Comments on Price Formation Enhancements
Scarcity Pricing Working Group Sessions

Department of Market Monitoring

February 27, 2025

Summary

The Department of Market Monitoring (DMM) appreciates the opportunity to comment on the ISO's *Price Formation Enhancements Working Group Sessions.*[1] The scarcity pricing working group meetings have been aimed at (1) exploring market mechanisms that allow prices to gradually rise as supply shortage risk grows, (2) considering potential enhancements to the ISO's real-time ancillary services procurement in the context of scarcity pricing, and (3) exploring approaches to better align pricing run penalty prices with market conditions and scarcity value of reserves.

As the ISO considers potential options to address these issues, DMM continues to suggest that the ISO place priority on two foundational enhancements to price formation:

(1) extending the 15-minute uncertainty time horizon of the flexible ramping product (FRP), or creating a new uncertainty reserve product that serves a similar purpose

(2) full re-optimizing of ancillary services in the real-time market

These two enhancements would allow the real-time market to better reflect real-time conditions and provide earlier price signals prior to a scarcity event.

Based on discussions in the working group, DMM believes creation of a new hour-ahead uncertainty product is worth serious consideration and may be the most practical path forward for implementing the first of these price formation enhancements. As noted in working groups, extending the uncertainty horizon of the current FRP may involve significant complexity. A simpler hour-ahead uncertainty product may be much more compatible with the current hour-ahead scheduling timeframe of the broader Western Energy Imbalance Market (WEIM).

This type of hour-ahead product would allow the costs of protecting against uncertainty to be reflected in real-time market prices, and reduce the potential for scarcity before it occurs. This product could also be designed to replace the need for the large load bias by ISO operators in the hour-ahead and 15-minute markets to account for uncertainty and create additional ramping capacity. DMM looks forward to further discussion of this possibility in future working group meetings.

With respect to scarcity pricing, DMM does not believe that anchoring penalty prices to the value of lost load would necessarily provide more accurate price signals to the market during

[1] *Price Formation Enhancements – Scarcity Pricing Working Group Sessions*, California ISO, December 16, 2024 - February 6, 2025: https://stakeholdercenter.caiso.com/StakeholderInitiatives/Price-formation-enhancements

times of potential scarcity. Using an estimated value of lost load could significantly inflate penalty prices to values several times greater than the current market bid cap, without a sound theoretical reason for doing so. DMM recommends the ISO first address more fundamental price formation issues before attempting to establish any higher value for prices during scarcity conditions.

Comments

Extending the flexible ramping product uncertainty time horizon (or creating a new product to serve a similar role) would allow earlier and more gradual price signals of upcoming scarcity

At the January 22 working group meeting, DMM presented its longstanding recommendation to extend the uncertainty time horizon of the flexible ramping product (FRP) to account for uncertainty over longer time horizons.[2] The FRP is designed to address net load forecast uncertainty between the 15-minute and 5-minute markets. However, the time horizon for which this uncertainty is considered is a single 15-minute interval into the future. In real time, grid operators face significant net load uncertainty over longer timeframes (e.g., 30, 60, and 120 minutes from the current market interval), and the range of uncertainty increases over those longer time horizons.

As summarized in DMM's presentation, the 15-minute uncertainty used by FRP is substantially less than what actual net load forecast uncertainty may be one to four hours in the future. Therefore, the real-time market software does not optimally position the resource fleet to meet potential high net load outcomes in these future time horizons. This prevents price impacts from potential tightening supply conditions from impacting prices in the binding interval, and contributes to operators needing to enter large load adjustments that can impact both the CAISO BAA and the entire WEIM.

DMM continues to suggest that extending the FRP uncertainty horizon, or creating a new uncertainty product to serve the same purpose, can provide three key benefits:

(1) Allow the optimization to better position resources to consider upcoming scarcity in further out market intervals.

(2) Improve flexible capacity and energy price formation ahead of a scarcity event by considering a longer time horizon for uncertainty.

(3) Procure capacity to meet net load uncertainty over longer time horizons in the market, reducing the need for operator interventions.

