SINGAPORE FORWARD CAPACITY MARKET

FCM Design Proposal

Third Consultation Paper

PREPARED FOR

Energy Market Authority

Smart Energy, Sustainable Future

PREPARED BY

Sam Newell

Judy Chang

Kathleen Spees

Walter Graf

John Imon Pedtke

Matt Witkin

Carson Peacock

May 2020

THE Brattle GROUP

Notice

- This report was prepared for the Energy Market Authority of Singapore (EMA), in accordance with The Brattle Group's engagement terms, and is intended to be read and used as a whole and not in parts.
- The report reflects the analyses and opinions of the authors and does not necessarily reflect those of The Brattle Group's clients or other consultants.
- There are no third party beneficiaries with respect to this report, and The Brattle Group does not accept any liability to any third party in respect of the contents of this report or any actions taken or decisions made as a consequence of the information set forth herein.

Table of Contents

I. Introduction

Purpose of This Document

The Energy Market Authority of Singapore (EMA) is proposing a Forward Capacity Market (FCM) to address concerns in the current Singapore Wholesale Electricity Market (SWEM). The EMA has retained The Brattle Group (hereafter "Brattle" or "we"), an international economic consulting firm, to assist in the design of a FCM. This document represents the third public version of a design proposal for the FCM, and it provides stakeholders an opportunity to provide feedback on each market design element.

Context and Objectives

The SWEM is currently an energy-only market (EOM) with ancillary services. Generation companies are remunerated primarily based on prevailing half-hourly spot prices for energy generated. By design, the EOM provides short-term price signals to guide both operations and investments in generation capacity. However, the concern is that wholesale electricity spot prices may not attract sufficient and timely investment in generation capacity to support resource adequacy, *i.e.*, to meet the required reserve margin corresponding to the reliability standard.

Other jurisdictions with similar concerns have implemented FCMs to ensure resource adequacy. The concept is to express the demand for capacity in a forward auction, and let suppliers compete to meet that demand at the lowest price. In combination, the real-time wholesale energy and ancillary services markets, and FCM, aim to meet the following objectives:

- Maintain resource adequacy by providing adequate incentives to existing and new resources; and
- Maximize economic efficiency to minimize long-run costs to consumers.

The components of the FCM jointly support these objectives by clearly expressing a demand for the capacity product and encouraging suppliers to compete to offer that product at lowest cost. The product definition in an FCM is simply a megawatt (MW) of capacity supply obligation (CSO) to be available and to offer into the real-time energy and/or ancillary services market, for a year, subject to penalties for failing to perform. Broadly, the three main components of the market are: (1) a demand curve for capacity, (2) the rules defining how suppliers participate and form a supply curve, and (3) the format of the auction in which supply and demand come together to determine which resources clear the market and the prices at which they are paid.

Demand for capacity expresses how much capacity to buy as a function of price. The FCM demand curve is developed to ensure sufficient capacity is procured to meet the reliability standard. It is designed to avoid procuring substantially more capacity than needed, and to allow prices to rise to attract new resources when necessary. It slopes upward to the left when supply is relatively scarce, and downward to the right in surplus, low-cost conditions.

In order to maximize competition and innovation to meet resource adequacy at least cost, **supply participation** should be open to existing and new resources across a wide range of technologies. Resources can qualify to participate if they pass certain eligibility criteria, and the qualified capacity each resource may offer reflects the marginal reliability value it provides (*e.g.*, derated from nameplate to the extent a resource is unavailable due to outages or intermittency). Each participating resource then provides an offer in terms of dollars per MW of **qualified capacity**, and the **supply curve** is formed by arraying the supply offers in ascending order. In addition, EMA has determined that it will place some limits on the quantities of different types of supply that may clear the auction.

Offer prices may be capped by the market monitor to mitigate the exercise of market power. Similar to real-time energy markets (during tight supply conditions), capacity markets are susceptible to the exercise of market power because available supply typically exceeds demand by small margins, such that even medium-sized suppliers could withhold capacity profitably, unless required to offer competitively. In principle, competitive offers would reflect resources' avoidable going-forward fixed costs after considering net revenues from selling energy and ancillary services.[1] In the long run, wholesale market revenues from the FCM, energy market, and ancillary services markets should be sufficient to recover the long-run marginal cost of capacity, including fixed costs. However, once certain fixed investment costs have been incurred, competitive market participants should exclude these costs from their offers (as they would be incurred regardless of receiving a CSO, so they are not marginal or additional). Resources' non-avoidable costs are recovered any time a resource is infra-marginal in the FCM—that is, when higher-cost capacity clears the FCM and all cleared resources receive the marginal clearing price.

The auction itself brings together the ascending supply and the descending demand curve in order to clear the market. The auction clears at the point where the supply and demand curves intersect. That clearing point determines which resources clear and accept a CSO—all those with offers at or below the clearing price.

The capacity auction must take place prior to the delivery year. Other jurisdictions vary considerably in how far ahead they conduct the auction. For Singapore, we propose a four-year forward period, corresponding to the lead-time for constructing a new combined cycle gas turbine (CCGT). This enables new generation to compete with existing resources. Such advance commitment also resolves uncertainties regarding the potential retirement of existing supply in time for new generation capacity to replace it. Subsequent to the forward auctions, rebalancing auctions would be held nearer to the delivery year to efficiently address changes in demand requirements or supply availability.

---

1 Net avoidable going-forward fixed costs are net costs that a resource could avoid if it did not have a capacity supply obligation. It is important to note that mothballing or retiring a generation resource may not avoid all fixed costs. For example, a take-or-pay fuel contract may be considered a non-avoidable fixed cost in that payment is required even if the generator does not produce electricity, and payment cannot be avoided by a retirement or mothball decision. In addition, property taxes and some insurance may be unavoidable for plants that mothball. Overall, any costs that are unavoidable would not vary depending on whether the plant stays online, and the capacity payment does not need to cover those costs in order to be willing to stay online.

Having considered stakeholder feedback and EMA's policy guidance, the proposal for each market design element in the Singapore FCM is presented in Table 1. Each element is discussed in more detail in the subsequent sections.

Table 1: Overview of FCM Market Design Proposal

Market Design Element	Design Proposal	Justification
Product Definition	• Capacity product defined in terms of MW-year of **capacity supply obligation (CSO)**; "qualified capacity" (QCAP) reflects expected availability, as addressed below (see Supply Participation). • A CSO entails a requirement to supply energy and/or ancillary services when needed, subject to penalties for being unavailable or otherwise not performing.	• Product definition must correspond to the MW "demand" for resource adequacy. • Product must have clear obligations consistent with reliability objectives.
Administrative Demand Curve	• Demand reflects the **peak load forecast plus required reserve margin** corresponding to the reliability standard (3 Loss of Load Hours). • **Downward-sloping straight line demand curve** with the quantity at the price cap set to the minimum acceptable reliability level, then sloping downward to the right; rest of the curve tuned to meet various price and quantity demand curve design objectives, under an assumed Net Cost of New Entry (CONE).[2] • **Price cap** established in the range of 1.5x to 1.75x estimated Net CONE; minimum on the cap set between 0.5x to 1× estimated Gross CONE to protect against Net CONE estimation error. • **Periodic comprehensive review** of Gross CONE, energy and ancillary services (E&AS) offset[3], and demand curve parameters. • Implement **annual updates** based on a formulaic approach. Update Gross CONE based on available public index, forward-looking E&AS offset with most recent market data, and demand curve parameters with new load forecasts and reliability analysis.	• The objective is to meet the reliability standard. • A downward-sloping demand curve reduces price volatility and recognizes incremental marginal reliability value at varying reserve margins. • Cap must be sufficiently high to express higher marginal value at low reserve margins, to mitigate the possibility of underestimating true Net CONE, and to shift the distribution of reserve margin outcomes rightward to express lower marginal value at excess capacity. • Net CONE parameters need to be adjusted to market conditions. • Demand curve performance should be evaluated in relation to design objectives (reliability, price rationality, price stability, and regulatory stability).
Supply Participation	• **Qualify all resources that can contribute to resource adequacy**, including demand response, imports, storage; both existing and new. • **Qualified MW ratings** account for unplanned and planned outage rates, intermittency, and energy-limits (applicable to storage & demand response). • **Supply curve** aggregates all supply offers in ascending order. • EMA will place a lower limit of minimally 9,000 MW on the amount of frequency responsive capacity from traditional generation resources, and a maximum limit on the amount of demand response and storage resources of 200 MW each.	• Enables efficiency, competition, and innovation. • "QCAP" is a uniform product, with all MW competing to provide the same marginal reliability value. • The proposed constraints on the amount of capacity from each technology type are intended by EMA to ensure stability of the power system.

[2] Net CONE is an administrative estimate of the long-run marginal cost of capacity ($/kW-year) from a reference resource based on the generation technology most likely to enter the market. It includes capital recovery plus the fixed and variable costs of operation for a new resource, net of expected revenues received from the energy and ancillary services markets.

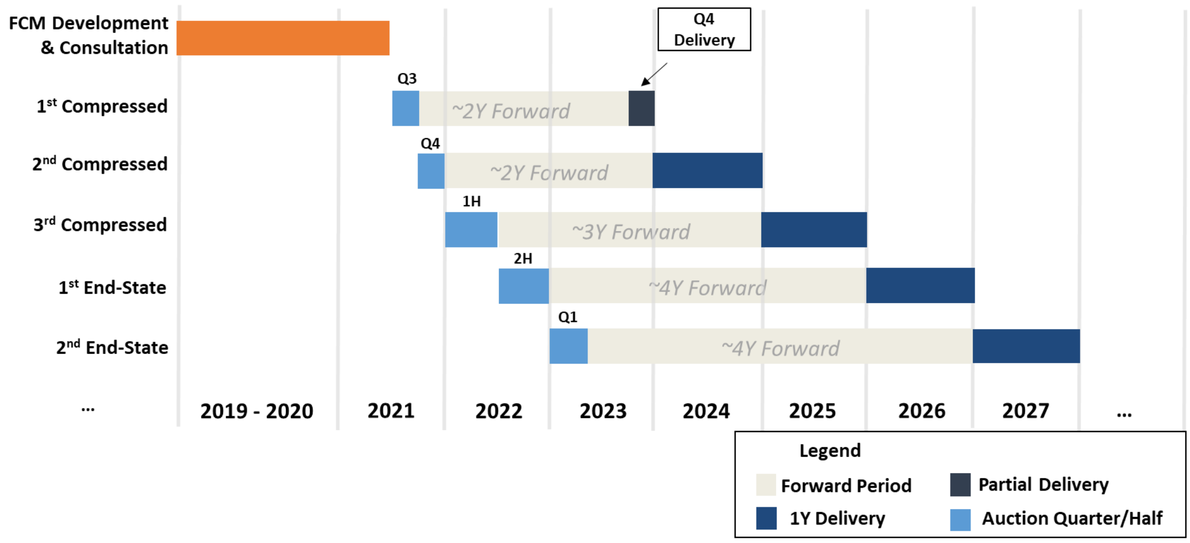
[3] The energy and ancillary services (E&AS) offset reflects the expected net revenues (or revenues minus variable costs) that the resource would expect to earn from participating in the E&AS markets.

