Harvard Electricity Policy Group

Session:
Market Reforms for Stressed Conditions

Scarcity Pricing in ERCOT

Sai Moorty
Principal, Market Design and Analysis

June 14, 2019
Where We Are Today

- Energy-only market
- System-Wide Offer Cap (SWCAP) is $9,000/MWh
- Value of Lost Load (VOLL) is $9,000/MWh
- Demand peaks in summer, approximately half of summer peak demand is from residential air conditioning
- Natural gas is the marginal fuel
- Scarcity pricing mechanism based on Operational Reserve Demand Curve (ORDC) and Power Balance Penalty Curve – system-wide
- ERCOT does NOT have a mandated Planning Reserve Margin
  - Periodically conducts studies to estimate Market Equilibrium Planning Reserve Margin (MERM) and Economically Optimum Planning Reserve Margin (EORM). Recent Brattle/Astrape estimates for 2022 are:
    - MERM = 10.25%
    - EORM = 9.0%
Where We Are Today Cont.

• Continued trends of retirements of thermal (‘dispatchable’) Generation Resources

• Persistent low average energy prices
  – Summer weather closer to historical norms, unlike 2011 extremes
  – No active demand response participation in the Real-Time Market
    • There has been growth in retail (‘passive’) price responsive demand, but true capability unknown due to lack of sufficient duration of high prices
  – Few scarcity pricing intervals
    • Less than 12 minutes of prices at $9,000/MWh since 2015 – primarily due to transient ramping issues
    • Last Energy Emergency Alert was in January 2014
Recent Events

- In May 2017, Dr. Hogan and Dr. Pope submitted paper to PUCT: “Priorities for the Evolution of an Energy-only Electricity Market Design in ERCOT”

  - Three recommendations:
    - System-wide price formation improvement
    - Adjust ORDC parameters and determination of available operating reserves: better represent renewable intermittency risk and renewable tax subsidy
    - Add marginal cost of losses to Day-Ahead and Real-Time Markets

  - Locational scarcity pricing improvement
    - Include locational pricing impact of reliability commitment on transmission constraint
    - Revise RTM mitigation of ERCOT-committed Resources
    - Locational reserves with energy and AS co-optimization in RTM

  - Revise transmission planning criteria and cost recovery
Recent Events

• Removed Reliability Unit Commitment (RUC) and Reliability Must-Run (RMR) capacity from available Real-Time reserves (2018)

• In January 2019, PUCT directed ERCOT to:
  – Make changes to the value of mean used in ORDC but did not make changes to:
    • Minimum contingency reserve value (X = 2,000MW)
    • Value of Lost Load (VOLL)
  – Implement Real-Time and Ancillary Service (AS) co-optimization

• Policy issues being discussed at the PUCT:
  – Setup of the individual AS product demand curves
  – Coordination of the values of SWOC, VOLL, Power Balance Penalty and the maximum value on the AS demand curve
  – Should Virtual AS Offers be allowed in DAM
  – Rules governing the offering of AS in the Real-Time Market

• ERCOT Board of Directors approved major revisions to the AS product set (Nodal Protocol Revision Request 863, approved in February 2019)
Recent Events: Changes to ORDC Parameters

PUCT expressed concern over the declining Planning Reserve Margin and issued a directive in January 2019 to adjust the LOLP parameters (Mean and Standard Deviation) used for ORDC.

- Use single blended statistics for the Mean and Standard Deviation instead of 4-hour blocks by season (impact is similar to shifting the mean by 0.25 Standard Deviation)
- Shift the mean by 0.5 Standard Deviation (implement in two steps of 0.25 Standard Deviation, once in 2019 and then in 2020)
Recent Events – Major Revision to the AS Product Set

- Inertia is the major consideration in determining RRS quantities required

**CURRENT FRAMEWORK**

<table>
<thead>
<tr>
<th>REGULATION</th>
<th>157 to 687 MW*</th>
</tr>
</thead>
</table>

**RESPONSIVE RESERVE SERVICE**

- 2,300 to 3,200 MW*  
  1. Primary Frequency Response  
  2. Load Resources on Under Frequency Relay (UFR)

**RESPONSIVE RESERVE SERVICE (RRS)**

| Fast Frequency Response (FFR) |  
| Load Resources on UFR |  
| Primary Frequency Response (PFR) |  
| 2,300 to 3,200 MW |

**ERCOT CONTINGENCY RESERVE SERVICE (ECRS)**

| 10 minute ramp |  
| Load Resources may or may not be on UFR |  
| 508 to 1,644 MW** |

**NON-SPIN**

| 967 to 2,361 MW* |

**REGULATION**

| 157 to 687 MW |

**NPRR 863**

| No Change |

**FFR**

- Triggered at 59.85 Hz and full response in 15 cycles  
- Once deployed, sustain for up to 15 mins. Once recalled, restore within 15 mins

**PFR**

- PFR capable capacity reserved on generators or Controllable Load Resources (CLR)  
- Minimum 1,150 MW must be provided by resources capable of PFR

**Load Resources on UFR**

- Triggered at 59.70 Hz and full response in 30 cycles  
- Sustain until recalled. Once recalled, restore within 3 hours  
- Beyond the minimum PFR, up to 60% of total RRS can come from Load Resources on UFR or FFR.

