Efficient Market Prices During Tight Supply Conditions:

Scarcity Pricing Market Design

Issues and Recommendations

Efficient Market Prices During Tight Supply Conditions

Contents

Efficient Market Prices During Tight Supply Conditions

1 Executive Summary

Accurate market prices are essential to achieving reliable operations, an economically efficient energy dispatch, and open and non-discriminatory access to the transmission system. Inefficient market prices cause operators to intervene in markets to maintain reliability, incentivize irrational conduct by both consumers and producers, increase administrative uplifts, and therefore may require departures from open access or non-discriminatory access to transmission grid.

The accuracy of the real-time price signal, appropriately high or low, is essential to efficient market design. Accurately high prices during tight supply conditions escalate the consequences for non-performing supply, induce additional resource commitment, and appropriately encourage load-serving entities to both forward-contract and accurately schedule demand. In addition, real-time pricing that corresponds closely to physical system conditions reduces the need to use non-market mechanisms to guarantee that the system operator has enough supply at the right time and location.[1]

Over the past decade, California has experienced rapid changes to the characteristics and composition of its energy supply as well as the ways in which demand is organized and participates in the energy market. This transition, as driven by the State's interest in lowering carbon emissions, has placed considerable strain on both the routine and extraordinary operation of the California Independent System Operator's (CAISO) system and fundamentally changed the operational characteristics that the CAISO values the most. With the transition comes a large and more diverse set of energy purchasers and sellers to the market, often with new views on risk and the value of energy. This evolution will continue for the foreseeable future and place a premium on accurate price signals. In fact, these new players are more, not less dependent on accurate, transparent, and efficient energy pricing especially during the tail events created by extraordinary conditions

Under its current design, the CAISO market's energy prices have not provided efficient signals during tight supply conditions. In the near term, inefficiently low market prices leading into and during reserve shortages do not sufficiently incentivize supply performance, do not induce additional real-time resource commitment, discourage day-ahead load scheduling, and do not appropriately allocate the costs of under-scheduling load. In the longer term, these suppressed prices make it harder for California to attract needed capacity from outside the State, blurs investment signals, and could lead to inefficient investment or under-investment in generator plant installations and upgrades.

[1] *PJM Reserve Markets: Operating Reserve Demand Curve Enhancements*, Published on March 21, 2019 by William W. Hogan and Susan L. Pope.

Efficient Market Prices During Tight Supply Conditions

To the extent that market prices are not reflecting the true reliability value of supply leading into and during tight supply conditions, the CAISO will face increasingly difficult operating conditions, suppliers will increasingly not be held accountable for the harm unavailable supply causes to the system, load-serving entities will increasingly be expected to cross-subsidize each other for inefficient demand scheduling, and CAISO will increasingly find it harder to procure power in the West.

There are five underlying issues with the current scarcity pricing market design that cause inefficient market performance during tight supply conditions:

- Real-time market prices do not reflect the high costs of extraordinary out-of-market actions leading into actual reserve shortages and do not rise with sufficient lead time to induce more supply,
- The real-time energy market price signal does not reflect the fact that CAISO must always maintain a minimum reserve requirement to meet reliability standards,
- Real-time prices are not able to rise sufficiently high in real-time to incentivize efficient market participant behavior,
- The current market optimization masks important scarcity pricing signals during demand curtailment, and
- The ancillary services market design and settlement does not incentivize overall supply performance.

This report recommends that the CAISO enhance scarcity pricing and modify its ancillary services market design to improve price formation and market incentives during tight supply conditions. CAISO's current market design requires several changes to drive the efficient scheduling and dispatch of its system to the benefit of both consumers and suppliers.

This report recommends the following market design enhancements:

- The market should use an operating reserve demand curve (ORDC) to ensure that market prices reflect the reliability value of supply leading up to and during tight supply conditions. This design should allow market prices to rise prior to contingency reserve shortages with sufficient lead time to induce more supply participation and ensure all suppliers contributing to the reliable operation of the system are appropriately compensated.
- The market should price reserve shortages equal to the price of maintaining power balance, given that CAISO must always maintain reserves to meet reliability standards and that CAISO would curtail demand to meet these requirements.

Efficient Market Prices During Tight Supply Conditions

- The CAISO should raise the real-time power balance constraint violation penalty price to a price that would provide a strong incentive for accurate day-ahead scheduling and real-time supply performance. When the day-ahead market clears at or near the bid cap, the real-time market penalty prices must increase beyond those used in the day-ahead market. Also, the CAISO should commission an independent consultant to study the value of lost load in California[2] to illuminate the on-going discussion about a fair administrative price to apply when supply is scarce.
- The market should ensure that the system marginal energy price remains at the power balance constraint violation price for the duration of compulsory demand curtailment. The CAISO should also ensure that the implementation of this recommendation does not disturb congestion management.
- Energy and ancillary services schedules should be re-optimized and settled-for-differences in the fifteen- and five-minute markets to ensure suppliers pay for the harm that unavailable supply causes to the system and ensure the efficient use of resources on the system. At a minimum, the five-minute market prices should reflect reserve shortages and no-pay provisions should appropriately penalize unavailable supply.

[2] Given recent actions by the Governor of California and the California Public Utilities Commission, there is reason to believe that this value is no lower than $2,000 per MWh and based on studies in other regions this value can be extremely high.

2 Changing supply and demand fundamentals and the expected prevalence of volatile conditions in the future drive the need for more efficient market prices

As described in detail in the ***Appendix***, over the past decade, California has experienced rapid changes in the composition of its energy supply and the ways in which demand is organized and participates in the energy market. These secular changes, driven in large part by the State's interest in carbon reductions, have challenged the CAISO market systems almost to the breaking point as manifest during the summers of 2020 and 2021. The combination of west-wide generation retirements, wildfires, and persistent drought thrust the West into physical capacity deficits. Arguably, the CAISO market systems did not allow energy and reserves prices to adequately reflect these stressed conditions.

Replacement resources across the west, while still in the planning phase, are yet to come into operations. Even when new resources emerge, most will have significant, binding energy limitations (e.g., wind, solar, and storage) that must be accommodated by grid operators of the future. In fact, the transparency, accuracy, and efficiency of energy and reserves pricing will be critically important to the rational and optimal dispatch of this changing grid. In particular, prices during tight supply conditions must send clear and unambiguous production and consumption signals to suppliers as well as consumers. Efficient market prices will promote and reflect the optimal dispatch of available resources, provide powerful scheduling and performance incentives, and accurately signal the need for participation of additional supply or reductions in consumption.

Without improvements to the CAISO's energy and reserves pricing, manual intervention in the market will increase, prices will not reflect the true costs of reliability, and California will struggle to attract the import resources it so desperately needs.

3 Scarcity Pricing in ISO/RTO Markets

Efficient market prices are essential to achieving reliable operations, economically efficient energy dispatch, and open and non-discriminatory access to the transmission system. Inefficient market prices cause excessive operator action to maintain reliability, incentivize irrational conduct by both consumers and producers, and therefore may require departures from open access or non-discriminatory to access to the transmission grid.

The accuracy of the real-time price signal, appropriately high or low, is essential to efficient market design. When there is sufficient supply to meet demand, the marginal price to serve load will reach equilibrium at the marginal cost of generation. Low prices are appropriate when there is excess supply. Likewise, high prices are appropriate when supply for energy and reserves is constrained. When there is not sufficient supply to meet demand and reserves, prices rise above the marginal cost of generation, reflecting the value that competing consumers place on reliable energy production and delivery.

Accurately high prices during tight supply conditions escalate the consequences for non-performing supply, induce additional resource commitment, and appropriately encourages load-serving entities to both forward-contract and accurately schedule demand. Real-time dispatch and pricing that correspond closely to physical system requirements reduce the need to invent alternative, non-market mechanisms to guarantee that the system operator has enough supply at the right time and location.[3]

3.1 Generally accepted scarcity pricing concepts

Ideally, energy and ancillary services prices would reflect the true marginal cost of production, considering all physical system constraints, and fully compensate all resources for the variable cost of providing service. If demand were fully price responsive and shortage pricing rules accurately reflected the value of avoiding compulsory demand curtailments, short-run energy prices would provide both an accurate price signal for short-term supply and demand behavior and facilitate long-term entry and exit.[4]

A failure to properly reflect in market prices the value of reliability to consumers, including necessary out-of-market operator actions, can lead to inefficient prices in the energy and ancillary services markets. This leads to inefficient system utilization and muted investment signals. When the market optimization is unable reach a solution that meets system needs, administrative pricing rules are needed. Such rules seek to include

[3] *PJM Reserve Markets: Operating Reserve Demand Curve Enhancements*, Published on March 21, 2019 by William W. Hogan and Susan L. Pope.
[4] *Price Formation in Organized Wholesale Electricity Markets*, Docket No. AD14-14-000, Published in October 2014 by Federal Energy Regulatory Commission staff.

the costs associated with the failure to meet minimum reserve requirements in market prices. Ideally, these prices would reflect the value consumers place on avoiding compulsory demand curtailment (commonly referred to as the Value of Lost Load, or VOLL). Under such conditions, prices should rise to induce performance of existing supply resources and encourage consumers to reduce consumption so that the system operator does not need to administratively curtail demand to maintain reliability.[5]

When there is not enough supply to meet energy demand, contingency reserves, and regulating reserve requirements, one expects prices to rise. This occurs due to several reasons. First, the marginal cost of the most expensive resource available and needed to serve demand during tight supply conditions is likely to be much higher than under normal conditions. Second, various studies suggest that consumers place a very high value on avoiding compulsory demand curtailment.[6] However, system operations on stressed days complicates rational price formation because most demand is not price responsive and therefore provides no price signal regarding its willingness to forgo consumption. Also, operators appropriately may take pre-emptive out-of-market actions to ensure reliability when the market optimization is not expected to reach a reliable solution.

ISO/RTOs employ scarcity pricing to more accurately price energy and operating reserves during tight supply conditions. Scarcity pricing is intended to achieve two primary objectives. The first objective is to send a short-term price signal to incent performance of existing resources and help to maintain reliability. In the short term, these prices should be high enough to induce existing resources (including imports) to be available to the market to the maximum extent possible and signal consumers to reduce demand. In this sense, scarcity pricing policies should allow market prices to reflect the value consumers place on energy. The second objective is to facilitate long-term economic entry through the construction of new supply resources and the exit of resources that are no longer economic. Asset owners and developers consider whether expected energy and ancillary service revenues will cover a resource's long-run revenue requirements. Where such revenues are insufficient, the asset owners and developers may determine to exit or not enter a market absent a high capacity payment.

Scarcity pricing also provides an incentive for suppliers to continue to offer their resources into the market at marginal cost. Concern about the exercise of market power is especially acute leading into and during shortage conditions. If prices are expected to rise leading into shortage conditions, suppliers do not have the same incentive to

---

[5] *Id.*

[6] *Estimating the Value of Lost Load*, Published on June 17, 2013 by London Economics International, Inc.; *An Estimate of the Value of Lost Load for Ireland*, Published in October 2010 by Eimear Leahy and Richard S.J. Tol; *An Overview of Selected Studies on the Value of Lost Load*, Published November 15, 2007 by Adriaan van der Welle and Bob van der Zwaan.

withhold, economically or physically, to realize high market clearing prices as they would if prices were capped at the marginal cost of some resource. Suppliers will offer their resources at marginal cost to ensure they clear into the market, knowing that a scarcity pricing mechanism will appropriately elevate prices above marginal costs. Even marginal resources are assured profits to cover their long-run revenue requirements.

While stakeholders in California generally agree that economic entry is driven more by the resource adequacy market and state mandates than by market prices, facilitating long-term economic entry through accurate spot market prices is still important. Today, asset owners and developers rely on bilateral capacity markets to cover the difference between the expected market revenues and their long-run revenue requirements. When market prices are allowed to accurately reflect consumers' valuation of energy, operating reserves, and planning reserves, the resulting long run entry and exit will be economically efficient because revenues will cover a resource's long-run revenue requirements.[7] In addition, the energy and ancillary services markets may provide more differentiated compensation to resources based on the value of the energy that they produce and their operating characteristics compared to capacity payments.

Better capacity prices will follow from more accurate energy market prices and it is more economically efficient to control long-run participation directly through spot-market pricing to the extent possible. In other words, energy market prices that accurately reflect consumers' valuations and the severity of actual system conditions lead to lower capacity prices.[8] On the other hand, capacity market constructs (bilateral or organized) must account for long planning horizons over which many more uncertainties could materialize. These realities lead to crude capacity market structures which come with less efficient revenue allocation to system suppliers. In addition, given the CPUC-led bilateral capacity procurement approach which has limited transparency on the marginal value of capacity, it is very important that the spot energy market prices are not suppressed during tight system conditions since the spot energy prices are transparent to all market participants.

[7] *Resource Adequacy Requirements: Reliability and Economic Implications*, Published in September 2013 by The Brattle Group and Astrape Consulting.

[8] Any remaining capacity revenues a market may provide to suppliers will represent either the additional planning margin that market requires beyond expected day-to-day operations or a supplier's other opportunities to sell energy and ancillary services directly into other markets.

3.2 Relevant scarcity pricing policies employed by other ISO/RTOs

3.2.1 Past scarcity pricing market design developments

In 2008, in Order No. 719,[9] the FERC found that existing RTO/ISO rules used to establish the price of operating reserves (i.e., Spinning Reserves and Supplemental Reserves) do not allow for prices to rise sufficiently during an operating reserve shortage to allow supply to meet demand, and are thus unjust and unreasonable.[10] The Commission specified four reforms that RTOs/ISOs could choose to pursue to remedy this issue: (1) increase the energy supply and demand bid caps above the current levels only during an emergency; (2) increase bid caps above the current level during an emergency only for demand bids while keeping generation bid caps in place; (3) establish an operating reserve demand curve (ORDC), which has the effect of raising prices in a previously agreed-upon way as operating reserves grow short; or (4) set the market clearing price during an emergency for all supply and DR resources dispatched equal to the payment made to participants in an emergency DR program.[11] All RTOs/ISOs ultimately adopted the third reform, i.e., implementation of an ORDC.[12] An ORDC establishes a predetermined schedule of energy or reserves prices according to the level of operating reserves, and increases the price as the availability of operating reserves decreases.[13] Some ORDC implementations directly affect energy prices (e.g., ERCOT) while others generally do not (e.g., CAISO).

Below, ***Table 6*** provides a high-level summary of the status and details of those scarcity pricing policies as they existed in 2018. It puts those policies in the context of each ISO/RTO's resource adequacy construct and describes related design criteria such as energy price ceiling, generator energy offer ceiling, how shortage reserves impact pricing, and how the scarcity pricing is related to Value of Lost Load.

This summary presents several trends. ERCOT, as an energy-only market in a deregulated state wholly reliant on efficient energy pricing, stands out as having the highest priced scarcity pricing design, the highest price caps, and usage of an operating

[9] *Wholesale Competition in Regions with Organized Electric Markets*, Order No. 719, 125 FERC ¶ 61,071 (2008), *order on reh'g*, Order No. 719-A, 128 FERC ¶ 61,059 (2009), *order on reh'g*, Order No. 719-B, 129 FERC ¶ 61,252 (2009).

[10] Order No. 719, 125 FERC ¶ 61,071 at P 192.

[11] *Id.* P 208.

[12] *See PJM Interconnection, L.L.C.*, 139 FERC ¶ 61,057 (2012); *N.Y. Indep. Sys. Operator, Inc.*, 129 FERC ¶ 61,164 (2009); *ISO New Eng. Inc.*, 130 FERC ¶ 61,054 (2010); *Cal. Indep. Sys. Operator Corp.*, 131 FERC ¶ 61,280 (2010); *Sw. Power Pool, Inc.*, 146 FERC ¶ 61,050 (2014). MISO adopted an ORDC prior to the issuance of Order No. 719. *See Midwest Indep. Transmission Sys. Operator, Inc.*, 122 FERC ¶ 61,172 (2008).

[13] Order No. 719, 125 FERC ¶ 61,071 at P 221.

reserve demand curve (ORDC).[14] Conversely, many of the other ISO/RTOs employ a penalty factor, step function approach, which is simpler than an ORDC. Finally, most of the ISO/RTOs developed a scarcity pricing design that has some connection to the VOLL. However, the level of VOLL employed has some disparity. Since 2018, many of the relevant scarcity pricing policies have continued to develop and evolve in response to industry events and changes. These evolutions are explored in further depth in the context of MISO and PJM.