Market Power Monitoring and Mitigation	• **All existing resources with an installed capacity of 10 MW and above must offer** capacity. • **Screen suppliers to detect supply-side market power** using one-pivotal supplier test. • **Mitigate offer prices of those that fail** the market power screen to a pre-defined threshold.	• Must-offer requirement and mitigated offers prevent supply-side market power abuse (buyer-side market power abuse unlikely in Singapore).
Forward Capacity Auction	• **Uniform price auction** whereby all cleared suppliers earn the same price. • **Single round, sealed bid** auction. • **Four-year forward period**. • Enable the use of multi-block offers and allow specification of the first block as non-divisible. • Clear the auction to maximize social surplus, subject to (i) minimizing consumer cost when the marginal offer is non-divisible; and (ii) procuring at least the minimum acceptable reliability level. • Set price at the value of the demand curve in cases where the entire cleared supply curve lies below the demand curve. • Provide a 10-year multi-year commitment for new/repowered CCGT resources with an economic lifespan of at least 25 years and which meets the proposed heat rate standard for power generation.	• Uniform price, single-round, sealed-bid auctions maximize competition; has a proven record of delivering efficient market outcomes. • Enabling flexible offer formats allows resources to accurately represent their costs. • EMA aims to facilitate investment in efficient CCGTs to meet energy demand more efficiently and to provide reliable online reserves.
Rebalancing Auctions	• **Rebalancing auction(s)** conducted between the base auction and delivery year. • Auction would be **cleared on a gross basis** (with all supply and demand included), and settled on a net basis (*i.e.*, changes in quantities from the base auction would settle at the rebalancing auction clearing price) • **Supply offers** would include: – Any incremental supply that did not clear in the base auction; – Any supply with an existing CSO that wishes to buy out at a non-zero price, including supply with a QCAP derate that is required to buy out at the price cap; and – Any supply with an existing CSO that does not want to change its position, participating as price takers. • **Demand** would include: – The updated auction demand curve reflecting an updated load forecast.	• Provides an opportunity to adjust capacity commitments with changes in demand and/or supply availability.
Bilateral Transactions	• Enable buyers and sellers to engage in bilateral exchange of CSOs during the forward period and delivery year.	• Facilitate market participants in managing their own risks and uncertainties.
Supply Obligations and Performance Penalties	• Suppliers are obligated to demonstrate availability consistent with their obligations, and face penalties for under-performance. • Penalty rates will be high enough to incentivize performance (but not so high as to impose undue costs that discourage participation).	• An appropriate penalty system will ensure capacity obligations are appropriately fulfilled and supply is available during shortage conditions.
Settlements and Cost Allocation	• Higher costs allocated to consumers in proportion to their consumption during peak (and potentially also mid-peak) hours of the year.	• Consumption during these hours drives the need for capacity, and cost allocation should reflect cost causation.
Reforms to Energy, Ancillary Services	• Consider conforming changes to the E&AS markets, including potentially mitigating energy offers more strictly to reflect competitive outcomes.	• Emulates a perfectly competitive market, with the FCM supporting recovery of fixed costs and E&AS markets supporting recovery of variable costs.

IMPLEMENTATION TIMELINE

EMA has reviewed the indicative timeline to develop and implement the proposed FCM. Considering feedback from stakeholders, including the Energy Market Company (EMC)[4], on the development timeline for the market rules and IT systems, as well as the need to conduct pre-auction processes (including resource qualification, market trials and market power mitigation), the updated indicative timeline will work toward conducting the first auction in Q3 2021 with a "compressed" forward period of two years. The forward period gradually extends over time until reaching the "end-state" of a four-year forward period, as illustrated in Figure 1. This design allows for a phased approach in the implementation of the FCM. With the experience gained from each auction, EMA intends to further adapt and incorporate enhancements or refinements where appropriate to the FCM's features and rules, to ensure alignment with policy objectives.

Figure 1: Implementation and Transition Timeline to Full FCM



II. Product Specification

The product definition specifies exactly what each resource in the market is obligated to provide if it clears the auction. Consistent with the concept of "capacity," the product should be 1 MW of capacity supply obligation (CSO) for a year. A CSO requires the resource to offer into the real-time energy market (and/or ancillary services markets) when available, subject to penalties for unavailability and non-performance.

We recommend defining the capacity product such that each unit of capacity transacted represents a MW of capacity, normalized for expected unavailability. In reality, all resources are affected by planned and unplanned outages, and for other reasons that they cannot always produce at their full capability, so the amount of capacity they qualify to sell will generally be lower than their installed capacity. Thus, each MW of qualified capacity will have the same reliability value per MW as another MW of qualified capacity. The discrepancy between

---

4 EMC is the intended operator and administrator of the FCM.

installed capacity and qualified capacity accounts for each resource's outage rates, intermittency, and other factors affecting reliability value, as described in Section IV. The qualified capacity naturally forms the basis for any performance penalties, discussed in detail in Section X. This creates a uniform product for which all resources can compete and be compensated fairly and be accounted for appropriately when procuring capacity to meet the reliability standard.

In general, capacity products could be more multi-faceted and varied to specify certain sub-products with specific characteristics (such as fast-start capacity), locational products, seasonal or time-of-day products. We recommend adopting a simpler approach with an annual product with no locational requirement and no additional specifications.

This proposal for a relatively simple product is suitable for the supply and demand dynamics in Singapore's electricity market:

- Locational capacity differentiation is not recommended at this time due to limited persistent transmission constraints during peak conditions that would preclude a unified market for capacity. This design choice can be re-evaluated in the future.
- Seasonal capacity product differentiation is unnecessary because load and supply availability do not differ greatly across the year.
- Resources clearing the auction will receive the obligation to supply capacity for a pre-defined period, the "commitment term." A commitment term of one year is consistent with other international jurisdictions; a shorter commitment term would not provide sufficient revenue certainty and a longer commitment term could disadvantage resources that are not able to commit to a longer period.[5]

If certain resource characteristics are absolutely needed to operate the system, one option is to specify the need for them as sub-products in the capacity market. But if those characteristics are merely more valuable or convenient than substitutes (such as fast-start versus spinning reserves) then we recommend recognizing that value only in the ancillary services markets and/or in capacity ratings, rather than specifying sub-products for capacity. This avoids inefficiently biasing the resource mix and complicating the mechanics for resource qualifications.

The EMA has determined that a minimum amount of reserves from frequency responsive traditional generation resources (*e.g.*, CCGTs) is required to ensure stability of the power system. Brattle has not evaluated whether this constraint is necessary as a resource adequacy requirement, and we believe it is possible that the need for frequency response could be met by other technologies, such as load response on under-frequency relays, energy storage systems (ESS), or other technologies in the Singapore system.[6] Nonetheless, we understand EMA plans

[5] However, the EMA is proposing to allow some suppliers to lock-in their clearing price for multiple years, to improve investment incentives; this is discussed in more detail in Section I.C.

[6] Other jurisdictions such as ERCOT recognize in their frequency response regime that fast-frequency resources might provide more than 10 times as much frequency response per MW of capacity than

to impose a minimum requirement of at least 9,000 MW[7] of installed capacity on the amount of frequency responsive traditional generation resources, until the frequency response capability and reliability of other technologies have been proven in the Singapore power system. This minimum requirement will be reviewed prior to each auction by EMA.

In addition, we understand EMA plans to impose a cap of 200 MW on the amount of cleared capacity of demand response (DR), and 200 MW on the amount of cleared capacity of ESS. These limits will be in place for the first compressed auction in Q3 2021 for delivery period Q4 2023. This maximum cap will be reviewed prior to each auction by EMA, taking into account the track record and operational experience with more DR and storage resources in the Singapore system.

While the FCM design could be more technology-neutral, Brattle understands EMA's current preference for technologies with which it has the most experience in the Singapore system until newer technologies are sufficiently tested and proven therein. EMA believes these constraints are necessary in the initial years of the FCM, to avoid unintended outcomes and gain confidence in operating its system with an evolving resource mix.

III. Administrative Demand Curve

The capacity market demand curve establishes the willingness to pay at each quantity of capacity. The demand curve is designed consistent with the primary objective to procure sufficient capacity to meet forecasted annual peak load plus the reserve margin required to meet the reliability standard. The shape and width can be adjusted consistent with other objectives such as mitigating price volatility and limiting the ability to exercise market power. Consistent with the economics that drive private investment, the demand curve should produce prices high enough to attract and retain capacity when supply is needed, while avoiding over-investment when additional supply is not needed.

To evaluate the likely performance of alternative demand curves, we use a Monte Carlo probabilistic simulation model. This model simulates many potential future capacity market scenarios and outputs distributions of price, cost, and reliability outcomes for a selected demand curve. After quantitatively and qualitatively analyzing a larger range of options, we have

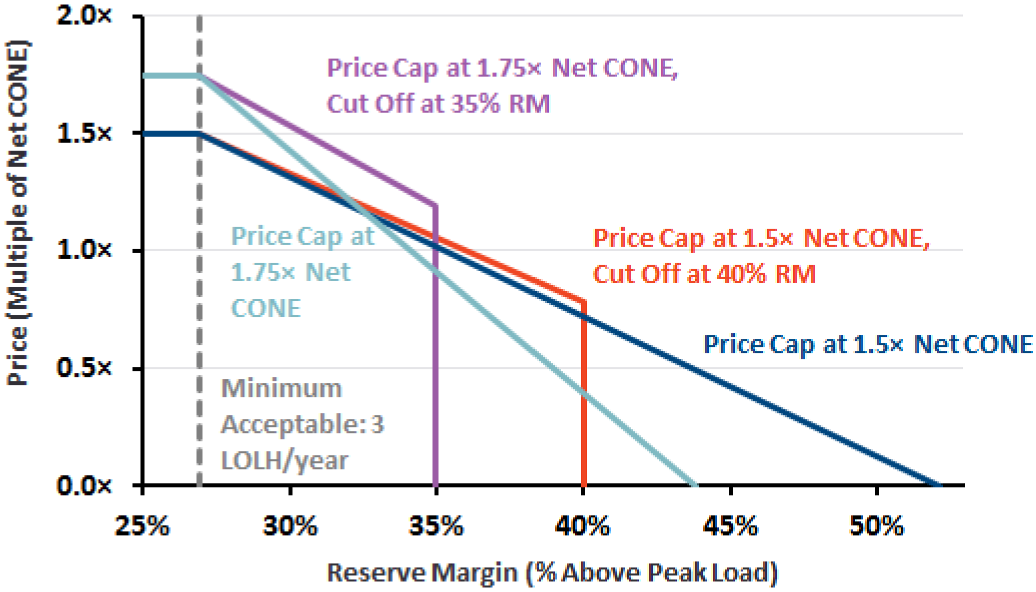
traditional generators (*i.e.*, 1/9%, or more considering the greater effectiveness of fast-frequency response in supporting system frequency post-contingency).

7 Based on EMA's Transmission Code, frequency sensitive generating unit are required to be capable of providing primary reserves of at least 9% of its rated MW capacity (when the unit's output is between its minimum stable load and 75% of its rated MW capacity). See EMA, "Transmission Code," Updated October 31, 2019. Available at: https://www.ema.gov.sg/cmsmedia/Licensees/Electricity/Transmission%20Code_31%20Oct%202019_04122019.pdf.

Considering the minimum primary reserves capability of 9% and the required reserve margin to cater for outages, EMA has determined that at least 9,000 MW of frequency responsive traditional generators would be required to provide spinning reserves to maintain the system frequency ≥ 49Hz, during a contingency event where the largest online generating unit (assumed to be 600 MW) trips.

shortlisted four candidate demand curves, shown in Figure 2 below. All of these curves would produce prices consistent with Singapore's reliability objectives and enable price formation consistent with Net CONE on a long-run average basis. Administrative estimates of Net CONE would be updated over time to maintain consistency with prevailing market conditions.