**Generation**

- Online or offline capacity that can be converted to energy within 10 minutes  
- Dispatched by SCED

**Load Resources (UFR not required)**

- Up to 50% of ECRS capacity can come from Load Resources with or without UFR  
- Once deployed, must respond within 10 minutes. Restoration within 3 hours

**Non-protocol changes.**

- Proposed methodology for Non-Spin Reserve Service quantities in this framework - quantities computed using 2018 A/S Methodology are reduced by ECRS quantities.

**Overall A/S: 3,807 to 5,958 MW*”

*Quantities computed/estimated using 2018 Ancillary Service Methodology. **Quantities estimated using this reference. ***Quantities estimated using this reference and method in box on far left. For Discussion Purposes Only. The intent of this slide is to represent NPRR 863 (with STEC comments from 10/1/2018). Protocol language prevails to the extent of any inconsistency with this one page summary.
Real-Time Energy and AS Co-optimization

ERCOT staff proposal for the individual AS product demand curves are derived from the disaggregation of the composite LOLP ORDC

- Implementation of energy and AS co-optimization will replace the current ORDC price adders

- To allow the prices to rise to VOLL when reserves are at or below the Minimum Contingency Reserve (X = 2,000 MW):
  - Power Balance Penalty = $11,000/MWh
  - SWCAP = $2,000/MWh
  - VOLL = $9,000/MWh
Other RTM Pricing Issues in Stakeholder Discussions

- Mitigation of RUC Resources
  - Incremental energy offers of RUC Resources input to SCED
    - Incorporate startup and minimum energy costs (similar to ELMP)
    - Set to values beyond market offers or just below transmission constraint shadow price cap (taking into account shift factor of Resources)
    - Combination of the above

- Mitigation of RMR Resources also under discussion

- Locational price impact of minimum output limit of RUC/RMR Resources
  - Tighten transmission limits by the amount of counterflow provided by RUC/RMR Resources at their minimum output limit
    - Modify above proposal to use a lower transmission penalty price/curve for the portion corresponding to change in transmission limit
      - Used in dispatch run – no uplift
      - Mitigate scenarios with oversold DAM congestion hedges leading to uplift
      - Mitigate extreme pricing outcomes
Future Outlook – Evolving Scarcity Pricing Mechanism

Scarcity pricing mechanism will continue to evolve

• Low prices most of the time with brief periods of scarcity pricing – “binary pricing?”

• Expected changes in generation mix likely to reduce time of scarcity pricing
  – Renewable (wind and solar)
  – Battery storage Resources
    • 3,500 MW of battery storage Resources in ERCOT’s interconnection queue
    • Could reduce transient price spikes due to their fast response
  – Distributed Energy Resources (DER)
    • ERCOT is seeing increasing amounts of fast-acting passive price responsive (thermal) DERs
      – Most are co-located with load to provide additional benefit of demand charge reduction
      – Could reduce transient price spikes due to their fast response
    • Small-scale solar increasing but not near penetration levels seen in other areas
Active Demand Side Participation in Price Formation – How?

• Passive price responsive demand response is growing
  – Small customer response enabled by new technologies (IOT),
    communication via social media
  – Many industrial customers on indexed pricing
  – Incentive (for industrials, municipalities and co-ops) is primarily demand
    charge reduction and no associated external compliance requirement

• Enabling active demand side resources in price formation would
  represent the holy grail of electricity markets
  – Why has it not taken off in ERCOT?
    • Low price with very brief periods of scarcity pricing – lack of incentive
    • Current rules for active participation have strong compliance metrics
    • ERCOT does not follow FERC order 745
      – Only benefit is avoided consumption charge – no payment for
        curtailment
    • No capacity market
Other Ideas?

• Review firm load shed procedures?
  – Firm load shed almost exclusively targets residential customers
  – With almost full deployment of smart meters with remote disconnect/reconnect feature, would it be possible to move to a more surgical load shed at premise level?
    • Firm load shed events in 1989, 2006, 2011
  – Can firm load shed procedure move from a reliability action to a market based one?
    • Customers can choose a rate with a value for reliability?