Table 1: 2018 comparison of scarcity pricing policy per ISO/RTO[15]

Market	Resource Adequacy Construct	Price Cap ($/MWh)	Generator Offer Cap ($/MWh)	Reserves Depletion Pricing	Relationship to VOLL
ISO-NE	Forward capacity market	Highest shortage price is $2,350	$1,000	Additive penalty factors by type	Price cap + capacity market performance incentives = VOLL
PJM	Forward capacity market	$3,700	$2,000	Additive penalty factors and step functions by type	Price cap + capacity market performance incentives = VOLL
NYISO	Prompt capacity market	None, but highest shortage price is $2,775	$1,000	Additive penalty factors and step functions by type	None
CAISO	Developed through regulatory process with ISO procurement backstop	None, but highest shortage price is $1,000	$1,000	Additive penalty factors and step functions by type	None
SPP	Reserve margin requirement for utilities	$50,000	$1,000	Additive penalty factors and step functions by type	None
MISO	Voluntary capacity market	$3,500	$1,000	Hybrid additive penalty factors and function of VOLLxLOLP	Price cap = residential VOLL
ERCOT	Energy only	Highest shortage price is $9,000	$9,000	Step function for regulation; economic demand curves for operating reserves	Price cap = VOLL

Some RTOs/ISOs have since proposed reforms to change the quantity of existing reserve products and the pricing of those reserves to address changing operational needs. Both

[14] For clarity, the SPP $50,000 price cap refers to an arbitrarily high parameter SPP uses to meet SPP's global power balance constraint. Presently, SPP prices have a hard energy offer cap of $2,000/MWh, scarcity price adders up to $1,700/MWh, and congestion / losses adders. In February 2021, SPP's prices peaked in the mid $4,000's/MWh.
[15] Shortage Pricing in North American Wholesale Electricity Markets prepared for AESO by the Brattle Group

NYISO and PJM have proposed ORDC revisions to: (1) increase shortage prices; and (2) procure reserves beyond the minimum reserve requirement. The Commission partially accepted NYISO's[16] and is still considering PJM's proposed ORDC revisions.[17] MISO is also considering ORDC enhancements that, according to MISO, are "intended to better price and manage growing uncertainty, incent flexibility, visibility, and availability needs, and address issues identified during recent emergency events."[18]

3.2.2 Recent MISO Scarcity Pricing Reforms

In response to emergency events over the past several years and in anticipation of future industry changes, MISO has been engaged in several initiatives to enhance its markets, planning, and operations. MISO has called these initiatives Resource Availability and Need ("RAN") and the Reliability Imperative. Within those, MISO has been exploring and implementing pricing reform, including to its scarcity pricing. The MISO Independent Market Monitor ("IMM") has also been advocating for scarcity pricing reform for several years. MISO reached a significant milestone in May 2021 by releasing a Scarcity Pricing Evaluation whitepaper.[19] That paper detailed MISO's identified drivers and need for reform as well as MISO's intended approach and solutions. Within that, MISO is planning significant overhauls to many facets of its scarcity pricing design including updating VOLL, implementing a new ORDC, implementing and revising shortage products, and updating its market price caps. MISO is currently in the process of implementing short-term enhancements and has laid out a timeline for pursuing longer-term enhancements.

The exact form of MISO's longer-term enhancements is still yet to be completely developed. However, the MISO IMM's analysis and recommendations in this area provide some insight into what might be expected. ***Figure 12*** illustrates both MISO's current ORDC as well as the IMM's recommendation. It should be noted that although the IMM's recommended curve reaches $10,000/MWh, the IMM has stated an appropriate VOLL for MISO would be much higher at $23,000/MWh (compared to MISO's current VOLL of $3,500/MWh).

---

[16] *N.Y. Indep. Sys. Operator, Inc.*, 175 FERC ¶ 61,241 (2021).
[17] The Commission originally issued an order on PJM's filing in May 2020. *See PJM Interconnection, L.L.C.*, 171 FERC ¶ 61,153, *reh'ng denied*, 173 FERC ¶ 61,123 (2020) (finding PJM's existing tariff unjust and unreasonable, largely adopting PJM's replacement rate as just and reasonable subject to modification and compliance, and reaching the same result on rehearing). The Court subsequently granted the Commission's motion for voluntary remand in the ensuing appeal. *See Am. Municipal Power, Inc. v. FERC*, Nos. 20-1372, et al. (D.C. Cir. Aug. 23, 2021) (granting Commission motion for voluntary remand).
[18] MISO, Scarcity Pricing Evaluation, at *i* (May 2021), *available at* https://cdn.misoenergy.org/20210513%20MSC%20Item%20XX%20Scarcity%20Pricing%20Evaluation%20Paper550162.pdf.
[19] [20210513 MSC Item XX Scarcity Pricing Evaluation Paper550162.pdf (misoenergy.org)](https://cdn.misoenergy.org/20210513%20MSC%20Item%20XX%20Scarcity%20Pricing%20Evaluation%20Paper550162.pdf)

Efficient Market Prices During Tight Supply Conditions

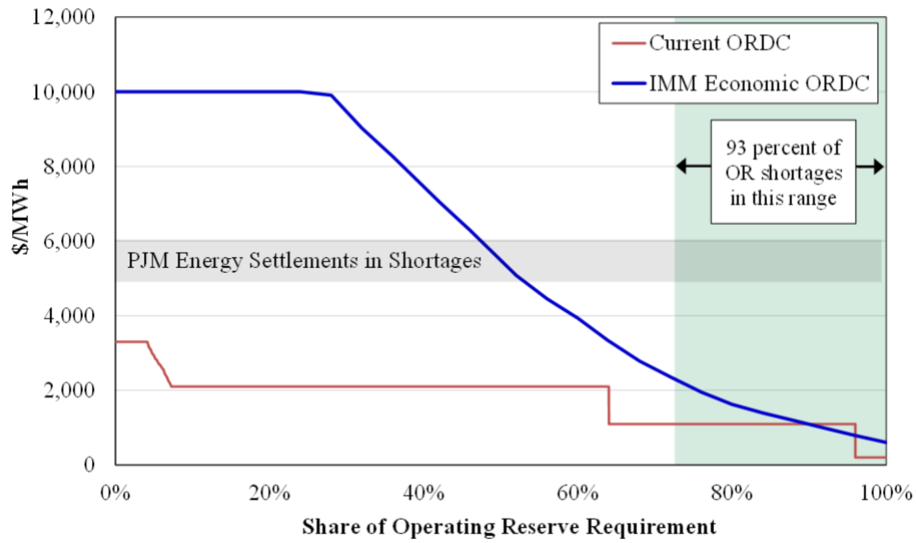

Figure 1: MISO IMM Recommended ORDC

The comparison between the existing MISO ORDC, proposed MISO ORDC, and the PJM shortage energy settlement shows that MISO IMM is concerned about MISO's ability to attract regional uncommitted capacity to offer into its balancing authority area.

A significant aspect of MISO's shortage pricing, in both its current form and from ongoing reform efforts, is the coordination of scheduling and interchange of energy with neighbors during critical times of need. MISO is well interconnected with several market (SPP, PJM, IESO) and non-market neighboring areas (AECI, TVA, SOCO). When MISO is experiencing challenging conditions, it signals its relative need and value for energy with its scarcity pricing, and neighboring entities respond to the prevailing economics by scheduling imports into MISO accordingly. This dynamic has been recognized by the MISO IMM several times recently. For example, in its most recent 2020 State of the Market Report, the IMM included a table of MISO reserve margins in the Summer of 2021 based on range of possible factors.[20] In doing so, the IMM' explains that: "… the table includes additional imports that reflect the average amount of additional imports during emergency conditions. This is conservative because the import levels would likely rise to much higher levels in response to shortage pricing in MISO." In other words, a realistic accounting of MISO's supply position during critical periods must include imports into MISO based on scarcity pricing incentives. As another more specific example, the IMM concluded that during a recent Maximum Generation Event in MISO that rising prices in MISO led to increased imports.[21] On a looking forward basis, the

---

[20] https://www.potomaceconomics.com/wp-content/uploads/2021/05/2020-MISO-SOM_Report_Body_Compiled_Final_rev-6-1-21.pdf
[21] https://www.potomaceconomics.com/wp-content/uploads/2021/09/IMM-Quarterly-Report_Summer-2021-OCT-MSC.pdf

IMM has cited improved incentives for scheduling imports as a benefit to its proposed scarcity pricing reforms.[22]

3.2.3 Recent PJM Scarcity Pricing Reforms

Not unlike MISO's current process, PJM undertook an extensive stakeholder process to significantly overhaul its shortage pricing practices that started in 2017, led to a FERC filing in 2019, with implementation scheduled for 2022.[23] Also similar to MISO, PJM released a whitepaper detailing its need and proposed changes.[24] As part of its overhaul, PJM proposed changes to the ORDC, implementation of a new reserve product, and an increase of penalty factors. ***Figure 14*** illustrates the revisions to the ORDC and the increase of the penalty factor for one of PJM's reserve products. Note that the pink dashed line represents PJM's old ORDC, the green line represents a short-term enhancement PJM made, and the blue line represents PJM's longer-term enhancement to be implemented in 2022. PJM is replacing its prior step-function approach with a smooth ORDC with a significantly higher ultimate penalty price.[25]

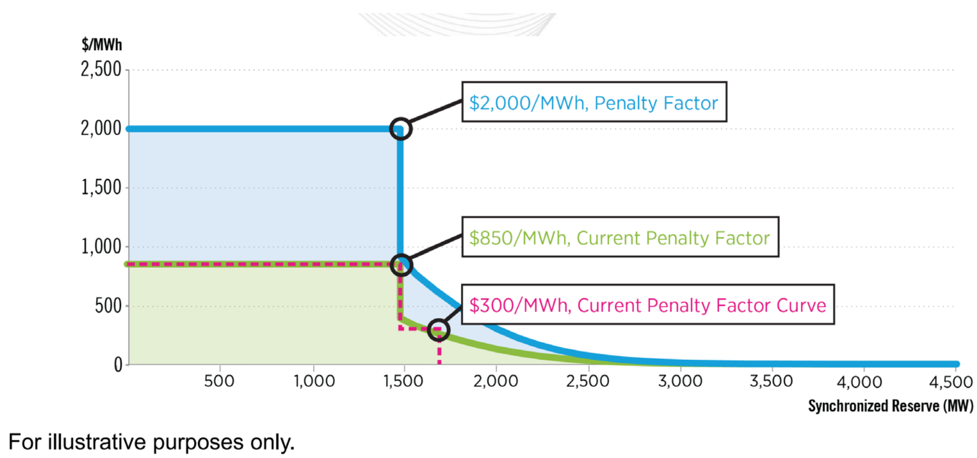

Figure 2: Comparison between existing and proposed PJM ORDC

---

[22] https://cdn.misoenergy.org/20210122%20Scarcity%20Pricing%20Evaluation%20Workshop%20Item%2002%20IMM%20Recommendations514944.pdf

[23] EL19-58-001

[24] 20171115-proposed-enhancements-to-energy-price-formation.ashx (pjm.com)

[25] PJM recently delayed implementing its extended ORDC. At issue is the extent to which it would affect market prices well beyond the "tight supply conditions" discussed in this report. PJM's market monitor estimates the extended ORDC could affect prices in 85% of all market intervals. The recommendations in this report are focused on CAISO's reliability needs leading up to and during compulsory demand curtailment (times during which CAISO already engages in active out-of-market energy and reserve procurement) and therefore are more similar to the scarcity pricing policies implemented by the MISO and supported by the MISO market monitor.

The PJM IMM has calculated the maximum possible LMP under a variety of shortage situations.[26] In the most extreme case, which assumes an energy offer at the cap of $2,000, shortage of each of PJM reserves in all zones (total of 5 additive penalties of $2,000), and a violated transmission constraint resulting in a penalty of $2,000, the maximum possible LMP is $14,000/MWh.

3.2.4 The direction of scarcity pricing reforms in MISO and PJM

Although conducted at different times and containing many unique idiosyncrasies, there are many similarities between MISO and PJM's shortage pricing reforms. They both involve implementation of a smooth ORDC curve that extends beyond minimum reserve requirements with a much higher ultimate price penalty as well as new reserve product types. The result in each case is the possibility for much higher prices in shortage events. It appears that MISO and PJM are each interested in attracting regional uncommitted capacity to offer into their balancing authority area rather than their neighboring balancing authority areas.

Beyond the design changes, both ISO/RTOs were also driven to reform for similar reasons. Both ISO/RTOs saw the need for efficient energy prices to signal the critical need for resources in scarcity situations as well as the industry changes that are increasing the frequency and likelihood of such situations. MISO recently published ***Figure 15*** to describe the key industry trends leading to needed price signal improvement.

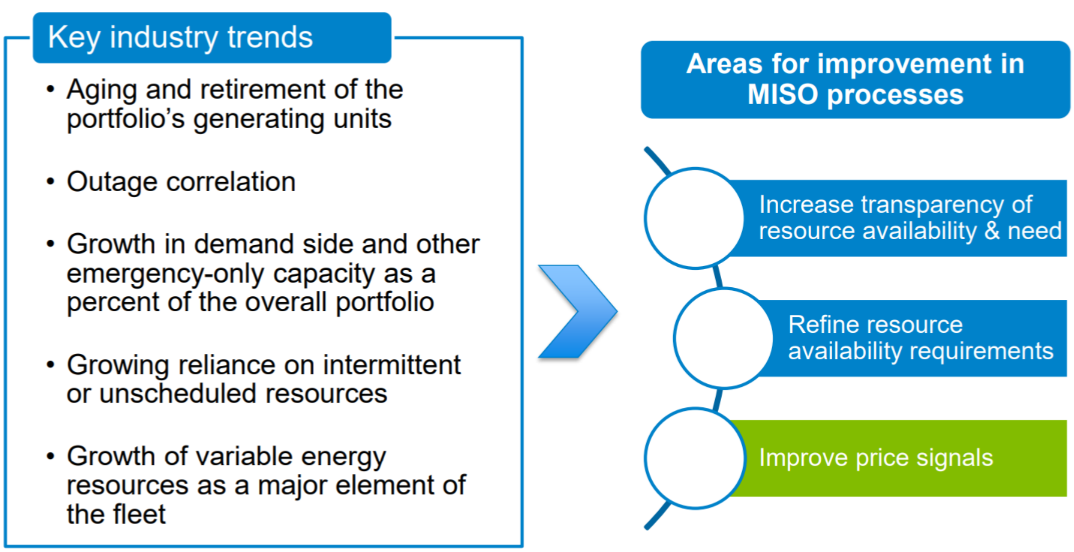

Figure 3: MISO Resource Availability and Need Drivers

[26] http://www.monitoringanalytics.com/reports/PJM_State_of_the_Market/2021/2021q2-som-pjm.pdf, see Table 3-71

4 CAISO's ancillary services market design and scarcity pricing policies

The CAISO first implemented its scarcity pricing policies in the form of co-optimized energy and ancillary service procurement including a scarcity reserve demand curve for ancillary services. This section describes CAISO's current ancillary services market design, reserve price formation, settlement, and how this market design impacts energy prices under certain circumstances.

4.1 Ancillary service products and requirements

CAISO primarily procures ancillary services[27] in its day-ahead market. It determines the total amount of regulating reserves and contingency reserves needed to meet WECC requirements.[28] The CAISO procures 100% of the forecasted requirement in the day-ahead market. The CAISO procures ancillary services in a manner such that higher quality products also count towards meeting lower quality ancillary service requirements (i.e., cascading procurement). Every regulation-up megawatt CAISO procures above the regulation-up requirement can count towards meeting the spinning and non-spinning reserve requirements. Likewise, every spinning reserve megawatt CAISO procures above the spinning reserve requirement can count towards meeting the non-spinning reserve requirement.

The CAISO procures at least 50 percent of the contingency reserves as spinning reserves and no more than 50 percent as non-spinning reserves. The total contingency requirements change each hour as any of four driving factors requires more reserves than the others. The CAISO procures contingency reserves each hour equal to the maximum of four calculated driving factors:

- The sum of three percent of hourly integrated load plus three percent of hourly integrated generation,
- The sum of photovoltaic resources at risk of tripping due to voltage fluctuations,
- Load forecast based requirements, or
- The single largest contingency

The CAISO procures regulation as a percentage of CAISO's demand forecast for the hour based on its need to meet WECC and NERC performance standards (primarily CPS1 and CPS2). The total requirement can change each hour based on the net demand forecast (system demand minus solar minus wind) and the hour-to-hour change in net-

[27] Regulation up, regulation down, spinning reserve, and non-spinning reserve
[28] CAISO sets its procurement target in accordance with WECC Minimum Operating Reliability Criteria requirements.

scheduled interchange between CAISO and its neighboring balancing authority areas. The CAISO maintains a minimum level of approximately 300 MW each hour.

4.2 Ancillary services deliverability

The CAISO achieves some locational diversity in its ancillary services procurement by requiring various regional minimum and maximum procurement requirements in its market. The CAISO establishes these values to account for expected transmission congestion inside the CAISO balancing authority area, as well as other system conditions that may prevent the market from converting reserves to energy, without exacerbating congestion on the paths inside the CAISO balancing authority area. The CAISO describes that the purpose of these constraints is to "increase the probability of deliverability of ancillary services to each region."[29] However, energy and ancillary service capacity do not directly compete for transmission within the CAISO balancing authority area.

In total, the CAISO manages minimum and maximum reserve requirements for 10 inter-related regions. The CAISO uses two broad regions to ensure that generators within the CAISO balancing authority area provide at least 50 percent of all ancillary services. The "expanded system region" encompasses all resources in the CAISO balancing authority area plus all resources participating in CAISO's markets from outside its balancing authority area. The CAISO sets the expanded system region minimum ancillary service requirements to 100 percent of the total requirements. The "system region" includes only the subset of resources in the expanded system region are located inside the CAISO balancing authority area. The CAISO sets the system region minimum ancillary service requirements to 50 percent of the total requirements. Because the system region is in the expanded system region, this ensures that at least 50 percent of the total procurement occurs on generators inside the CAISO balancing authority area.

The CAISO then uses eight other sub-regions[30] to account for expected congestion on Path 15 and Path 26.[31] Depending on whether operators expect congestion on Path 15 and/or Path 26 and the direction of the congestion on those paths, the CAISO calculates minimum and maximum ancillary service requirements for each applicable region.

[29] *Business Practice Manual for Market Operations*, Section 4, CAISO.
[30] The eight other regions are South of Path 15, Expanded South of Path 15, South of Path 26, Expanded South of Path 26, North of Path 15, Expanded North of Path 15, North of Path 26, and Expanded North of Path 26.
[31] Path 15 and Path 26 are major transmission corridors in the CAISO balancing authority area that limit power flow between northern and southern California.