Figure 2: Candidate Singapore Demand Curves



In addition to achieving the minimum resource adequacy objective, the candidate Singapore demand curves have several features that will support a sustainable capacity market design. The curves have quantities higher than the minimum acceptable level for any price below the price cap, which ensures that the minimum procurement volume will be secured in the auction in a large majority of all forward auctions.

They also have price caps high enough to attract and retain supply when the market is tight, with a price cap minimum based on Gross CONE to prevent the curve from collapsing if the estimated net revenues from selling energy and ancillary services become very high (causing the Net CONE estimate to fall).

The curves' downward sloping shapes are consistent with the diminishing reliability benefits of incremental capacity at higher quantities. The curves are wide enough to control excessive price volatility and limit opportunities for the exercise of market power, but steep enough to limit over-procurement.

A. Singapore's Design Objectives

The primary objective of a demand curve is to support reliability by appropriately reflecting the reliability requirement, in addition to other objectives described in Table 2 below. We understand that EMA has established a reliability standard of no more than three expected Loss of Load Hours (3 LOLH) per year.[8] This is defined as the minimum acceptable reliability level

8 The reserve margin, corresponding to 3 LOLH per year, may fluctuate over time as fleet and load characteristics evolve.

for the Singapore market, meaning that the quantity procured from the capacity auction needs to be at or above this level in each year.

The demand curve adopted for Singapore must be consistent with these objectives while balancing trade-offs among consumer cost, price volatility, and quantity volatility. Some curves should be ruled out based on the inability to meet these objectives (*i.e.*, those that do not meet the primary objective of delivering reliability). However, there is a range of workable demand curves that align with these objectives, and in many cases, selecting a demand curve requires weighing the tradeoffs between steeper curves (that provide smaller risk of over-procurement, more quantity certainty, and lower consumer costs) and wider curves (that provide lower price volatility and reduced susceptibility to exercise of market power).

Table 2: Overview of Singapore Demand Curve Design Objectives

Design Objective	Description
Primary Objective: **Deliver Reliability**	• Ensure sufficient supply to meet the reliability standard of 3 LOLH expressed in QCAP terms. This is interpreted as a "minimum acceptable" reliability level
Send Efficient Price Signals	• Send efficient price signals to attract entry when the market is short, and discourage entry when the market is long[9]
Minimize Consumer Costs	• Ensure reliability but avoid over-procurement relative to target capacity
Mitigate Price Volatility	• Reduce price impact from small changes in supply and demand • Reduce the impact of lumpy entry/exit on market outcomes
Mitigate Susceptibility to Market Power	• Complement market power mitigation mechanisms to limit structural susceptibility to market power
Reflect Singapore's Unique Market	• Account for unique characteristics of Singapore's market (*e.g.*, smaller market size)

B. Approach and Key Assumptions

Demand Curve Approach

The three main approaches to designing a demand curve are described below and illustrated in Figure 3 (note that the figure is schematic; for example, some of the jurisdictions' demand curves that are characterized as "downward-sloping" are actually curved or kinked, even though it appears as a straight line):

- **A Vertical Demand Curve** establishes the exact quantity of capacity that is needed based on the reliability standard.[10]

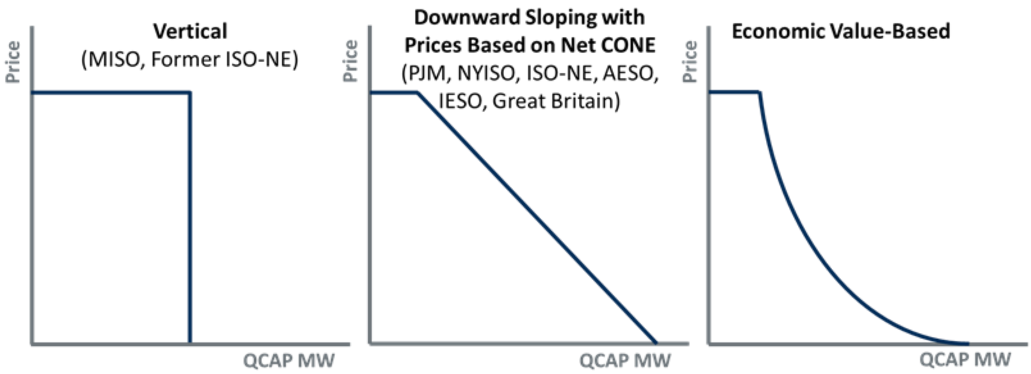
9 A "long" market describes a market that has an oversupply of capacity in comparison to its peak load and reliability needs. A "short" market describes a market that has an undersupply of capacity.

10 Although MISO still uses this simple approach, the Independent Market Monitor has recommended implementing a sloped demand curve. This is being considered as part of a suite of issues in the 2020 MISO Integrated Roadmap. See MISO, "Sloped Demand Curve in the Capacity Market (IR084)," November 6, 2019. Available at: https://www.misoenergy.org/stakeholder-engagement/issue-tracking/sloped-demand-curve-in-the-capacity-market/.

- **A Downward-Sloping Demand Curve with Prices Based on Net CONE** is designed around the reliability standard and estimated long-run marginal prices at Net CONE.
- **A Marginal Economic Value-Based Demand Curve** is based on a probabilistic analysis of marginal system costs at varying reserve margins. At each reserve margin, the analysis estimates the value of loss of load, the cost of emergency actions, and production costs. From that cost function, one can derive the demand curve as the marginal change in cost per MW of change in reserve margins. The shape of such a curve is convex to the origin, with diminishing marginal value as reserve margins increase.

None of these approaches will directly set the capacity price; that is done in combination with the supply side as resources represent the marginal cost of meeting demand in the near and long term.

Figure 3: Approaches to Determining Capacity Demand Curve
(Adopted or Proposed)



The advantage of a **vertical demand curve** is that it is simple, but that simplicity comes at the expense of greater price volatility and susceptibility to the exercise of market power because small changes in supply or demand quantities can result in significant price swings.[11] It also fails to recognize any marginal value beyond the reliability target. These and other disadvantages of the vertical demand curve drove ISO-NE to switch to a downward-sloping demand curve in 2015.[12]

The **downward-sloping demand curve** with prices tied to Net CONE is most closely aligned with the primary objective of meeting the reliability standard. This design provides price signals that support the reliability standard by increasing prices to reflect the higher reliability value of supply as the reserve margin tightens and decreasing prices when the market has excess supply. In addition, compared to a vertical demand curve design, small changes in supply and

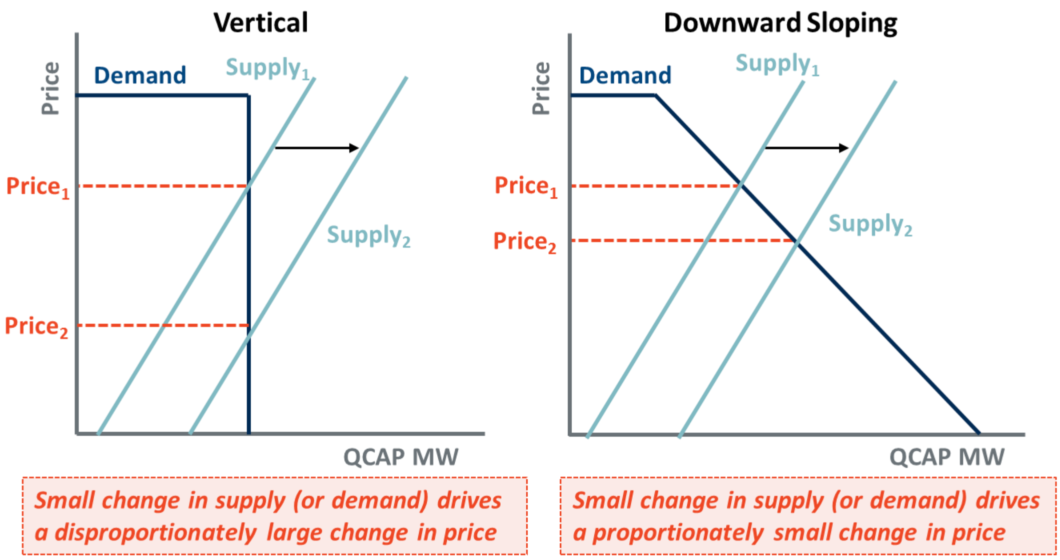
---

11 A vertical curve would also have to be shifted to the right of the minimum reserve margin in order to meet the same reliability outcomes as a downward-sloping demand curve (as some years the auction may clear at the cap and yield unacceptable reliability otherwise).

12 See ISO-NE, "FCM Sloped Demand Curve Key Project." Note that ISO-NE first switched to a linear downward-sloping curve, then transitioned to a convex "relative value-based curve" that is shaped like the economic value-based curve but is fundamentally still a downward-sloping curve indexed to the reliability standard and Net CONE, hence its characterization is as such in Figure 3.

demand produce more modest changes in prices and limit price volatility, as illustrated in Figure 4.

Figure 4: Illustrative Price Clearing Outcome with Change in Supply



The **marginal economic value-based curve** is grounded in economic value and enables the capacity auction to maximize economic efficiency. It can procure the economically optimal quantity of capacity, clearing a higher optimum reserve margin under conditions where the marginal cost of capacity is low (when there is excess supply or there are low-cost sources of new capacity); or it will clear at a lower optimum reserve margin when capacity is scarce.

Among these three primary demand curve concepts, only the downward-sloping curve tied to the reliability standard and Net CONE appears consistent with Singapore's demand curve design objectives. Specific parameters of such a curve can be adjusted to manage tradeoffs among other design objectives. We describe the balance between price and quantity certainty across the four candidate Singapore demand curves later in Section III.I.C.

Demand Curve Parameters

Long-term performance of the capacity market relative to the objectives is determined by how all aspects of the demand curve design jointly support reliability by supporting prices that attract entry when needed. Individual features of the demand curve including its price cap, quantity at the cap, width and steepness, and shape can influence that performance as summarized below and in Figure 5:

- **Price cap** defines the maximum willingness to pay for in-market supply and is often set at a multiple of Net CONE. During tight supply conditions as LOLH increases above target, the reliability value of additional resources exceeds the long-run marginal cost. As a result, market operators should be willing to pay substantially above Net CONE to procure supply under these tight market conditions. This also allows for high price outcomes that can offset low prices during surplus market conditions, such that investors can earn Net CONE on average over the long run. In other markets, the demand curve price caps range from 1.5× Net CONE to around 2× Net CONE.[13] High price caps are generally associated with less out-of-market intervention, and less reliability risk from underestimating Net CONE, but also tend to result in higher price volatility and are more susceptible to exercise of market power. We recommend Singapore's demand curve have a price cap in the range of 1.5× to 1.75× Net CONE to prevent price volatility and limit the exercise of market power. We also recommend a backstop minimum price cap in the range of 0.5× to 1× Gross CONE to prevent estimation error from artificially collapsing the price cap and the entire demand curve, should the estimated energy and ancillary services offset be close to estimated Gross CONE.

