation in late 2020 of SCR811, *Addition of Intra-Hour PhotoVoltaic Power Forecast to GTBD Calculation*, updating generation to be dispatched again to include a predicted five-minute solar ramp component. On March 1, 2020, Phase 1 of NPRR 863, *Creation of ERCOT Contingency Reserve Service and Revisions to Responsive Reserve* became effective, implementing Fast Frequency Response (FFR), the automatic self-deployment and provision by a resource of their obligated response within 15 cycles after frequency meets or drops below a preset threshold, or a deployment in response to an ERCOT Verbal Dispatch Instruction (VDI) within 10 minutes.[25]

1. Ancillary Services Requirements

Since June 2015, ERCOT has calculated responsive reserves requirements based on a variable hourly need. This requirement is posted in advance for the year. ERCOT procures non-spinning reserves such that the combination of non-spinning reserves and regulation up will cover 95% of the calculated net load forecast error. ERCOT will always procure a minimum quantity of non-spinning reserves greater than or equal to the largest generation unit during on-peak hours. In 2019, ERCOT removed the 1,375 MW floor on non-spinning quantities during on-peak hours, which slightly reduced the average quantity of reserves held by ERCOT.[26] There were no changes to the methodology for determining Ancillary Services amounts in 2020. ERCOT did place a limit of 450 MW on Resource providing Fast Frequency Response (FFR) when phase 1 of NPRR 863 was implemented.

The average total ancillary services requirement in 2020 was just shy of 4,800 MW, although the quantity of reserves held varies hour to hour. For example, on average ERCOT held roughly

---

25 Resources capable of automatically self-deploying and providing their full Ancillary Service Resource Responsibility within 15 cycles after frequency meets or drops below a preset threshold and sustaining that full response for at least 15 minutes may provide Responsive Reserve (RRS).

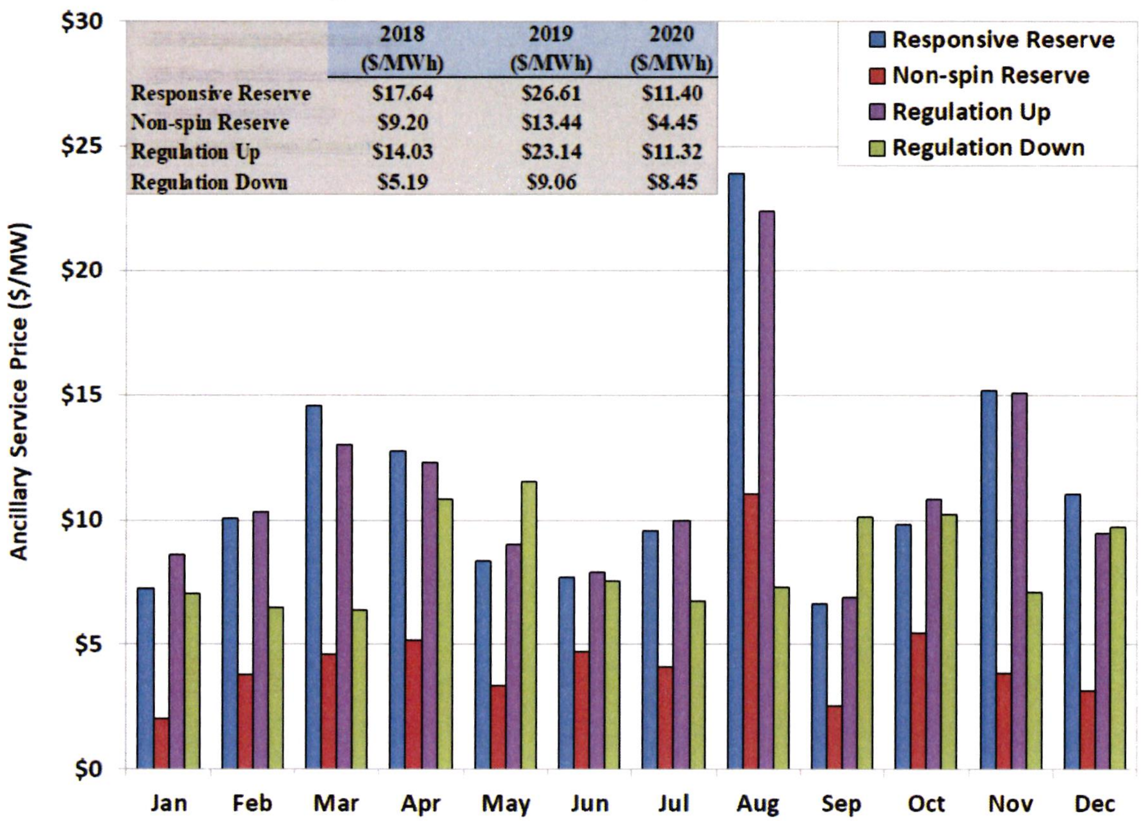
26 2020 Methodology for Determining Minimum Ancillary Service Requirements (approved by the Board on December 11, 2019).

5,400 MW of total reserves in the hour ending at 6 a.m., while it held less than 4,500 MW of reserves in hour ending 10 p.m. The primary reason ERCOT holds more reserves in some hours is that the demand for resources change output (i.e., to ramp up) is higher in some hours than others, which can cause the system to be more vulnerable to contingencies. Figure A22 and Figure A23 in the Appendix shows ERCOT's average monthly and hourly ancillary service requirements in 2020.

2. Ancillary Services Prices

Figure 28 below presents the monthly average clearing prices of capacity for the four ancillary services in 2020, while the inset table shows the average annual prices over the last three years. The prices for ancillary service were highest in August. This outcome is consistent with the higher clearing prices for energy in the day-ahead market for August because ancillary services and energy are co-optimized in the day-ahead market. This means that market participants need not include expectations of forgone energy sales in their ancillary service capacity offers. Because ancillary service clearing prices explicitly account for the opportunity costs of selling energy in the day-ahead market, ancillary service prices should generally be correlated with day-ahead energy prices.

Figure 28: 2020 Ancillary Service Prices



	2018 ($/MWh)	2019 ($/MWh)	2020 ($/MWh)
Responsive Reserve	$17.64	$26.61	$11.40
Non-spin Reserve	$9.20	$13.44	$4.45
Regulation Up	$14.03	$23.14	$11.32
Regulation Down	$5.19	$9.06	$8.45

The decrease in ancillary services prices caused the average ancillary service cost per MWh of load to decrease to $1.00 per MWh in 2020 from $2.33 per MWh in 2019. Figure A24 in the Appendix shows the monthly total ancillary service costs per MWh of ERCOT load.

3. Provision of Ancillary Services by QSEs

Day-ahead ancillary services are procured by resource, but the responsibility to provide them is aggregated up to the QSE. Table 4 shows the share of the 2020 ancillary services that are procured from the top ten QSE providers of ancillary services, in terms of volumes, compared to last year. This allows us to determine how concentrated the supply is for each product. The table also shows the total number of QSEs that represent resources that can supply each ancillary services product.

Table 4: Share of Reserves Provided by the Top QSEs in 2019-2020

	2019				2020			
	Responsive	Non-Spin	Reg Up	Reg Down	Responsive	Non-Spin	Reg Up	Reg Down
# of Suppliers	43	39	30	32	46	32	30	30
QLUMN	2%	37%	14%	43%	3%	27%	13%	40%
QLCRA	11%	6%	4%	3%	12%	7%	3%	4%
QNRGTX	8%	2%	1%	0%	11%	4%	6%	5%
QEDF26	1%	1%	5%	1%	2%	0%	18%	4%
QBRAZO	4%	6%	13%	3%	3%	6%	10%	2%
QAEN	2%	7%	3%	7%	3%	7%	4%	7%
QCALP	1%	2%	1%	3%	1%	3%	4%	10%
QOCCID	12%	0%	2%	5%	10%	0%	4%	3%
QFPL12	0%	0%	9%	4%	0%	0%	9%	4%
QEXELO	4%	0%	13%	5%	2%	0%	6%	4%

During 2020, 46 different QSEs self-arranged or were awarded responsive reserves as part of the day-ahead market. The number of providers has been roughly the same for the past five years, with three additional providers in 2020 from the previous year. A breakdown of ancillary service providers by QSE, by type of service provided, can be found in Figure A25, Figure A26, Figure A27, and Figure A28 in the Appendix.

Regarding the concentration of the supply for each product, Table 4 shows that in 2020:

- The supply of responsive reserves has not been highly concentrated, just as in 2019, with the largest QSE providing only 12% of ERCOT's responsive reserves (QLCRA as opposed to QOCCID in 2019).
- The provision of non-spinning reserves is still more concentrated than responsive reserves, but less so than 2019. A single QSE (Luminant, shown above as "QLUMN") bore almost 40% of the requirements in 2019 but only 27% in 2020. Luminant's share has continued to fall from a high of 56% in 2017. The change in composition of Luminant's generation fleet, due to mergers and retirements, likely explains this trend.

- Regulation up is provided by many different QSEs and the supply is not concentrated because, in general, many different units can ramp up to provide regulation.
- Regulation down in 2020 exhibited similar concentration to regulation down (and non-spinning reserves) in 2019. Luminant remained the dominant supplier, selling 40% of all regulation down in 2020, with Calpine also providing 10%.

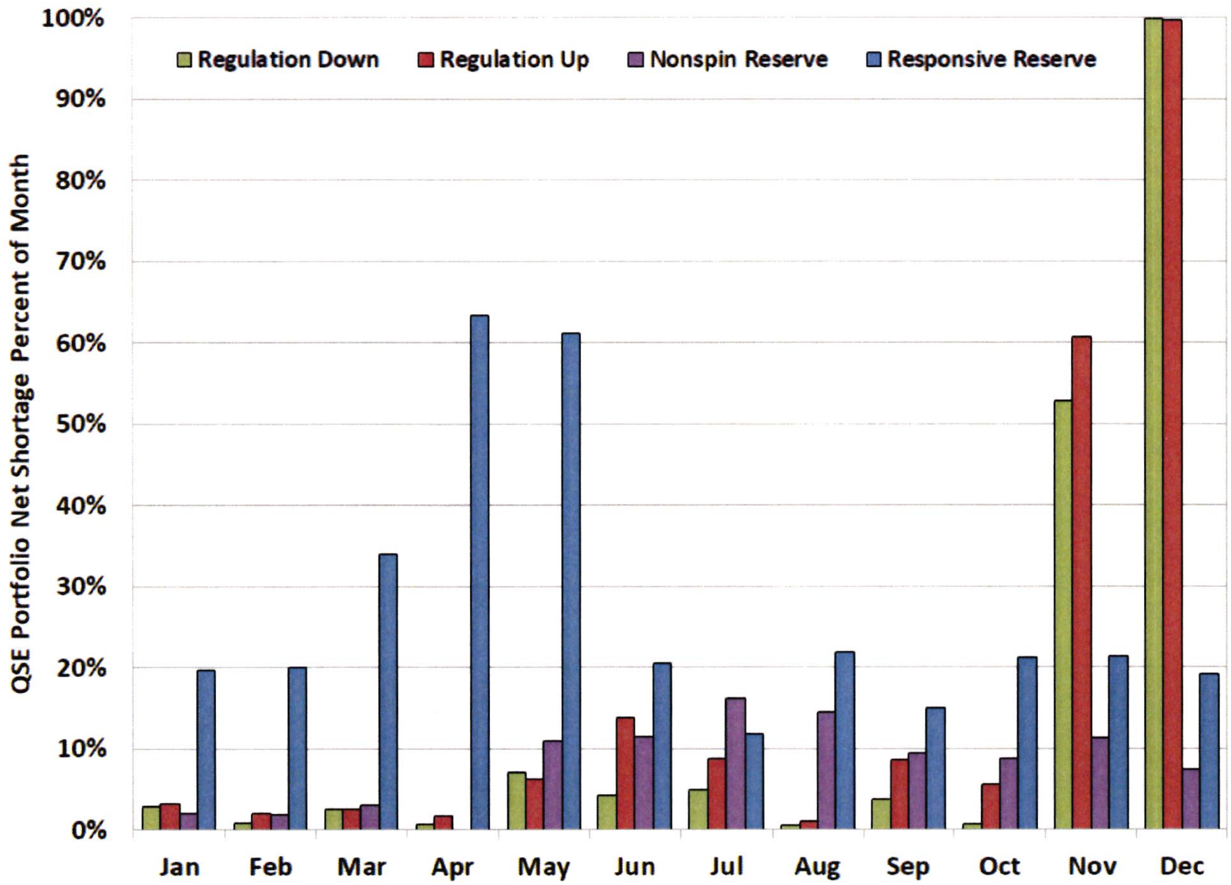
The ongoing concentration in the supply of non-spinning reserves and regulation down highlights the importance of modifying the ERCOT ancillary service market design and implementing real-time co-optimization. Jointly optimizing all products in each interval will allow the market to substitute its procurements among units on an interval-by-interval basis to minimize costs and set efficient prices. Doing so will reduce the competitive advantage of larger entities and should reduce concentration in these markets. Additionally, the use of ancillary service demand curves in the day-ahead co-optimization rather than absolute requirements will improve the efficiency of the day-ahead purchases by allowing those curves to set prices when there is a relative shortage of offers.

In addition to the procurement of ancillary services discussed above, our final evaluation relates to QSEs' delivery of the ancillary services sold in the day-ahead market. Between the time an ancillary service is procured and the time that it is needed, a QSE with multiple units may review and adjust the resources that will provide its ancillary services, presumably to reduce the costs of providing the ancillary service. However, when all ancillary services are continually optimized in response to changing market conditions, the efficiencies will be much greater than can be achieved by QSEs acting individually. These efficiencies will be achieved through real-time co-optimization.

Further, QSEs without large resource portfolios are effectively precluded from selling ancillary services because of the replacement risk faced in having to rely on a supplemental ancillary services market (SASM). If there is a forced outage in a small portfolio, the replacement risk is substantial because the clearing prices for ancillary services procured in SASM can be up to 200 times greater than annual average clearing prices from the day-ahead market. Large portfolios can often replace ancillary services without a SASM. Real-time co-optimization will address these issues. Because real-time co-optimization is set to be implemented in 2025 and will obviate the need for SASMs, we will not discuss SASM deficiencies and issues here, but we have discussed these issues in previous reports. See Section III of the Appendix for more information on SASM activity in ERCOT in 2020.

Finally, QSEs do not always provide the ancillary services that they are obligated to, due to a combination of day-ahead awards, self-arrangement, or trades. Figure 29 below shows the percentage of each month during which there was at least one QSE that did not satisfy its full ancillary services obligation. A shortage is defined as greater than 0.1 MW of obligation not being provided for at least 15 minutes out of an hour.

Figure 29: QSE-Portfolio Net Ancillary Service Shortages



Deficiencies of QSEs in meeting their ancillary service responsibilities were pervasive in 2018. However, that trend reversed over the course of 2019, most notably for regulation down service. In 2020, this trend reversal did not continue, perhaps because NPRR 947 was withdrawn. The positive effective from ERCOT's altered approach to ancillary shortages was muted due to the lack of automation of the process. The NPRR had refined the ERCOT process for determining when a QSE has failed on its ancillary service supply responsibility and, relatedly, ERCOT's process for charging QSEs for a failed ancillary service quantity, creating a mechanism to reduce payment for ancillary service awards in situations when the QSE has not fully met the award.[27] We note that there were significant shortages for RRS in April and May with multiple responsible QSEs and that for November and December, one QSE accounted for almost all the regulation shortages.

27 See NPRR 947: Clarification to Ancillary Service Supply Responsibility Definition and Improvements to Determining and Charging for Ancillary Service Failed Quantities, later withdrawn in August 2020.

IV. Transmission Congestion and Congestion Revenue Rights

An essential function of any electricity market is to efficiently manage power flows on the transmission networks. Congestion management occurs as the markets coordinate the dispatch of generation to ensure that the resulting power flows do not exceed the operating limits of the transmission facilities. This coordination occurs through the real-time market dispatch software, which optimizes based on each generator's energy offer curve and how much of its output will flow across the overloaded transmission element. The result of this market dispatch is a set of locational prices that vary at different locations across the network and resulting congestion costs that are collected from participants. Congestion exists most of the time; at least one constraint was binding (the flow at the constraint's limit) in real time during three-quarters of 2020.

The locational difference in prices caused by congestion can result in costs or risks for parties in long term power contracts who are liable for the price differences between the location of the generator and the location of the load. CRRs are economic property rights that are funded by the congestion collected through the day-ahead market. CRR markets enable parties to purchase the rights to locational price differences in monthly blocks as much as three years in advance. Hence, CRRs provide a hedge for day-ahead congestion, and can easily be converted into a real-time congestion hedge.