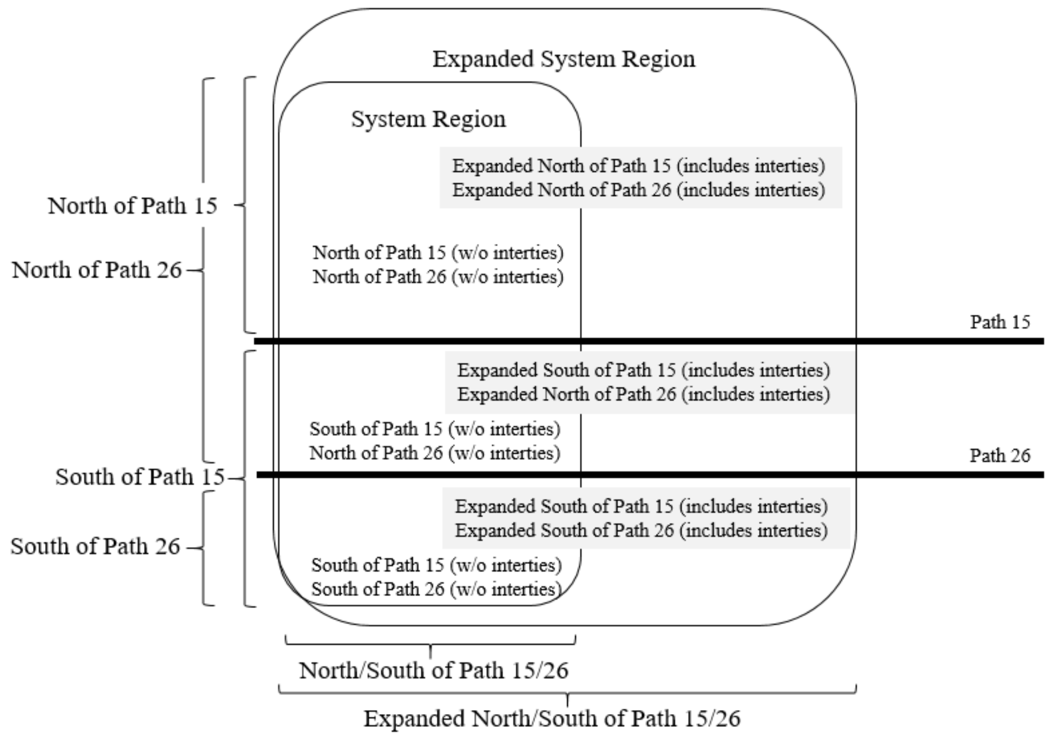

Figure 4: Graphic depicting the two AS system regions and the eight AS sub-regions

4.3 Ancillary services procurement in the day-ahead and real-time market

Suppliers submit energy and ancillary service offers to CAISO in both the day-ahead and real-time markets. In the real-time market, ancillary service offers are limited to the portion of capacity that does not already have a day-ahead ancillary service award.

The CAISO procures 100% of its projected ancillary service needs in its day-ahead market. The day-ahead market co-optimizes energy and ancillary services over the 24-hour trade day. This ensures that available capacity is scheduled to meet demand and reserve obligations over the course of the day in the most economically efficient manner. The co-optimization in the day-ahead market also allows ancillary service costs to appropriately influence energy and ancillary service market clearing prices.

Many situations may occur after the day-ahead market that require the CAISO to procure incremental reserves in its real-time markets. Additional ancillary services could be needed if actual system conditions vary from day-ahead forecasts or if non-trivial amounts of ancillary services become unavailable due to resource or transmission limitations.[32]

The CAISO operates an un-settled hour-ahead scheduling process, a settled fifteen-minute market, and a settled five-minute market during the operating day. It procures *incremental* ancillary services in the hour-ahead scheduling process and the fifteen-minute market, but not in the five-minute market. When ancillary services are unavailable from a given resource in real-time, the CAISO rescinds day-ahead payments from scheduling coordinators but does not charge them the cost of the replacement.

The CAISO procures *incremental* ancillary services in its hour-ahead scheduling process only from resources outside[33] the CAISO balancing authority area that submit hourly block offers.[34] The hour-ahead scheduling process does co-optimize energy and ancillary service offers, but only considers the additional reserves that are needed, if any (i.e., it is not a re-optimization of all reserves). The CAISO settles these awarded ancillary services at the prices generated in the corresponding fifteen-minute market where they are treated as self-scheduled ancillary services. When the hour-ahead scheduling process procures ancillary services to meet incremental needs, fifteen-minute ancillary services prices may be below the marginal cost of providing ancillary services.[35]

The CAISO also procures *incremental* ancillary services when needed in its fifteen-minute market from internal resources. The fifteen-minute market co-optimizes energy and ancillary service offers to meet incremental requirements. Ancillary services awarded in the day-ahead market are self-scheduled and as result not re-optimized in the fifteen-minute market.

The CAISO does ***not*** procure operating reserves in the five-minute market. To the extent that the fifteen-minute market co-optimization raises energy clearing prices in the fifteen-minute market, this added value is missing from the five-minute energy market clearing prices. Thus, the absence of operating reserves procurement and the related co-optimization creates a structural price difference between the fifteen-minute and five-minute market that results in inefficiently low five-minute energy prices during ancillary services shortages.

---

[32] The market software adjusts resource offer curves to represent resource limitations. Such offer curve limitations may cause the real-time market to procure incremental ancillary services from other resources to meet the regulation or contingency reserve requirement. Also, operators can manually adjust the ancillary services requirements to represent other known resource and transmission limitations as such limitations become apparent to the operators

[33] Slow demand response resources offering hourly blocks into the market may also certify to provide ancillary services, but there is very little.

[34] The CAISO refers to these resources as "Non-Dynamic System Resources that submit Hourly Block Bids."

[35] Like energy, this occurs when the ancillary services procured in the hour-ahead scheduling process are marginal.

4.4 Ancillary services pricing

The day-ahead market produces day-ahead ancillary service schedules and computes ancillary services marginal prices. The resulting ancillary services marginal prices represent the cost of procuring an increment of a particular ancillary service in that ancillary service region. Resources in each common area of overlapping ancillary service regions (see discussion on nested zones above) will receive the same ancillary service marginal price for providing the service. Each common area's ancillary services marginal price incorporates the marginal costs from all overlapping regions.

Ancillary services marginal prices reflect any lost opportunity costs associated with keeping the resource capacity unloaded for ancillary services instead of scheduling that capacity as energy. As such, a supplier with an ancillary service award is indifferent between providing energy or ancillary services in the day-ahead market.

Unlike the day-ahead market, the real-time market computes ancillary services marginal prices only when incremental ancillary services procurement is needed, either due to unavailable day-ahead schedules or more changes in system conditions.

The CAISO markets (other than the 5-minute market) establish prices for reserve shortages using a scarcity reserve demand curve. If supply is insufficient to meet the minimum ancillary services procurement requirements in an ancillary service region or sub-region, the scarcity reserve demand curve allows the ancillary service marginal prices in the scarce region or sub-region to rise automatically to administratively determined values. The scarcity reserve demand curve increases the ancillary service clearing prices by $500 to $1,000 depending on which products are scarce and the severity of the scarcity. This scarcity reserve demand curve also has the effect of raising overall energy clearing prices by the same amount, due to the co-optimization of energy and ancillary services.

The scarcity reserve demand curve has five steps and importantly, does not escalate to $1,000 until there is an actual shortage of ancillary services capacity to meet the total regulation, spinning, and non-spinning reserve requirements.

> ***Non-spinning reserve shortages.*** When there is not enough supply to provide non-spinning reserves, the market will forgo non-spinning reserve procurement and incur between $500/MWh and $700/MWh cost. When there is not enough regulation, spinning reserve, and non-spinning reserve offers to provide up to 70 megawatts of non-spinning reserves, reserve shortages are priced at $500/MWh. When there is not enough regulation, spinning reserve, and non-spinning reserve offers to provide up to 210 megawatts of non-spinning reserves, reserve shortages are priced at $600/MWh. Finally, when there is not enough regulation, spinning reserve, and non-spinning

reserve offers to provide more than 210 megawatts of non-spinning reserves, reserve shortages are priced at $700/MWh.

Spinning reserve shortages. When there is not enough supply to provide all the required spinning reserves, the market will forgo spinning reserve procurement and incur $100/MWh cost. Spin shortage prices will be between $100/MWh and $800/MWh depending on the particulars of the overall reserve shortage. For instance, if the spin shortage occurs while there are still sufficient regulation, and non-spin offers available, spinning reserves will be priced at $100/MWh. However, if the spinning reserve shortage occurs concurrent with a regulation and non-spinning reserve shortage, spinning reserves will be priced at $800/MWh. The scarcity prices are additive because the higher quality reserves also count towards the lower quality reserve requirements.

Regulation and energy shortages. When there is not enough supply to provide upward regulation, the market will forgo regulation procurement and incur $200/MWh cost. Regulation shortage prices will be between $200 and $1,000 depending on the particulars of the overall reserve shortage. For instance, if the regulation shortage occurs while there are still sufficient spin, and non-spin offers available, regulation will be priced at $200/MWh plus the marginal cost of spin and non-spin reserves. However, if the regulation shortage occurs concurrent with a spin and non-spin shortage, regulation will be priced at $1,000/MWh. The scarcity prices are additive because the higher quality reserves also count towards the lower quality reserve requirements.

The figure below is a simplified representation of the scarcity reserve demand curve. Comparing the various ancillary services scarcity prices on the vertical axis and the reserve shortage megawatt quantities on the horizontal axis. Theoretically, this curve can affect ancillary services prices in the day-ahead market and the fifteen-minute market. Realistically, this curve will never affect day-ahead market prices because scheduling run penalty parameters dictate the market to relax all self-scheduled demand before relaxing minimum reserve requirements.

Efficient Market Prices During Tight Supply Conditions

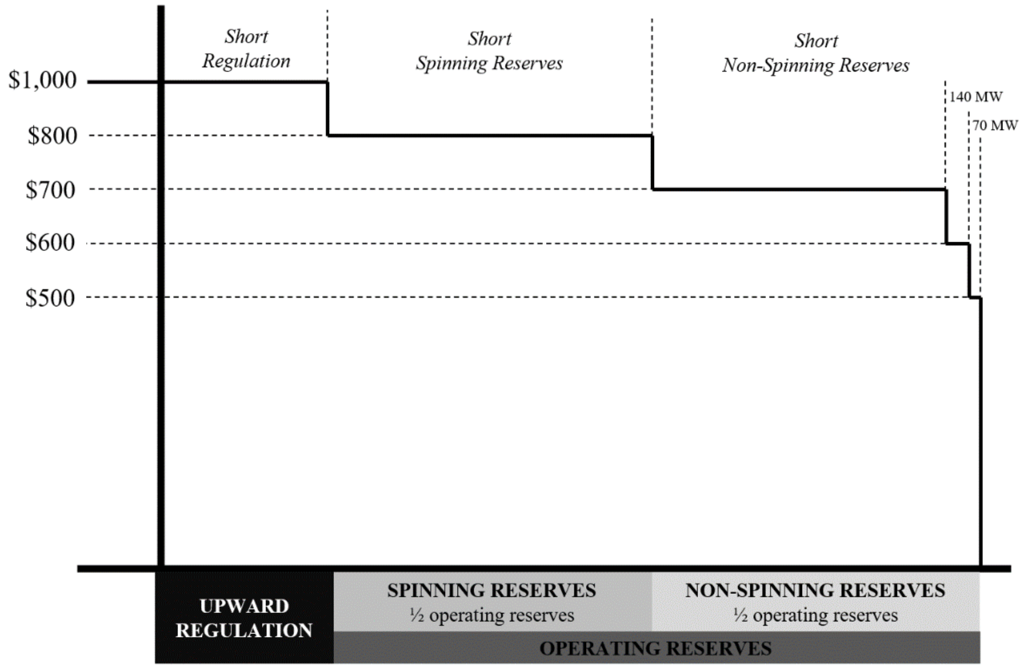

Figure 5: Simplified representation of the scarcity reserve demand curve

Noticeably, the important price signal associated with violating reserve requirements only begins to escalate after the market is unable to clear its minimum reserve requirements. In this situation, system operators are likely already taking out-of-market action to preserve contingency reserves to continue to meet reliability standards. If the deficiency could be addressed through the market, one would see prices above the marginal resource cost representing the reliability value of these necessary actions.

Theoretically, a demand curve should allow energy prices to rise when there are trade-offs between scarce ancillary services and energy. However, in practice, this pricing result would rarely occur because there are many resources with energy bids that can meet the system energy requirements and not meet ancillary services requirements in the real-time market footprint.

Realistically, *energy* prices will either be marginal cost or $1,000/MWh at the time when *reserves* are priced based on the scarcity reserve demand curve discussed above. The ancillary services shortage prices discussed above occur when there are not enough ancillary services offers to meet ancillary services requirements. Ancillary services shortages can happen independently or concurrent with a lack of energy offers to meet energy needs. When there are ample energy offers to meet energy needs, energy prices stay relatively unchanged (at marginal cost) even as the ancillary services prices rise. However, when there are not enough energy offers to meet energy needs, concurrent with the ancillary services shortage, energy prices will rise from marginal energy cost to $1,000/MWh.

Energy prices rise from marginal cost to $1,000 because the scheduling run prioritizes clearing ancillary services over energy. In the scheduling run, ancillary services carry a $1,350-$1,450/MWh penalty price while the first step of a power balance violation carries only a $1,100/MWh penalty price.[36] This means that the optimization would choose to violate the power balance constraint before it would back down an ancillary services schedule that is scarce in order to provide an additional MW of energy. The pricing run does not allow a higher level of ancillary services shortage than scheduled in the scheduling run, so this result carries through to the pricing run even though the ancillary services penalty prices are lower than the power balance penalty prices in the pricing run.

4.5 Pre-emptive out-of-market operator actions

CAISO Operating Procedure 4420 mandates that operators take pre-emptive, out-of-market actions to preserve reliability prior to and during and reserve shortages. These actions further reduce the opportunity for the CAISO's reserve scarcity demand curve to have any effect on prices. CAISO operators regularly add (or reduce) demand that must be balanced with supply to manage the uncertainty inherent in near-term forecasts of load, wind generation, and solar generation (or for unexpected plant outages). They also take other out-of-market actions to preserve reliability such as manually procuring additional supply from importers, calling on emergency assistance from neighboring balancing authority areas, and dispatching emergency demand response resources. These operator actions clearly have an effect on the reserves and energy markets, but the out-of-market actions are not accounted for in clearing prices. It should always a concern that these actions are not reflected in market prices. But it is especially a troubling when such actions prevent the market from pricing what would otherwise be reserve shortages. Operator actions needed to manage uncertainty or prevent emergency situations result in muted energy and reserve market price signals and do not compensate all resources for the value they provide during this critical time.

In preparation for Summer 2021, the CAISO implemented two fairly inconsequential, yet appreciated market pricing enhancements. It identified and fixed a shortage pricing issue when CAISO operators manually "arm load" to meet its reserve requirements. It also

[36] Using penalty prices, the market forgoes energy procurement up to the quantity of a pre-determined seasonal regulation requirement. The energy prices increase from marginal cost to $1,000/MWh because of the interaction between the $1,100 power balance constraint penalty price step (associated with the quantity of seasonal regulation) and the $1,350-$1,450 ancillary services penalty prices in the market's scheduling run. After the total quantity of the seasonal regulation requirement is exhausted at a $1,100 penalty price, the power balance constraint penalty price then rises to $1,450.

partially fixed a pricing issue that occurs during emergency conditions when CAISO calls on reliability demand response resources.

4.6 Ancillary services settlement

The CAISO "claws-back" payments for reserves awarded in the day-ahead market which are unavailable in real time. Separately, it settles replacement reserves acquired subsequently at the fifteen-minute real-time market ancillary service clearing price. However, resources with day-ahead ancillary service schedules are not charged the replacement cost (i.e., settled for differences at real-time prices).

In 2019, the majority of scarcity events were triggered by decreases in available resource capacity in real-time from day-ahead market schedules.[37]

When real-time conditions are stressed, California consumers pay a premium to replace this capacity.

When real-time system conditions become stressed, this settlement arrangement reduces a scheduling coordinator's incentive to ensure their ancillary services are available in the real-time market and California consumers must purchase replacement capacity at real-time ancillary service prices.

In addition to resource unavailability, transmission congestion may limit a resource's ability to provide ancillary services in real-time. If operators observe that the energy and ancillary services would violate transmission line ratings if the resource were to supply both awarded energy and potentially converted reserves, they may block ancillary service awards on the resource and require incremental real-time procurement.[38] This scenario likely requires more expensive ancillary service procurement. Again, the CAISO does not charge the resource the fifteen-minute market ancillary service clearing price to replace its undeliverable capacity. California consumers continue to pay the resource for ancillary services that they will never be able use while also paying to purchase replacement ancillary services.

The CAISO may contract out-of-market for ancillary services in real-time.[39] In addition to procuring incremental ancillary services in the hour-ahead scheduling process and the

---

[37] *2019 Annual Report on Market Issues and Performance,* CAISO Department of Market Monitoring

[38] For instance, according to the CAISO Department of Market Monitoring's 2019 annual report, there were 24 ancillary service scarcities across real-time intervals on November 20, 2019. This was the result of manually blocked ancillary service awards, which were blocked in the real-time market but not in the day-ahead market for this day. This led to a shortage of regulation in real-time.

[39] The CAISO calls this "manual real-time contracting for ancillary services or unloaded resource capacity."

fifteen-minute market, the CAISO enters into real-time contracts for ancillary services and unloaded resource capacity. If a real-time need for ancillary services arises, the CAISO may individually contact scheduling coordinators to secure ancillary services or unloaded resource capacity to serve as contingency reserves. Although these ancillary services were not included in the fifteen-minute price formation, scheduling coordinators receive payment at the applicable fifteen-minute market ancillary service prices.

5 Other CAISO market product interactions with scarcity pricing policy

The CAISO, as part of any consideration of enhanced scarcity pricing should evaluate both the intended and unintended consequences of other key market pricing formulations. In this section we review several of the more significant programs and highlight possible interactions or exclusiveness.