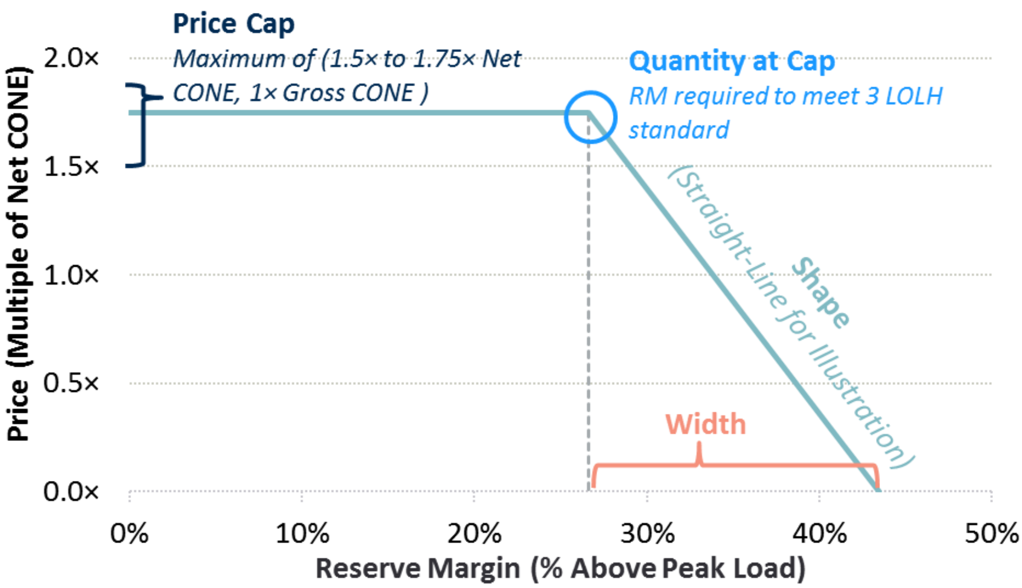
- **Quantity at the cap** determines the level of supply at which prices reach the cap. This quantity should be equal to the minimum acceptable reliability in order to ensure all in-market supply is procured before any out-of-market backstop procurements are considered. Across some (but not all) other markets, the quantity at the cap is set to the minimum acceptable quantity.[14] We recommend setting the quantity at the cap equal to Singapore's minimum acceptable quantity corresponding to 3 LOLH, given the load forecast for the delivery year.
- **Demand curve width and steepness** affect performance metrics such as average reliability outcomes, price volatility, opportunity for the exercise of market power, and consumer costs. Wider and flatter curves generally mitigate the opportunity to exercise market power and lead to outcomes with lower price volatility, but may also lead to over-procurement of supply and produce higher quantity uncertainty, which could lead to higher consumer costs. Tighter and steeper curves generally reverse these trade-offs. In a small market such as Singapore, a relatively wider curve may be needed so that entry or exit of one resource does not introduce extreme price volatility or susceptibility to market power.
- **Demand curve shape** ranges in complexity from vertical curves (MISO) to downward-sloping straight line curves (NYISO), to two-part convex kinked curves (PJM, AESO), to smoothed multi-point curves (ISO-NE), and to two-part concave curves (Great Britain, prior PJM). Vertical curves are simple to implement, but they suffer from high price volatility and susceptibility to exercise of market power. Straight line curves are also relatively simple to implement while typically resulting in stable price outcomes. Convex curves are slightly more complicated to implement but are more consistent with diminishing reliability value of incremental supply. Concave curves help to mitigate price volatility but may understate the value of reliability at high reserve margins. We tested a range of these demand curve shapes and found that the downward sloping straight line curves performed well for Singapore's capacity market.

---

[13] PJM's price cap is set as the maximum of 1.5× Net CONE or 1× Gross CONE. ISO-NE's price cap is set as the maximum of 1.6× Net CONE or 1× Gross CONE. NYISO's price cap is set to 2× Net CONE, AESO's proposed curve had a price cap set to 1.75× Net CONE or 0.5x Gross CONE, and Great Britain's price cap is set to 1.53× Net CONE. See Table 4 for sources.

[14] The quantity at the cap ranges from 97-100% of the reliability requirement or minimum acceptable reliability in other markets. NYISO's demand curve is the exception and has a very low quantity at the cap, set to around 92% of the reliability requirement. See Figure 11 for sources.

Figure 5: Illustration of Key Demand Curve Parameters



Some stakeholders have expressed the importance of including a price floor on the demand curve. However, while this is inconsistent with efficient operation of the capacity market[15] and we would not recommend implementing a price floor, we understand that EMA has considered this feedback and are proposing a price floor at 0.2× Net CONE to reduce market uncertainty in the transitional period. This price floor will be in place from the first compressed auction up till the auction for delivery year 2028 and will be removed thereafter.

Modelling Approach

To take these concepts and develop specific demand curve parameters to fit Singapore's unique market, we use a Monte Carlo simulation model to test the performance of a range of alternative demand curves. This approach has been used across several other markets such as PJM, ISO-NE, MISO, Ontario, and Alberta when designing capacity market demand curves and/or conducting periodic reviews of the demand curve performance.[16] At a high level, this model represents the capacity auction under long-run equilibrium conditions, determining market clearing prices and quantities and expected reliability by intersecting supply and demand curves. The model simulates many auction clearing outcomes, representing a realistic range of supply and demand conditions for Singapore's market.[17] By simulating the capacity auction many times, we develop distributions of cleared prices and quantities. A stylized depiction of

---

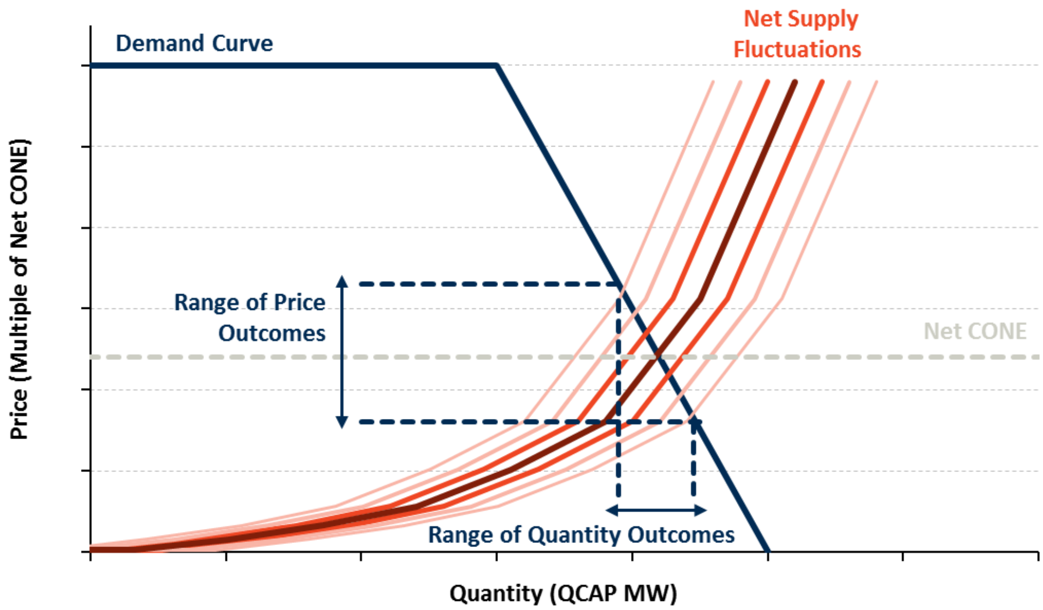
15 Beyond a certain reserve margin, the incremental reliability value of additional capacity is negligible, and the demand curve should express this value. Artificially limiting downward movement in the price could inefficiently retain capacity that is not needed, increasing overall system costs.

16 Examples include: Spees, Kathleen, *et al.*, "Alberta's Capacity Market Demand Curve," Prepared for AESO, January 2019. Newell, Samuel, *et al.*, "Fourth Review of PJM's Variable Resource Requirement Curve," Prepared for PJM, April 19, 2018.

17 The model runs 10,000 Monte Carlo simulation draws. The first 9,000 draws are used to calibrate the supply and demand balance, and the last 1,000 are used to evaluate the performance of the demand curve.

the price and quantity distributions resulting from our Monte Carlo model is shown in Figure 6, with the intersection of supply and demand curves determining the different price and quantity outcomes across simulation draws.

Figure 6: Stylized Depiction of Market Clearing Outcomes in Monte Carlo Analysis

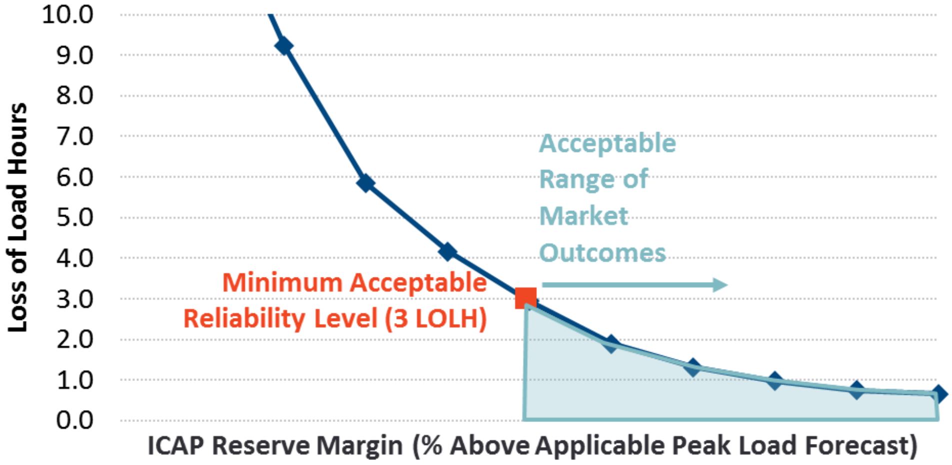


MODELLING ASSUMPTIONS

To come up with inputs for our Monte Carlo modeling, we relied on historical and future projected Singapore data, as well as data from other jurisdictions with established forward capacity auctions. Our analysis considers a 2026 modeled year, the delivery year of the first four-year forward End-State Auction. Some of the main inputs are listed below:

LOLH curve was provided by EMA and used to calculate the reliability outcome from each modeled draw that represents a market clearing. For a given draw we took the cleared quantity and found the associated LOLH value using the LOLH curve. Figure 7 shows the asymmetrical relationship between the reserve margin and LOLH. As shown in the figure, LOLH outcomes deteriorate drastically at reserve margins below the reliability standard and improve gradually at reserve margins greater than the reliability standard.

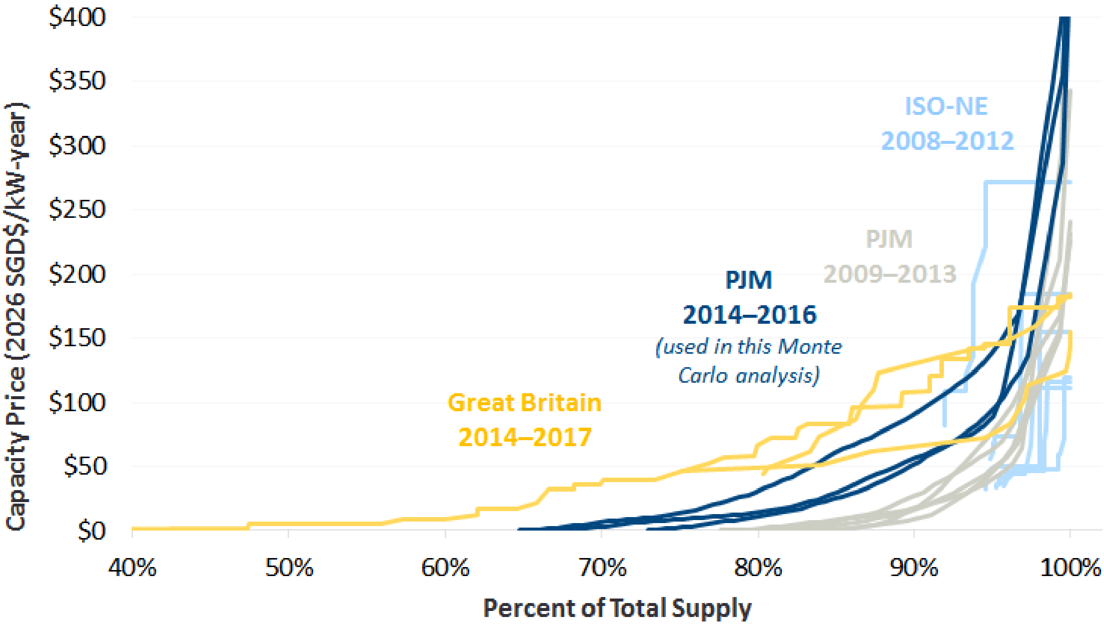
Figure 7: Loss of Load Hours as a Function of Reserve Margin



Supply curves represent merchant supply offers in the FCM and combined with the demand curve determine the market clearing price and quantity. We use supply offer curves that fall in the middle of the range of supply curves across jurisdictions with capacity markets. We also conducted sensitivities that showed differences between different supply curve shapes were small relative to the effect of other parameters, including the net supply fluctuations (see next paragraph). Figure 8 shows the supply curves used in our modeling compared to other jurisdictions' historical supply curves.