• Ancillary Service for other services we currently get for free from dispatchable resources?
  – Inertia
  – Availability
Appendix
Wind and solar values are based on nameplate capacity (not adjusted for peak capacity contribution).
Private Use Network capacity not included.
Trend of declining of “dispatchable” generation.
Implied prices at 9000$/MWh for August 2019

- 22 on-peak days in August 2019
- 120 $/MWh August 2019 on-peak forward price cost = $42,240 = 120 * 22 * 16

- If “normal” RT price during August 2019 on-peak averages 60$/MWh
  - 349.64 hours @ 60 $/MWh
  - 2.36 hours @ 9,000 $/MWh
Summer Reserve Margins (Prior Year May CDR Reports)

ERCOT Summer Reserve Margins

Planning Reserve Margin (%)


ERCOT Public
Peaker Net Margin (PNM)

- PNM is a calculation designed to measure the annual net revenue of a hypothetical peaking unit.
- If the PNM for a year reaches a cumulative total of $315,000 (3xCONE), the System-Wide Offer Cap is then reduced to the higher of $2,000 per MWh or 50 times the daily natural gas price index.
All-In Energy Price

Average All-In Energy Price

Historical Load Weighted Monthly Price

Month and Year

- Average FIP
- Uplift Price
- Reliability Adder
- Energy Price
- AS Price
- Online/Offline Adders
- Total Price

Electricity Price ($/MWh)

Natural Gas Price ($/MMBtu)
Real-Time Energy Prices 2016-2018

Prices exceeded $75/MWh in 1.6% of intervals over 3 years

<table>
<thead>
<tr>
<th>Intervals</th>
<th>2016</th>
<th>2017</th>
<th>2018</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;$0</td>
<td>569</td>
<td>235</td>
<td>160</td>
</tr>
<tr>
<td>$0-$30</td>
<td>31,562</td>
<td>30,568</td>
<td>27,446</td>
</tr>
<tr>
<td>&gt;$30-$75</td>
<td>2,605</td>
<td>3,685</td>
<td>6,691</td>
</tr>
<tr>
<td>&gt;$75</td>
<td>400</td>
<td>552</td>
<td>743</td>
</tr>
</tbody>
</table>

HUB Avg. 15-min. Settlement Point Prices

- 2016 SPP
- 2017 SPP
- 2018 SPP
Revenue Neutrality Allocation to Load

Key factors influencing high RENA
- Outage difference between DAM and RT
- Special PTP (settled as options)
- Oversold PTP MW in DAM relative to Real-Time
- Load Distribution Factors used in DAM
- Remedial Action Scheme modeling in DAM

RENA is a measure of DAM->RTM performance
Overview of Today’s Real-Time Market (RTM)

- Security Constrained Economic Dispatch (SCED) – no commitment
  - Single, 5-minute interval optimization horizon
  - Marginal cost of transmission losses not considered
  - Optimizes energy dispatch only (SCED does **not** co-optimize energy and Ancillary Services, nor does it commit Resources)
    - Day-Ahead Ancillary Service (AS) capacity procured is protected in SCED
- Treatment of Quick Start Generation Resources (QSGR)
  - Qualified QSGRs are 10-minute start Resources
  - Can submit special status while physically off-line for SCED to treat them as on-line for dispatch
    - Incremental Energy Offers can incorporate startup and minimum energy costs
- Active price formation by demand side participation (including aggregated demand) enabled but no participation to date
  - Passive price responsive demand is growing
- Dispatch run produces binding dispatch
Overview of Today’s Real-Time Market (RTM) Cont.

- Certain out-of-market actions trigger a pricing run which can produce a system-wide price adder (Reliability Deployment Price Adder) that is added to all the locational prices (LMP) from the dispatch run.

- Scarcity pricing based on Operating Reserve Demand Curve (ORDC) and Power Balance Penalty Curve (PBPC).
  - Implementation is a “poor man’s co-optimization” of energy and operating reserves.
  - Post-processing step after dispatch run where operating reserve value is determined and two price adders determined:
    - Off-Line Operating Reserve Price Adder: This adder is paid to qualified off-line Resources (cold start time is 30 minutes or less).
    - On-Line Operating Reserve Price Adder: This price adder is added to all the locational prices (LMP) from the dispatch run and is used to pay the unloaded on-line capacity of all Resources.
Background on Real-Time Price Adder Based on Operating Reserve Demand Curve

- Off-Line Operating Reserve Price Adder \( (P_{ns}) \)
  \[
P_{ns} = 0.5 \times (VOLL - \lambda) \times LOLP_{sns}(\mu, \sigma, R_{sns} - 2000)
\]

- On-Line Operating Reserve Price Adder \( (P_{s}) \)
  \[
P_{s} = P_{ns} + 0.5 \times (VOLL - \lambda) \times LOLP_{s}(0.5\mu, 0.707\sigma, R_{s} - 2000)
\]

\( \lambda \) is the Power Balance Shadow Price