5.1 Flexible ramping product

CAISO also procures real-time ramping capability through its flexible ramping product. The purpose of the flexible ramping product is to address uncertainty caused by demand and variable energy resources that materializes between real-time market runs. Flexible ramping capacity is procured in both an upward and downward direction according to a demand curve. The product is co-optimized with energy and ancillary services. As less ramping capability is available to meet ramping demand, the flexible ramping product demand curve raises flexible ramping product prices and energy prices by up to $247/MWh. Given the low maximum clearing price and the specific nature of ramping capability, this product was not built to signal overall supply scarcity discussed in this paper.

The flexible ramping product addresses a significant price formation issue resulting from having a multi-interval optimization but a single interval settlement. If the CAISO settled the entire optimization horizon, then all marginal prices would be consistent with the energy bids of resources. Prior to implementing the flexible ramping product, a resource could have been held back out-of-merit in the binding interval to provide energy in an advisory interval yet not be compensated for this opportunity cost through the advisory interval's eventual binding price. In these unfortunate scenarios, the binding interval price was inappropriately lower when the resource was held back to provide upward ramping and inappropriately higher if dispatched to provide downward ramping. The CAISO addressed this price formation issue by introducing the settlement of forecasted movement at the shadow price of the upward ramping constraint separately from energy prices.[40]

The market procures an uncertainty requirement through a demand curve. The demand curve ensures that the expected value of holding ramp capability to meet future net load is less than the energy cost impact in the binding interval. For example, if there was a 10% chance of incurring a power balance constraint violation which would result in $1,000/MWh energy costs in the future if the ramp capability was released, then if energy

[40] CAISO's August 18, 2017 FERC filing regarding Fast-Start Pricing in Markets Operated by Regional Transmission Operators and Independent System Operators discusses this in detail. http://www.caiso.com/Documents/Aug18_2017_SupplementalComments-Fast-StartPricingNOPR_RM17-3.pdf

costs in the current interval exceed $100, it is more economic to release the capacity for energy now than to hold it for uncertainty that may materialize later.

The addition of uncertainty requirements results in the price for energy being slightly elevated all the time to minimize spurious price spikes in a few intervals. The flexible ramping product demand curve is implemented by including surplus variables that take value as the demand curve levels are reached. The surplus variables add supply to meet the requirement which has the effect of reducing the actual requirement to the point where the expected value of flexible ramp procured exceeds the probability that the uncertainty will materialize.

Three key elements of the flexible ramping product design should be recognized in the scarcity pricing context:

1. Forecasted movement addresses an energy price formation issue where the financially binding marginal price is inaccurate. While the addition of uncertainty requirements will in general raise the marginal price of energy in the binding interval, it reduces the probability of future financially binding spurious price spikes.
2. The uncertainty demand curve for the flexible ramping product determines if the market should use the ramp capability now as energy or potentially later as energy. This is different than an operating reserve demand curve, where it determines if the market should schedule capacity as energy or ancillary services in this interval with no regard for future intervals.

5.2 Cost-conditional offer cap and penalty price scaling (FERC Order 831)

CAISO recently implemented higher offer caps and penalty prices which are effective when supply is more costly, but not necessarily scarce. CAISO implemented these changes to meet the FERC Order 831 requirements to allow suppliers to bid higher than $1,000/MWh when cost justified. It also implemented cost-conditional penalty parameter scaling which raises the scarcity reserve demand curve prices in certain circumstances.[41] While FERC found these changes just and reasonable to achieve the goal of allowing prices to represent actual costs when those costs are greater than $1,000/MWh, the changes inadvertently dampen the important scarcity price signal during supply shortages.

A recent scarcity event illustrates the issue with cost-conditional penalty parameter scaling. On July 9, 2021, CAISO experienced what it calls a “genuine” scarcity event, but none of the cost conditions allowed the scarcity reserve demand curve to scale up to

[41] When the CAISO accepts and validates offers greater than $1,000/MWh or when the maximum import bid price exceeds the soft energy bid cap for any trading hour of the integrated forward market.

$2,000/MWh. As such, the market valued supply scarcity on this day less than on other days. At 16:06 due to a loss of resources and fire threat to the transmission system, the CAISO forecasted a resource deficiency with all available resources in use or forecasted to be in use between 17:00 and 22:00. At 17:44 the CAISO declared EEA2, and at 18:32 the CAISO declared EEA3 and began arming load that could be reduced in 10 minutes. Between 18:00 and 20:00, the CAISO operators made available contingency spinning reserve and contingency non-spinning reserve in the market for energy dispatch at the bid cap. The market dispatched the spinning reserve and non-spinning reserve capacity for energy. The culmination of these events brought market clearing prices to near $1,000/MWh.

It is not evident that demand should value lost load differently depending on supplier costs. The conditional nature of the higher scarcity reserve demand curve prices implies that demand values scarce supply differently on different days. On the CAISO's most stressed day of the year, while CAISO operators were on the cusp of ordering compulsory demand curtailment and struggling to maintain system balance and reserves, prices only rose to around $1,000/MWh. If the same conditions occurred on another day, but with higher gas costs, the scarcity prices would be allowed to rise to $2,000/MWh.

Cost-conditional penalty parameter scaling obfuscates the value of lost load, dampens the incentives for supply performance in the real-time market, rewards under-scheduled demand, and provides no incentive for suppliers to bid true to marginal cost leading into supply shortages. Integration of more or different scarcity pricing mechanisms could drive reconsideration of these approved, cost-driven Order 831 conditions.

5.3 Minimum state-of-charge requirement for energy storage resources

As energy storage resources play an increasingly significant role on the CAISO system, it is imperative that CAISO ensure competitive participation by these resources.[42]

As part of its Summer 2021 readiness initiative, the CAISO proposed imposing a constraint that batteries that are contracted for resource adequacy attain a level of charge specified by the CAISO prior to the evening peak. A core driver for proposing this requirement was a concern that these resources would not be charged sufficiently to meet day-ahead obligations to discharge. In the absence of scarcity pricing, the low prices that may occur even during tight system conditions will not incent charging and discharging actions that meet the reliability goals of the CAISO.

This concern is reasonable and will become particularly important as the number of small, distributed energy storage resources rises. While operators can use the minimum state of charge constraint and exceptional dispatch to override dispatch instructions, this

[42] FERC Order ER21-2779-000, October 26, 2021.

will become more of an operational challenge for operators and more likely to lead to unintended outcomes as the number of storage resources on the system increases and their importance in meeting CAISO load increases. The CAISO's Market Surveillance Committee noted that implementation of a scarcity pricing system that sends an efficient price signal for decentralized operation of energy storage could also help eliminate the need for and cost of such as constraint.

Scarcity pricing reform would go a long way towards avoiding the illogical storage charging and discharging patterns that the CAISO observed in the summer of 2020 while also providing improved storage incentives to respond consistent with system conditions throughout the day.

5.4 Emergency Load Reduction Program

In March 2021, the California Public Utilities Commission (CPUC) created the Emergency Load Reduction Program (ELRP) to pilot a new demand response approach to help avoid rotating outages. The program is designed to help eliminate the need for rotating outages while minimizing costs to ratepayers associated with building additional standby resources that are seldom used. The CPUC describes it as an insurance layer on top of existing resource adequacy reliability planning.

The ELRP provides consumers a $1,000/MWh payment contingent on demonstrated load shed. Consumers may also use any self-generation technology to offset their energy consumption during an ELRP event. California's investor-owned utilities can only call an ELRP event following a CAISO Alert, Warning, or Emergency (AWE) declaration.

The ELRP program operates outside of CAISO's wholesale market. To the extent that ELRP demand reductions occur, California's investor-owned utilities purchase demand reduction for $1,000/MWh, but paradoxically, the actions will depress CAISO's wholesale prices due to less system demand. These $1,000/MWh demand reduction purchases are not represented in CAISO's market clearing prices. At a time when CAISO is trying to attract more wholesale supply to resolve system emergencies, demand reduction purchases for greater than wholesale market rate are reducing its wholesale prices. The CPUC recently raised this price to $2,000/MWh for 2022 and 2023.[43]

5.5 Proclamation of a State of Emergency

More recently, the Governor of California created a similar program through a *Proclamation of a State of Emergency*.[44] Like ELRP, consumers are paid to reduce consumption or increase on-site generation. This program pays participants up to

[43] A December 2, 2021 CPUC Decision raised this price to $2,000/MWh for 2022 and 2023 (Decision 21-12-015).
[44] Executive Department, State of California, Proclamation of State of Emergency on July 30, 2021.

$2,000/MWh. This program also is not priced in CAISO's wholesale market and as such will have a suppressive effect on prices when in operation. However, it does indicate that the State of California is willing to pay up to $2,000/MWh to avoid compulsory demand curtailments during system emergencies.

6 Recommendations

The recommendations in this section require the CAISO to enhance scarcity pricing policy and modify its ancillary services market design to improve price formation and market incentives. As discussed in further detail below, the CAISO's current market design includes several elements that must be changed to drive the efficient scheduling and dispatch of its system to the benefit of both consumers and suppliers during tight supply conditions.

The CAISO needs a properly functioning scarcity pricing market design to maintain important market incentives during tight system conditions. Unfortunately, during these conditions, the current market design results in:

- Energy and reserves prices that do not reflect the reliability value of extraordinary out-of-market operator actions as the system approaches reserve shortages,
- Inefficiently low energy prices during compulsory demand curtailment,
- Insufficient day-ahead scheduling incentives, and
- Energy and ancillary service supply performance penalties that do not follow cost-causation.

These problems reflect five underlying issues with the current scarcity pricing market design:

- Real-time market prices do not reflect the reliability value of supply leading into actual reserve shortages and do not rise with sufficient lead time to induce more supply participation,
- The real-time market price signal does not reflect the fact that CAISO must always maintain a minimum reserve requirement to meet reliability standards,
- Real-time prices are not able to rise sufficiently high in real-time to incentivize efficient market participant behavior,
- The current market optimization may be masking important scarcity pricing signals during demand curtailment, and
- The ancillary services market design and settlement design does not incentivize overall supply performance.

6.1 Real-time market prices do not reflect the reliability value of supply leading into actual reserve shortages and do not rise with sufficient lead time to induce more supply participation

The current market pricing is ineffective at signaling reliability threats as evidenced by reserve capacity shortages. Practically, the market design only raises prices after a shortage occurs and market operators have taken extreme, out-of-market actions. And even if prices could rise, they would do so in a discontinuous, sporadic, and ineffective manner. In short, market clearing prices do not reflect the reliability value that every available generator provides to the system leading into and during stressed conditions.

As discussed above, during a reserve shortage, the current market design will rarely raise energy prices above the marginal cost of generation. This being the case, while the CAISO does have a scarcity reserve demand curve to raise reserve prices, it's role in the overall scarcity pricing design is quite limited.

In addition, long before a reserve shortage, system operators begin to execute emergency procedures[45] that are ultimately aimed at avoiding reserve shortages and compulsory demand curtailment.[46] Admittedly, it is not practical or implementable to ensure each individual out-of-market action is appropriately and automatically incorporated into the market optimization to calculate wholesale market prices in a coordinated manner. The system operators may execute steps of their emergency procedures in a different order based on their judgement of the evolving situation and they generally seek out extra supply and emergency assistance sequentially rather than simultaneously. Nonetheless, the important take-away is that these actions are valuable to consumers because they avoid catastrophic consequences and all suppliers providing energy (or ancillary services) when operators take extraordinary actions should be compensated for the value they are providing.

Even if CAISO could overcome deficiencies with its current scarcity reserve demand curve and early operator intervention, the price-related effects of the demand curve do not impact prices until an actual (not imminent or forecast) reserve shortage has occurred. At that point, prices might rise (depending on the combination of specific product deficiencies) but would do so in a discontinuous manner. This late and sporadic binary

[45] CAISO Operating Procedure 4420 outlines the extraordinary actions operators take. These actions include calling for voluntary demand reduction, coordinating use of the Emergency Load Reduction Program, enabling Reliability Demand Response Resources, using exceptional dispatch, reducing participating pumping load, dispatching available unloaded generation capacity without real-time energy bids, dispatching excess operating reserves, dispatching Reliability Must Run Units, calling on out-of-market demand response and UDC interruptible load programs, arming firm load to be counted as operating reserves, procuring out-of-market operating reserves as available, and procuring emergency assistance from neighboring balancing authority areas.

[46] Such actions are justified to avoid the extreme costs associated with loss of load and to avoid bad reliability outcomes associated with minimum operating reserve deficiencies.

movement in prices is – and has been – highly unlikely to result in the behavioral changes necessary to bring more supply or less demand into the market.

Finally, in leading up to a reserve capacity shortage, it is difficult to distinguish between legitimate high-cost offers reflecting the scarcity value of resource capacity and the exercise of market power. However, allowing higher energy offers from generators is not the only way to effectuate higher market-clearing prices. Instead, allowing prices to rise following a prescribed curve that reflects the reliability value of supply on the system prior to and during tight supply conditions would reduce incentives for suppliers to inflate their variable offers to achieve higher prices reflecting the scarcity value of their capacity.[47] Consequently, this design could provide the market monitor an important tool to distinguish price increases resulting from scarcity from the attempted exercise of market power.

Recommendations: The CAISO should consider a market design that would allow market prices to always reflect the reliability value of supply, especially leading into and during reserve shortages. Having such a design in place would create proper price signals to attract additional supply to, or reduce demand in, the market at this critical time as well as incentivize resources to continue to offer into the market at marginal cost.

The CAISO should implement an enhanced operating reserve demand curve (ORDC) that begins to raise prices *prior to* reserve shortages. An operating reserve demand curve would be an administrative intervention in the market, but this is already true of the administrative requirement and reserve scarcity demand curve for operating reserves. In the presence of a necessary and persistent operating reserve requirement, a superior administrative rule would be a better model of the demand for operating reserves that goes beyond the fixed quantity requirement.[48]

The basic outline of an operating reserve demand curve marries the value of lost load with the probability that demand will be curtailed. It is a decreasing curve extending beyond the minimum reserve requirement that can be designed to terminate at a specific quantity of excess reserves. The value of an increment of operating reserves beyond the minimum requirement would be the value of lost load multiplied by the probability that net demand would increase enough in the coming interval to reduce reserves to the minimum reserve requirement where operators would order compulsory demand curtailment.

The CAISO will need to design the operating reserve demand curve to work with its market. There are many design considerations to explore, including the flexible ramping

[47] William W. Hogan & Susan L. Pope, "PJM Reserve Markets: Operating Reserve Demand Curve Enhancements," March 21, 2019.
[48] William W. Hogan, "Electricity Scarcity Pricing Through Operating Reserves: An ERCOT Window of Opportunity," November 1, 2012.

product demand curve and existing reserve scarcity demand curve. The CAISO should consider the following discussion of key design elements while designing an operating reserve demand curve.

Value of lost load. Some ISO/RTO's use a strict calculated value of lost load, while others use a maximum cost at which resources could be procured based on market offers. At a minimum, as discussed below (***Section 6.3***), the CAISO should consider setting this value at least as high as the chosen value for power balance constraint violations that is sufficient to incentivize supply performance during tight supply conditions, incentivize accurate day-ahead demand scheduling, and incentivize load-serving entities with short real-time positions to seek out additional supply during the day.

Loss of load probability. The fundamental calculation of loss of load probability is not new or unique. The CAISO itself does a similar calculation to define its flexible ramping product demand curve. The objective is to determine the probability that a power balance constraint violation will occur given demand forecast error, intermittent supply forecast error, and generator forced outage uncertainty over a defined time horizon between real-time market runs. The calculation will produce various levels of MW quantity shortfalls that occur at various frequencies. For instance, the CAISO may find that a 100 MW shortfall would occur 5 percent of the time, a 200 MW shortfall would occur 2 percent of the time, and so on. This shows how often an additional 100 MW or 200 MW of reserves would have been needed to avoid compulsory demand curtailment. The objective of the operating reserve demand curve is to determine the probability that a load shed event would occur and levels of excess reserves are reduced.

It may be useful for the CAISO to define this horizon as the time between the hour-ahead scheduling process run for the upcoming hour and the fifteen-minute market. The CAISO may find this a desirable horizon because it uses the hour-ahead scheduling process to incorporate all offers for the upcoming hour to determine the intertie schedules going into the hour. It would also allow CAISO to resolve for the uncertainty between when offers are due and when uncertainty materializes.

Like other ISO/RTOs, the CAISO could determine these probabilities using three-years of historical data and could further refine its methodology to produce hourly loss of load probabilities for different seasons.

Demand curve construction. The operating reserve demand curve has two parts. The first part represents the minimum reserve requirement, which the system operator cannot violate, priced at the determined value of lost load. This first part of the curve is a horizontal line at the value of lost load from zero MW of reserves to the minimum reserve requirement quantity. The second part is a decreasing price slope given the product of value of lost load and loss of load probability for any given level of additional reserves. This second part of the curve represents the expected cost of marginal demand curtailment during actual operation corresponding to any given level of scheduled reserves. The CAISO should consider subtracting the variable cost of generation at the margin from the value of lost load in this calculation, which would yield the equation: Price equals the value of lost load minus marginal cost, multiplied by the loss of load probability. Doing so would ensure that the marginal willingness to pay for an increment of operating reserves at the moment of load curtailment will not exceed the value of lost load.[49] Mechanically, it reduces the scarcity adder as system resources become more costly.

Product and Counting. It is standard practice to create operating reserve demand curves for each ancillary service product/localized-region and allow qualified ancillary services capacity to count toward meeting those operating reserve demand curves. Depending on its needs, the CAISO may weigh the value of this standard approach versus a generalized design where *any* capacity that can meet its desired reserve performance can count toward meeting one system-wide (or multiple regional) operating reserve demand curve(s) in addition to other capacity products. In addition, consideration will need to be paid to the cascading rules of the upward reserve products.