The supply curve shape is consistent with the expected fleet-wide resource economics in Singapore, given the cost structure of highly capital-intensive resources with long economic lives. With this cost structure, we anticipate that a majority of resources are likely to offer at relatively low prices, consistent with the offer of a resource whose investment costs are sunk and avoidable going-forward costs are largely offset by anticipated energy and ancillary service net revenues. These resources would not be likely to retire or mothball even if capacity prices were zero. The remaining fleet would be made up of aging resources, demand response, imports, and new resources that are expected to offer at higher prices, and that may enter or exit the market depending on that specific year's capacity prices.

Figure 8: Modeled Supply Curves

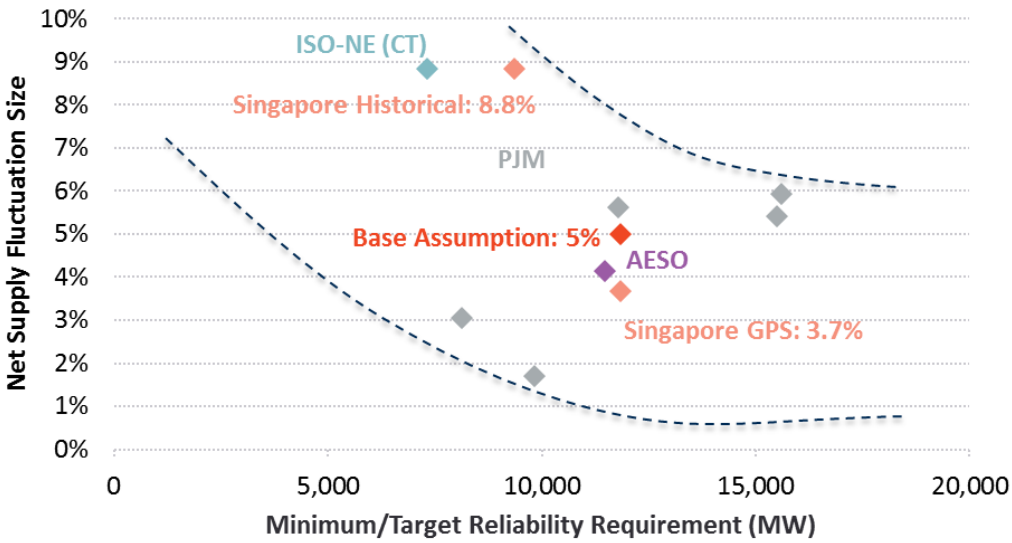


Sources and Notes: Years represent year of auction for other markets. PJM data from Fourth Review of VRR curve. ISO-NE data from testimony of FCA1- FCA7. Great Britain data from annual capacity auction results reports: National Grid, T-4 Capacity Market Auction, 2014, 2015, 2016, and 2017. For ISO-NE and Great Britain, only a portion of the supply curve is shown because only the offers above the clearing price was made publicly available in those auctions. Converted from nominal $USD to 2018$SGD using historical exchange rates and inflation rates posted by the Monetary Authority of Singapore. Inflated to 2026$SGD using 1.5% inflation rate as a mid-point between 1% and 2% after discussion with EMA.

Net supply fluctuations (supply offers minus the required reserve margin) represent year to year variation in the FCM due to lumpy entry and exit decisions, cost fluctuations, and variation of load forecast (and required reserve margin). Incorporating this variation into the demand curve modeling is crucial to obtain realistic estimates of the distribution of clearing outcomes, including reserve margin and prices. We model net supply fluctuations of 5% standard deviation because it falls in the range of Singapore-specific net supply fluctuations calculated using historical and EMA's forecast of supply and demand data, and is in the range of other similarly-sized markets, as shown in Figure 9 below.[18]

18 Net supply fluctuations should be less than historical values (because the market dampens historical fluctuations from an EOM and increases correlation to demand), but more than EMA's forecast which assumes perfectly well-behaved entry in response to exit and load growth.

Figure 9: Net Supply Fluctuations



Sources and Notes: ISO-NE FCA1 – FCA8, PJM 2009/10– 2019/20 BRA, and AESO Annual Market Statistics. Singapore data provided by EMA. Demand fluctuation observations for each market represent zones within that market. Since Singapore is small, we treat it as a single zone in this analysis. For "Singapore Historical" value 2018 peak load is used to calculate the minimum target, and for "Singapore GPS" value 2030 peak load is used to calculate the minimum target since that is projected future supply and demand.

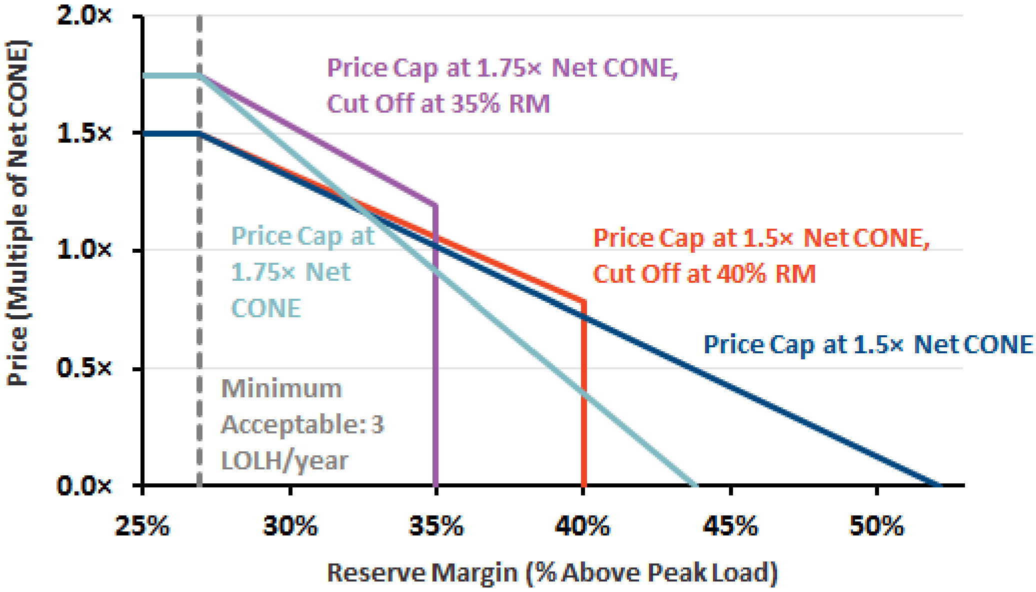
C. Demand Curve Shape

Candidate Singapore Demand Curves

Considering Singapore's unique context and drawing on experience from other markets, we assessed a variety of possible demand curves with a range of price caps that varied in shape: curves that follow a straight line to zero price, straight line curves with cut offs at varying reserve margins, vertical, and kinked curves. In collaboration with EMA, we narrowed the range of curves down to four candidate curves whose qualities result in an acceptable balance of tradeoffs between reliability, price and quantity volatility, and other design objectives.

Among these curves are two straight line curves and two straight line curves with a quantity cutoff. We offer straight line curves over convex kinked curves in our final candidate curve selection, because we find that kinked curves do not meaningfully improve performance relative to straight line curves. Additionally, no vertical curves are included in the candidate curve selection due to the increased price volatility when administering vertical curves. Of our four final candidate curves, two follow a straight line to a zero price (**blue** and teal), and the two other curves are straight line curves that cut off at 35% (purple) and 40% reserve margins (**red**). The curves with a straight-line cutoff are chosen to explore the advantages and disadvantages of limiting the occurrence of overcapacity situations. The price cap for all four candidate curves is set at either 1.5× or 1.75× Net CONE in order to limit opportunities for the exercise of market power. The four candidate demand curves are depicted Figure 10 below.

Figure 10: Candidate Singapore Demand Curves



All curves assessed are tuned to ensure that the frequency below the minimum acceptable level never goes above 5%, and all curves meet this standard with different tradeoffs on other objectives.[19] We describe the tradeoff between price and quantity uncertainty for the candidate curves in Table 3. The simulated performance of the candidate curves with higher price caps at 1.75× Net CONE (teal and purple lines) produce slightly lower average clearing reserve margins, but with higher price and cost volatility. On the other hand, the candidate curves with lower price caps at 1.5× Net CONE (**blue** and **red** lines), lead to lower price and cost volatility but slightly higher average clearing reserve margins.[20] Cut off curves (**red** and purple lines) were considered because they help avoid extreme over procurement. However, we see that they only slightly decrease the average clearing reserve margins (and consumer costs) but lead to significantly higher price volatility.

---

19 The demand curves are "tuned" through our analysis to ensure that the quantity cleared does not drop below the minimum acceptable reliability in more than 5% of the 1,000 model simulations. If tuned to 0% frequency below minimum acceptable, there would be significant over-procurement under normal system conditions. Tuning to 5% frequency below minimum acceptable standard in the base auction does not mean reliability will be compromised in the delivery year, due to the opportunity to procure more capacity, in the forward period, through the rebalancing auctions and/or out-of-market actions.

20 Importantly, note that demand curves with a higher price cap are narrower and steeper, reaching zero price at a lower reserve margin; those with a lower price cap are correspondingly wider. This allows each of the four candidate curves to meet the minimum reliability standard at lowest cost; a curve that was both higher and wider would tend to clear more capacity at increased consumer costs, violating one of the design objectives.

Table 3: Simulated Performance of Candidate Demand Curves

	Price and Cost		Reliability		
	Std. Dev. of Price	Std. Dev. of Cost	Avg. LOLH	Avg. Reserve Margin	Freq. Below Min. Acceptable
Demand Curve	*($/kW-year)*	*(% of avg)*	*(hours)*	*(%)*	*(%)*
Straight Line Curves to Zero Price					
Price Cap at 1.5× Net CONE, Straight Line to Zero Price	**$54**	**27%**	**0.88**	**35%**	**5.0%**
Price Cap at 1.75× Net CONE, Straight Line to Zero Price	**$75**	**39%**	**0.93**	**34%**	**5.0%**
Straight Line Curves with Cut Off					
Price Cap at 1.5× Net CONE, Cut Off at 40% RM	**$59**	**31%**	**0.89**	**35%**	**5.0%**
Price Cap at 1.75× Net CONE, Cut Off at 35% RM	**$92**	**49%**	**0.93**	**33%**	**5.0%**

EMA is recommending the straight line curve with a price cap at 1.5× Net CONE (**blue** line), finding that it presents the best trade-offs between the design objectives and is least sensitive to changing and uncertain conditions (explored in various sensitivity analyses studied). We believe this curve is in the range of reasonable curves to meet EMA's objectives.