[2] *Recommendation to increase the FRP uncertainty horizon,* Department of Market Monitoring, January 22, 2025: https://stakeholdercenter.caiso.com/InitiativeDocuments/Presentation-Department-of-Market-Monitoring-Flexibility-Ramping-Product-Jan-22-2025.pdf

The examples provided in DMM's January 22 presentation were constructed within the context of Short-Term Unit Commitment (STUC) and a 270 minute forward-looking timeframe. However, DMM does not suggest that the FRP horizon must be extended to this specific timeframe in order for improvement or benefit to take place. Extending the FRP uncertainty time horizon to even one hour, or creating a new hour-ahead uncertainty product, could improve the current real-time market design.

A one-hour time horizon for uncertainty could align with the timeline for submission of base schedules by WEIM entities, allowing such schedules to be considered when calculating future interval net load uncertainty. This type of product would allow costs of protecting against uncertainty to be reflected in real-time market process, and could essentially replace the need for the large load bias that is currently used by ISO operators in the hour-ahead and 15-minute markets to account for uncertainty and create additional ramping capacity.[3]

Given the need for an approach that is compatible with the broader WEIM and the potential complexity added to the existing FRP by extending the uncertainty horizon, DMM believes creation of a new hour-ahead uncertainty product is worth serious consideration and may be the most practical path forward. DMM looks forward to further discussion of this possibility in future working group meetings.

DMM supports enhancements to ancillary service procurement to improve overall price formation

The scarcity pricing working group discussions contemplated changes to ancillary services procurement as one approach to implement scarcity pricing. These potential changes included elimination of cascading ancillary services procurement, full real-time re-optimization of ancillary services procured in the day-ahead market, and modification of the existing ancillary services pricing mechanism to establish a new operating reserve demand curve (ORDC) that would better inform scarcity pricing during tight system conditions.

DMM agrees that re-optimizing ancillary services in the real-time market would be beneficial. The real-time market only procures ancillary services incremental to day-ahead ancillary services awards, which may not fully capture the extent of scarcity in real-time. A full ancillary services re-optimization in real-time could increase efficiency and allow real-time energy prices to better reflect real-time ancillary services conditions. This could be especially helpful when reserve capacity is scarce in tight real-time conditions.

DMM also generally agrees that ancillary service pricing could be enhanced to establish a more robust ORDC. However, reliance on ancillary service pricing mechanisms to establish scarcity pricing has two significant limitations: they only apply to the 15-minute real-time market, and

[3] Initial WEIM base schedules are due 75 minutes before the trading hour (T-75). While base schedules may be revised slightly at the T-55 and T-40 timelines, an hour-ahead lookout should capture the bulk of base scheduling activity.

they would not apply to the broader WEIM footprint since ancillary services are only procured by the market for the CAISO BAA.

Elimination of cascading ancillary service procurement would likely increase prices of both energy and ancillary services during extreme system conditions

In the absence of cascading ancillary service procurement, the potential increase in prices during tight conditions would result from the reduced fungibility of ancillary service products, and the increased extent to which there is a tradeoff with energy to procure additional quantities of each product. Further, elimination of cascading ancillary service procurement may lead to more ancillary services being held as contingency-only reserves. When such reserves are released as energy during tight supply conditions, they would be released at the bid cap, thus acting as a type of scarcity pricing.

The extent to which higher prices from eliminating cascading ancillary service procurement is an efficient and desired outcome depends on the true fungibility of different ancillary service products. If from an operational perspective there is an acceptable degree of substitution between some ancillary service products, eliminating cascading procurement may lead to artificial constraints and unnecessary price inflation. Cascading ancillary service procurement should not be eliminated solely for the purpose of exaggerating potential ancillary service shortfalls to drive higher prices.

Ancillary service procurement enhancements have implications beyond scarcity pricing

DMM views enhancements to ancillary service procurement – particularly the real-time re-optimization of ancillary services – as a foundational price formation enhancement that extends beyond its implications for scarcity pricing. DMM views this important enhancement as being appropriately in scope for the current scarcity pricing initiative. However, DMM would not oppose ancillary service topics being moved to a separate, more focused initiative, if the ISO chooses to implement a scarcity pricing approach that does not directly involve ancillary services. DMM's support of such a move would also be conditional on this topic continuing to receive the same level of serious consideration as it has in the scarcity pricing context.