This section of the Report evaluates congestion costs and revenues in 2020. We first discuss the congestion costs in the day-ahead and real-time markets, which totaled $1.3 billion and $1.4 billion respectively, in 2020. We then discuss the CRR markets and funding in 2020.

A. Day-Ahead and Real-Time Congestion

As the day-ahead market clears financially-binding supply, demand and PTP obligation transactions, it does so while also respecting the transmission system limitations. This can result in widely varying locational prices and associated congestion. This congestion can be affected by planned transmission outages, load, and renewable forecasts, which also inform market participants' decisions on how to hedge portfolios before real-time. In real-time, congestion costs represent the cost of managing the network flows resulting from physical dispatch of generators. Figure 30 and Figure 31 summarize the monthly and annual congestion costs in the day-ahead and real-time markets. The values are aggregated by geographic zone.

Figure 30 shows that the total day-ahead congestion costs in 2020 were roughly 19% higher than costs in 2019; similarly, real-time congestion costs increased 11%. Most of the differences in congestion costs between day-ahead and real-time were in the West zone, which constituted approximately half of all the congestion in ERCOT. The differences in these costs in the West zone reflect the uncertainty surrounding outages and severity of constraints in the area. Congestion costs were much higher in the first quarter of 2020.

Figure 30: Day-Ahead Congestion Costs by Zone

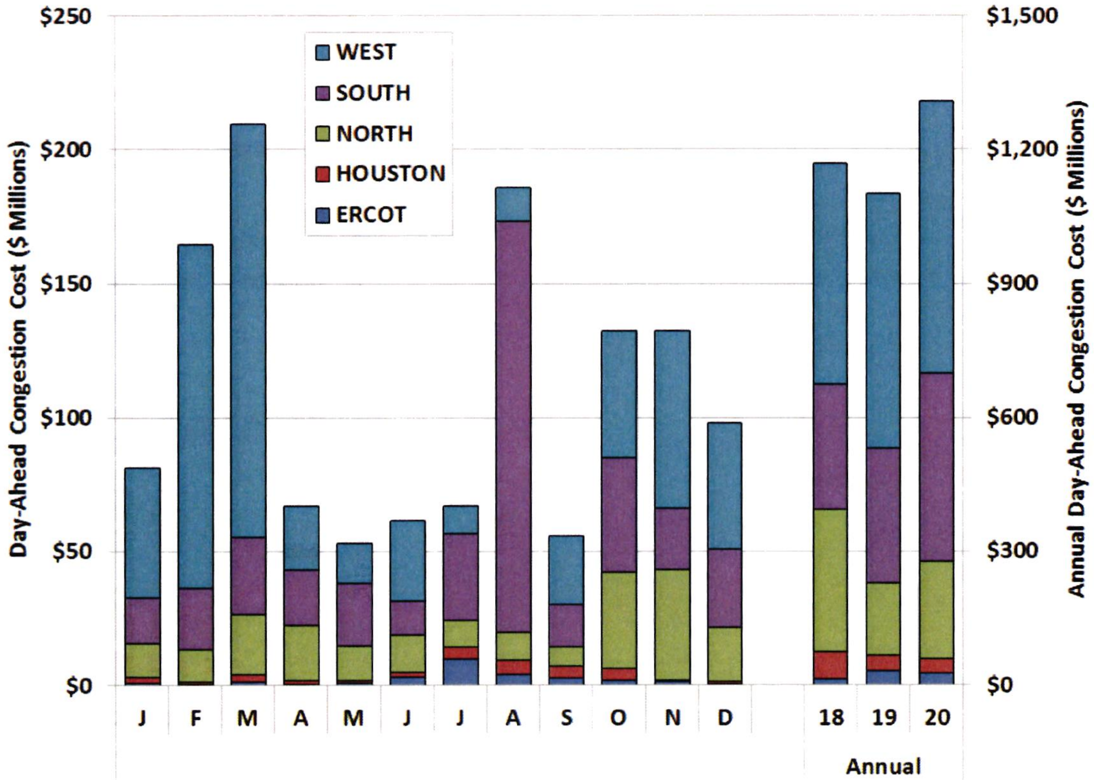
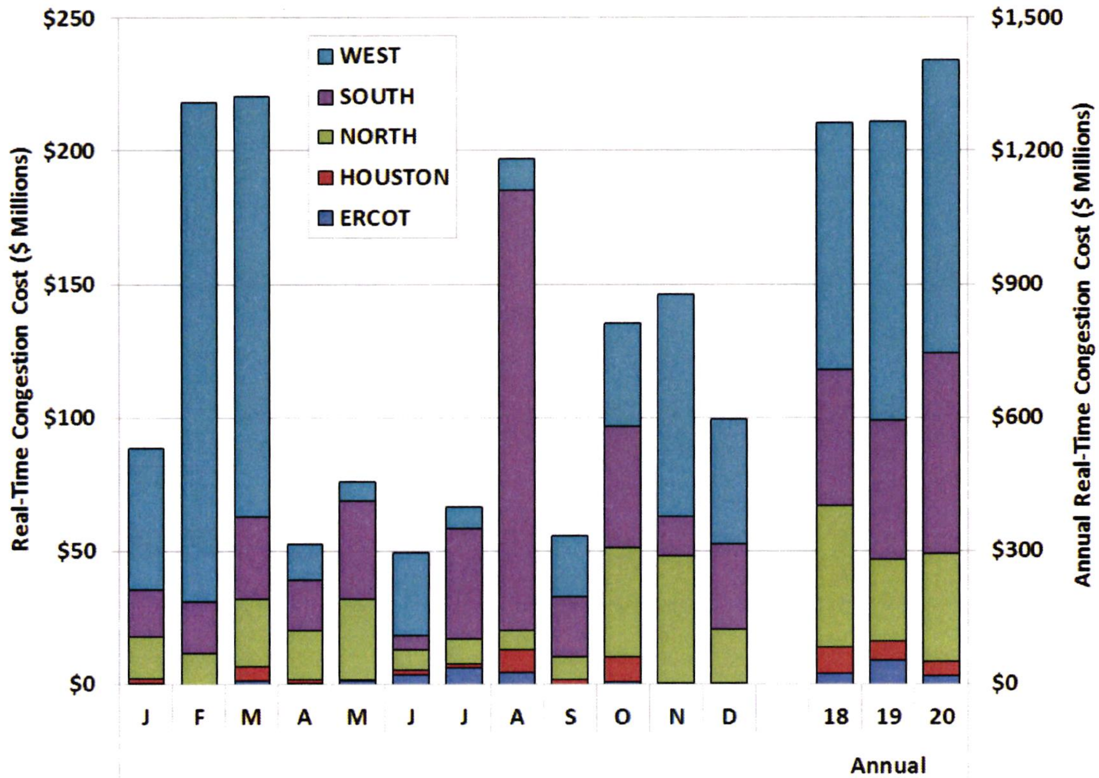


Figure 31: Real-Time Congestion Costs by Zone



The 2020 monthly congestion profile shows that congestion was highest in the winter and fall, which is an expected pattern. Shoulder months are typically when most transmission and generation outages for maintenance and upgrades occur. The increased congestion in January through May was likely due to an increase in significant transmission and generation outages, some of which were postponed to increase resource availability in the summer.

The ERCOT cross-zonal and Houston zone saw a decrease in congestion in 2020 because of the continued benefits of the North to Houston transmission project completion in April of 2018. The largest contributor to congestion costs in 2020, as was in 2019 with similar totals, was the congestion in the West zone. The congestion continued to be north of Odessa in 2020, a result of the high load caused by oil and gas development in the Permian Basin, concurrent to transmission outages for maintenance, new construction, and upgrades in the far west. The south zone experienced some weather-related outages due to Hurricane Hanna in July 2020, which led to the congestion costs in August and September. Specific top constraints in terms of dollars contributing to the real-time congestion costs are described in the next subsection.

B. Real-Time Congestion

While the expected costs of congestion are reflected in the day-ahead market, physical congestion occurs only in the real-time market. ERCOT operators manage power flows across the network as physical constraints become binding in real time. Therefore, any review of congestion must focus on the real-time constraints and resulting congestion, which we evaluate and discuss in the section.

1. Types and Frequency of Constraints in 2020

Constraints arise in the real-time market through:

- Real-Time Contingency Analysis (RTCA) that runs on an ongoing basis; and
- Generic Transmission Constraints (GTCs) that are determined by off-line studies, with limits determined prior to the operating day.[28]

RTCA is the process that evaluates the resulting flows on the transmission system under many different contingency scenarios. A base-case constraint exists if the flow on a transmission element exceeds its normal rating. A thermal contingency constraint exists if the outage of a transmission element (i.e., a contingency) would result in a flow higher than the rating of an in-

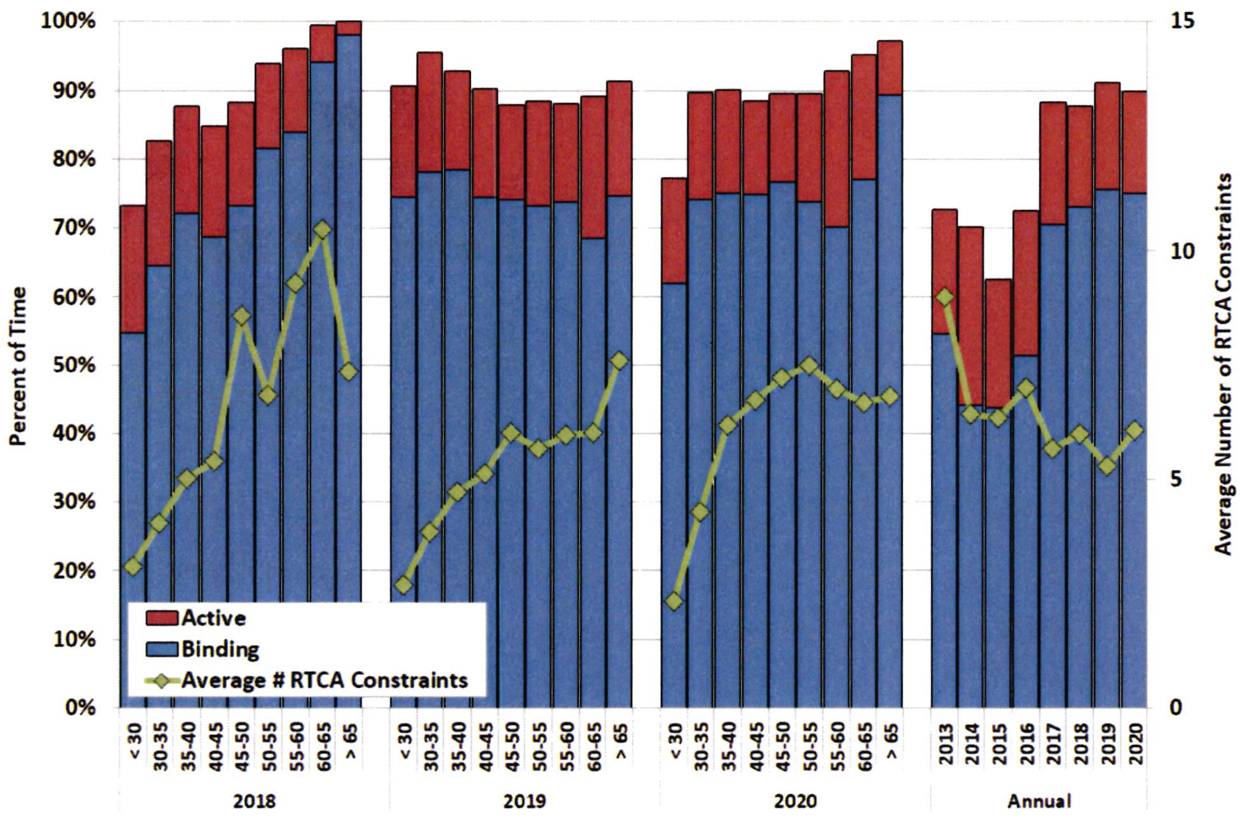
28 A Generic Transmission Constraint (GTC) is a transmission constraint made up of one or more grouped Transmission Elements that is used to constrain flow between geographic areas of ERCOT for the purpose of managing stability, voltage, and other constraints that cannot otherwise be modeled directly in ERCOT's power flow and contingency analysis applications and are based on offline studies (i.e. RTCA will not provide indication of encroaching concerns.)

service element.[29] Active transmission constraints are those that are passed by the operator to the dispatch software and that evaluated them, whereas some constraints are identified but not activated by the operator for various reasons. The active constraints are "binding" when positive dispatch costs are incurred to maintain transmission flows below the constraint limit and "not binding" when they do not require a re-dispatch of generation and thus have no effect on prices.

Our review of the active and binding constraints during 2020, Figure 32, shows the following:

- The ERCOT system had at least one binding constraint 75% of the time in 2020, a slight decrease from 76% in 2019.
- On average, slightly more than seven state estimator constraints were identified for each load bucket, up from approximately six in 2019.
- The state estimator RTCA constraints were relatively consistent across above the 35 GW load levels, although the average number of state estimator constraints were highest when load was in the 50 to 55 GW load bucket and lowest when load was less than 30 GW.

Figure 32: Frequency of Binding and Active Constraints



29 Typically, a contingency constraint is described as a contingency name plus the name of the resulting overloaded element. This section will refer to a constraint based solely on the overloaded element to identify the bottleneck in the electric grid.

GTCs doubled in binding intervals since 2019, increasing from 16% of the time to 33% in 2020 likely due to the increase in inverter-based generation in certain areas. GTCs are used to ensure that the generation dispatch does not violate a transient or voltage stability condition. Certain GTC limits are determined in real-time using the Voltage Stability Assessment Tool (VSAT) or the Transient Stability Assessment Tool (TSAT). These tools are used continuously to evaluate the North to Houston and the Rio Grande Valley Import limits, which provides a more accurate real-time limit than could be achieved through offline studies. ERCOT, Inc., has been working on getting better data for the full range of inverter technology, which over time will allow all GTC limits to be calculated in real-time rather than using offline studies. This should result in less generation curtailment. Apart from the North to Houston, Rio Grande Valley Import, and East Texas constraints, all GTCs resulted from issues identified during the generation interconnection process. As more renewable generation and energy storage resources comes online in the ERCOT region, the benefits of these dynamic models will grow.

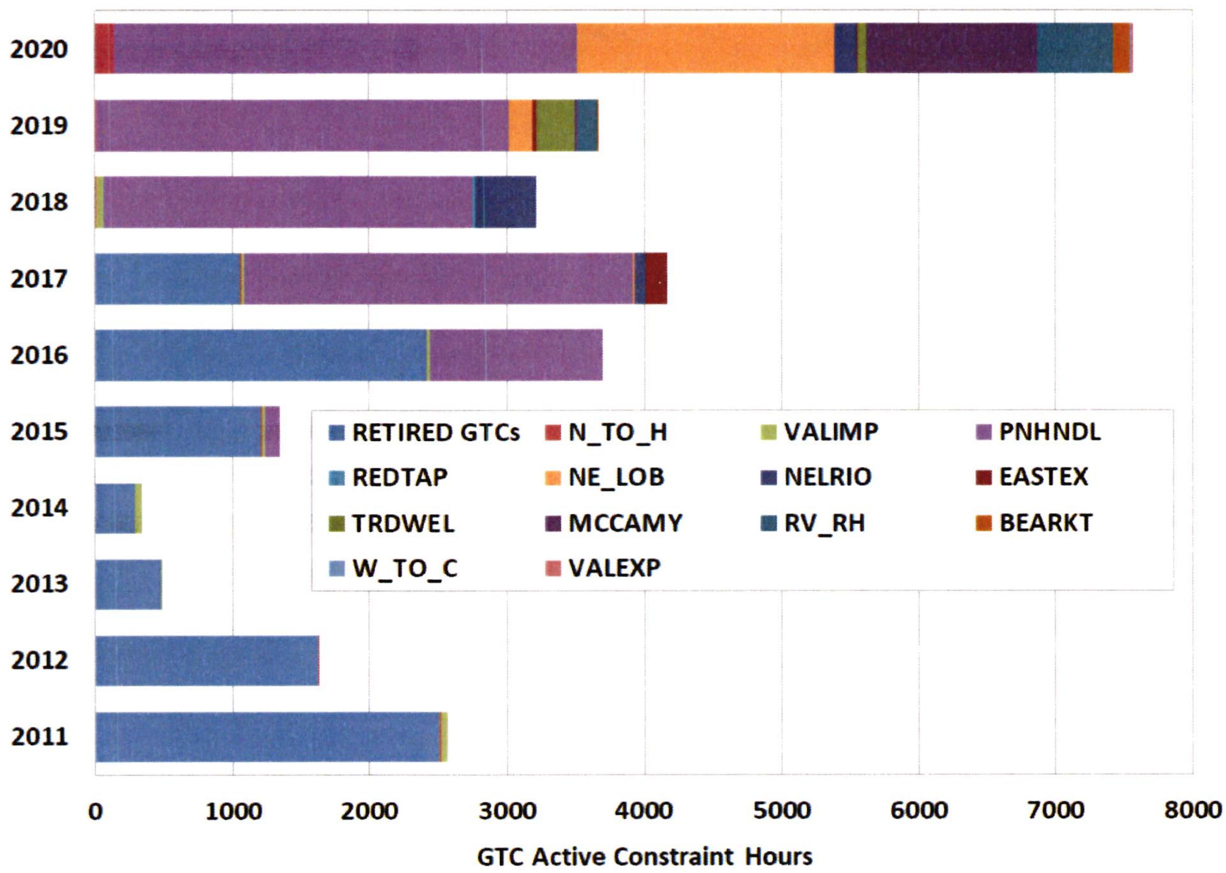
Table 5 below shows GTCs that were implemented and the number of binding intervals during 2019 and 2020.