Comparison to Other Jurisdictions

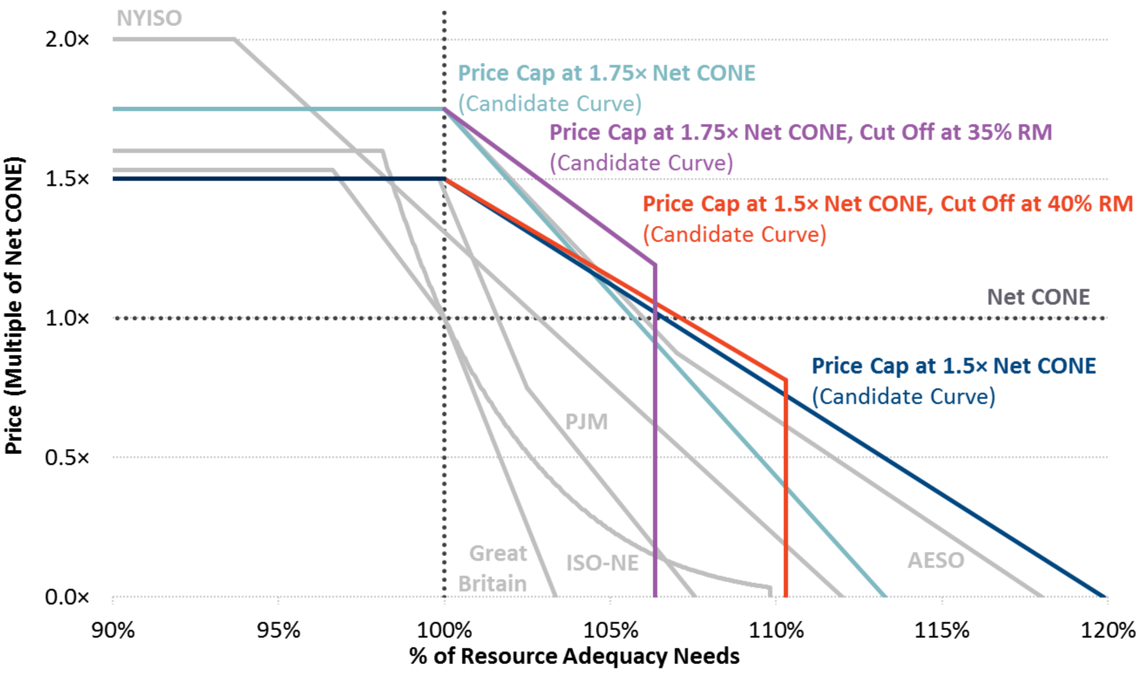
Other jurisdictions use a variety of approaches including qualitative analysis, market-specific considerations, and similar probabilistic modeling efforts to determine their demand curve parameters and/or conduct periodic reviews.[21] Figure 11 below shows a variety of demand curves used across other jurisdictions, illustrating a range of demand curves found to be workable in different circumstances.[22] Each of these curves is tailored to their specific market conditions. As shown in the figure below, even the cut off candidate curves tend to be wider than those seen in most other markets. We view a somewhat wider curve as sensible for Singapore given that the small market will be more susceptible to price volatility and exercise of market power, both of which can be partly mitigated through a wider demand curve.

---

21 Examples include: Spees, Kathleen, *et al.*, "Alberta's Capacity Market Demand Curve," Prepared for AESO, January 2019. Newell, Samuel, *et al.*, "Fourth Review of PJM's Variable Resource Requirement Curve," Prepared for PJM, April 19, 2018.

22 Although Alberta ultimately decided not to pursue a capacity market, the Alberta Electricity System Operator (AESO) had previously developed a detailed market design. This report includes information on their design choices, analysis, and rationale where useful.

Figure 11: Demand Curves in Other Markets



Sources and Notes: PJM Interconnection, "2021/2022 RPM Base Residual Auction Planning Period Parameters," February 2018. ISO New England, "Forward Capacity Market (FCA 12) Result Report," May 2018. New York Independent System Operator, "ICAP Translation of Demand Curve (Summer 2018)," March 2018. Spees, Kathleen, et al., "Alberta's Capacity Market Demand Curve," Prepared for AESO, January 2019.

McNamara, Fergal, "Capacity Market," United Kingdom Department of Energy & Climate Change, June 25, 2014. NYISO ICAP market differs from other markets in that it does not procure capacity multiple years forward, instead holding multiple auctions within one year.

Table 4: Demand Curve Parameters in Other Jurisdictions

	Unit	PJM	ISO-NE	NYISO	AESO	GB
Price Cap	Multiple of Net CONE	1.5	1.6	2.01	1.75	1.53
Minimum Price Cap	Multiple of Gross CONE	1.0	1.0	N/A	0.5	N/A
Quantity at Cap	Percent of Requirement of Minimum	99.8%	98.1%	92.1%	100.0%	97.0%
Demand Curve Shape	-	Kinked	Curved	Straight	Kinked	Kinked

Sources: PJM Interconnection, "2021/2022 RPM Base Residual Auction Planning Period Parameters," February 2018. ISO New England, "Forward Capacity Market (FCA 12) Result Report," May 2018. New York Independent System Operator, "ICAP Translation of Demand Curve (Summer 2018)," March 2018. Spees, Kathleen, et al., "Alberta's Capacity Market Demand Curve," Prepared for AESO, January 2019. (Note that AESO market was cancelled in 2019 due to a change in government.) McNamara, Fergal, "Capacity Market," United Kingdom Department of Energy & Climate Change, June 25, 2014.

D. Net Cost of New Entry

The pricing points on the demand curve will be based on the Net CONE, reflecting the long-run marginal cost of capacity. Net CONE is an administrative estimate of the long-run marginal cost of capacity based on the reference technology most likely to enter the market. Tying pricing points to Net CONE enables the demand curve to adjust as needed to remain consistent with market conditions and the cost of attracting enough supply to meet the reliability standard. The rules will include an approach to establish the follow parameters for developing the Net CONE:

- **Reference Technology** that is the assumed marginal resource type that will be attracted into the market;
- **Gross CONE** reflecting the total annual capital costs, ongoing fixed costs, and financing costs required to bring a resource online, after levelizing these costs over the economic asset life;
- **Energy and Ancillary Services (E&AS) Offset** reflecting the expected net revenues (or revenues minus variable costs) that the resource would earn from participating in the E&AS markets; and
- **Approach to Updating Net CONE** over time, including a formulaic approach for updating the parameter for each calendar year and a more comprehensive review of the parameter periodically (as discussed in Section III.I.E below).

Estimating Net CONE at the present time of market transition poses several challenges given that future market outcomes should not be expected to be similar to recent history. The introduction of a capacity market and any concurrent changes to the energy market could change the estimated value of Net CONE (especially the E&AS offset), the effects of which will not be observed through energy market prices or participant behavior until after a few years' experience with the new market. The market rules may therefore adopt two different approaches to estimating the Net CONE. First, we recommend the market incorporate a Transitional Net CONE parameter to be used in the early years of the FCM auction with a compressed forward period. For the first End-State base auction in 2022 for delivery year 2026, the Net CONE can be updated in a periodic study using a methodology as prescribed in the market rules. In both cases, the goal of the Net CONE estimate will be the same: to develop an unbiased estimate of the price needed to attract new supply into the market, subject to any limitations of unresolvable estimation uncertainties.

Table 5 below summarizes the approaches adopted in other capacity markets to estimate and update the Net CONE parameter. We discuss the merits of the various approaches in the following subsections as applied in the Singapore context, both during the market transition and in the long term.

Table 5: Approaches to Estimating Net CONE in Other Capacity Markets

	PJM	New England	New York	Ontario *(Proposed)*	Great Britain	Alberta *(Cancelled)*
Reference Tech	Frame 2x1 Gas Combustion Turbine (CT)	Frame CT	Frame CT	Aero CT, Frame CT, CCGT, Battery storage	CCGT	Aero CT, Frame CT, CCGT
Gross CONE	Capital, fixed & financing costs, level nominal	Bottom-up engineering costs, level real	Bottom-up engineering costs, level real	Capital, fixed & financing costs	Bottom-up engineering costs	Capital, fixed & financing costs, level nominal
E&AS Offset	Three-year historical average simulated	Forward looking prices derived from simulation of future energy market revenues	Simulation of revenues using rolling three-year historical locational energy and reserve price average, with adjustment	Forward looking market methodology	Forward looking multi-year dispatch simulation	Forward looking approach, dispatches reference technologies against a forecast of hourly market prices
Annual Net CONE Updates	Based on weighted index E&AS: three-year rolling average	Escalating cost components and revenues offsets according to indices E&AS: annual updates to reflect futures prices	Updates based on single state-wide technology specific escalation factor	Updated based on weighted average of public indices E&AS: annual update	Updated regularly based on electricity prices	Prior to each subsequent capacity auction based on applicable cost indices
Periodic Reviews	Full CONE study and methodology review every four years	Full re-evaluation of Net CONE every three years	Full review of reference resource, Gross CONE and demand curve every four years	Full review every three years	Net CONE and reference technology annually reviewed	Update estimated CONE values every four to five years

Sources and Notes: PJM Interconnection, "2021/2022 RPM Base Residual Auction Planning Period Parameters," February 2018. PJM: Review of PJM 's Variable Resource Requirement Curve; New England: ISO NE Filing of CONE; New York: NYISO Order Accepting Tariff Filing; Ontario: IESO Incremental Capacity Auction High-Level Design; Great Britain: Setting Capacity Market Parameters; Alberta: AESO Calculation of Demand Curve Parameters. The Alberta (AESO) capacity market was cancelled due to a change in government in Alberta. See source.

Reference Technology

Net CONE is the estimate of the long-run marginal cost of capacity, or the average capacity price that should prevail in a long-run equilibrium condition when market entry is needed to support the reliability standard. The reference technology used as the basis for estimating Net CONE should therefore be a resource that is most likely to be attracted into the merchant capacity market. We anticipate that a wide variety of resource types will be likely to participate and clear in Singapore's capacity auction including existing and new gas-fired generation plants, solar photovoltaics (PV), battery storage, demand response, and others. By definition, all of these cleared resources can be considered an economic portion of the resource mix, but some resource types would be more appropriate than others to adopt as the reference technology for estimating the administrative Net CONE.

The most appropriate resource type to select as the reference technology should:

- **Be economic to build when new capacity is needed.** The reference technology should be one that developers are likely to build when new supply is needed in the market. The determination of which technologies are likely to be economic in the long-run equilibrium can be determined by estimating the Net CONE across multiple technologies and identifying the least cost and supplementing this with evidence of commercial interest through recent developments and proposed projects.
- **Be feasible to develop given anticipated technical limitations and regulations.** The reference technology should be a technically feasible and proven technology, ideally as demonstrated through widespread adoption and development. The technology cannot be prohibited through any legal means, such as environmental regulations that might prevent the development of power plants without proper emissions controls.
- **Be possible to build in relatively large quantities at uniform cost.** The reference technology should be a resource type that could be developed in large quantities at relatively similar prices. This criterion rules out certain resource types that may be limited in their total available quantity, such as unique projects that face idiosyncratic circumstances (*e.g.*, demand response, and cogeneration projects), and thus would not be appropriate to adopt as a reference technology.
- **Be possible to estimate costs with relatively low uncertainty.** The Net CONE of the reference technology should be possible to estimate with as much accuracy as possible. This criterion introduces a preference to use the costs of a better-known technology type with more available data on costs and anticipated revenues.