Other Quantity and Price Considerations. The CAISO will need to consider its operating practice and market dynamics when evaluating if the total quantity of reserves implied by the loss of load probability calculation[50] is sufficient for its needs.[51] The quantity should be representative of the point in which system operators would normally begin to be concerned about supply conditions, which may be prior to officially invoking emergency procedures. The price magnitude should be representative of the actions the operator may take at this time. The purpose of evaluating this is to ensure that if system operators would normally start taking out of market actions to prevent reserve shortages, the reliability value will be represented in market clearing prices. These reserves will also improve reliability because more capacity will be available to respond to changing system conditions. The quantity should also be set to allow sufficient lead-time to enable additional supply participation in the market. In other words, the CAISO should set the quantity considering the uncertainty that materializes during the time in which incremental supply can actually make itself available to CAISO. Setting the quantity with this in mind would allow market prices to signal the need for more supply offers prior to when supply offers are due to the market (the real-time bid submission deadline is 75 minutes prior to the start of the operating hour). This quantity can be different

[49] PJM did find that in its implementation, it did not need to subtract the marginal cost of generation at the margin because it employs other market rules which preclude use of any energy offers for price setting that rise above a current system-wide offer cap.

[50] Theoretically, a general property of an operating reserve demand curve is that the demand is not vertical and price does not drop to zero. Scarcity pricing would arise to some degree for all hours. However, practically, the CAISO should be primarily concerned during times where the reliability value of extraordinary operator actions are not reflected in market prices.

[51] Calculated probabilities are likely not completely aligned with operating practice.

hourly, daily, or seasonally. The CAISO could support its selection (or selection method) of this quantity with a general analysis of operating practices during tight supply conditions or through discussions with its system operators.

6.2 The real-time market price signal does not reflect the fact that CAISO must always maintain reserves to meet reliability standards

CAISO operating reserves should never fall below the minimum required reserve level. The minimum level of operating reserves that the CAISO must maintain is prescribed in NERC/WECC standards and depends on various factors (discussed in detail in ***Section 3.3.1***). If the CAISO fails to maintain this minimum reserve requirement, not only is it a violation of NERC/WECC standards, but the system is in a precarious reliability state. The minimum reserve requirement reflects the level of remaining reserves at which the system operator will begin to direct compulsory demand curtailment to preserve reliability. The extraordinary actions that operators take to avoid reserve shortages indicates the high importance the CAISO places on maintaining operating reserves to meet minimum requirements.

Current market penalty prices ensure that only a portion of the minimum reserve requirements are protected from shortages. During tight supply conditions, the market optimization allows energy bids that overlap with ancillary services bids to be used to meet minimum reserve requirements up to the quantity of the seasonal regulation requirement.[52] This occurs when there are not enough energy and ancillary services bids to meet incremental ancillary services needs plus unavailable day-ahead ancillary services awards. When the market does this, system energy prices rise to $1,000/MWh while reserve prices remain at their respective marginal reserve price. Beyond the seasonal regulation requirement quantity, the market begins to incur ancillary services shortages (procurement below the minimum reserve requirement) to maintain power balance. This outcome conflicts with the straightforward regulatory and reliability requirement to always maintain a minimum level of reserves. It also results in inefficiently low ancillary services prices.

***Table** 7*, shows that the market's scheduling run will always forgo energy for a $1,100/MWh cost up until a certain point, after which, it will begin to forgo reserves for a cost between $1,350/MWh and $1,450/MWh.

Table 2: Scheduling run penalty parameters dictate the energy and ancillary services prices in the market

Scheduling Run Penalty Price ($/MWh)	Pricing Run Penalty Price ($/MWh)	Description
$1,100	$1,000	First step of power balance constraint violation up to the ***quantity of the seasonal regulation requirement***

[52] CAISO Market Operations Business Practice Manual, Section 6

$1,350	$500-$700	Non-spinning reserve violations
$1,400	$100	Spinning reserve violations
$1,450	$200	Regulation violations
$1,450	$1,000	Second step of the power balance constraint violation for power balance constraint violation in excess of the quantity of the seasonal regulation requirement.

Recommendations: Whenever there is compulsory demand curtailment and the system has just the minimum of operating reserves, then any increment of reserves would correspondingly reduce the demand curtailment. Hence the price of operating reserves should be set at the value the CAISO attributes to demand curtailment. An efficient market design would price reserve shortages equal to the price of maintaining power balance, given that CAISO must always maintain reserves to meet reliability standards and that CAISO would curtail demand to meet these requirements.

6.3 Real-time prices are not able to rise sufficiently high in the real-time market

Excluding congestion effects, real-time prices generally are capped at either $1,000/MWh or $2,000/MWh during tight supply conditions. When day-ahead market prices are near these levels, suppliers face limited risk for real-time unavailability and demand faces limited-risk for under-scheduling day-ahead demand.

When day-ahead prices clear near the offer cap, the market design currently provides deficient incentives for accurate virtual supply, virtual demand, physical demand, and VER scheduling. The market pricing mechanism should support day-ahead supply and demand schedules that are as close to expected real-time needs as possible. However, the market pricing incentives break down during tight system conditions when the day-ahead market prices are expected to approach the market offer cap.

For instance, if virtual demand expects day-ahead prices to approach the market offer cap, they face a lower expected profit. This incentivizes less virtual demand bidding which ultimately decreases competition among consumers that would otherwise correct for potential under-scheduled physical demand.

While incentivizing less virtual demand, the market design also incentivizes physical demand to under-schedule in the day-ahead market. When prices approach the market offer cap in the day-ahead market, demand sees little risk in waiting until real-time to purchase supply, resulting in under-scheduled demand. In this situation, even given equal probability of higher or lower prices in real-time, they can only rise a few dollars per MWh from the day-ahead price, while they can fall by hundreds of dollars per MWh. This phenomenon incentivizes physical demand to under-schedule in the day-ahead market. Not only will demand under-schedule in the day-ahead market, but if tight supply conditions begin to materialize in real-time there is little incentive to take action to correct their position. Allowing real-time prices to rise much higher than day-ahead prices during tight supply conditions would incentivize load-serving entities to appropriately schedule demand in the day-ahead market and to the extent that shortages remain, engage in additional intra-day bilateral supply purchases to hedge their real-time price exposure. Load-serving entities covering their exposure to real-time prices aligns their actions with system operators' efforts to secure additional real-time offers.

Likewise, if virtual suppliers expect day-ahead prices to approach the market offer cap, they face an increasingly diminishing loss potential. This changes the risk and reward for virtual supply bidding[53] which can result in under-scheduled physical supply to meet demand and a corresponding increase in exports not supported by physical supply. In practice, this increases the burden on the residual unit commitment process to correct the financial result. Ideally, the integrated forward market would get the incentives right so that prices in the integrated forward market are representative of anticipated real-time conditions and the system operators would not have to rely on the administrative residual unit commitment process.

Furthermore, the real-time market price signal is not sufficient to incentivize *real-time supply performance* during tight supply conditions. Suppliers with day-ahead energy schedules and ancillary services awards during tight system conditions face a minimal loss potential if they are unavailable in real-time, while their unavailability may cause catastrophic consequences. Allowing real-time prices to rise much higher than day-ahead prices during tight supply conditions would ensure suppliers with unavailable supply are appropriately penalized commensurate with the harm they cause to the system.

Recommendations: An efficient market design would incentivize supply performance during tight supply conditions, incentivize accurate day-ahead demand scheduling, and incentivize load-serving entities with short real-time positions to seek out additional supply during the day. The most targeted way to achieve these outcomes is to update the real-time market administrative pricing parameters.

The CAISO should raise the real-time power balance constraint violation penalty price, which is the reference point all other penalty parameters are set, to a price that would provide a strong incentive for accurate day-ahead scheduling and real-time supply performance. When day-ahead market clears at or near the bid cap, the real-time market penalty prices must increase beyond those used in the day-ahead market.

The CAISO should commission an independent consultant to study the value of lost load in California. Given recent actions by the Governor of California and the California

[53] For example, if day-ahead prices are $997 and there is a 95% chance that prices will go to $1,000 in real-time and a 5% chance prices will go to $500 in real-time, the virtual supplier will find a limited loss potential and submit virtual supply. The virtual supplier sees an expected loss of $2.85 ($3.00 x 95%) and an expected gain of $24.85 ($497 x 5%).

Public Utilities Commission,[54] there is reason to believe that this value is no lower than $2,000 per MWh and based on studies in other regions this value can be extremely high. The study can illuminate the on-going discussion about what is a fair administrative price to apply when supply is scarce.

6.4 The current market optimization may be masking important scarcity pricing signals during demand curtailment

During compulsory demand curtailment periods, market clearing prices do not reflect the real demand for energy. The market optimization only sees the energy demand from connected circuits, rather than the true consumer demand for energy at that time. When system operators curtail demand, the market *optimization* sees that demand for energy has decreased and therefore can clear the market lower in the supply stack. However, the reality is that most consumers would like to have functioning air-conditioning.[55] During compulsory demand curtailment, the system marginal energy cost should reflect the fact that the latent consumer demand still exists, even though the market optimization cannot see it.

Recommendations: The CAISO should consider ensuring that the system marginal energy price remains at the power balance constraint violation price for the duration of compulsory demand curtailment. It should also not undermine congestion management through market prices at these times.[56] The CAISO could investigate implementing this concept by modifying the offer output of the market power mitigation pass to set unconstrained resource offers to the offer cap. This type of modification would be like the enhancement the CAISO recently made[57] regarding contingency reserve deployment while it relies on demand curtailment to meet its reserve requirements.

6.5 The ancillary services market design and settlement does not incentivize overall supply performance

To the extent that suppliers cannot provide ancillary services in real-time, it is appropriate for them to pay the full replacement cost. As discussed in ***Section 3.3.1***, CAISO procures 100% of its forecasted reserve requirements in the day-ahead market. When conditions change in real-time (e.g., unavailable resources, transmission constraints, higher than anticipated demand), the CAISO procures incremental reserves in the fifteen-minute

[54] Executive Department, State of California, Proclamation of State of Emergency on July 30, 2021
[55] One such individual consumer in Southern California called into an emergency CAISO board meeting on Monday, August 17, 2020 to personally express his hardship in withstanding 120 degree heat with no air-conditioner.
[56] Simple price floors may undermine congestion management by sending too high a price signal to constrained generators.
[57] CAISO's Market Enhancements for Summer 2021 Readiness initiative

market. It does not procure ancillary services in the five-minute market. To the extent that underlying capacity supporting day-ahead ancillary services is unavailable, consumers purchase replacement ancillary service in the fifteen-minute market at fifteen-minute market prices, while suppliers pay consumers back at the day-ahead market price.[58] During tight supply conditions, it is highly likely that the refunds are much less than the additional cost of scarce real-time ancillary services. The price differences that lead to this additional consumer cost can be even greater on days where the CAISO also requires additional real-time procurement, as generally higher ancillary services demand drives higher real-time prices.

Suppliers with day-ahead energy schedules and ancillary services awards during tight system conditions face a minimal loss potential if they are unavailable in real-time, while their unavailability may cause catastrophic consequences. Suppliers with day-ahead energy schedules that are unavailable in real-time face losses if the real-time prices rise higher than the day-ahead prices, as they need to buy-back the energy at the higher real-time price. If day-ahead prices were near the bid-cap, these potential losses are not commensurate with the ultimate harm they may cause to the system (i.e. energy emergencies, extraordinary out-of-market operator actions, and compulsory demand curtailment). Likewise, when ancillary services become scarce in real-time, the real-time ancillary services prices rise much higher than the day-ahead ancillary service prices. However, given the current market design, ancillary service providers "buy-back" their award at the day-ahead ancillary service price. During tight system conditions and reserve shortages, ancillary service providers do not incur financial loss commensurate with the harm they may cause in real-time.

Supply can be unavailable for a variety of reasons and may become more common in the future. A resource on forced outage is a commonly cited reason for unavailability. However, resource use-limitations also constrain output. For instance, multi-stage generators may not be able to provide the required output in real-time. Likewise, resources with ramping limitations, batteries with state-of-charge limitations, or resources with limited starts may also be unavailable. CAISO's interconnection queue is full of resources that may find themselves with constrained real-time supply. Resources in the real-time market should pay for the harm caused to the system should they be unable to fulfill day-ahead obligations.

In addition, during tight supply conditions, the CAISO's five-minute energy prices are inefficiently low due to lack of ancillary services procurement. When the CAISO cannot meet its minimum reserve requirements, the fifteen-minute market prices incorporate this deficiency using the scarcity reserve demand curve. Generally, the scarcity reserve demand curve prices are incorporated into both the energy and ancillary services clearing

[58] CAISO Business Practice Manual Configuration Guide: No Pay Spinning Reserve Settlement, Charge Code 6124

prices because these products are co-optimized.[59] This means reserve shortages also cause higher energy prices in the fifteen-minute market. However, the CAISO does not procure ancillary services in the five-minute market and therefore the reserve scarcity demand curve does not also cause higher five-minute market energy prices. This results in inefficiently low five-minute market energy prices that may incentivize suppliers to negatively deviate from dispatch to their benefit. Ideally, these prices should accurately reflect the real-time value for energy on the system.

Finally, the real-time market could provide additional value to consumers with a more efficient real-time utilization of capacity. The real-time market protects day-ahead ancillary services awards with an extremely high penalty price and only procures *incremental* reserves when needed. This practice locks a sizable portion of the resource fleet out of potentially providing greater value as energy rather than ancillary services in the real-time market (or vice versa) depending on the system conditions. The market design should put resources to the most efficient use in real-time time. For example, the market may be better off if use-limited resources with changing opportunity costs in real-time are re-purposed from ancillary services to energy. Additionally, if real-time minimum reserve requirements are less than day-ahead minimum reserve requirements, the market would benefit from the automatic and efficient re-purpose of ancillary services to energy.

Recommendations: An efficient market design would incentivize performance by holding suppliers responsible for replacement costs, holding consumers harmless for resource unavailability, not incentivize negative dispatch deviations, and putting resources to the most efficient use in the real-time market.

The CAISO should re-settle ancillary services awards in the fifteen-minute and five-minute market to incentivize supply performance and hold consumers harmless for resource unavailability. This can be accomplished through a wholesale re-design of the ancillary services settlement structure to follow the same design as the energy product settlement or through changes to the ancillary services no-pay provisions.

Given the ancillary services re-settlement design recommendation, the CAISO should consider either implementing a nodal ancillary services product or else hold suppliers harmless for unavailable capacity *due to real-time transmission constraints*. Like nodal energy prices, if transmission constraints could be fully incorporated into individual generator ancillary service prices, the overall result would not harm suppliers whose capacity is unavailable due to transmission constraints.

---

[59] There are instances where ancillary services deficiencies are not incorporated into energy clearing prices. This occurs when there is no energy opportunity cost. For example, when the CAISO has enough energy bids to meet demand, but not enough ancillary services bids to meet minimum reserve requirements.

Efficient Market Prices During Tight Supply Conditions

The CAISO should incorporate reserve shortage prices into five-minute market prices to ensure five-minute market energy prices accurately reflect the real-time value for energy during tight supply conditions. This can be accomplished by either re-procuring incremental ancillary services in the five-minute market or incorporating an administrative price floor in the five-minute market that accounts for the level of reserve shortage experienced in the fifteen-minute market.

The CAISO should re-optimize all real-time energy and ancillary bids in the fifteen-minute and five-minute market to ensure resources are put to the most efficient use. This design recommendation could also be employed in place of the previous recommendation to incorporate reserve shortage prices into the five-minute market clearing prices. The CAISO can implement this recommendation by no longer protecting day-ahead ancillary services awards at an extremely high penalty prices and allowing real-time ancillary service bid curves to cover each resource's entire output range. CAISO operators may be concerned that complete real-time re-optimization could affect status quo reliability as the market re-purposes ancillary services to energy and vice versa depending on real-time conditions. The concern is that operators would find it too difficult to follow exactly where ancillary services will be coming from at a given time and therefore ancillary services could become inaccessible. More granular ancillary services procurement or a nodal ancillary services product could go a long way towards alleviating these concerns. In addition to economic benefits, it is important to note that complete ancillary services re-optimization also brings reliability benefits as the market automatically adjusts ancillary services procurement up and down to meet system needs without operator intervention.

Appendix

Over the past decade, California has experienced rapid changes in the composition of its energy supply and the ways in which demand is organized and participates in the energy market. This evolution has placed considerable strain on the CAISO system and fundamentally changed the operational characteristics that the CAISO values the most. This evolution also brought a large and diverse set of energy purchasers to the market, each with varying objectives and levels of risk-aversion. Without improved scarcity pricing, energy prices are not sending appropriate price signals to both supply and demand.

The California supply evolution

When the CAISO implemented its nodal markets in 2009, it could rely on imports, a fleet of nuclear, hydro-electric, and gas-fired generators to reliably meet demand. Since then, developers added 19,000 megawatts of grid-scale renewable generation and 11,000 megawatts of behind-the-meter solar. During the same time, as a result of California energy policy, carbon-emitting generators, generators that employ once-through-cooling technology, and California's nuclear generators have been retiring. Also, hydro-electric generator availability has been decreasing due to more frequent and persistent drought. To manage this supply evolution, CAISO operators largely rely on CAISO's remaining dispatchable resources and imports. CAISO operators also rely on out-of-market actions to make additional supply available to the CAISO dispatch algorithm. Now, more than ever, the CAISO needs efficient market price signals to attract supply, incentivize performance, and appropriately compensate resources with valuable capabilities.

In 2009, gas-fired resources provided just over 25,000 megawatts of capacity, hydro-electric generating facilities added another 6,000 megawatts, and nuclear resources supplied close to 5,000 megawatts spread across both northern and southern California. The CAISO relied on imports, qualifying facilities, and other non-dispatchable generators for the remainder of its needs.[60] Notably, CAISO relied on 8,800 megawatts of expected import supply to meet summer planning criteria in a "low import" scenario.[61] At the time, renewable resource development was still in its infancy. One could count installed wind and solar capacity in the *hundreds* of megawatts. This fleet of resources produced energy throughout the day and afforded CAISO enough dispatch flexibility to meet its needs.