Table 5: Generic Transmission Constraints

Generic Transmission Constraint	Effective Date	# of Binding Intervals in 2019	# of Binding Intervals in 2020
North to Houston	December 1, 2010	-	37
Rio Grande Valley Import	December 1, 2010	-	-
Panhandle	July 31, 2015	15,352	24,762
Red Tap	August 29, 2016	-	-
North Edinburg - Lobo	August 24, 2017	59	8,230
Nelson Sharpe - Rio Hondo	October 30, 2017	-	524
East Texas	November 2, 2017	155	34
Treadwell	May 18, 2018	1,539	239
McCamey	March 26, 2018	3	5,660
Raymondville - Rio Hondo	May 2, 2019	385	1,703
Bearkat	November 20, 2019	14	354
West to Central	June 24 to Oct 1, 2020	-	-
Westex (replacing West to Central)	October 1, 2020	-	-
Zapata - Starr	Novemeber 5, 2020	-	-
Valley Export	Novemeber 5, 2020	-	65
Pigcreek Solstice	November 16, 2020	-	-

The frequency in which GTCs are binding is better shown in Figure 33 depicting the aggregate total of GTC binding constraint hours from 2011 to 2020. GTCs were binding more frequently in 2020 than in previous years.

Figure 33: GTC Binding Constraint Hours[30]



The next subsection describes where and some reasons why these constraints occurred.

2. Real-time Constraints and Congested Areas

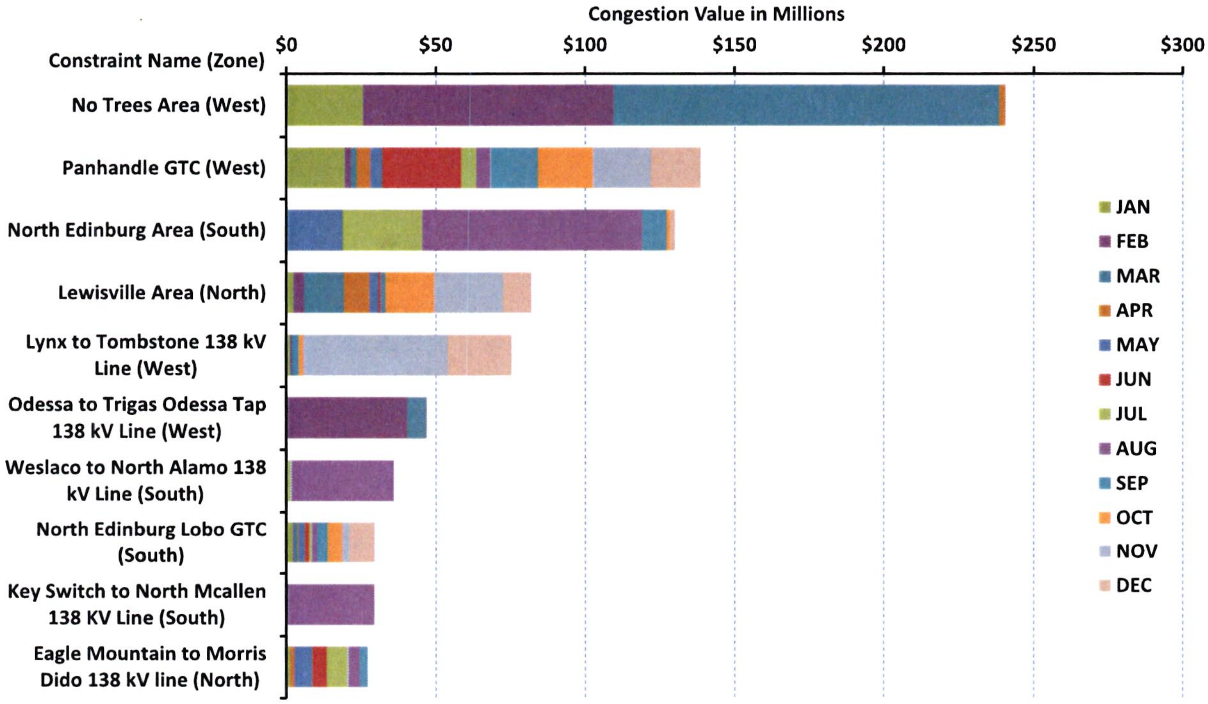
Our review of congested areas starts with describing the areas with the highest financial impact from real-time congestion. For this discussion, a congested area is identified by consolidating multiple real-time transmission constraints if the constraints are determined to be similar because of geographic proximity and constraint direction. We calculate the real-time congestion value by multiplying the shadow price of each constraint by the flow over the constraint. This gives the total dollar amount of the associate re-dispatch, where the shadow price represents the per-MW

30 Retired GTCs are Ajo to Zorillo, Bakersfield, Laredo, Liston, Molina, North to West, SOP110, West to North, and Zorillo to Ajo.

redispatch cost, defined as the marginal cost of the constraint (i.e., the dollar amount that would be avoided if the transmission element limit was 1 MW larger). Multiplying the shadow price by the flow over the transmission element itself gives that total cost of the constraint. The flow over the transmission element will be equal to the transmission element limit when the constraint is binding but may be over the limit if the constraint is violated.

There were 450 unique constraints that were either binding or violated at some point during 2020, with a median financial impact of approximately $220,000. In 2019, there were 450 unique constraints with a median financial impact of $197,000. Figure 34 displays the ten most costly real-time constraints with their respective zone measured by congestion value.

Figure 34: Most Costly Real-Time Congested Areas



The constraint with the highest congestion value in 2020 ($240 million) was the No Trees Area, consisting of the 138 kV lines Dollarhide to No Trees Switch and Andrews County South to Amoco Three Bar Tap. Much of the congestion value was generated on the line between Dollar Hide and No Trees Switch, which accounted for $193 million of real-time congestion. The congestion value associated with the No Trees Area in 2020 was $30 million less than the same congestion within the area in 2019. Most of this congestion occurred in January through April and was resolved with the 138 kV line upgrades in the area. However, the load growth from oil and gas development in Permian Basin, in conjunction with variable renewable output and outages required for transmission upgrades, continues to cause other congestion in the far west, such as Lynx to Tombstone 138 kV line and Odessa to Trigas Odessa Tap 138 kV line.

The second most costly constraint in 2020 was the Panhandle GTC constraint, which was mostly caused by planned outages, including ETT maintenance outages, in the area. The Panhandle constraint caused $140 million of congestion in 2020, a 30% increase from 2019. By the end of 2020, there was almost 4.6 GW of generation capacity in the Panhandle area, about 90% of which was wind generation. The GTC limit average for 2020 was approximately equivalent to 2019 at 3,200 MW. This average Panhandle GTC limit is attributable to the continued maintenance activity performed by Electric Transmission Texas (ETT) on its transmission structures located in the Panhandle, starting in 2017 and continuing through 2021. ETT continually monitors structures to find any additional damage and ETT has been providing updates to the market participants via the outage scheduler and market notices.

The congestion in Lewisville and Eagle Mountain has been a consistent concern as output from the Panhandle is deployed to meet the continuing load growth in the DFW area. ERCOT highlighted the aforementioned areas in the 2020 Long-Term System Assessment (LTSA) report within the ERCOT Constraints and Needs Report.[31] The report also mentions that congestion resulting from renewable output is linked to policy discussions around regional differences between the geographic location of generators and large loads. The congestion occurring in the South zone was due to the forced outages in the Rio Grande Valley from Hurricane Hanna in July 2020.

Day-ahead congestion costs were highest on the top three paths discussed above, with day-ahead congestion costs totaling roughly $456 million, somewhat less than the $510 million that accrued in the real-time market. This difference generally reflects the difference between expectations in the day-ahead market and actual real-time outcomes, and the fact that less wind generation is scheduled in the day-ahead market. Figure A34 in the Appendix presents additional detail on real-time congested areas with their respective zones in 2020.

3. Irresolvable Constraints

The shadow price of a constraint represents the marginal cost of managing a constraint (i.e., the cost of achieving the last MW of needed relief through the real-time dispatch). However, because some constraints are more costly to manage than the reliability cost of allowing them to be violated, ERCOT caps the shadow price. Without the cap, the dispatch costs and shadow price could theoretically rise to infinity, resulting in unreasonable prices. When the dispatch model cannot find a solution to manage the constraint at a marginal cost less than the shadow

31 See Report on Existing and Potential Electric System Constraints and Needs, December 2020; http://www.ercot.com/content/wcm/key_documents_lists/89026/2020_Report_on_Existing_and_Potential_Electric_System_Constraints_and_Needs.pdf

price cap, the constraint will be "irresolvable" or "in violation" in that interval, and the shadow price will be set at the cap.[32] The shadow price caps are:

- $9,251 per MW for base-case (non-contingency) constraints or voltage violations;
- $4,500 per MW for 345 kV constraints;
- $3,500 per MW for 138 kV, and
- $2,800 per MW for 69 kV thermal violations.

GTCs are considered stability constraints (for voltage or transient conditions) with a shadow price cap of $9,251 per MW.

Figure 35: Percentage Overload of Violated Constraints

presents the distribution of the percentage overload of violated constraints between 2019 and 2020. Violated constraints continued to occur in a small fraction of all the constraint intervals, 8% in 2020, down from 10% in 2019.

32 Shadow price caps are intended to reflect the reduced reliability that occurs when a constraint is irresolvable. *See* Methodology for Setting Maximum Shadow Prices for Network and Power Balance Constraints.

Figure 35: Percentage Overload of Violated Constraints

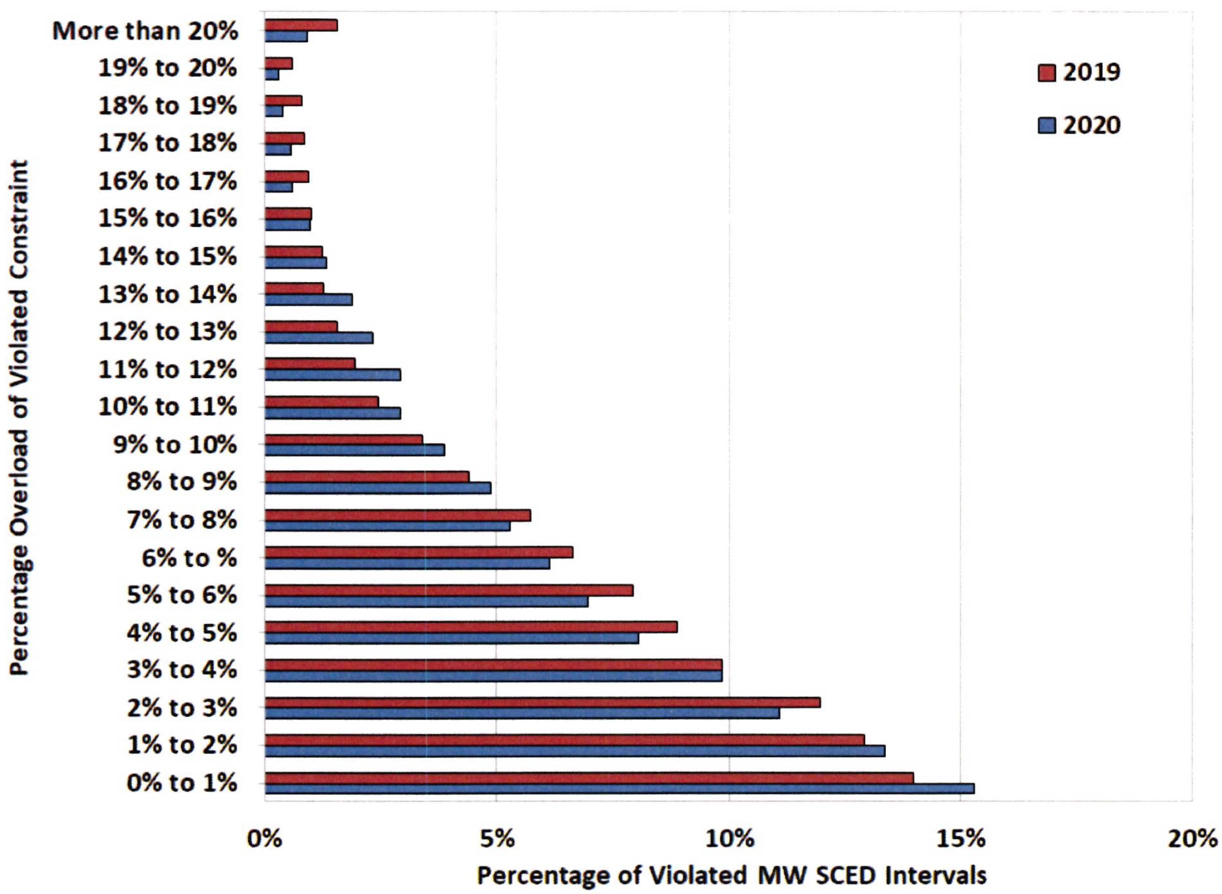


Figure 35: Percentage Overload of Violated Constraints

shows that the share of violated constraints in 2019 and 2020 were similar, with similar levels of severity of violations in both years as well. For example, none of the violated constraints levels in 2020 deviated by more than 1%, plus or minus, from the previous year. This suggests a maintained level of quality in ERCOTs ability to manage the flows in 2020.

Finally, 15% of the constraints were only slightly in violation (less than 1% of the rating), yet they are priced at the shadow price cap like the more severe violations. Almost 30% of the constraints are in violation by only small amount (between 0-2% of the transmission element rating) and these violations should be targeted for reduced shadow price caps. Implementing a well-designed transmission demand curve would recognize that the reliability risk of a post-contingency overload increases as the overload amount increases. Small violations should have lower shadow prices than large violations. Hence, we filed a revision request to implement transmission constraint demand curves.[33]

In general, violations can be resolved in subsequent intervals as generators ramp to provide relief. Nonetheless, a regional peaker net margin mechanism is applied such that once local price increases reach a predefined threshold, the constraint is deemed irresolvable and the constraint's shadow price cap is recalculated based upon the mitigated offer cap of existing resources and their ability to resolve the constraint.[34] A more detailed review of the number of violated constraints can be found in Figure A33 in the Appendix. Table A4 in the Appendix shows that 16 elements were deemed irresolvable in 2020 and had a shadow price cap imposed according to this methodology.

C. CRR Market Outcomes and Revenue Sufficiency

As discussed above, CRRs are valuable economic property rights entitling the holder to the day-ahead congestion payments or charges between two locations. CRRs are modeled as a power flow injection at the "source" and a withdrawal at the "sink." In this subsection, we discuss the results of the CRR auctions, the allocation of the revenues from the CRR auctions, and the funding of CRRs from the day-ahead market congestion.

1. CRR Auction Revenues

CRRs may be acquired in semi-annual and monthly auctions while Pre-Assigned Congestion Revenue Rights (PCRRs) are allocated to certain participants (Non Opt-In Entities or "NOIEs") based on generation units owned or contracted for prior to the start of retail competition in Texas. Parties receiving PCRRs pay only a fraction of the auction value of a CRR between the same locations. To summarize the CRR market results, Figure 36 shows the revenues, calculated by multiplying the shadow price by the flow on binding constraints in the CRR auctions.