Other markets have applied these or similar criteria with differing emphasis depending on their unique circumstances, and have ultimately chosen either CCGT or open-cycle plants as the most appropriate reference technologies (as summarized in Table 5). The Ontario market operator has also proposed to consider battery storage alongside other options as the potential reference technology; batteries may become a more relevant resource type to consider in regions that are aiming to phase out fossil fuel plants as part of their supply mix.

In Singapore, we recommend applying these principles to select the reference technology for both the Net CONE parameter for the first End-State base auction, as well as re-evaluating the reference technology in periodic reviews (see Section III.I.E below). In both cases, this evaluation should consider the best available data on resource costs, recent and anticipated net market revenues, recent project developments, and proposed developments.

Gross CONE

The Gross CONE parameter should reflect the annualized costs associated with building and maintaining the reference technology. The development of the Gross CONE in other markets is typically calculated through an independent bottom-up engineering cost study accounting for the following components:

- **Overnight capital expenditures** necessary to construct the plant including project development, permitting, engineering, procurement, construction, labor, materials, major equipment, transmission interconnection, gas pipeline interconnection, backup or onsite fuel storage (if relevant), the expected value of contingencies, taxes, capitalized inventories, working capital, and interest during construction.
- **Annual fixed operations and maintenance costs** necessary to maintain the plant on an ongoing basis over the asset's life including labor, asset management, regular maintenance, major overhauls, the firm/fixed portion of any fuel contracts (excluding any variable fuel costs), property tax, and insurance. These costs would exclude any variable costs that are anticipated to be incurred on an incremental basis as a function of how often the plant runs (such as start-up and variable running costs).
- **Financing costs** necessary to serve debt and equity. The financing cost analysis would consider the after-tax weighted-average cost of capital (ATWACC) consistent with attracting merchant power investments in Singapore, the relevant tax rate that would be applied to any earnings, and the asset's anticipated economic life.

Bringing this to the Singapore context, the Gross CONE parameter is essentially the same as the building and maintenance costs developed for the purposes of setting vesting contract prices.[23] That vesting parameter analysis has been developed for a different purpose, but is similar and recent enough that we recommend considering it appropriate to adopt the same or a slightly adjusted parameter for the purposes of establishing the transitional Net CONE parameter for the capacity auction. This approach would have a number of advantages including expedience, simplicity, transparency, and familiarity to market participants, but limits the selection of the reference technology to only consider a CCGT. For the initial auctions, we recommend using this vesting price parameter for Gross CONE, as suggested by several stakeholders. Periodic future reviews may include alternative potential reference technologies to ensure that the parameter can evolve with market conditions.[24]

Energy and Ancillary Services Offset

The E&AS offset is a parameter used to calculate Net CONE and reflects the expected net revenues (or revenues minus variable costs) that the reference resource would earn from participating in the E&AS markets. There is no single, commonly accepted approach for

---

23 These parameters are developed by the EMA to calculate the vesting contract price for gencos with vesting contracts and are reviewed biennially (or when deemed necessary), with a mid-term review of the capital cost parameters, in accordance with the published procedures. See EMA, "EMA's Procedures for Calculating the Components of the Vesting Contracts," July, 2019. Available at: https://www.ema.gov.sg/cmsmedia/Version%202%207%20-%20Vesting%20Contract%20Procedures.pdf.

Based on EMA's "Review of the Long Run Marginal Cost Parameters for Setting the Vesting Contract Price for 2019 and 2020", the Gross CONE would be about S$222/kW-year for a 432.2 MW F-Class CCGT (on an installed capacity basis). As the vesting parameters are reviewed periodically, per the published procedures, EMA intends to use the latest updated values at the time of the relevant auction, to determine the Gross CONE.

24 EMA intends to conduct this review in time for the first End-State Auction to be held in 2022 for delivery year 2026.

estimating the E&AS revenue offset given the unique issues of data availability, market context, and underlying uncertainties that affect each market region. However, there are some useful underlying principles that can be used to develop a reasonable approach for any market. To the extent possible, the approach should:

- Be an accurate representation of expected net revenues for the reference technology (considering expected average revenues across weather-driven and other uncertainties);
- Be simple, replicable, and transparent, using trusted and reliable sources and procedures;
- Reflect future market conditions and/or market equilibrium conditions as currently perceived; and
- Be validated against the historical net revenues earned by representative existing units that are similar to the reference technology.

These principles should be interpreted as an ideal to strive toward. However, data limitations, uncertainty surrounding the market outlook, and trade-offs among these principles make it challenging to achieve all of these outcomes simultaneously. Key choices and considerations include the following:

- **Observed Net Revenues versus Simulated Dispatch:** Estimated E&AS margins can be derived from those of representative existing resources historically observed in the marketplace. This approach can be simple and straightforward, but requires a sample of representative generating resources, is backward looking, and tends to be more volatile compared to forward-looking approaches. Alternatively, E&AS margins can be estimated based on a simulated dispatch of the particular reference technologies. This approach allows reference resources to be dispatched against either historical or future prices, and the method of dispatch simulation can take different levels of complexity. Both backward- and forward-looking approaches using a simulated dispatch can be further validated by comparing to the observed outcomes for representative existing plants.
- **Historical versus Future versus Equilibrium Market Prices:** E&AS margins can be estimated based on historical, future, or equilibrium-based market prices. Historical prices can be readily observed but can be volatile and do not capture expectations about the future. Futures-market-based prices are observable, and when based on liquid futures markets, provide a reasonable reflection of market participants' expectations for near-term (and weather-normalized) changes in market fundamentals. Near-term futures (1-year forward) can be used as a proxy for longer-term futures as they will account for some, but not all, of the changes in market conditions going forward. Forecasts of future prices derived from market simulation models can explicitly incorporate expectations about the future but developing price forecasts through market simulation models (1) requires agreement on reasonable simulation assumptions, (2) can be very sensitive to modelling inputs and assumptions, and (3) are often less transparent to market participants.

  However, both historical and forecast methods can destabilize the reserve margin by perpetuating disequilibria. Both methods decrease Net CONE when supplies are tight (by increasing the E&AS offset) and increase Net CONE when supply is long (by reducing the E&AS offset). Using prices from a simulated equilibrium, with the future reserve margin adjusted to the target level, helps to stabilize the E&AS offset. For maximum stability and efficiency, this target reserve margin should reflect the average reserve margin expected from the demand curve simulations described earlier.

Each of these approaches offers advantages and disadvantages that usually depends more on local context, such as data availability and current market conditions, than underlying principles. For example, historical approaches may provide simpler, more transparent, and more replicable means of estimating the E&AS offset, even if a forward-looking methodology is otherwise desirable. In practice, most approaches utilized in other regions with capacity markets apply a blend of forward- and backward-looking features. Therefore, there is not a consensus on the best practices approach to estimating the E&AS offset, as illustrated by the variety of approaches adopted in other markets as summarized in Table 6.

Table 6: Method to Estimate E&AS Offset in Other Jurisdictions

Market	E&AS Methodology
PJM	• Three-year average of simulated E&AS values based on virtual dispatch against historical hourly prices • Calculated zonally to get a zone-specific Net CONE
ISO-NE	• Simulate future energy revenues over 20 years using a market pricing model to develop a price forecast and using a dispatch model to estimate revenues
NYISO	• Simulation of revenues using rolling three-year historical market prices and reserve prices, fuel and emission prices, and variable operations and maintenance costs
IESO *(Proposed)*	• Forward-looking market methodology to estimate E&AS offset reflecting the expected market fundamentals that will affect revenues available to the reference resource
AESO *(Cancelled)*	• Forward-looking methodology, assuming a stand-alone resource which assesses options to maximize its offset • Would initially exclude ancillary service revenues

For the Singapore FCM, we see the forward-looking, equilibrium-based simulation of E&AS offsets as most accurate and appropriate. A methodology reflecting historical prices and observed net revenues would likely be incorrect due to the current overcapacity situation. Thus, we recommend using prices and E&AS margins from a simulation model, with the reserve margin adjusted to the equilibrium level. EMA should develop forward-looking values for various input assumptions to a market simulation model, including forecasted gas prices, demand forecast, and solar adoption forecast. The simulation modeling will take these input assumptions together with the heat rates curves submitted by generation companies to EMA into account to forecast energy, reserve and regulation prices for every dispatch period. Further, as mentioned earlier, to maximize stability in the forecasted prices, the equilibrium reserve margin obtained from the demand curve modelling will be assumed. Pulling from these simulated prices, EMA will be able to calculate an estimated E&AS offset for a new entry reference technology, and the offset will be netted off Gross CONE to calculate expected Net CONE.

Importantly, this analysis must be conducted assuming the entry of a new reference resource as the marginal FCM resource. This is to ensure that the E&AS offset is accurately calculated and reflected in Net CONE. To accomplish this, in some years EMA will need to adjust the input assumptions in the simulation modeling, *i.e.*, the level of existing capacity to enable the entry of a new reference resource while maintaining equilibrium levels of capacity.[25] As discussed above, this equilibrium-based analysis of the E&AS margins will maximize stability of market outcomes. The high-level assumptions proposed for the development of the E&AS offset methodology are represented in Table 7.

Table 7: Proposed High-Level E&AS Offset Assumptions

Key Model Inputs	Proposed Assumptions
Demand	• EMA's Electricity Demand Forecast, converted to half-hourly demand in the delivery year based on historical load profiles.
Resource Mix	• Entry of new reference resource (*i.e.*, CCGT) to be assumed, with target (or equilibrium) reserve margin based on demand curve simulations and achieved by retiring the least efficient thermal resource(s). • Quadratic solar growth up to 2GWp in 2030, after 2030 under review. • Embedded generation under review.
Fuel Prices	• The International Energy Agency's forecast of Brent and high sulfur fuel oil (HSFO) prices to forecast liquified natural gas (LNG) and piped natural gas prices (PNG) respectively; based on the correlation between historical Brent and LNG prices, and historical HSFO and PNG prices.
Heat Rates	• Existing thermal resources to be based on market participants' submission of original equipment manufacturer data to EMA. • New thermal resources to be based on vesting contract parameters, until comprehensive Gross CONE review is conducted.
Dispatch	• Economic dispatch based on short-run marginal costs.

Note: EMA intends to use PLEXOS, a energy market simulation software, to conduct the modelling.

In the long term, we recommend annual formulaic updates to Net CONE in addition to conducting a full Net CONE study every few years based on updated data. This study would result in a recommended E&AS offset estimate and methodology for performing annual formulaic updates to reflect evolving supply and demand conditions. See Section III.I.E below for further discussion on demand curve review and updates.