Today, renewable generators produce up to 19,000 megawatts. On April 24, 2021, renewable resources boasted a 94.5 percent instantaneous generation penetration rate. Solar grew from a few hundred megawatts to 13,200 megawatts, wind grew to 7,000

---

[60] *2009 Annual Report on Market Issues and Performance,* CAISO Department of Market Monitoring

[61] *2009 Summer Loads and Resources Operations Preparedness Report*, CAISO

megawatts, and demand response grew to 3,800 megawatts. While solar, wind, and demand response capacity additions have exceeded reductions in gas and nuclear capacity, they have a very different set of operating characteristics.[62]

In addition to grid-scale solar generation, California commercial and residential developers have added approximately 11,000 megawatts of solar generating capacity behind-the-meter.[63] This supply does not participate in the wholesale market, but it does alter the traditional daily demands that system operators must meet. On a given day, when the sun sets, grid-scale generators must replace behind-the-meter solar production and grid-scale solar production at the same time that consumers require additional energy in their transition from work to home. The CAISO net demand curve[64] commonly depicts this phenomenon. Below, ***Figure 1*** compares the traditional demand curve (including the netting effects of behind-the-meter solar production) with the net demand curve on February 22, 2021. In the evening, system operators increased non-solar (and wind) grid-scale production by over 17,000 megawatts in three hours.

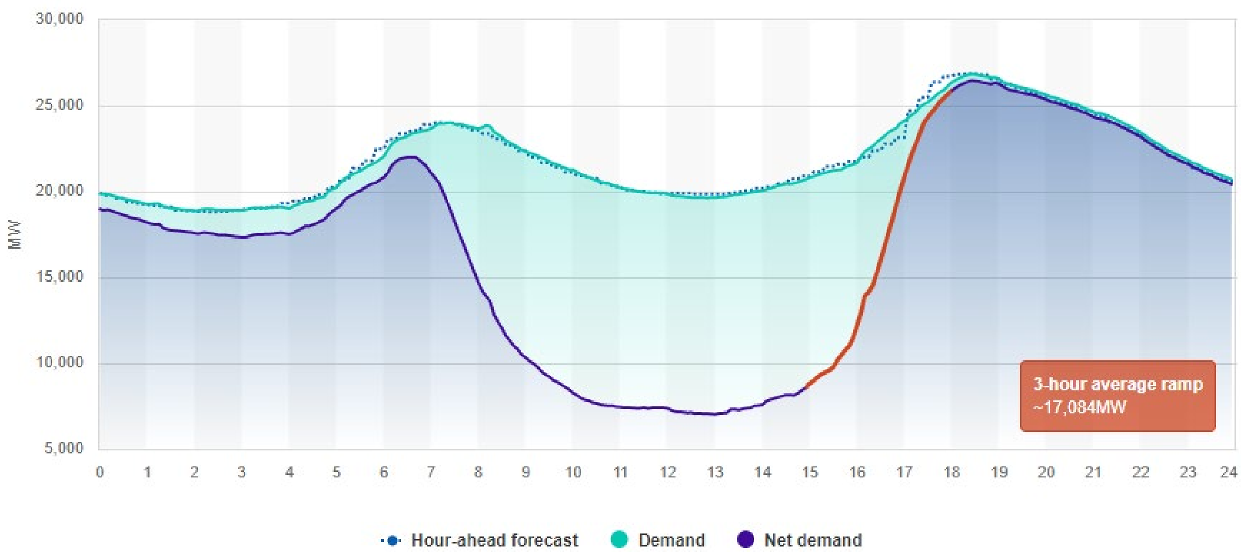

Figure 6: Demand versus net demand on February 22, 2021[65]

California energy policy has led to the retirement of gas and nuclear capacity, especially capacity reliant on once-through-cooling. Overall gas-fired capacity fell from about

[62] *2019 Annual Report on Market Issues and Performance*, CAISO Department of Market Monitoring
[63] *2022 Flexible Capacity Needs Assessment*, Presentation, Published on April 22, 2021 by CAISO
[64] The net demand curve is the traditional demand curve which includes the netting effects of behind-the-meter solar minus low-cost grid-scale renewable production that needs to be replaced by other energy each evening.
[65] *Today's Outlook, February 22, 2021*, CAISO

35,000 megawatts in 2012 to 30,000 megawatts today. The California State Water Resources Board regulations require generator owners and operators to phase-out 20,500 megawatts of power production from generators using once-through-cooling technology. 10,400 megawatts have already been retired since 2010[66] and all but 3,800 megawatts will be retired by the end of 2023.[67] In January 2012, Southern California Edison shut down the San Onofre Nuclear Generation Station due to premature wear on recently installed replacement steam generators and high regulatory hurdles preventing it from producing again, eliminating approximately 2,250 megawatts of baseload southern California power supply.[68] Finally, in June 2016, Pacific Gas & Electric announced that it plans to close the Diablo Canyon nuclear power plant by 2025, eliminating another 2,250 megawatts of baseload supply.[69]

In addition, hydro-electric generator availability is changing. Due to more frequent and persistent drought, California may not be able to depend on the full capacity of its hydro-electric generation fleet from year to year. In 2021, California hydro conditions were below normal. The statewide snow water content for the California mountain regions peaked at 60 percent of average on March 31, 2021 and statewide snow water content was lower than 2020 when the statewide snow water content peaked at 63 percent of average. On April 1, 2021, California's major reservoir storage levels were at 70 percent of average.[70]

In 2021, hydro-electric generating conditions in the Pacific Northwest were also alarming. The CAISO's own hydro-electric generation planning metric was out of date by the time it published its summer planning report.[71] The CAISO used the Northwest River Forecast Center's projected reservoir storage at The Dalles Dam on the Columbia River as a marker for expected summer hydro-electric energy production in the Pacific Northwest. The CAISO's studies assumed the April through September reservoir level would be 89 percent of average, but by the time it published the report, this level had dropped another five percentage points to 84 percent of average.[72] These levels were markedly worse than in 2020 and reservoir inflows are down 30.3 percent year over year.[73]

---

[66] https://www.energy.ca.gov/sites/default/files/2019-12/once_through_cooling_ada.pdf
[67] https://www.waterboards.ca.gov/press_room/press_releases/2020/pr09012020_otc_amendment.pdf
[68] https://newsroom.edison.com/releases/southern-california-edison-announces-plans-to-retire-san-onofre-nuclear-generating-station
[69] https://www.pge.com/en_US/safety/how-the-system-works/diablo-canyon-power-plant/diablo-canyon-power-plant/diablo-decommissioning.page
[70] *2021 Summer Loads and Resources Assessment*, CAISO
[71] *Id.*
[72] April through September Dalles Dam Reservoir Level, Northwest River Forecast Center
[73] U.S. Army Corps of Engineers

Efficient Market Prices During Tight Supply Conditions

The California supply evolution has placed considerable strain on system operators. Evening ramping needs have increased from a leisurely 1,500 megawatts per hour in 2012 to over 5,753 megawatts per hour today.[74] The new supply mix requires operators to hold additional dispatchable resource capacity to manage increased uncertainty. At any given time, cloud cover could suddenly remove solar production, which has a similar system impact as a generator contingency that requires replacement energy. Cloudy days can be particularly challenging for another reason: clouds can swing solar production off *and on* over very short time periods.

To manage this supply evolution, CAISO operators largely rely on CAISO's remaining dispatchable resources and imports. Below, ***Figure 2*** shows CAISO meeting a large evening net load ramp with approximately 8,500 megawatts from natural gas resources (orange), 5,500 megawatts from imports (red), and 2,000 megawatts from large hydro-electric generators (blue).

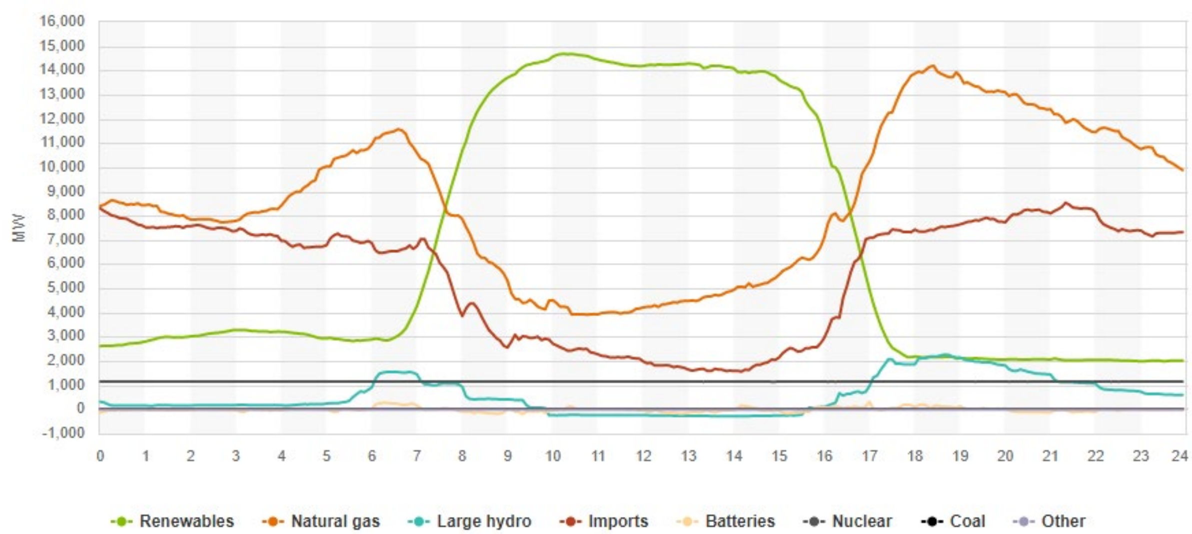

Figure 7: CAISO supply sources on February 22, 2021[75]

Below, ***Figure 3*** shows cloud cover swinging solar production (orange) off and on over very short time periods throughout the day. This type of variability emphasizes the CAISO's need for a resource fleet that will quickly respond in both the 5-minute dispatch and under 4-second automated generation control. In addition to its own resources, system operators rely on five-minute imbalance energy from their neighbors to resolve these types of variability challenges.

[74] In its *Draft 2022 Flexible Capacity Needs Assessment*, the CAISO showed the actual February 2021 maximum monthly three-hour ramp was 17,259 megawatts. This translates to 5,753 megawatts per hour. The CAISO duck curve shows an approximately 4,500 megawatt evening ramp over a 3-hour period in 2012.

[75] *Today's Outlook, February 22, 2021*, CAISO

Efficient Market Prices During Tight Supply Conditions

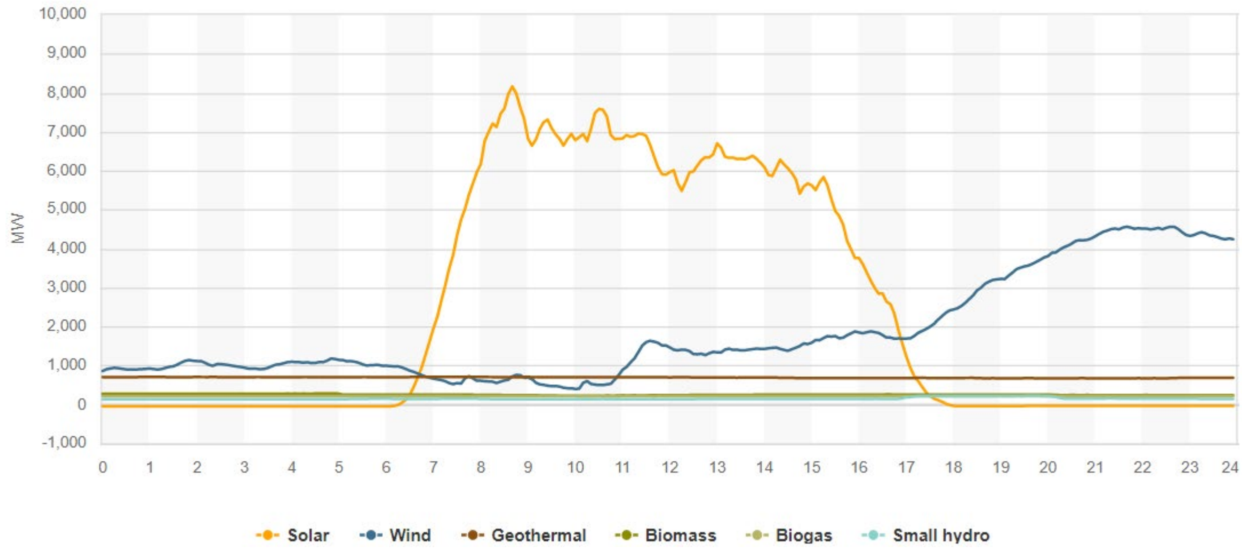

Figure 8: Renewable resource production on March 3, 2021

During tight supply conditions, system operators must currently make out-of-market adjustments to avoid reserve shortages and accompanying compulsory demand curtailment. System operators frequently rely on out-of-market actions to commit additional supply which is then available to the CAISO dispatch algorithm. When CAISO operators determine that market results do not or will not meet system reliability needs, they may make load adjustments, issue exceptional dispatches, or issue manual dispatches of intertie energy. The necessary out-of-market action can suppress prices which masks the system needs during periods of tight supply.

CAISO operators primarily and routinely use load adjustments in the hour-ahead and 15-minute market in a manner which helps to increase the supply of ramping capacity within the CAISO balancing authority area during morning and evening hours by increasing hourly imports and committing additional units within the CAISO.[76] Beginning in 2017, CAISO sharply increased load adjustments during the steep morning and evening net load ramp periods in the hour-ahead and 15-minute markets. In 2019, operators made positive load adjustments in over half of all real-time market intervals.[77] The sharp increase in load adjustments continued into the first quarter of 2021 with the average hourly load adjustments in these markets peaking at just about 1,100 MW.[78] While not as dramatic, CAISO operators also use out-of-market *exceptional dispatches* to acquire ramping capacity. On average in 2020, operators issued approximately 110 MW per hour of exceptional dispatches for unit commitment and energy, of which, approximately 40% were related to acquiring ramping capacity.[79]

---

[76] *2019 Annual Report on Market Issues & Performance*, Section 9.3, Published on July 2, 2021 by the CAISO Department of Market Monitoring

[77] *Id.*

[78] *Q1 2021 Report on Market Issues and Performance*, Published on June 9, 2021 by the CAISO Department of Market Monitoring

[79] *Id.*, Section1.12

CAISO operators can also issue exceptional dispatches on the interties, which the CAISO calls "manual dispatch imports." In 2017, CAISO operators significantly increased their use of this method, although they pared back their use in 2018 and 2019. Between May and September 2017 there were over 100 manual dispatches over a span of 12 days totaling about 15,300 MWh. The largest daily quantities occurred on September 1, 2017, and September 2, 2017, about 6,100 and 5,900 MWh, respectively. These dispatches occurred between hours ending 16 and 22, but were concentrated in hours ending 18, 19 and 20. The single largest hour of manual dispatch occurred on hour-ending 19 on September 1, 2017 totaling 1,700 MW.[80] In 2019, there were approximately 50 hours of manual dispatches for imports accounting for less than 6,000 MWh.[81]

As shown in ***Figure 4***, CAISO operators heavily relied on out-of-market actions in August and September 2020. During hours 17 through 22, CAISO operators used an average of 250 MW to 3,400 MW per hour of out-of-market energy and export curtailments. On the days in which the CAISO curtailed demand, it was relying on an average of approximately 300 MW per hour of imports from emergency assistance, 300 MW per hour of manual dispatch imports, and between 250 MW per hour and 500 MW per hour of exceptional dispatches issued to internal generators.[82] On August 14 through 16, the CAISO primarily relied on out-of-market imports from Palo Verde, Malin, and Sylmar. Later in the heat wave, CAISO began to rely on out-of-market imports from Sacramento Municipal Utility District (SMUD) and export curtailments. CAISO operators used approximately 3,400 MW per hour of out-of-market energy and export curtailments on August 18, 2020. These exceptional dispatches focused primarily on internal generation and export curtailments. Import assistance that appeared to be readily available in excess of 300 MW to 1,200 MW on other days only made up approximately 250 MW of the 3,400 MW on this day.[83]

---

[80] *2017 Annual Report on Market Issues & Performance*, Section 9.2, Published on June 11, 2018 by the CAISO Department of Market Monitoring

[81] *2019 Annual Report on Market Issues & Performance*, Section 9.2, Published on July 2, 2021 by the CAISO Department of Market Monitoring

[82] *Report on System and Market Conditions, Issues and Performance: August and September 2020*, Section 3.7.1, Published on November 24, 2020 by the CAISO Department of Market Monitoring

[83] *Id.*

Efficient Market Prices During Tight Supply Conditions

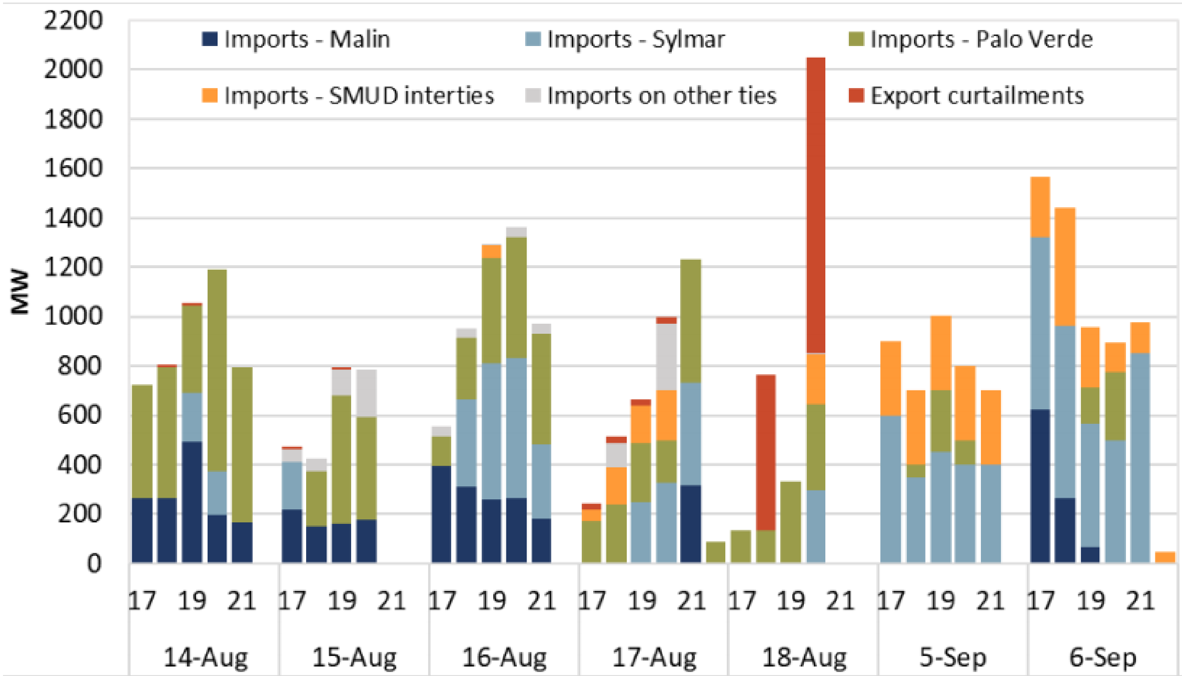

Figure 9: Hourly out-of-market imports, emergency assistance, and market export curtailments in hours 17 through 22[84]

In addition to managing intra-day variability, system *planners* are relying on 9,500 megawatts of import energy to meet CAISO's 2021 peak-demand summer planning criteria.[85] However, the CAISO studies warn that reduced levels of net imports during high-demand conditions significantly affects system reliability. Accordingly, it reported that it could face capacity shortfalls under "more extreme widespread, high load conditions that both drive up California loads and also restrict availability of imports from other systems due to the high demand across the West."[86] If such conditions materialize, CAISO plans to use additional extraordinary measures accessed under extreme or emergency conditions to minimize the risk of actual firm load shedding.[87]

Since these necessary out-of-market actions can suppress real-time prices during tight system conditions, it is important that efficient price formation is not undermined which is a key driver in the need for improved scarcity pricing in the CAISO.