Our calculation of the zonal CRR revenue is based on the binding constraint location, which is different from the method used to allocate CRR revenues to loads. The costs are separately shown by whether they were incurred in a monthly auction (labeled "monthly") or one of the six-

---

33 Filed on January 21, 2020 by the IMM, OBDRR026, *Change Shadow Price Caps to Curves and Remove Shift Factor Threshold*, makes certain congestion management changes for contingency constraints. This OBDRR 1) changes the default Shadow Price caps to curves (the change lowers the value for small violations and raises the value for large violations); and 2) removes the Shift Factor threshold as a factor for determining eligibility for Security-Constrained Economic Dispatch (SCED) consideration. Currently, a constraint is only eligible for resolution by SCED if at least one Resource exists that has a Shift Factor of greater than 2% or less than negative 2%. This OBDRR also proposes minor cleanup items and simplifications to Section 3, Elements for Methodology for Setting the Network Transmission System-Wide Shadow Price Caps.

34 See Section 3.6.1 of the business practice document, Setting the Shadow Price Caps and Power Balance Penalties in Security Constrained Economic Dispatch, which can be found in the Other Binding Document (OBD), Methodology for Setting Maximum Shadow Prices for Network and Power Balance Constraints.

month long-term auctions ("forward"). The "ERCOT" category contains costs associated with constraints having sources and sinks in different zones (for example North to Houston).

Figure 36: CRR Revenues by Zone

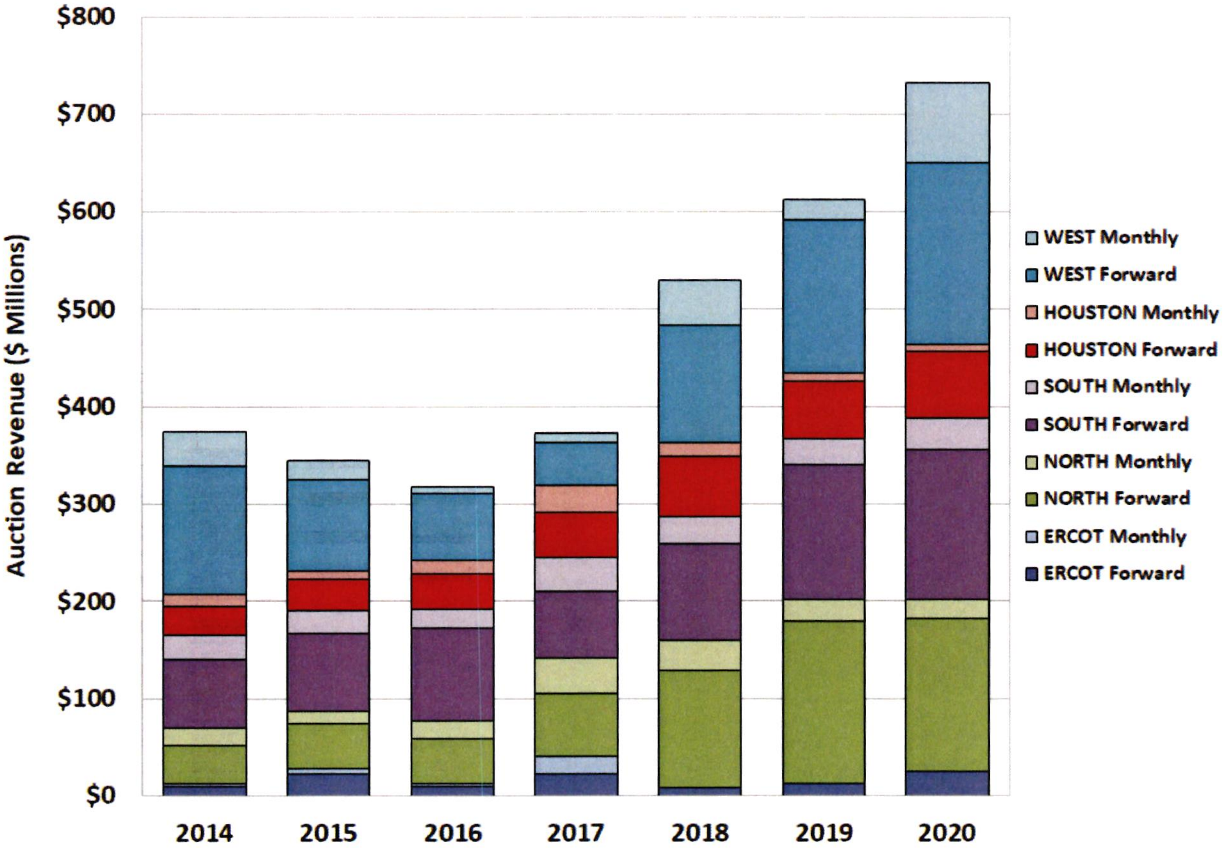


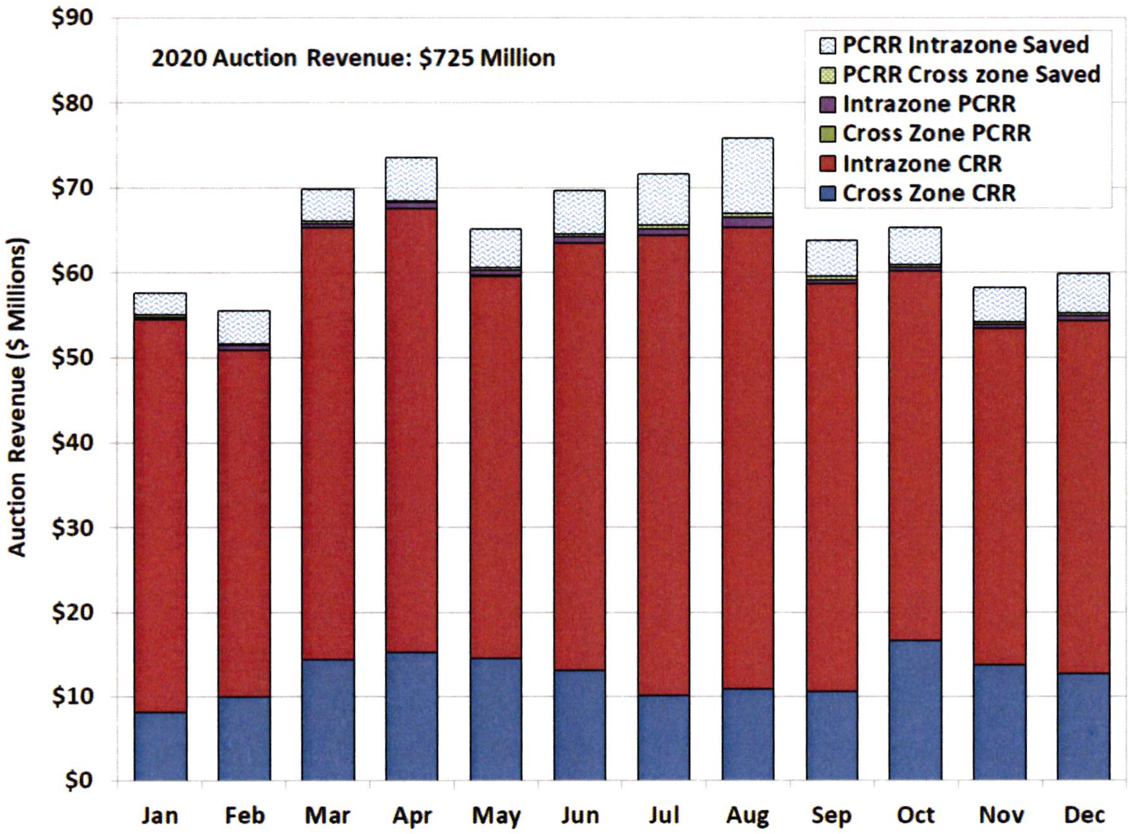
Figure 36 shows that aggregate CRR revenues have risen steadily since 2016. We note that all forwards for each of the categories increased between 2019 and 2020 except the North zone, whereas the monthly auction revenues either increased or decreased depending on zone. In general, monthly auctions will produce prices that reflect the most accurate expectations of actual congestion because they are closest to the operating horizon.

From early 2018 to early 2020, ERCOT was implementing third year CRR auctions for the first time.[35] These new auctions caused more of the transmission capacity to be sold in advance of the monthly auctions. Opportunities to purchase CRRs earlier improve forward hedging and add liquidity. However, earlier purchases can also increase differences between CRR auction revenue and day-ahead payouts because more of the CRRs are sold when there is higher uncertainty regarding the status of transmission elements, generator availability, and load levels.

35 See NPRR 808: *Three Year CRR Auction*. Approved on April 4, 2017 and implemented on September 1, 2017, this NPRR extended the CRR Auction process into the third year forward; revised the percentages sold in the CRR Long-Term Auction Sequence; and made aligning changes to the timetable for modifying load zones. The first block containing months three years in the future was posted in April 2018, and the first full cycle completed in April 2020.

ERCOT distributes CRR auction revenues to loads in one of two ways. First, revenues from cross-zone CRRs are allocated to loads ERCOT-wide. Second, revenues from CRRs that have the source and sink in the same geographic zone are allocated to loads within that zone. Figure 37 summarizes the revenues collected by ERCOT in each month for all CRRs, including both auctioned and allocated. We also show the amount of the discount provided to the PCRR recipients: the PCRR discount ("PCRR Intrazone Avoided" and "PCRR Cross Zone Avoided") is the difference between the auction value and the value charged to the purchaser.

Figure 37: 2020 CRR Auction Revenue



The total amount of CRR auction revenue increased to $725 million in 2020 from $612 million in 2019, while the total PCRR discount increased from $45 million in 2019 to $61 million in 2020. These increases reflect a yearly trend of an increased expectation of congestion in 2020.

2. CRR Profitability

CRRs are purchased well in advance of the operating horizon when actual congestion revenues are uncertain. Therefore, they may be purchased at prices below their ultimate value (based on CRR payments) and referred to as "profitable," or may be purchased at prices higher than their ultimate value and be "unprofitable". Historically, CRRs have tended in aggregate to be profitable. Although results for individual participants and specific CRRs varied, this trend continued in 2020 with participants again paying much less for CRRs they procured than their ultimate value. To evaluate these results, Figure 38 shows the monthly CRR auction revenue, the day-ahead congestion rent collected to fund the CRRs, and the payout to the CRR owners.

Figure 38: CRR Auction Revenue, Payments and Congestion Rent

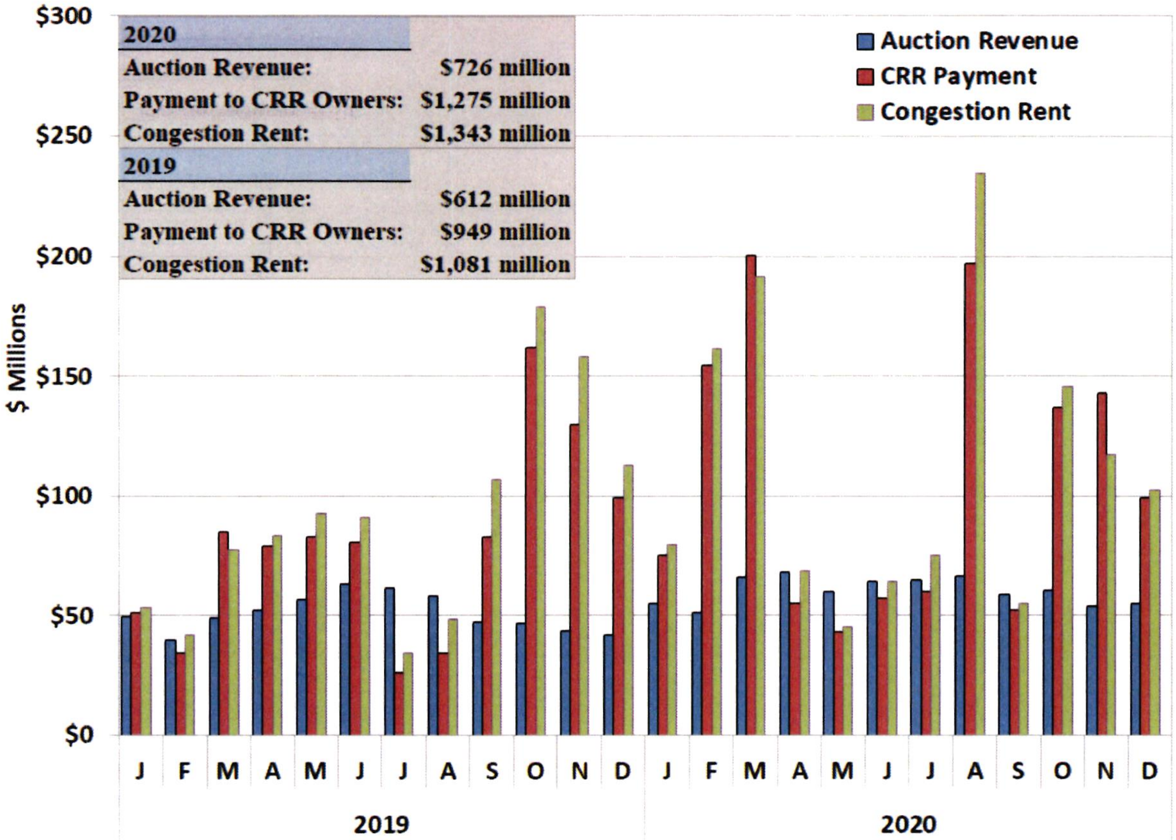
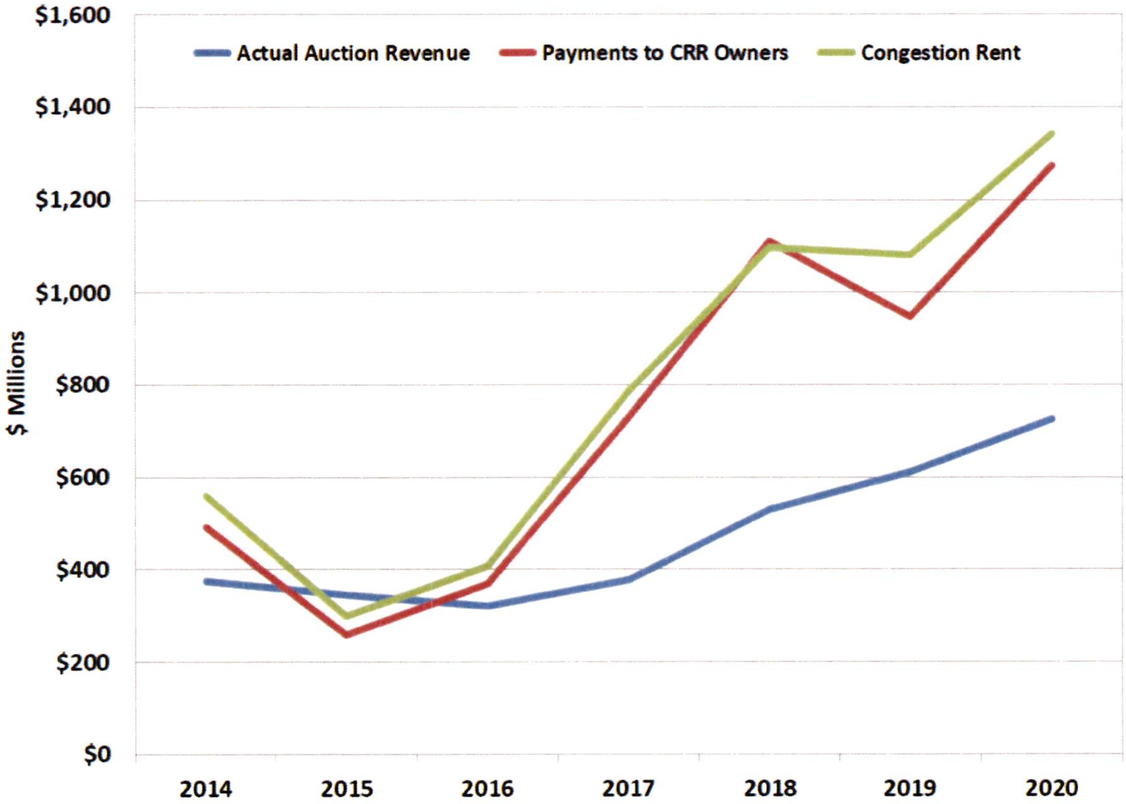


Figure 38 shows that for the entire year, participants spent $726 million to procure CRRs and in aggregate received $1,275 million, as shown in above. In general, this difference occurred because of the increase in congestion that occurred in 2020 was not foreseen by the market in the forward auction periods. The period of congestion that accounted for most of this difference was February, March, and August, which resulted in CRR payments that were $368 million higher than the auction revenue. Prices paid for CRRs represent the market expectations as of the time of the auction. Because many CRRs are purchased months (if not years) in advance, the load growth in far West that drove up the congestion costs was likely not apparent. Conversely, the CRR auction revenue in some months was higher than the CRR payouts when congestion was milder than expected. This occurred in April through July and in September in 2020.