Value of lost load (VOLL) estimates will inflate scarcity prices, but may still have weak theoretical underpinnings and may be difficult to apply uniformly in all areas

At the January 22, 2025 scarcity pricing working group meeting, the ISO presented the concept of value of lost load (VOLL) as a potential anchor for pricing run penalty prices during a supply shortfall or reliability event.[4] Maximum energy prices are currently anchored to the market bid cap. The ISO suggested that VOLL-based penalty prices could better reflect the true economic

[4] *Price Formation Enhancements – Scarcity Pricing: Anchoring Penalty Prices to the Value of Lost Load*, California ISO, January 22, 2025: https://stakeholdercenter.caiso.com/InitiativeDocuments/Presentation-Price-Formation-Enhancements-Session-2-Jan-22-2025.pdf

cost of load shedding, and improve market price signals and system reliability. Implementing VOLL-based penalty prices would anchor the operating reserve demand curve (ORDC) to the expected value of lost load (EVLL), which is the product of an estimated VOLL value and estimated loss of load probability (LOLP).[5]

The appeal of using a VOLL estimate to establish prices under extreme system conditions is that this value is thought to be a more accurate estimation of load's willingness to pay to avoid curtailment. However, there are many possible approaches and assumptions when choosing to establish prices based on a VOLL estimate. Ultimately, the use of a VOLL approach to pricing can lead to much higher prices under extreme conditions that may still lack a strong theoretical underpinning, and may not attract any additional supply in true scarcity situations.

DMM is concerned with the idea of applying static VOLL and LOLP values to establish penalty prices for all loads across all regions. These two values can vary significantly across different customer classes, regions, and different points in time.

VOLL estimates are typically developed from preference studies that estimate the economic cost of outages for different customer classes and outage durations. The VOLL is then administratively set at either a load-weighted average for a specific outage duration, or at a level focused primarily on a specific customer class (such as residential customers). In either approach, the VOLL is left over-valuing and under-valuing the real VOLL for certain loads.

The true range of VOLL across the system may be very large, with demand distributed at both extremes. Load associated with the lowest willingness to pay should be the first to be curtailed, and may be sufficient in quantity to resolve reliability issues. However, penalty prices are applied non-discriminatorily during extreme system conditions, meaning there is no way to apply a class-based VOLL to class-specific load shed. For example, a residential-based VOLL penalty price would apply to *all* load shed (not just residential load), which would inaccurately quantify the real economic cost of the outage. Even within customer classes, the true VOLL may vary significantly across regions, adding further complication to establishing accurate values across the entire real-time market footprint.

The LOLP is the probability of system load exceeding the available generating capacity during a given time period. Because the LOLP is constantly changing with market conditions, there is no way to calculate a single standardized LOLP value that accurately represents real-time conditions in every hour of every day. In theory, to account for real-time system conditions, this means that the system LOLP would need to be re-calculated relatively frequently (such as every hour) in order for it to remain accurate.

[5] EVLL = VOLL * LOLP. Page 10, *Price Formation Enhancements – Scarcity Pricing: Anchoring Penalty Prices to the Value of Lost Load*, California ISO, January 22, 2025: https://stakeholdercenter.caiso.com/InitiativeDocuments/Presentation-Price-Formation-Enhancements-Session-2-Jan-22-2025.pdf

Additionally, without extensive hourly data from all WEIM BAAs (for both participating and non-participating resources), there is no way to accurately calculate a LOLP value for the regional market footprint during all hours of the day. While LOLP could be estimated under different sets of defined system conditions, such estimates are likely to be imprecise and inaccurate representations of the true LOLP at a given point in time.

Therefore, DMM does not believe VOLL-based penalty prices are likely to provide more accurate price signals to the market. DMM is concerned that VOLL-based penalty prices could significantly inflate penalty prices to values potentially several times greater than the current market bid cap, without a sound theoretical underpinning to do so.[6]

[6] Gorman and Callaway (2024) found an average VOLL estimate of $10- 14/kWh ($10,000-$14,000/MWh) for California residential customers. Page 15, *Price Formation Enhancements – Scarcity Pricing: Anchoring Penalty Prices to the Value of Lost Load*, California ISO, January 22, 2025: https://stakeholdercenter.caiso.com/InitiativeDocuments/Presentation-Price-Formation-Enhancements-Session-2-Jan-22-2025.pdf