---

[84] *Report on system and market conditions, issues, and performance: August and September 2020*, Published on November 24, 2020 by the CAISO Department of Market Monitoring
[85] This figure includes non-Resource Adequacy imports.
[86] 2021 Summer Loads and Resources Assessment, CAISO
[87] *Id.*

Efficient Market Prices During Tight Supply Conditions

The California demand evolution

Competition among California load-serving entities is the highest it has ever been in California history. Relative to 2009, there is now a large and diverse set of energy purchasers with varying levels of risk-aversion, procurement prerogatives, and financial backing. In addition, there is now a large and growing segment of price-sensitive demand participating in the CAISO markets. Also, entities with an interest in exporting energy from California add thousands of megawatts of competitive demand to the market. Now, more than ever, California needs efficient market price signals to incentivize efficient supply contracting and demand scheduling.

In 2009, three large investor-owned utilities represented the vast majority of consumers participating in the CAISO's wholesale market. Smaller electric service providers and municipal electric providers represented the remaining consumers. At the time, California's first community choice aggregator, Marin Clean Energy, would not start operations for another year. Demand response and convergence bids (i.e. virtual demand bids) did not exist and price-sensitive demand in the CAISO markets was limited to a relatively small amount of water pumping loads.[88]

Today, the demand-side of the market is much more diverse, and each load-serving entity has a different energy and capacity procurement prerogative. The share of CAISO demand represented by the three large investor-owned utilities has fallen to 45 percent.[89] There are 25 community choice aggregators that represent over 200 communities and 27 percent of CAISO demand.[90] Community choice aggregators have differentiated themselves from the investor-owned utilities by focusing on providing consumers a much more renewable-centric supply mix at lower rates and by offering their communities innovative programs.[91] Although each load-serving entity has a different energy and capacity procurement prerogative, purchasing enough energy to keep their customers' lights on is a common objective.

Although demand is now more diverse, the price at which load-serving entities would be willing to forgo consumption remains largely unknown. A 2019 report published by the CAISO Department of Market Monitoring shows day-ahead market demand curves dominated by non-price-sensitive self-schedules on relatively stressed days. This

---

[88] *2009 Annual Report on Market Issues & Performance*, Section 2.1.3, Published on April 19, 2010 by the CAISO Department of Market Monitoring

[89] *California Energy Demand Forecast Update*, 2020-2030, Mid Demand case.

[90] *Id.*

[91] Recently, Central Coast Community Energy launched rebate programs to support the purchase or lease of new or used electric vehicles, and to support the installation of electric vehicle home charging equipment. Silicon Valley Clean Energy is building a local marketplace for load flexibility from distributed energy resources such as battery storage and smart thermostats.

indicates that the price at which demand would be willing to be curtailed is higher than consumers can bid into the market at this time.[92]

Nonetheless, developers added close to 5,000 megawatts more price-sensitive demand to the CAISO's markets since 2009. Demand response programs have grown to over 3,800 megawatts. Also, developers have been installing energy storage resources at a rapid pace: over the past five years, developers added 1,000 megawatts of energy storage to the grid and will install 800 megawatts more by the end of summer 2021. While the CAISO markets incorporate demand response and energy storage consumption bids in its aggregate supply curve, these resources represent close to 5,000 megawatts of *price-sensitive demand* in the day-ahead and real-time markets.

Demand response programs allow market participants to indicate a price at which they will reduce consumption. These programs are valuable to typical load-serving entities because they reduce capacity costs and provide load-serving entities an energy hedge. Demand response programs are primarily used by load-serving entities according to its energy procurement strategy leading to different objectives for demand response programs between load-serving entities.

Energy storage operators have a different prerogative than typical load-serving entities. Typical load-serving entities value price stability and seek to purchase energy at the lowest cost to meet their consumption needs. These needs are well-known and largely inelastic. On the other hand, energy storage operators value price volatility and will purchase energy at any price if they forecast a profitable opportunity to return the energy to the grid. Energy storage operators compete with typical load-serving entities to purchase energy in the marketplace.

In 2011, the CAISO added convergence bidding features to its market that allow market participants to bid price-sensitive demand (i.e., virtual demand) directly, whether or not they serve end-use customers. Over the years, the average hourly quantity of virtual demand has been around 3,000 megawatts and recently increased to over 4,000 megawatts.[93] Entities that submit virtual demand bids in the market directly compete with others in California to purchase energy in the marketplace.

With increased demand diversification, growing levels of price-sensitive demand, and convergence bidding, it is critical that energy prices appropriately and accurately signal the true value of energy, especially during tight supply conditions. California needs efficient market price signals to incentivize efficient supply and demand contracting and scheduling. Without accurate real-time prices the penalty for load under-scheduling is

[92] *Report on day-ahead market competitiveness for September 25, 2019*, Published October 30, 2019 by CAISO Department of Market Monitoring

[93] *2011 through 2019 Annual Reports on Market Issues and Performance*, CAISO Department of Market Monitoring

reduced, the efficiency convergence bidding can add to the market is reduced, and the incentives for storage to participate in a manner aligned with system needs is reduced.

Competition in the Western Interconnection

Market participants that wish to export energy from the CAISO market can also add thousands of megawatts of demand bids. Recent reports published by CAISO Department of Market Monitoring show that there were up to 2,000 megawatts of price-sensitive export bids and another 3,500 to 4,500 megawatts of non-price-sensitive self-scheduled export bids submitted to CAISO in critical hours on August 17 through 19, 2020.[95] Entities that wish to export energy from California directly compete with others in California to purchase energy in the marketplace.

"Who has the supply and how can we get it to stay in California? Everybody is fighting for the same megawatts. It's been a tough journey."

- Marci Palmstrom, Director of Trading and Market Operations, Southern California Edison[94]

In the past, the CAISO has been able to rely on import energy from its neighbors in the West to supplement in-state supply, but there are strong indications that it will be challenging to attract regional supply in the future. Several studies, discussed further below, find growing capacity shortages across the West. These shortages will necessitate more competitive prices to induce neighboring systems to provide additional supply when it is needed. However, CAISO's market prices are falling short of bilateral energy prices in the West. The results of this price disparity can be readily observed: A diverse set of resource owners controlling 6,000 megawatts of supply have already found it more attractive to commit supply to Arizona load-serving entities than to California load-serving entities.

Supply margins across the Western Interconnection are shrinking

Several studies conclude that the Western Interconnection is short capacity today or will be short soon. Consumers in the Western Interconnection are actively navigating large thermal unit retirements and acquiring significant amounts of renewable generation with less certain availability. The Northwest Power and Conservation Council found the Northwest supply would likely become inadequate by 2021 and the shortage would

---

[94] July 8, 2021 joint agency workshop with the California Public Utilities Commission, the California Independent System Operator and other energy industry representatives, As reported by NewsData, "Officials Reduce Hydro Projection, Find Potential Grid Power Shortage This Summer"

[95] *Report on System and Market Conditions, Issues and Performance: August and September 2020*, Published November 24, 2020 by the CAISO Department of Market Monitoring

worsen in the following years. This study assumed a conservative level of available Southwest market import supply but noted that decreasing Southwest supply would raise the need for over 1,000 megawatts of additional capacity.[96] The Pacific Northwest Utilities Conference Committee found that annual summer capacity decreases will leave the northwest barely resource adequate through 2022 with growing deficits over the following eight years.[97] The Bonneville Power Administration found that the Federal system in the Northwest and the region as a whole face annual energy deficits throughout their 10-year study period.[98] Other studies come to similar conclusions.[99]

The Central, Desert Southwest, and California-Mexico regions will remain locked in an intense competition for supply for the foreseeable future. The Western Electricity Coordinating Council found each of these regions may experience 30 to 425 hours in which demand is at risk of not being served with internally available resources. In 2021 and beyond, even with all planned resource additions, these regions need external assistance to maintain resource adequacy. Accounting for potential external assistance, Southern California and the Desert Southwest still fall below the study's resource planning threshold.[100]

Notably, each study clearly demonstrates that each Western region will need to rely on external assistance from other regions to avoid compulsory demand curtailment. When unfavorable regional conditions collide, compulsory demand curtailment will occur, and supply will be rationed to the consumers that value it most.

During tight supply conditions consumers outside of California are increasingly willing and able to pay more than the CAISO to secure western supply.

Over the past four years, bilateral prices at the Palo Verde Hub have become more attractive than prices in CAISO. Both the magnitude and the frequency of higher prices at Palo Verde have generally been trending up. To the extent that CAISO's prices during critical time periods do not reflect CAISO's needs, its current market design may not attract sufficient energy to serve CAISO load.

---

[96] *Pacific Northwest Power Supply Adequacy Assessment for 2023*, Published on June 14, 2018 by the Northwest Power and Conservation Council

[97] *Northwest Regional Forecast of Power Loads and Resources 2021 through 2031*, Published in April 2021 by the Pacific Northwest Utilities Conference Committee

[98] *2019 Pacific Northwest Loads and Resources Study*, Published in October 2020 by the Bonneville Power Administration.

[99] *2021 Long-Term Summer Reliability Assessment*, Published in May 2021 by the North America Electric. Reliability Corporation. *Long-Term Assessment of Load-Resource Balance in the Pacific Northwest*, Published on October 31, 2018 by E3 and Portland General Electric. *Resource Adequacy in the Pacific Northwest*, Published in March 2019 by E3. *The Western Assessment of Resource Adequacy Report*, Published on December 18, 2020 by the Western Electricity Coordinating Council.

[100] *The Western Assessment of Resource Adequacy Report*, Published on December 18, 2020 by the Western Electricity Coordinating Council.

Efficient Market Prices During Tight Supply Conditions

Table 1 below compares the number of days when Palo Verde Hub prices were greater than SCE default load aggregation point (DLAP) prices by more than 10% over the past four summers. It also compares the number of days when Palo Verde Hub prices were lower than SCE DLAP prices by more than 10% over the past four summers.

Table 2 below compares the average magnitude of these price differences over the past four summers. Consumers outside of California are increasingly willing to pay more than the CAISO to secure Western supply.

Table 3: Percent of peak days where Palo Verde Hub prices are higher or lower than integrated SCE DLAP prices

Year	Percent of peak days where PV Hub Prices less than integrated SCE DLAP Prices by more than 10%	Percent of peak days where PV Hub Prices greater than integrated SCE DLAP Prices by more than 10%
June through September 2018	71%	8%
June through September 2019	35%	33%
June through September 2020	28%	35%
June through September 2021	22%	27%

Table 4: Average price differences on summer peak days each year

Year	Average peak day price difference where PV Hub Prices are less than integrated SCE DLAP Prices by more than 10%	Average peak day price difference where PV Hub Prices are greater than integrated SCE DLAP Prices by more than 10%
June through September 2018	($18)	$14
June through September 2019	($10)	$9
June through September 2020	($10)	$36
June through September 2021	($17)	$45

Most recently, the growing premium for energy outside of California can be seen by comparing prices in June 2021. The graphs below compare CAISO's integrated 16-hour peak prices to nearby bilateral hub prices during tight western supply conditions. During tight supply conditions in the west, CAISO's prices lag prices outside California.

Figure 9 compares PG&E DLAP prices to Mid-Columbia hub bilateral prices in June 2021. In this month, there were 8 out of 22 peak days where Mid-Columbia bilateral prices were higher than PG&E DLAP prices by more than 10%. There were very large premiums for energy at Mid-Columbia over three distinct time periods during the month.

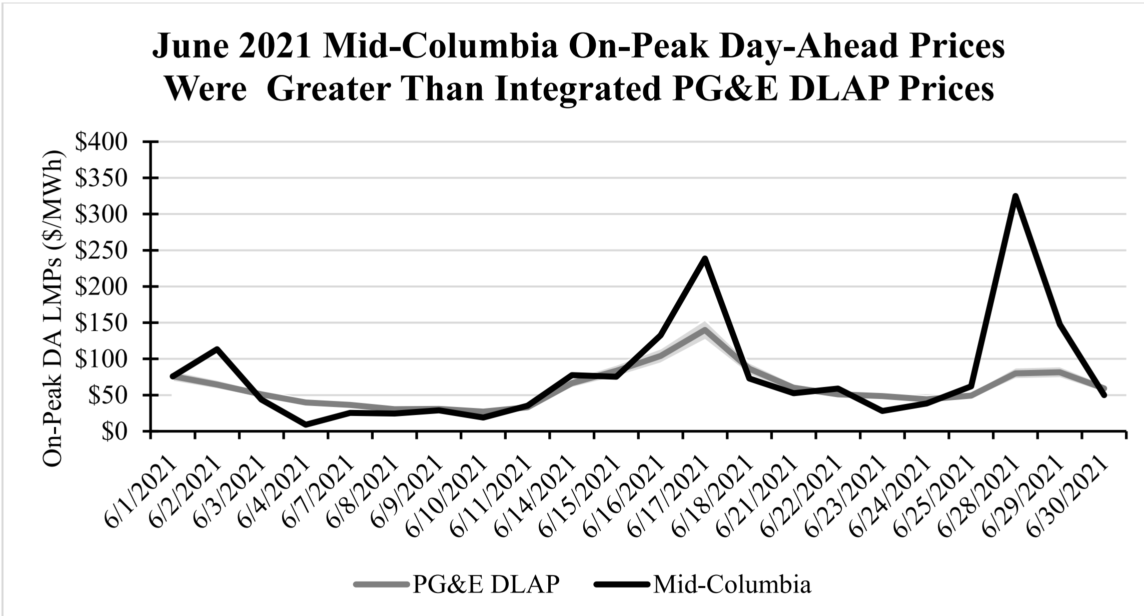

Figure 10: Comparison of On-Peak Day-Ahead Prices at Mid-Columbia and PG&E DLAP in June 2021

Figure 10 compares SCE DLAP prices to Palo Verde hub bilateral prices in June 2021. While prices remained fairly aligned, on the tightest supply days Palo Verde bilateral prices were 247% higher than SCE DLAP prices. There were very large premiums for energy at Palo Verde mid-month. Palo Verde Hub bilateral prices were only less than SCE DLAP prices by more than 10% on three days this month.