Finally, the payout can be less than the congestion rent collected in the day-ahead market when the quantity of CRRs sold is less than the day-ahead network flows. This occurred in 2020, when the payout in aggregate was approximately $68 million less than the day-ahead congestion rent. One reason this occurs in ERCOT is that the CRR network model uses line ratings that are 90% of a conservative estimate of the lowest line ratings for the month. Therefore, CRRs tend to be a little undersold. Excess congestion rent will be discussed in the next subsection. It is instructive to review these three values over a longer timeframe, so Figure 39 provides the annual CRR auction revenues, payments to CRR owners and day-ahead congestion rent.

Figure 39: CRR History



In 2020, like the three years prior, CRRs were profitable in aggregate because of unanticipated factors that led to much higher congestion. Note that this "profit" does not account for the time value of money, which is notable because a CRR is paid for at the time of the auction and that auctions can be as much as three years in advance.

Figure 39 shows that actual congestion continues to rise more quickly than CRR auction revenues, although these revenues have been increasing in recent years. This is not unexpected because the markets must forecast the actual revenues and, even after the congestion has begun to materialize, must determine whether it will be sustained.

Figure A35 in the Appendix shows the price spreads between all hub and load zones as valued at four separate points in time: at the average of the four semi-annual CRR auctions, monthly CRR auction, day-ahead, and real-time.

3. CRR Funding Levels

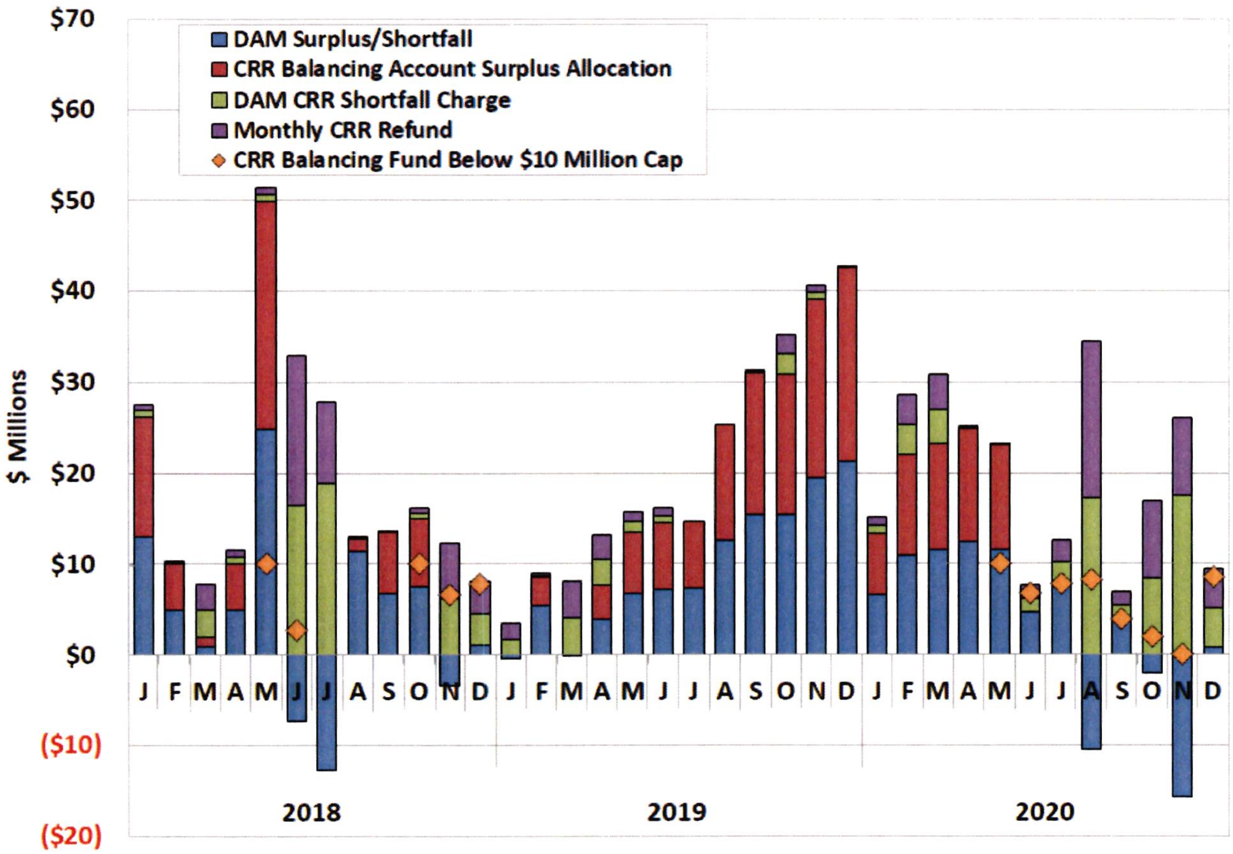
The target value of a CRR is the quantity of the CRR multiplied by the price difference between sink and source. It is desirable for the payout to fully equal the target value because it makes the CRR more valuable to the holder and ultimately will increase the CRR auction revenues. While the target value is paid to CRR account holders most of the time, ERCOT will pay less than the target value when the day-ahead congestion rent is insufficient (i.e., CRRs are not fully funded). This occurs when the CRRs' network flows exceed the capability of the day-ahead network. This is generally the result of unforeseen outages or other factors not able to be modeled in the CRR auction but that are modeled in the day-ahead market, reducing the network's transfer capability.

If this occurs on specific line or transformer (i.e., the flows on the line or transformer are "oversold"), CRRs that sink at resource nodes (generator locations) that affect the flows on the oversold transmission element have the potential to be "derated" based on the day-ahead capability of the element. Here, derated means that the CRR owner is not paid the full target value. After this deration process, if there are residual shortfalls then all holders of positively valued CRRs will receive a prorated shortfall charge. This shortfall charge has the effect of lowering the net amount paid to CRR account holders in the day-ahead settlement.

Sometimes there is excess day-ahead congestion rent that has not been paid out to CRR account holders at the end of the month (undersold hours). In that case, the excess congestion rent is tracked in a monthly settlement process referred to as the balancing account. Excess congestion rent residing in this balancing account is used to make the CRR account holders that received shortfall charges whole, i.e., they are refunded their shortfall charges. If there is not enough excess congestion rent from the current month to refund all shortfall charges, the rolling CRR balancing fund from prior months can be used to fully pay CRR account holders that received shortfall charges. Figure 40 shows the CRR balancing fund since the beginning of 2018. The CRR balancing fund has a $10 million cap, beyond which the remaining is dispersed to load.

The fact that ERCOT's processes are designed to only sell 90% of the forecasted transmission capability makes funding shortfalls less likely. Figure 40 shows that in 2020, despite this design, CRR holders experienced shortfalls in the latter half of the year due to outages that were not reflected in the CRR model. The total day-ahead surplus was nearly about $42 million, much lower than the surplus of $115 million in 2019. From the perspective of the load, the monthly CRR balancing account allocation to load totaled amount of $53 million at the end of the year.

Figure 40: CRR Balancing Fund



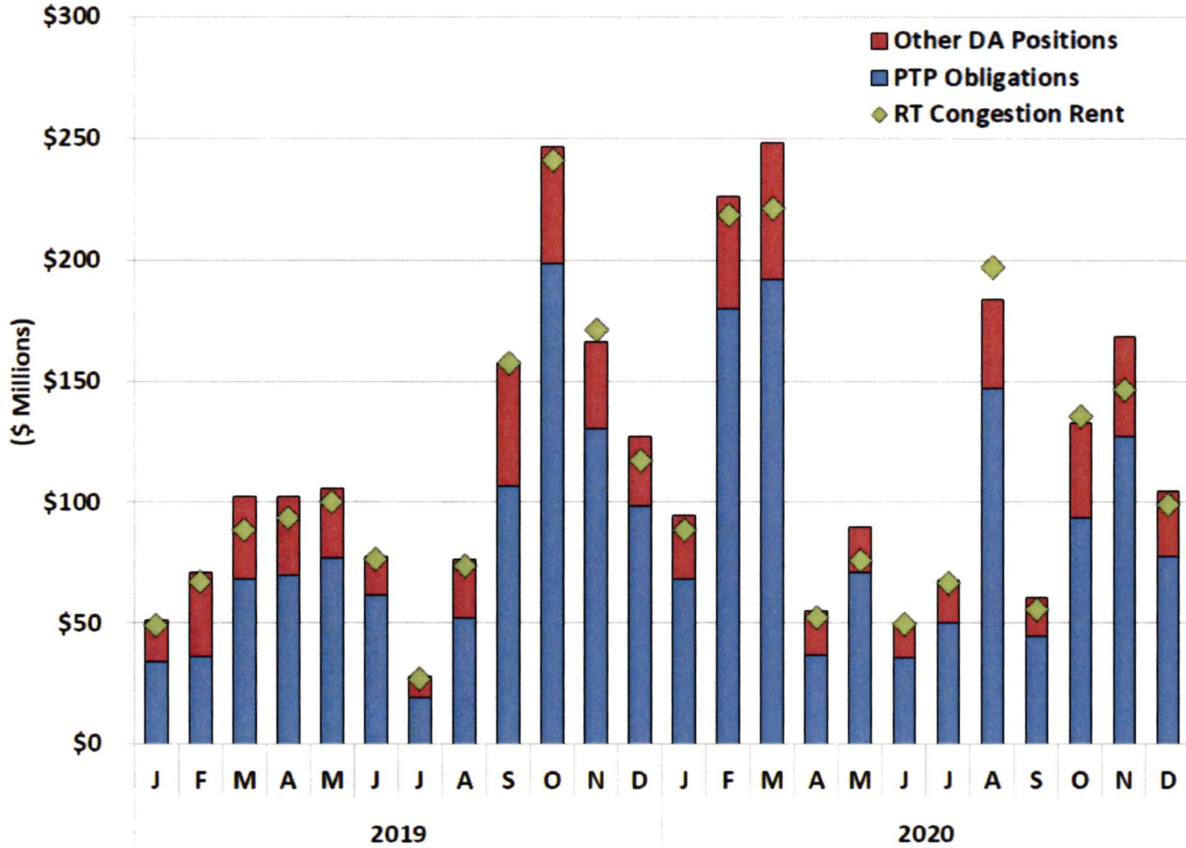
Importantly, even though the day-ahead market produced more than enough revenues to fully fund the CRRs, many CRRs were derated in 2020 and not paid the full target value due to the mandatory deration process. In total, CRR deratings resulted in a $24 million reduction in payments to CRR holders. These deratings reduced ERCOT's overall funding percentage to 98%, slightly higher than the previous year. ERCOT's deratings and shortfalls are shown on a monthly basis in Figure A36 in the Appendix. Derating CRRs, especially when the market is producing sufficient revenue to fully fund them, introduces unnecessary risk to those buying CRRs, which ultimately results in lower CRR auction revenues.

4. Real-Time Congestion Shortfalls

Just as reductions in network capability from the CRR auctions to the day-ahead market can result in CRR shortfalls, reductions in the network capability between the day-ahead market and the real-time market can result in real-time congestion shortfalls. In addition to outages or limit changes, a binding real-time constraint that is not modeled in the day-ahead market can result in real-time congestion shortfalls. In summary, if ERCOT schedules more flows in the day-ahead market over the network than it can support in real time, it will incur cost to "buy-back" the flow. These real-time congestion shortfall costs are paid for by charges to load as part of the uplift charge known as "RENA".

The day-ahead schedule flows are comprised of PTP obligations and other day-ahead positions that generate flows over the network. Figure 41 shows the combined payments to all these day-ahead positions compared to the total real-time congestion rent.

Figure 41: Real-Time Congestion Rent and Payments



In 2020, real-time congestion rent was $1,406 million, while payments for PTP obligations (including those with links to CRR options) were $1,125 million and payments for other day-ahead positions were $355 million. This resulted in a shortfall of $74 million for the year.

By comparison, payments for PTP obligations and real-time CRRs were $954 million in 2020 and payments for other day-ahead positions were $359 million, resulting in a shortfall of approximately $49 million for the year. This represents an increase over 2019 but was still lower than 2018. Higher congestion cost can tend to also drive higher shortfall amounts; in general, ERCOT has improved in coordinating the network capability in its day-ahead and real-time market. Continuous improvement in this area should be the goal of all RTOs.

V. Reliability Commitments

One important characteristic of any electricity market is the extent to which market dynamics result in the efficient commitment of generating resources. Under-commitment can cause shortages in the real-time market and inefficiently high energy prices, while over-commitment can result in excessive production costs, uplift charges, and inefficiently low energy prices.

The ERCOT market does not include a mandatory centralized unit commitment process. The decision to start-up or shut-down a generator is made by the market participant. ERCOT's day-ahead market informs these decisions but is only financially binding. That is, when a generator's offer to sell is selected (cleared) in the day-ahead market, there is no corresponding requirement to actually start that unit, although it must buy back the energy at real-time prices if it does not. Hence, this decentralized commitment depends on clear price signals to ensure an efficient combination of units are online and available for dispatch. In its role as reliability coordinator, ERCOT has the responsibility to commit units it deems necessary to ensure the reliable operation of the grid. In this way, ERCOT bridges the gaps between the economic decisions of its suppliers and the reliability needs of the system. In the event of these gaps, ERCOT uses its discretion to commit additional units to ensure reliability.

When ERCOT makes these reliability unit commitments (RUCs), the units become eligible for a make-whole payment, but also forfeit any market profit through a "clawback" provision. Generators complying with a RUC instruction are guaranteed to recover their costs, but any market revenue received over these costs are either partially or fully taken away. However, suppliers can opt to forfeit the make-whole payments and waive the clawback charges, effectively self-committing the resource and accepting the market risks.

From a market pricing perspective, ERCOT applies an offer floor of $1,500 per MWh the resource and calculates a Real-Time On-Line Reliability Deployment Adder (reliability adder) based on the low sustained limit of that resource that we described in Section I, which is intended to negate the price-lowering effects of the RUCs. In the past three years, ERCOT has made several improvements to the RUC process relating to fast-starting generators and switchable generators that are dually connected to other control areas. These improvements have caused the number of RUCs to drop dramatically, a trend that is expected to continue. For a complete list of the historical changes in the RUC processes and rules, see Section V in the Appendix.

In this section, we describe the outcomes of RUC activity in 2020. We also describe the Current Operating Plan data submitted by Qualified Scheduling Entities (QSEs) and used by ERCOT to determine the need for a RUC, whether for capacity or local congestion.

A. RUC Outcomes

ERCOT continually assesses the adequacy of market participants' resource commitment decisions using the RUC process, which executes both on a day-ahead and hour-ahead basis. Additional resources may be needed for two primary reasons:

- To satisfy the forecasted system-wide demand (0% of RUC commitments in 2020); or
- To make a specific generator available resolve a transmission constraint (100% of RUC commitments in 2020).

This is the first year since the start of nodal market that the RUC commitment reasons were all issued to manage transmission congestion. However, the number of RUC instructions in 2020 was almost identical to the number in 2019:

- The 224 unit-hours of RUC instructions were issued in 2020, down only slightly from the 228 unit-hours in 2019 and again the lowest number of instructions since the start of the nodal market.
- 82% of the RUC instructions of 2020 were issued to generators in the south zone in late July and August as a result from the damage caused by Hurricane Hanna in July 2020.
- The balance of the RUC instructions were issued as follows: 88% in the South zone, 7% in the West zone, and the remaining 5% were issued in the North zone.

The low number of RUC instructed hours had minimal make-whole payments and clawback revenues. Table 6 displays the total annual amounts of make-whole payments and clawback charges attributable to RUCs since 2011. There are two sources of funding for RUC make-whole payments. The first is from QSEs that do not provide enough capacity to meet their short real-time position, rendering them capacity short. If those charges are insufficient to cover all make-whole payments, the remaining make-whole amount is uplifted to all QSEs on a load-ratio share basis.

Table 6: RUC Settlement

	Claw-Back from Generator in millions	Make-Whole to Generator in millions
2011	$8.54	$27.80
2012	$0.34	$0.44
2013	$1.15	$2.88
2014	$2.81	$3.83
2015	$0.34	$0.48
2016	$1.41	$1.24
2017	$1.20	$0.54
2018	$3.07	$0.61
2019	$0.90	$0.05
2020	$0.48	$0.40

Table 6 shows that the make-whole payments rose to roughly $400,000 in 2020, an average level since the start of the market in 2011 (the average being about $380,000). This increase from 2019 was likely due to increased transmission congestion for which specific resources were needed for resolution. The clawback amount was slightly higher than the make-whole payment in 2020. In theory, the clawback amount should be low because units that are economic (and therefore subject to the clawback provision) would generally benefit by opting out of the RUC instruction, if such profitability is foreseeable. In 2020, approximately 8% of RUC units opted out, much lower than past years because of the high amount of congestion from forced outages in the aftermath of Hurricane Hanna, which made it difficult to predict the profitability of the "opt-out" ahead of time.

RUC Generators with Day-Ahead Offers. Generators that participate in the day-ahead market forfeit only 50% of markets revenues above cost through the clawback, rather than 100%. Given this incentive to offer in the day-ahead market, it is somewhat surprising that all units do not submit day-ahead offers. In 2020, 87% of the total RUC unit-hours had day-ahead offers, a sharp increase from 2019 when only 25% of the total RUC unit-hours had day-ahead offers, likely attributable to reliability needs of the grid after Hurricane Hanna in the Rio Grande Valley in July 2020.

Funding of RUC Payments. There are two sources of funding for RUC make-whole payments. The first is from QSEs that do not provide enough capacity to meet their short real-time position, rendering them capacity short. If those charges are insufficient to cover all make-whole payments, the remaining make-whole amount is uplifted to all QSEs on a load-ratio share basis. RUC make-whole payments in 2020 were collected almost exclusively from QSEs that were capacity short, while the amount of make-whole that was uplifted to load was de minimis.

Section V in the Appendix provides more detail on the RUC activity, showing total activity by month, statistics on day-ahead offers and decisions to opt-out of the RUC instruction, as well as the RUC instructions issued to individual generating resources. Section V also summarizes the dispatch levels of the RUC resources, which is generally at their low dispatch limit (LDL) given the $1,500 per MWh offer floor. However, RUC resources were dispatched above their LDLs in 2020 because of the mitigation of some of the resources committed to resolve non-competitive constraints. That mitigation can effectively eliminate the $1,500 per MWh offer floor for those resources in those RUC intervals.

B. QSE Operation Planning

The Current Operating Plan (COP) is the mechanism used by QSEs to communicate the expected status of their resources to ERCOT. After aggregating COP information about the amount of capacity that QSEs expect to be online every hour, ERCOT then evaluates any potential locational or system-wide capacity deficiency. If such a deficiency is identified and there is insufficient time remaining in the adjustment period to allow for self-commitment, ERCOT will issue a RUC instruction to ameliorate the shortfall.

The accuracy of COP information greatly influences ERCOT's ability to effectively perform supplemental commitment using the RUC process. COPs are updated on an ongoing basis by QSEs, providing multiple views of their expectations for a particular operating hour. Presumably, QSE expectations about which units will be online in a particular hour are most accurate for the COP submitted just before the operating hour. Figure 42 evaluates the accuracy of the COPs by showing the average difference between the actual online unit capacity and the capacity represented in the COPs in the peak hours (hour ending 12-20) in July and August, as submitted each of the 24 hours leading up to the close of the adjustment period. We show these differences for each of the past two years.

Figure 42: Capacity Commitment Timing – July and August Hour Ending 12 through 20

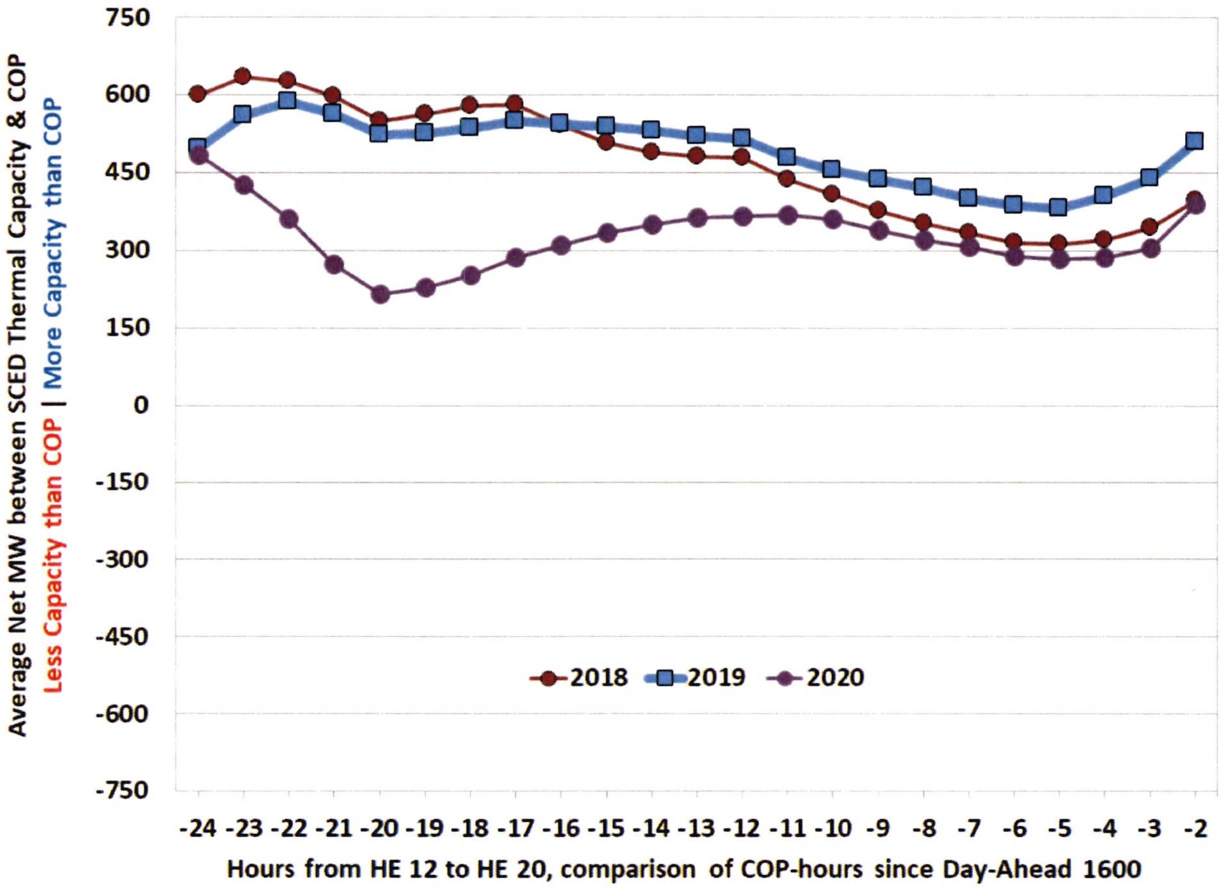


Figure 42 shows that the amount of online capacity needed exceeded the thermal capacity represented in COPs at the end of the adjustment period, signifying that generators changed their commitment decisions within the operating period. Commitment of resources for hours ending 12 to 20 show that 2020 had earlier commitments on average, approximately 20 hours before the operating hour. The difference in the last COP on average decreased from 500 MW in 2019 to 350 MW in 2020.

An average of the hours from hour-ending (HE) 12 through HE 20 masks the changes market participants may make closer to real-time. In 2019, when we focused on HE 17 during July and August, it was apparent that two QSEs (one a large supplier and one a NOIE) tended to make large changes to capacity commitments relative to their size shortly before the operating hour. This creates additional uncertainty for ERCOT operators as they fulfill their responsibility to ensure that sufficient capacity is available in the right locations to meet real-time requirements.

However, only a small portion of total RUC instructions were issued to ensure system-wide capacity sufficiency. This is testament to the restraint exhibited by ERCOT operators, allowing market participants to make their own commitment decisions, including the nearly 500 MW of near real-time thermal capacity commitments. The commitment decisions of both QSEs in 2020 indicate that they were able to represent COP capacity more accurately than they were in 2019, with COP capacity more closely aligning the with real-time capacity.

Additional analysis on COP behavior is presented in the Section V of the Appendix, which includes the analysis of hour ending 17 discussed above.

VI. RESOURCE ADEQUACY

One of the primary functions of the organized wholesale electricity market is to provide economic signals that will facilitate investment needed to maintain a set of resources adequate to satisfy the system's needs. Without revenue contributions from an installed capacity market, energy and reserve prices provide the only funding for compensation to generators. To ensure that revenues will be sufficient to maintain resource adequacy in an energy-only market, prices should rise during shortage conditions to reflect the diminished reliability and increased possibility of involuntary curtailment of service to customers. The sufficiency of revenues is a long-term expectation and will not necessarily be met in any one year: actual revenues may vary greatly from year to year.

The ERCOT market has seen many years of sufficient generation, with revenues less than estimated costs of investing in new generation (known as the "cost of new entry" or "CONE"). If long-term expectations of revenues sufficient to support resource adequacy are to be met, revenues that far exceed the CONE must occur in some years as well. This principle of cyclical revenue sufficiency to maintain resource adequacy is applied in the evaluation in this section.

This section begins with our evaluation of these economic signals in 2020 by estimating the "net revenue" that resources received from the ERCOT real-time and ancillary services markets and providing comparisons to other markets. Next, we review the effectiveness of the Scarcity Pricing Mechanism.[36] We present the current estimate of planning reserve margins for ERCOT, followed by a description of the factors necessary to ensure resource adequacy in an energy-only market design. Finally, we conclude with a discussion of the Reliability Must-Run (RMR) process in ERCOT in 2020.

A. Net Revenue Analysis

We calculate net revenue by determining the total revenue that could have been earned by a generating unit less its variable production costs. Put another way, it is the revenue in excess of short-run operating costs that is available to recover a unit's fixed and capital costs, including a return on the investment. In ERCOT's energy-only market, the net revenues from the ancillary services and real-time energy markets alone provide the economic signals that inform suppliers' decisions to invest in new generation or, conversely, to retire existing generation. To the extent that revenues are available through the day-ahead market or other forward bilateral contract markets, these revenues are ultimately derived from the expected ancillary service and real-time energy prices. Although most suppliers are likely to receive the bulk of their revenues through bilateral contracts, the spot prices produced in the real-time energy market should drive bilateral

[36] See 16 TAC §25.505(g). This report generally employs the more accurate "shortage pricing" terminology in place of "scarcity pricing", except in cases where Scarcity is part of a name.

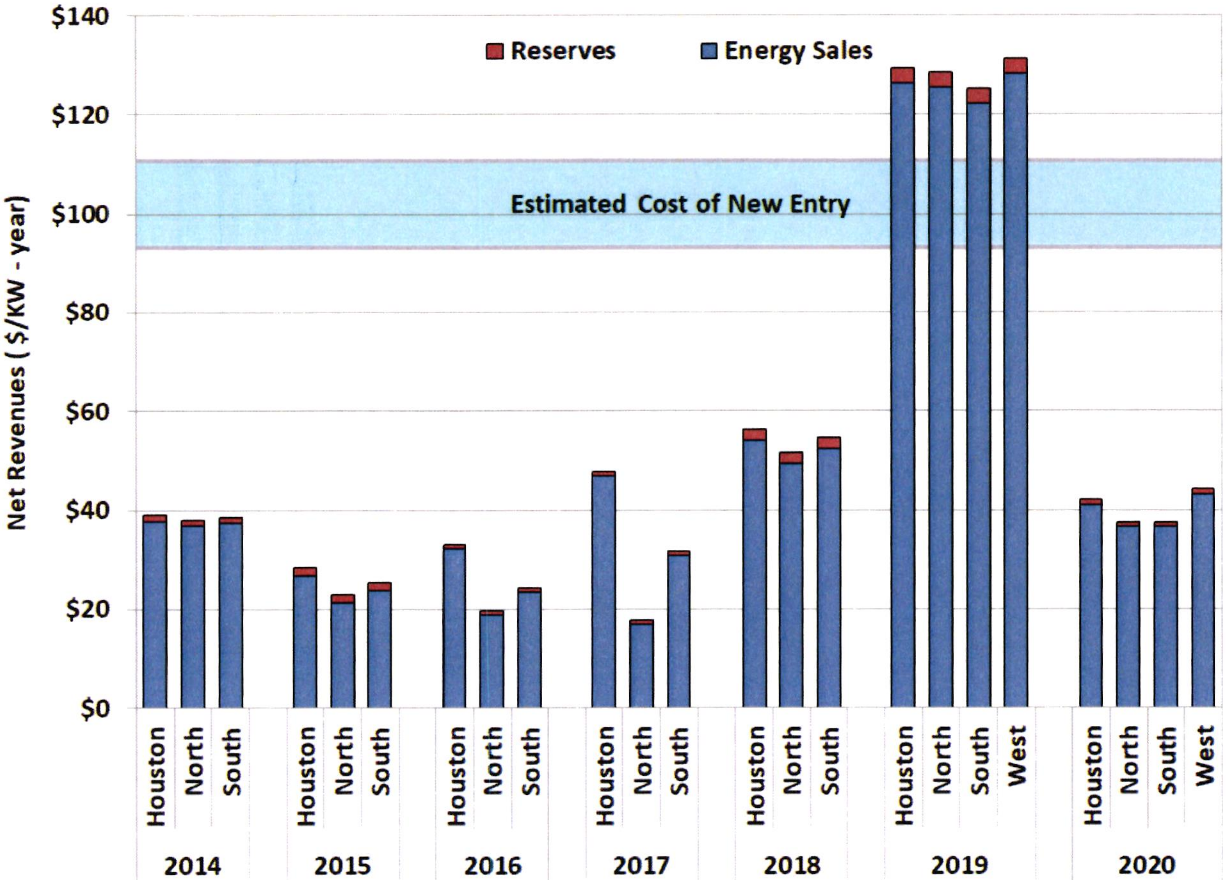
energy prices over time and thus are appropriate to use for this evaluation. It is important to note that this net revenue calculation is a look back at the estimated contribution based on actual market outcomes. Suppliers will typically base investment decisions on expectations of future electricity prices. Although expectations of future prices are informed by history, they also factor in the likelihood of shortage pricing conditions that may or may not actually occur.

In this analysis, we compute the energy net revenues based on the generation-weighted settlement point prices from the real-time energy market.[37] The analysis may over-estimate the net revenues because it does not include: 1) start-up and minimum energy costs; or 2) ramping restrictions that can prevent generators from profiting during brief price spikes. Despite these limitations, the analysis provides a useful summary of signals for investment in ERCOT.

The next two figures provide an historical perspective of the net revenues available to support investment in a new natural gas combustion turbine (

Figure 43) and combined cycle generation (Figure 44), which we selected to represent the marginal new supply that may enter when new resources are needed.