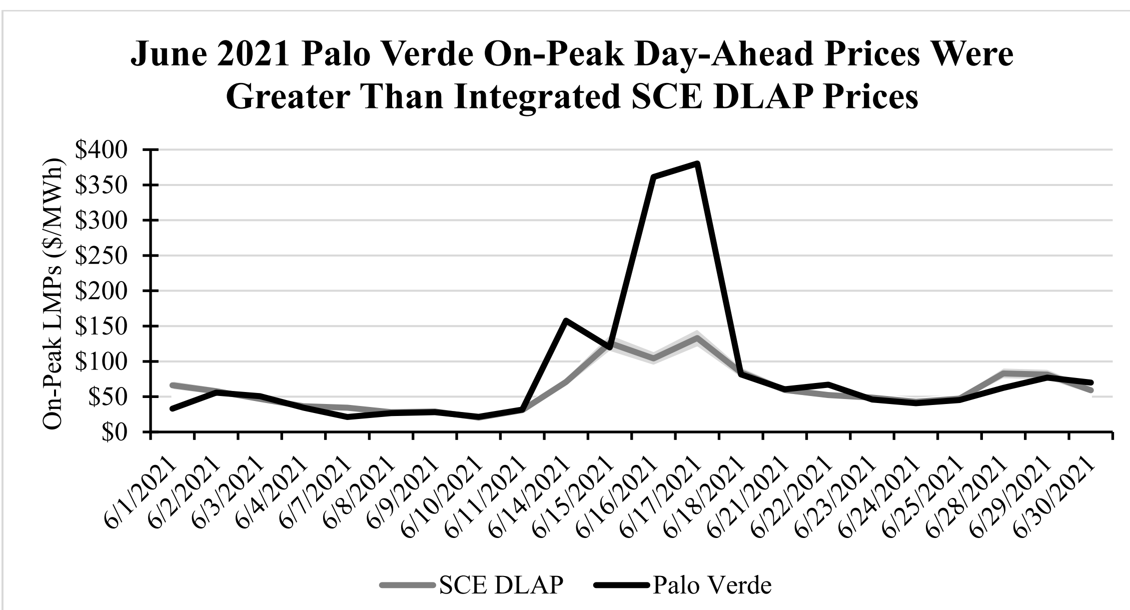

Figure 11: Comparison of On-Peak Day-Ahead Prices at Palo Verde and SCE DLAP in June 2021

In summary, on days where CAISO most needs supply to meet demand and reserve requirements, its prices are falling short of competitive bilateral prices outside of

California. On days like these, prices are so attractive outside California that CAISO may not be able to attract import supply, let alone retain supply from in-state non-resource adequacy resources. To the extent that CAISO's prices during critical time periods do not reflect the value of supply, it may lead to the misallocation of energy across the region.

Efficient Market Prices During Tight Supply Conditions

It is becoming more difficult for CAISO to attract in-state and out-of-state resource adequacy capacity due to higher out-of-state firm energy premiums

While this report focuses on energy market reforms to address real-time operational issues, the CAISO should not foreclose additional resource adequacy reforms to address planning issues. It is important to consider the distinction between the role of resource adequacy capacity and the role of the energy in the overall marketplace. These roles operate at different timeframes and are commonly described as serving a planning purpose and an operations purpose. Planning functions are performed over months, seasons, or years and operations functions are performed over weeks, days, and hours. Planning functions deal with more uncertainty and a wider range of possible outcomes, and operations functions deal with near-term realities. These two functions are related in that planning and its outcomes eventually translate to the operational timeframe. Put simply, the goal in the planning horizon is to ensure sufficient resources (and necessary attributes) exist and can be committed, whereas the goal of the operations horizon is to ensure that resources committed in the planning horizon are available and can be used at the right times and in the right places to meet the instantaneous, evolving needs to the system.

The scarcity pricing concepts in this paper are an operational solution to address an operational need. While evaluating issues and potential solutions, it is important to understand the relationship between planning and operations, identify which

Performance-based Resource Adequacy

In response to the 2014 Polar Vortex, which caused widespread generation failures during critical winter weather, PJM adopted a new capacity product called the Capacity Performance Product ("CPP"). CPP imposes technology-agnostic, no excuses (besides approved planned outages), and year-round performance requirements. Those obligations are tied to performance during any emergency action by PJM with steep, net CONE-based, financial penalties for underperformance. CPP has been in place for several years and is now the only available capacity product in PJM.

MISO, as a part of its Resource Availability and Need ("RAN") initiative, has been pursuing capacity accreditation changes that it is close to filing called Seasonal Accredited Capacity ("SAC"). The key differences with PJM's CPP are two-fold. First, MISO's changes will only apply to thermal generation and will not directly impact intermittent or use-limited resources. Second, SAC will impose a reduction in accredited capacity based on an individual resource's availability during tight margin hours (80% of accreditation from the top 3% of hours) rather than impose a financial penalty. Planned outages will only be excused under specific circumstances.

What these two schemes have in common is that they are both capacity market reforms targeted at resources with resource adequacy commitments. In each case, the goal of the design is to provide resources a capacity-based incentive to perform during critical times.

horizon observed issues impact, and select appropriate reforms to address the issue. In other words, operations issues should not be addressed with a planning solution. Efficient energy pricing coordinates the efficient allocation and dispatch of resources on an operational timeframe. That is not to say that CAISO may not have planning issues that require planning reform. For example, other RTOs have and are actively pursuing capacity accreditation or performance reform. However, if such reforms were adopted by CAISO with consultation with the CPUC and CEC, these planning improvements would not substitute the need for scarcity pricing. Similarly, scarcity pricing reforms does not substitute the need for possible capacity accreditation or performance reform.

Nonetheless, CAISO's current energy price formation is making it more difficult for CAISO to attract in-state and out-of-state resource adequacy capacity due to higher out-of-state firm energy premiums. Because the provision of import RA capacity entails a commitment to provide energy to California, suppliers of import RA capacity generally require compensation for the foregone opportunity to sell energy outside of California.

Supply margins outside California are shrinking driving elevated firm energy prices in those markets.

Two obvious, but important, concerns arise in this environment. First, out-of-state supply requires much higher capacity compensation to take on a commitment to CAISO. Second, in-state supply not owned by in-state load-serving entities requires much higher capacity compensation to continue committing their production to the CAISO. Recent examples illuminate these concerns.

Attracting out-of-state supply. Based on July 2021 forward firm energy price spreads, out-of-state suppliers required at least $68.48/kW-month to commit themselves to CAISO. This is nine times higher than a low estimate of the cost of new entry[101] and four and a half times higher than the current costs of failing to meet resource adequacy requirements.

As of May 3, 2021, July 2021 forward firm energy prices at Palo Verde were $253.25/MWh and forward firm energy prices at SP15 were $125.50. Suppliers with physical resources to sell at Palo Verde also receive a physical resource premium of approximately $25/MWh. Naturally, a supplier committing to import supply at SP15 rather than selling it at Palo Verde faces an opportunity cost of $278.25/MWh. It would also face approximately $10 of GHG costs and an average of $2 of intertie congestion costs. The difference between its proceeds and costs for committing to this arrangement would be a loss of $164.62/MWh. To accommodate this loss, the supplier would require

[101] The average gross cost of new entry across all zones in the Midcontinent Independent System Operator is $7.64/kW-month.

$68.48/kW-month in resource adequacy payments to commit its supply as import resource adequacy in CAISO.

Table 5: Calculation of required capacity proceeds to attract out-of-state supply

	Costs	Expected Proceeds
Palo Verde forward firm energy (opportunity cost)	$253.25/MWh	
Palo Verde physical resource premium (opportunity cost)	$25/MWh	
Greenhouse Gas Emissions	$10/MWh	
Intertie congestion costs	$2/MWh	
SP15 forward energy		$125.50/MWh
Total $/MWh	($164.62/MWh)	
Required capacity proceeds to attract out-of-state supply	$68.48/kW-month[102]	

In ***Figure 11*** below, observe the required capacity proceeds to attract out-of-state supply for the summer of 2021 based on forward firm energy prices[103] at Palo Verde and SP15.

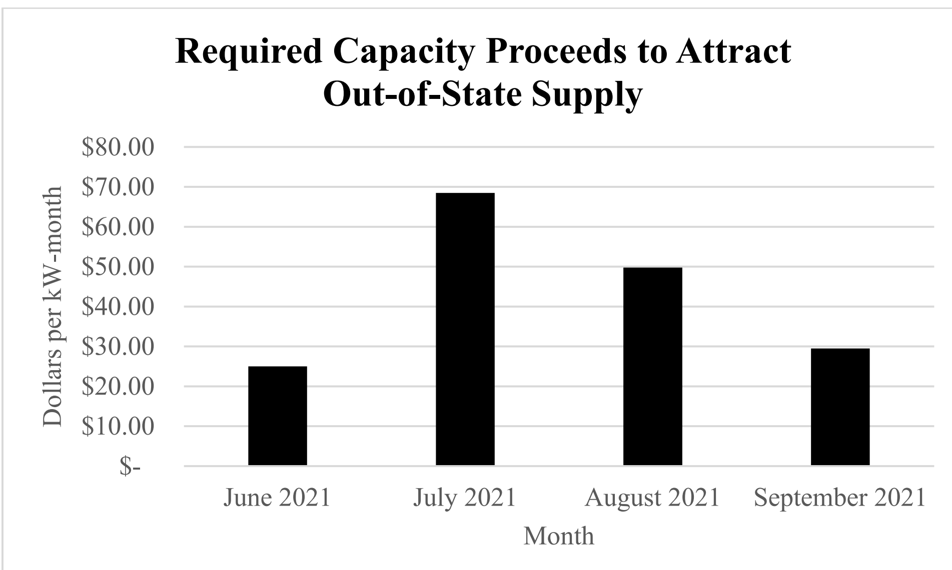

Figure 12: Required capacity proceeds to attract out-of-state supply

Retaining in-state supply. Given current forward firm energy price spreads, the CPUC resource adequacy program penalties and CAISO backstop procurement offer caps would have to rise to at least $56.89/kW-month for rational suppliers to continue committing in-

[102] Assumes 416 peak hours per month.
[103] Forward prices were captured as of the first business day of the month two months prior to the delivery month.

state generation to the CAISO. This is seven and a half times higher than a reasonable estimate of the cost of new entry[104] and three and a half times higher than the current costs of failing to meet resource adequacy requirements.

As of May 3, 2021, July 2021 forward firm energy prices at Palo Verde were $253.25/MWh and forward firm energy prices at SP15 were $125.50. In-state suppliers with physical resources to sell at Palo Verde also receive a physical resource premium of approximately $25/MWh. A supplier seeking to export its energy to Palo Verde would face approximately $14/MWh in export fees and an average of $2 of intertie congestion costs. The difference between its proceeds and costs for committing to this arrangement would be a gain of $136.75/MWh. The supplier requires $56.89/kW-month in resource adequacy payments to continue committing its supply as resource adequacy in CAISO.

Table 6: Calculation of required capacity proceeds to retain in-state supply

	Costs	Expected Proceeds
Palo Verde forward firm energy		$253.25/MWh
Palo Verde physical resource premium		$25/MWh
Export fees	$14/MWh	
Intertie congestion costs	$2/MWh	
SP15 forward energy (opportunity cost)	$125.50/MWh	
Total $/MWh	$136.75/MWh	
Required capacity proceeds to retain in-state supply	$56.89/kW-month[105]	

From a purely economic standpoint, suppliers should not expect California load-serving entities to pay much more than $15.19/kW-month for resource adequacy capacity. California load-serving entities currently face $8.88/kW-month in regulatory penalties if they fail to meet their resource adequacy requirements.[106] They also are responsible for paying their share of CAISO's backstop capacity procurement costs which are soft-capped at $6.31/kW-month and offers above that value cannot include a supplier's going forward costs. At capacity prices greater than $15.19/kW-month, it is economically

[104] The average gross cost of new entry across all zones in the Midcontinent Independent System Operator is $7.64/kW-month.
[105] Assumes 416 peak hours per month.
[106] The California Public Utilities Commission recently issued a decision (Decision 21-06-029 on June 24, 2021) that escalates this penalty to double or triple this value for repeated violations of minimum resource adequacy requirements over a two-year period. Even so, load-serving entities could be deficient 5 out of the 6 peak summer months and still only face a $8.88/kW-month regulatory penalty.

prudent for load-serving entities to forgo meeting their resource adequacy requirements and focus on energy hedging and procurement.[107]

In ***Figure 12*** below, observe the required capacity proceeds to retain in-state supply for the summer of 2021 based on forward firm energy prices[108] at Palo Verde and SP15.

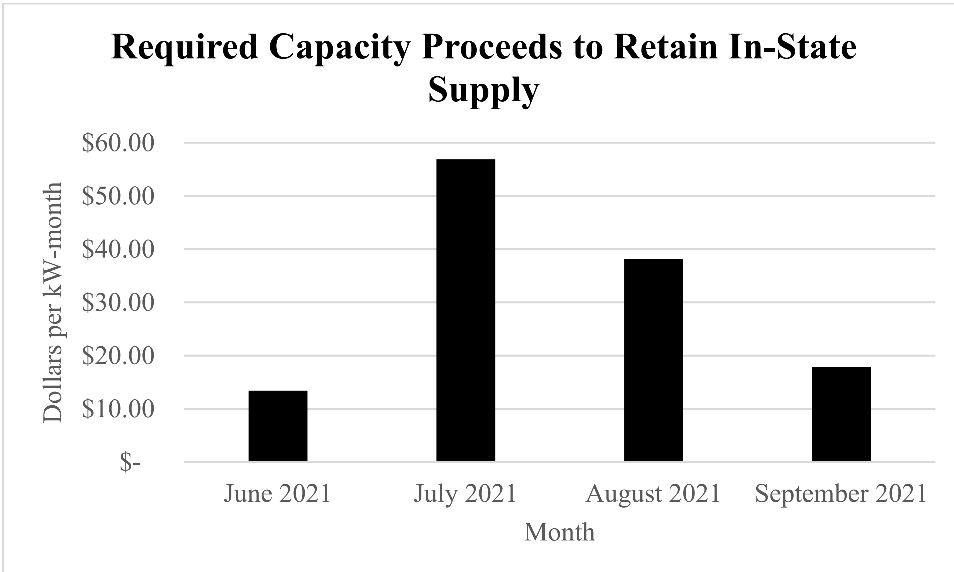

Figure 13: Required capacity proceeds to retain in-state supply

By themselves, penalties and backstop procurement offer caps would have to rise much higher to allow California to keep supply in state. Allowing capacity prices to rise much higher (e.g., a multiple of the cost of new entry) could be a solution to this issue but has been considered and rejected in the past. However, some combination of energy pricing reforms as well as capacity pricing reforms is necessary to allow California consumers to compete successfully with out-of-state consumers. Some entities will attempt to build themselves out of the capacity shortage conditions. New natural gas plants, renewable generators, and batteries are planned to be built across the West and California is warming up to a 20 percent planning reserve margin. If the west-wide capacity shortage proves to be transitory this new mix of supply and demand still requires accurate and reliable price signals which will require energy pricing reforms.

[107] It should be noted that some load-serving entities may be willing to procure in-state resource adequacy capacity at higher prices due to other factors such as the risk of bad press or the risk that the CPUC may take action to revoke their load-serving entity status. Also, a load-serving entity that directly owns supply may not wish to sell firm energy out-of-state to avoid risk of compulsory demand curtailments, which they would value at the value of lost load.

[108] Forward prices were captured as of the first business day of the month two months prior to the delivery month.

Previously available merchant import capacity no longer available

Unfortunately, because of lower energy prices and insufficient capacity revenues in the CAISO, there is an emerging trend for resource owners to commit their energy production to consumers outside of California, rather than offer the energy as imports into CAISO as they have done in the past. For example, a diverse set of resource owners controlling 6,205 megawatts of supply in Arizona have already found it more attractive to commit supply to Arizona consumers than to market their energy to California consumers. Below, ***Table 5*** shows the shift in energy commitment for various resources between 2015 and 2021.

Table 7: Comparison of merchant supply to committed supply between 2015 and 2021

Resource	Owner	2015 Merchant Capacity	2021 Commitments	2021 Merchant Capacity
Gila River	Panda	2,200 MW	1,650 MW SRP 550 MW TEP	0 MW
Mesquite	Sempra	1,250 MW	625 MW SRP 625 MW SWPPR	0 MW
Harquahala	PG&E NEG	1,100 MW	366.66 MW TEP 366.66 MW SRP 366.66 MW APS	0 MW
Griffith	PPL/Duke	570 MW	570 MW APS	0 MW
Arlington Valley	Duke	565 MW	565 MW APS	0 MW
SouthPoint	Calpine	520 MW	520 MW APS	0 MW
Total		6,205 MW	6,205 MW	0 MW

The CAISO must make substantial market pricing reforms to respond to the marketplace evolution and meet its reliability imperative

The California supply evolution has placed considerable strain on the CAISO system and fundamentally changed the operational characteristics that the CAISO values the most. The CAISO must now respond to net demand variability throughout the day. It must now replace large amounts of grid-scale and behind-the-meter solar in the evenings. It must also operate resources to avoid many different resource energy limitations. To meet these challenges, the CAISO must now attract and retain fast-ramping supply that can quickly change directions, supply with wide continuous operating ranges that can sustain upward and downward ramps, and supply that can be online and available at a moment’s notice. The CAISO needs more attractive scarcity prices to signal when it values these capabilities.

The California demand evolution brought a large and diverse set of energy purchasers to the market, each with varying objectives and levels of risk-aversion. There has been considerable fragmentation among the large energy purchasers, yet the price at which each would be willing to forgo consumption is still unknown. Load-serving entities, demand response providers, energy storage operators, out-of-state energy purchasers, and financial entities all compete to purchase energy in the market. With increased demand fragmentation, growing levels of price-sensitive demand, and an interest in California export energy, the CAISO needs an economically rational way to allocate supply to consumers that value it the most.

The Western Interconnection is facing capacity shortfalls for the foreseeable future. Each western region will need to rely on external assistance from other regions to avoid compulsory demand curtailment. When unfavorable regional conditions collide, demand curtailment will occur, and supply will be rationed to the consumers that value it most. The CAISO needs more attractive scarcity prices to allow its consumers to compete for energy in a competitive western interconnection.

The CAISO cannot rely on status quo pricing to attract and retain supply when it is needed. To the extent that market prices are not reflecting the true reliability value of supply leading into and during tight supply conditions, the CAISO will face increasingly difficult operating conditions, suppliers will increasingly not be held accountable for the harm unavailable supply causes to the system, load-serving entities will increasingly be expected to cross-subsidize each other for inefficient demand scheduling, and CAISO will increasingly find it harder to procure power in the west.

The best way to resolve these issues is to pursue *efficient* day-ahead and real-time market pricing. Efficient market prices will promote and reflect the optimal dispatch of available resources, provide powerful scheduling and performance incentives, and accurately signal the need for participation of additional supply or reductions in consumption. Achieving efficient market prices will bring long-term benefits to consumers in California.