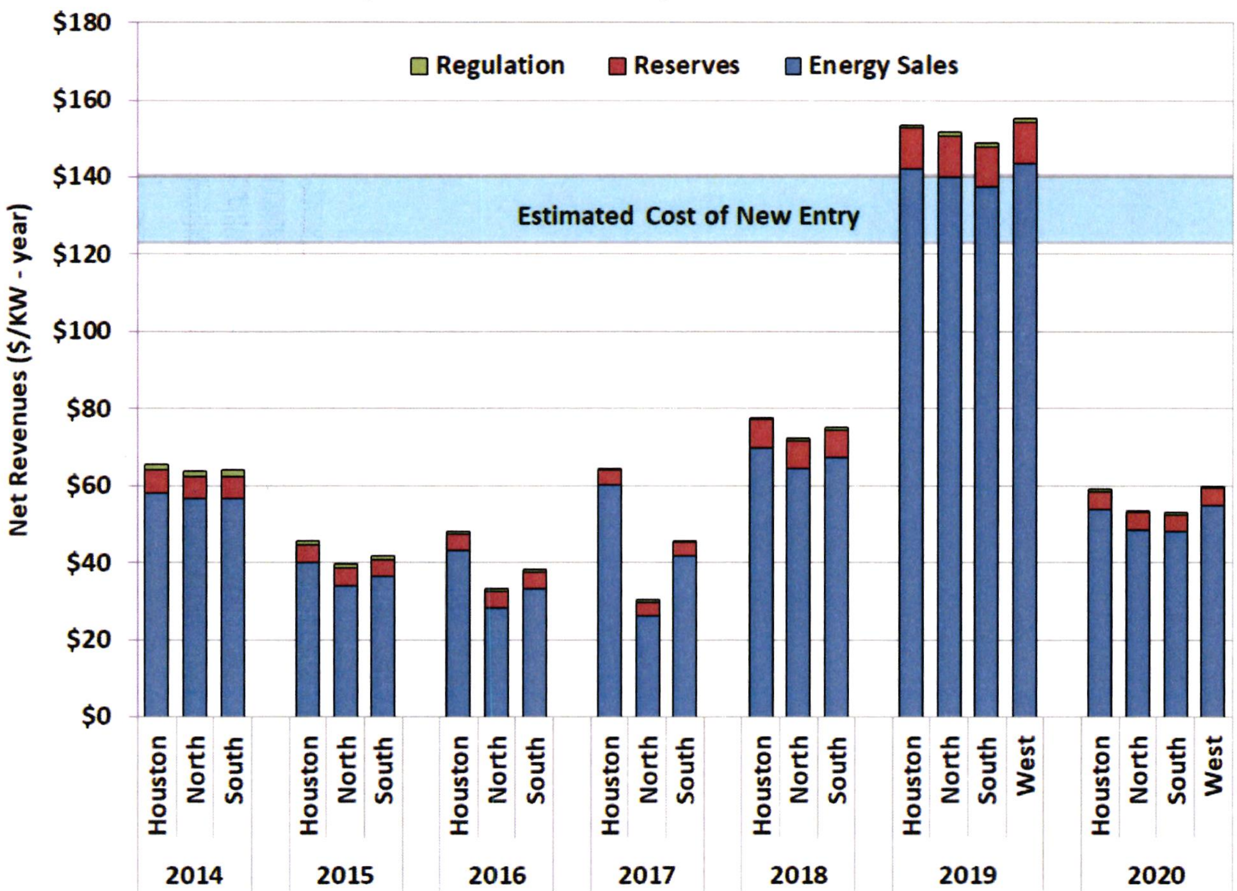
Figure 43: Combustion Turbine Net Revenues



[37] This can mask the effects of unusually high or low prices at a specific generator location.

We calculate net revenues for these units by assuming they will produce energy in any hour for in which it is profitable to do so. We further assume that when they are not producing energy, that both types of units will be available to sell spinning or non-spinning reserves in other hours, and that combined cycle units can provide regulation.[38] The figures also show the estimated CONE for each technology for comparison purposes.

Figure 44: Combined Cycle Net Revenues



In 2020, the estimated CONE values for both types of resources increased, with the CONE values for natural gas combustion turbines ranging from $70 to $117 per kW-year. The ERCOT market did not provide net revenues above the CONE level needed to support new investment in 2020:

- Net revenues for combustion turbines fell to less than $37 per kW-year in the South zone to roughly $41 per kW-year in Houston; while
- Net revenues for combined-cycle units ranged from approximately $48 to $54 per kW-year, depending on the zone.

---

38 For purposes of this analysis, we used the following assumptions: heat rates of 7 MMBtu per MWh for a combined cycle unit, 10.5 MMBtu per MWh for a gas turbine, and $4 per MWh in variable operating and maintenance costs. A total outage rate (planned and forced) of 10% was assumed for each technology.

These sharp decreases in net revenues back to 2018 levels were primarily caused by the absence of significant shortages in 2020, even with the additional adjustment to the ORDC in 2020. The decreases in the frequency of sustained shortages is consistent with the improving reserve margin going into the summer of 2020. In an energy-only market, shortages play a key role in delivering the net revenues an investor needs to recover its investment. Such shortages will tend to be clustered in years with unusually high load or poor generator availability, neither of which were present in 2020.

The figures above also show that average net revenues were highest in the West zone in 2020 as congestion led to higher prices in that zone. Variations in fuel prices were also an important factor in the West zone. Fuel prices are a substantial determinant of net revenues because they are the primary offset from market revenues when calculating net revenues. In 2020, we saw a continuing trend of the separation in natural gas prices between the Waha and Katy locations in the West. Increased drilling activity in the Permian Basin has produced a glut of natural gas and consequently, much lower prices at the Waha location, coupled with the COVID-19 oil demand shock in the spring of 2020. Waha prices dipped below $0 several times throughout 2020 and were much more volatile than prices at Katy.

Because of this lower fuel cost, generators served by the Waha location would have significantly higher net revenues than those procuring gas at Katy. In Section VI of the Appendix, we show the fuel price trends at these locations and the differences in net revenues that they would produce for the two new resources. This analysis shows that the new resources would produce net revenue ranging from $70 to $82 per KW-year at the Waha location, compared to net revenues of $43 to $55 per KW-year at Katy.

B. Net Revenues of Existing Units

Given the continuing effects of low natural gas prices, we evaluate the economic viability of existing coal and nuclear units that have experienced falling net revenues. Non-shortage prices, which have been substantially affected by the prevailing natural gas prices, are the primary determinant of the net revenues received by these baseload units. Low natural gas prices tend to lead to lower system-wide average prices, but it is the prices at these units' specific locations that matter; the prices at these locations have tended to be lower than the ERCOT-wide average prices.

As previously described, the load-weighted ERCOT-wide average energy price in 2020 was $25.73 per MWh. Table 7 shows the output-weighted average price by generation type based on the generator's specific locational price in 2020.

Table 7: Settlement Point Price by Fuel Type

Generation Type	Output-Weighted Price		
	2018	2019	2020
Coal	$33.31	$43.92	$24.84
Combined Cycle	$35.53	$47.06	$24.60
Gas Peakers	$71.64	$126.16	$60.26
Gas Steam	$66.09	$135.16	$41.90
Hydro	$34.40	$42.90	$23.88
Nuclear	$29.00	$35.38	$20.31
Power Storage	$103.19	$154.80	$80.50
Private Network	$34.41	$46.16	$24.08
Renewable	$39.84	$141.09	$35.23
Solar	$35.37	$61.45	$25.49
Wind	$19.26	$20.54	$11.45

Table 7 shows that the prices and associated net revenues were lower at all resources' locations in 2020 than the previous two years. This is again explained by the absence of significant shortage pricing in 2020.

Nuclear Profitability. According to data published by the Nuclear Energy Institute, the total generating cost for nuclear energy in the U.S. was $30.41 per MWh in 2019.[39] The 2019 total generating costs were 7.6% lower than in 2018, and nearly 32% below the 2012 costs. Assuming that operating costs in ERCOT are similar to the U.S. average, and that nuclear operating costs have either continued to be stable or declining, ERCOT's 5 GW of nuclear capacity should have costs less than $31 per MWh. The table above shows an average price for the nuclear units of approximately $20 per MWh making it likely that the nuclear units in ERCOT are not profitable in 2020.

Coal Profitability. The generation-weighted price of all coal and lignite units in ERCOT during 2020 was $24.84 per MWh, a decrease from $43.92 per MWh in 2019. Although specific unit costs may vary, index prices for Powder River Basin coal delivered to ERCOT were approximately $2.55 per MMBtu in 2020, similar to 2019. At these average fuel prices, coal units in ERCOT are likely receiving more than enough revenue to cover operating costs.

Natural Gas-Fired Resource Profitability. Figure 45 shows the net revenues at different locations for a variety of technologies. Because natural gas prices can vary widely, the revenues for natural gas units are shown only for the Houston zone to reflect Katy hub prices and the West zone for Waha. This figure also underscores the effects of the increase in natural gas production in the Permian Basin with insufficient transportation capacity to export the natural gas. This has resulted in low gas prices at the Waha location, and much higher net revenues for these gas resources. New transportation projects have been identified and are currently underway so it is unclear how much longer the large basis difference in natural gas prices will continue.

39 https://www.nei.org/resources/reports-briefs/nuclear-costs-in-context

Figure 45: Net Revenues by Generation Resource Type

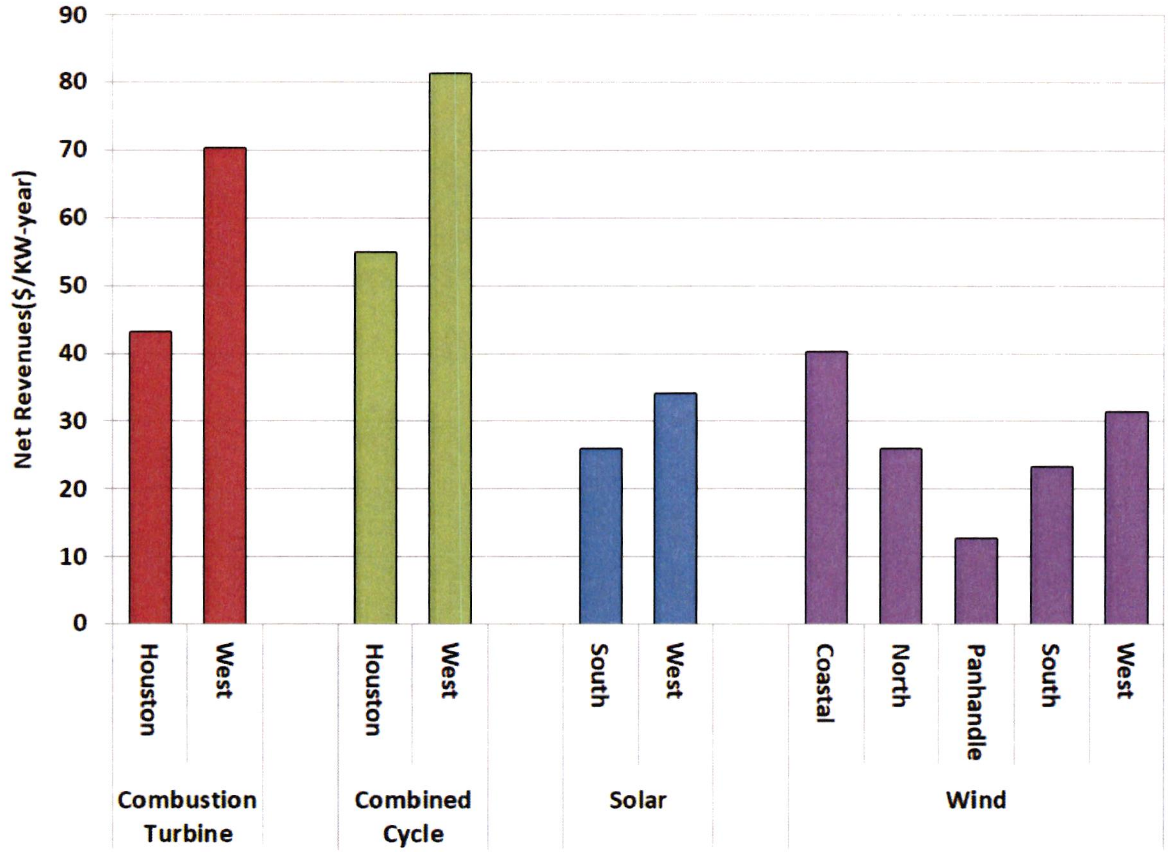


Figure 45 also shows the net revenues for wind and solar generation at multiple locations. As the cost to install wind or solar does not vary much by location, the profitability of those resources is chiefly determined by the available natural resource and the prevailing price to be received. Net revenues for wind and solar were less than gas technologies in 2020 in all areas. This is partly because intermittent technologies cannot maximize its output and associated revenues during shortage conditions. This is particularly true for wind resources that tend to produce less output during hot summer conditions.

Interpreting Single-Year Net Revenues. These results indicate that on a stand-alone basis during 2020, the ERCOT markets did not provide sufficient revenues to support profitable investment in combustion turbine and combined cycle technologies. Net revenues were down as result of lower shortage pricing in 2020 than in 2019. Investors' response to these prices will depend on whether they expect them to reoccur in the future. Additionally, investors may invest instead in new technologies, such as battery energy storage or load-flexible renewables, which have different value propositions from traditional generation. Ultimately, investment decisions are driven by multiple factors:

- Historical net revenue analyses do not provide a view of the forward price expectations that will spur new investment, which can vary widely by supplier. For example, small differences in expectations about the frequency of shortage pricing can greatly influence revenue expectations.

- Bilateral contracts may offer additional revenue because they allow risk-averse buyers to hedge against high shortage pricing.
- Prices and revenues over multiple years may fluctuate in a manner than causes average expected net revenues to be quite different than the net revenues in any one year.
- The CONE for any particular project may be quite different than the generic CONE values we have derived based on average development costs in the Texas market on undeveloped greenfield sites. Companies may have opportunities to build generation at much lower cost than these estimates because of lower cost equipment, access to an existing site, or access to superior financing.

For all these reasons, it is important to be cautious in interpreting single-year net revenues and projecting their long-term effects. Please see Section VI of the Appendix for additional detail and discussion of the net revenue results presented in this subsection.

C. Planning Reserve Margin

Ultimately, the importance of the market signals discussed above is that they facilitate the long-term investment and retirement decisions by market participants that will maintain an adequate resource base. This subsection discusses the trends in the planning reserve margin, which is one measure of the adequacy of the resource base.

Prior to the summer of 2018, there were expectations by many market participants of shortage driven prices in the ERCOT market that mainly went unrealized. Significant shortages were not realized until 2019, due in some part to the impact of the first step of the ORDC change. There are many ways that the market can respond to high prices, all of which result in rising planning reserve margins:

- Building new generation facilities;
- Increasing investment in existing resources, including more maintenance to improve availability, as well as capital investment to increase the capability of the resource;
- Loads investing in systems and procedures to enable non-consumption during shortage pricing events (demand response).

In 2020, there were no such expectations of shortage conditions, and that expectation bore out. There were also circumstances that were unique to 2020, such as the COVID-19 pandemic quarantine and relatively moderate summer weather conditions. Similar to the analysis of net revenues year over year above, it is important to be cautious in interpreting single-year lack of shortage pricing and projecting the long-term based on planning reserves, as shortages can occur in peak net load intervals that may be different than those studied in the planning horizon.

Planning reserves take a more holistic and long-term view of market conditions and may not indicate the frequency of shortage conditions in any given year.

In the December 2019 Capacity, Demand, and Reserves (CDR) report, the 2020 summer reserve margin was projected to be 10.6%, up slightly from 10.5% from the May 2019 CDR report.[40] ERCOT adjusted its peak load forecast to 75,200 MW to account for economic impacts related to COVID-19 and the planning reserve margin for summer 2020 ultimately increased to 12.6% based on the resource updates in the final summer 2020 SARA report.[41] Recent market outcomes and pre-existing investment plans are causing expected increases in the planning margins. Figure 46 shows ERCOT's current projection of planning reserve margins.

Figure 46: Projected Planning Reserve Margins

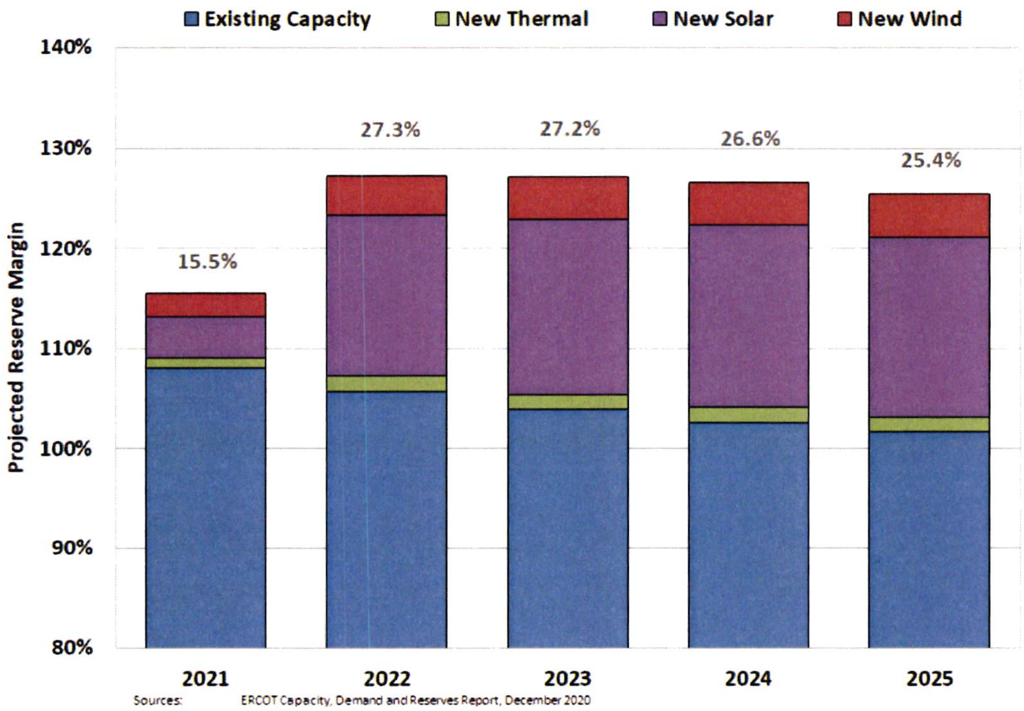


Figure 46 indicates that Texas heads into the summer months of 2021 with an improved reserve margin of 15.5%, higher than the 12.6% reserved margin for 2020. It is worth noting that the current methodology of performing the CDR does not consider power storage resources (e.g., batteries). Including storage resources would increase the reserve margin, potentially by a greater amount than planned thermal generation. Ensuring that the market can efficiently price and dispatch energy from newer technologies will become increasingly important. In addition,

---

40 See Report on the Capacity, Demand and Reserves in the ERCOT Region, 2019-2028 (December 5, 2019), http://www.ercot.com/content/wcm/lists/167023/CapacityDemandandReserveReport-Dec2019.pdf

41 See Seasonal Assessment of Resource Adequacy (SARA) (May 13, 2020), http://www.ercot.com/content/wcm/lists/197378/SARA-FinalSummer2020.pdf.

the CDR relies solely on hour ending 5 p.m. (as the peak hour), when the peak net load hour is likely a more accurate predictor of scarcity conditions, particularly as solar generation continues to be added to the ERCOT system.

The range of planning reserve margins going into 2020 and beyond are consistent with expectations for ERCOT's energy-only market. On December 1, 2020, ERCOT filed a draft report with the Commission titled "*Estimation of the Market Equilibrium and Economically Optimal Reserve Margins for the ERCOT Region for 2024*."[42] The report estimates the Market Equilibrium Reserve Margin (MERM) and Economically Optimal Reserve Margin (EORM) for ERCOT's wholesale electric market with projected system conditions for 2024. ERCOT retained Astrapé Consulting to perform a study, and Astrapé calculated a MERM of 12.25% under projected 2024 market conditions. This was higher than the MERM projection of 10.25% in the 2018 study, however, the projections of system reliability were nearly identical at 0.5 Loss of Load Expectation.

Finally, with growing installed reserve margins for summer of 2020, the retirement of uneconomic generation should be viewed as essential to resource adequacy. Facilitating efficient decisions by generators to retire uneconomic units is nearly as important as facilitating efficient decisions to invest in new resources. With expectations for future natural gas prices to remain low, the economic pressure on coal units in ERCOT is not expected to subside soon. American Electric Power's (AEP) 650 MW Oklaunion coal unit was permanently decommissioned on October 1, 2020, which accounted for 5% of ERCOT's summer coal capacity.

D. Effectiveness of the Shortage Pricing Mechanism

One of the primary goals of an efficient and effective electricity market is to ensure that, over the long term, there is an adequate supply of resources to meet customer demand plus any required installed or planning reserves. Generators earn revenues from three sources: energy prices during non-shortage, energy prices during shortage and capacity payments. Without a long-term capacity market in ERCOT, suppliers' revenues are derived solely from energy prices under shortage and non-shortage conditions. Revenues during non-shortage conditions tend to be more stable as planning margins fluctuate, but shortage revenues are the primary means to provide investment incentives when planning margins fall (or incentives to keep existing units in operation). Therefore, the performance of shortage pricing in the ERCOT market is essential, which we evaluate in this subsection.

---

42 The final version of the report, *Estimation of the Market Equilibrium and Economically Optimal Reserve Margins for the ERCOT Region for 2024*, was published on January 15, 2021; http://www.ercot.com/content/wcm/lists/219844/2020_ERCOT_Reserve_Margin_Study_Report_FINAL_1-15-2021.pdf.

1. Background on Shortage Pricing in ERCOT

Shortage pricing refers to the price escalation that occurs when supply is not sufficient to satisfy all the system's energy and operating reserve requirements. In these cases, prices should reflect the reliability risks borne by the system as the shortage deepens. Ideally, the value of the shortage should be priced based on the loss of load probability at varying levels of operating reserves multiplied by the value of lost load.

Shortage pricing in ERCOT occurs through the ORDC, implemented in 2014 to ensure electricity prices more accurately reflect shortage conditions. The ORDC is described above in Section I: Review of Real-Time Market Outcomes. Over the time it has been in effect, ORDC has had an increasingly material impact on real-time prices, especially in 2019 when reduced installed reserves led to higher expectations of shortage pricing. For a variety of reasons discussed throughout this report, the impact on 2020 real-time prices was more muted.

The ORDC automatically increases the price of power as reserves get tighter. The ORDC adder reflects the Value of Lost Load (VOLL), which was set to $9,000 per MWh in June 2014. The real-time prices determined by Security Constraint Economic Dispatch (SCED) are increased by the Real-Time Reserve Price, which is determined based on the value of the remaining reserves in the system as specified by the predefined ORDC.

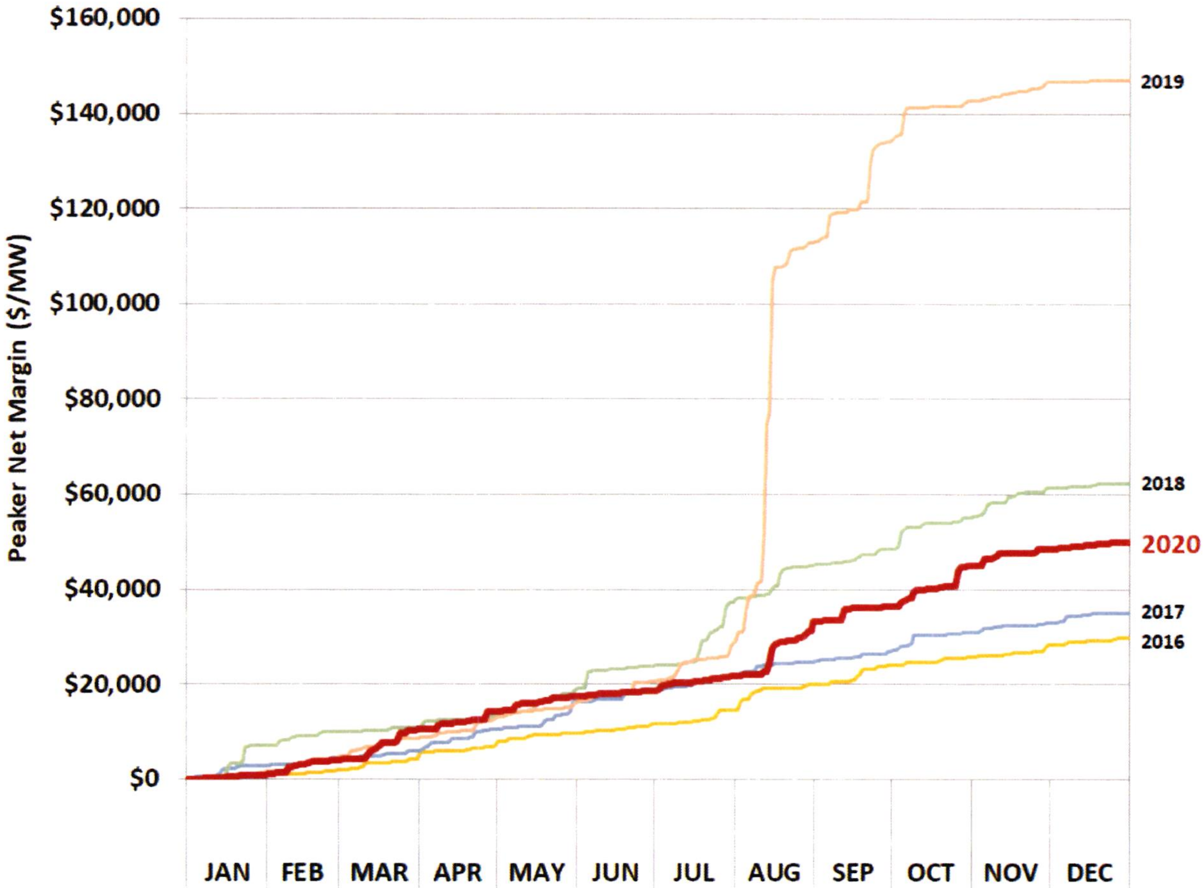
The Scarcity Pricing Mechanism includes a provision termed the Peaker Net Margin (PNM) that is designed to provide a pricing "fail-safe" measure. If the PNM is exceeded, the system-wide offer cap is reduced. PNM also serves as a simplified measure of the annual net revenue of a hypothetical peaking unit.[43] Section I contains several summaries and discussions of the shortage pricing that occurred in 2020. The next section, however, reviews pricing in 2020 showing the PNM in 2020 compared to prior years.

2. Peaker Net Margin in 2020

Figure 47 shows the cumulative PNM results for each year since the creation of the Scarcity Pricing Mechanism. This figure shows that PNM in 2020 was middling, higher than both 2016 and 2017 but far below the high of 2019. PNM was initially defined to provide a "circuit breaker" trigger for lowering the system-wide offer cap. However, as of the end of 2020, PNM had not approached levels that would dictate a reduction in the system wide offer cap, even after 2019, when it reached the highest level to date. The PNM outcomes in 2020, significantly lower than 2019, only reinforce that position.

[43] The proxy combustion turbine in the Peaker Net Margin calculation assumes a heat rate of 10 MMBtu per MWh and includes no other variable operating costs or startup costs.

Figure 47: Peaker Net Margin



3. Changes to the ORDC

The Commission directed a significant change to the ORDC in 2019. The Commission considered proposals modifying various defining aspects of the ORDC, including shifting the LOLP portion of the curve.[44] The LOLP portion of the curves used to determine the ORDC price adder has typically been constructed using normal probability distributions defined by two factors: a) the average of historical differences between expected and actual operating reserves ("MU"), and b) the standard deviation in those values ("SIGMA").[45] On January 17, 2019, the Commission approved a two-part process to modify the ORDC by implementing a .25 standard deviation shift in the LOLP calculation and transitioning to a single blended ORDC curve, and a second step of .25 in the spring of 2020. The second step of the ORDC change was implemented on March 1, 2020 and we have estimated the effects in 2020. These results are shown below in Table 6.

---

44 See PUCT Project No. 48551, *Review of Summer 2018 ERCOT Market Performance.*

45 MU and Sigma are separately calculated for each of the twenty-four curves currently used (six time of day blocks and four seasons).

Table 8: Effect of ORDC Shift on Price

	Average RT price $ per MWh	ORDC contribution $ per MWh	ORDC Price increase $ per MWh	Percent increase %	Total RT Market Cost $ in Millions	RT Market Cost Increase $ in Millions
January	17.82	0.02	0.01	0.04	516	0
February	25.28	0.24	0.11	0.45	706	3
March	31.14	1.21	0.52	1.66	874	15
April	21.01	2.84	1.18	5.64	564	32
May	20.73	1.22	0.53	2.56	645	17
June	16.13	0.07	0.03	0.21	569	1
July	21.65	1.45	0.68	3.15	869	27
August	43.13	13.64	5.17	11.98	1,751	210
September	23.23	1.96	0.79	3.39	759	26
October	34.18	4.56	2.03	5.92	1,065	63
November	26.65	0.63	0.28	1.06	725	8
December	21.32	0.08	0.04	0.18	645	1
Total	25.48	2.62	1.06	4.15	9,688	402

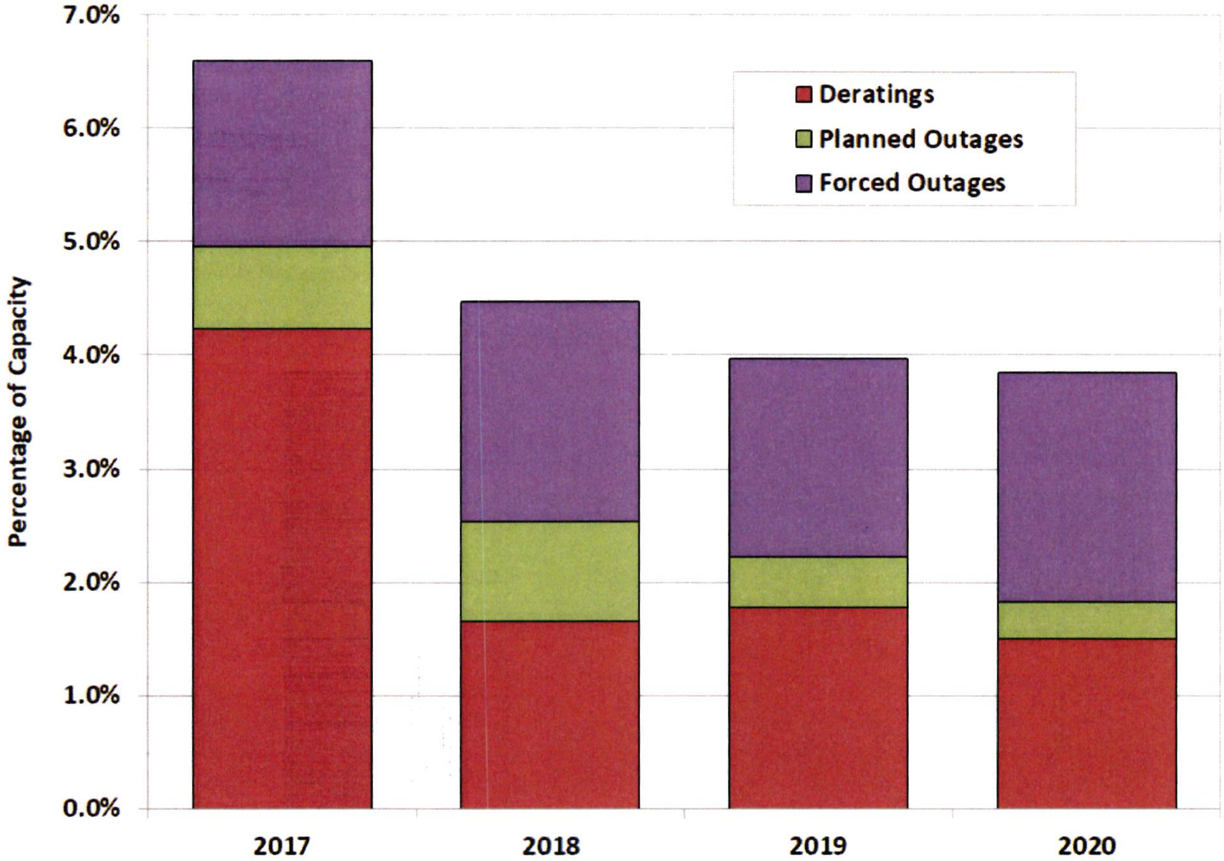
Table 8 above shows that the 2020 second step ORDC change increased the effects of shortage pricing by an estimated 4% -- increasing the total impact of the ORDC on average prices by $1 per MWh. This led to increased market costs and revenues to generators of roughly $400 million in 2020. Because planning reserve margins rose as projected in 2020, shortage pricing fell well short of 2019 levels. The first step of the ORDC change, particularly blending the curves, was a significant driver of the higher impact realized in 2019. No further changes to the ORDC are scheduled prior to the implementation of real-time co-optimization, when the ORDC will be retired.

4. Short-Term Effects of Shortage Pricing in 2020

In addition to the long-term incentives that shortage pricing creates to facilitate investment and retirement decisions, it also creates important short-term incentives. For example, it creates a strong incentive for generators to be available at the times when they are expected to be needed most. Figure 48 shows the level of outages and deratings that have occurred during summer peak conditions over the past four years.

This figure shows that as expectations of shortages remain strong, outages have decreased substantially, even as the planning reserve margin rebounded in 2020. Most of these outage reductions were in planned outages and deratings, the class for which the suppliers have the most control. These results demonstrate that the suppliers in ERCOT respond to price signals and associated incentives.

Figure 48: Summer Month Outage Percentages



E. Reliability Must Run and Must Run Alternatives

Reliability-Must-Run procedures are essential for determining and addressing the need for generation units to support grid reliability.[46] A Reliability Must Run (RMR) Unit is a resource operated under the terms of an agreement with ERCOT that would not otherwise be operated except that it is necessary to provide voltage support, stability or management of localized transmission constraints under credible single contingency criteria where market solutions do not exist. If ERCOT determines a resource is needed to maintain electric stability, it can enter into an RMR agreement to pay the plant an "out-of-market" payment to continue operating. ERCOT also has a process to consider other resources, known as Must-Run Alternatives (MRA). In lieu of paying an uneconomic to stay open to ensure grid reliability, ERCOT may issue a Request for Proposals for alternative solutions that can address the specific reliability concern.

[46] http://www.ercot.com/content/wcm/lists/89476/OnePager_RMR_May2016_FINAL.pdf