E. Demand Curve Review and Updates

Singapore's capacity market rules will need to incorporate a process for updating and reviewing demand curve parameters. These periodic reviews provide the opportunity to evaluate the performance of the demand curve relative to the design principles and make any changes necessary to improve its design. These reviews are important to ensure the demand curve is

---

25 In years where the existing capacity is above the equilibrium reserve margin, the resource(s) with the highest net avoidable going-forward fixed costs would be assumed retired, to accommodate a new reference resource in the simulations; in years with existing capacity below the equilibrium reserve margin, new reference resource(s) would be assumed to enter in the simulations.

adjusting to the market's changing needs and cost of supply. These updates can be conducted in two timeframes:

- **Annual formulaic updates** that require minimal administrative effort but are necessary to maintain consistency with market demand and supply costs over time. The annual updates normally focus on updating demand curve quantity points based on new load forecasts and reliability analysis, as well as updating Gross CONE and E&AS offset with the most recent market data to get a more accurate Net CONE. These updates ensure the pricing points on the administrative demand curve maintain consistency with market conditions and the auction procures sufficient capacity while avoiding significant over-procurement. We recommend Gross CONE updates to be based on the most recently available public index and E&AS updates to be based on either recent historical or futures-based market price data, as applied using a formulaic updating approach.
- **Periodic comprehensive reviews** to address longer-term trends and fundamental shifts to technology. These comprehensive reviews are a detailed evaluation of demand curve parameters and methodologies used to calculate Gross CONE and E&AS offset. Often, they review:
  - *Reference technology.* Evaluate which technologies are economic to build when new merchant supply is needed. Account for changes in policy regulations and technology cost trends.
  - *Gross CONE.* Evaluate change in costs of technology, labor, and land, as well as updates to tax rates and deductions and policy incentives and regulations to more closely align with observed and anticipated market conditions. Assess methodology used to calculate Gross CONE (*i.e.*, using level-real or level-nominal approach to calculate annualized costs).
  - *E&AS Offset Methodology.* Evaluate forward- or backward-looking methodology and whether to use simulated or actual market data. Review changes in fuel prices, energy and ancillary service prices, generation resource mix, and policy regulations.
  - *Demand Curve Parameters.* Evaluate performance relative to the reliability standard and whether the standard needs to be updated. Determine whether the shape of the demand curve, width, price cap, or any set of price and quantity points of the demand curve need to be adjusted based on any observed or anticipated challenges to the market.

Details on comprehensive periodic reviews of reference technology, Gross CONE and E&AS Offset estimates, and demand curve parameters will be determined at a later stage. The practices in other markets are shown in Table 8 below.

Table 8: Capacity Market Comprehensive Review Cycles in Other Jurisdictions

	PJM	ISO-NE	NYISO	Great Britain
Frequency	4 years	3 years	4 years	5 years
Scope	CONE estimate, E&AS offset methodology, demand curve	CONE estimate, E&AS offset, resource type mitigation levels	CONE estimate, demand curve performance	Assess market performance relative to objectives, review market objectives

Sources and Notes: PJM's major reviews were initially on a three-year cycle and included a broader scope. See, for example, Pfeifenberger, J., Newel, S., Spees, K., Hajos, A., Madjarov, K., *Second Performance Assessment of PJM's Reliability Pricing Model*, August 26, 2011. Great Britain's Energy Act of 2013 calls for a comprehensive review of the market 5 years from passage of the Act.

In line with practices in other jurisdictions, we recommend EMA conduct comprehensive periodic reviews on a three- to five-year cycle. In the early years of the market, it may be advantageous to conduct the review more often, at the lower end of (or perhaps even below) this range. This determination is based on tradeoffs between:

- The desire for certainty for market participants and investors in new resources, and reduced administrative burden, by a less frequent review cycle; versus
- The benefit of reflecting latest market conditions, reducing the risk that the prevailing demand curve would fail to procure sufficient reliability or would increase costs by consistently over-procuring capacity relative to the reliability standard.

F. Recommendations for Singapore

Recommendations and Next Steps
Reliability Standard • Maintain Singapore reliability standard (currently defined as 3 LOLH). This quantity will be translated into the equivalent QCAP procurement volume based on the load forecast (used to set the minimum quantity points on the demand curve)
Net Cost of New Entry • Estimate a transitional Net CONE parameter based on currently available data, from the vesting contract parameters, to apply over the Q4 2023 to 2025 delivery periods (with auctions to be conducted in a compressed forward period in initial years)
Demand Curve Parameters • Price cap to be in the range of 1.5× to 1.75× Net CONE • Minimum price cap to be in the range of 0.5× to 1× Gross CONE • Quantity at the price cap set at a minimum acceptable reliability level of 3 LOLH • Downward-sloping straight line shape, without a cutoff at high reserve margins
Demand Curve Review • Conduct demand curve review on a three- to five-year cycle
Next Steps • Solicit stakeholder feedback on four candidate Singapore demand curves • Solicit stakeholder feedback on the method to estimate the E&AS offset

IV. Supply Resource Qualification and Capacity Ratings

The FCM is intended to admit a broad range of resources that can contribute to supply adequacy, including existing and planned conventional thermal resources, demand response, imports, solar, and storage.[26]

A resource qualification process is needed to validate that offered resources will be online and able to operate in the delivery year. As part of this process, the capacity value or rating for each resource is also determined. This rating gives the MW quantity each resource is qualified to offer into the auction, given its demonstrated availability and any operating limitation. This is needed both to ensure that resources are compensated fairly and consistently with their value, and to ensure sufficient capacity is procured in the auction to meet resource adequacy requirements.

26 Some stakeholders have asserted that the FCM would favor resources with lower avoidable going-forward costs, unfairly discriminating against newer resources with higher unavoidable fixed costs. We disagree to the extent that the real-time energy and ancillary services markets recognize the value higher fixed cost resources might provide through higher net revenues. In that case, economic efficiency will be maximized when all resources are treated fairly and can compete to offer the same product.

We evaluate the qualification and rating approaches with the following key criteria:

- Inclusivity in encouraging broad participation from potential resource types;
- Accuracy in representing contribution of resources to reliability objective (including durability with evolving future system conditions); and
- Simplicity of design and feasibility of implementation (in qualification stage).

In this section, we recommend adopting a QCAP rating approach and describe how to qualify and rate the relevant resource types. We also provide considerations and recommendations for establishing a qualification timeline in Singapore.

A. QCAP Rating Approach

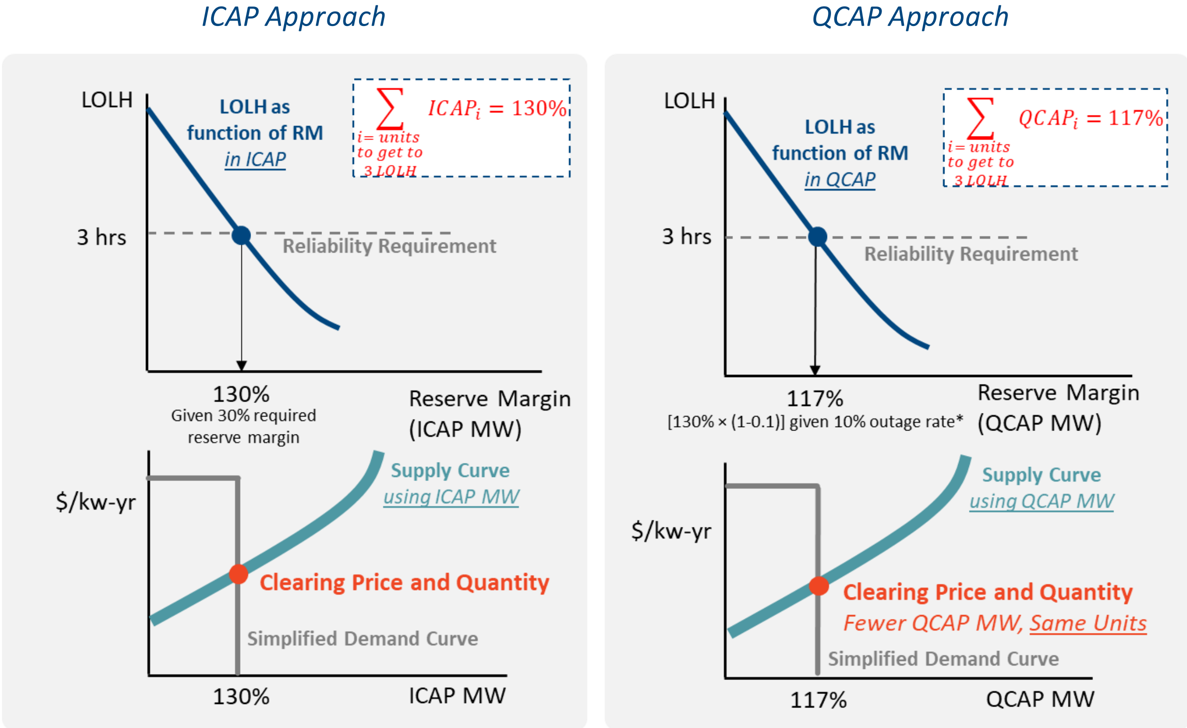
There are two primary approaches to determine capacity ratings. The first reflects the maximum output of a given resource, or a resource's installed capacity (ICAP). The second, which we refer to as the "qualified capacity" (QCAP) approach, discounts the resource's ICAP to reflect the capacity that it is expected to be able to provide during potential shortages, expressed in MW-equivalents of always-available capacity. While the ICAP approach is the most straightforward and simple, we recommend the QCAP approach for its superior reliability and economic benefits. We understand that various stakeholders support the proposed approach.

Mechanics of the QCAP and ICAP Approaches

Most other jurisdictions with capacity markets rely on either an approach similar to the QCAP approach, which discounts a resource's nameplate capacity, or an ICAP approach, to reflect its expected marginal contribution to system reliability. The difference between the ICAP and QCAP approaches does not directly impact the amount of cleared capacity and system reliability. Under both approaches, the market seeks to procure enough capacity to achieve the reliability requirement and reserve margin calculated in a loss of load hours (LOLH) study.

An ICAP-based reliability requirement is higher on a MW-basis since the capacity counted under a QCAP approach reflects the lower number of MW of always-available capacity that provides equivalent contribution to reliability. However, the reliability requirement in both ICAP and QCAP approaches describe the same underlying level of reliability established in the LOLH study. This is captured in Figure 12, which shows two examples of the same level of capacity being procured to meet the same LOLH of three hours, yet different ways of accounting for the procured capacity. The differences between the QCAP and ICAP approaches are summarized in Table 9.

Figure 12: ICAP vs QCAP Illustrative Reserve Margin



Notes: In this figure we make the simple assumption that all resources have a 10% outage rate, so there is no "merit order" switching between ICAP and QCAP approach. Below we discuss the more complex cases where the fleet differs from expected, and when resources vary from each other.

Table 9: Summary of ICAP versus QCAP Design Elements

Design Element	ICAP	QCAP
Capacity for Thermal Resources	Maximum rated output of the supply resource, *i.e.,* nameplate capacity	Maximum output rating adjusted for expected outages that reduce resources' resource adequacy value
Capacity for Intermittent or Use-Limited Resources	For resources whose nameplate capability may be materially different from their reliability value (*e.g.*, wind, solar, storage, hydro), special accounting rules are often employed	Like ICAP, different approaches are required to estimate QCAP MW. Guiding principle is that 1 MW of QCAP should provide equivalent reliability value across resource types
Reliability Requirement	Traditional reserve margin standard consistent with reliability requirement, expressed in ICAP MW terms	Same reliability requirement, but expressed in lower QCAP MW terms based on the fleet mix and associated outage rates modeled in the LOLH study

Approaches in Other Jurisdictions

Other jurisdictions have varying approaches to account for planned and unplanned outages in rating capacity. PJM, MISO, NYISO, and Alberta all rely upon an "unforced capacity" (UCAP) methodology, which is very similar to the QCAP approach discussed here in that it accounts for expected performance during potential shortage events, except it does not consider planned outages. Ireland and the UK use a derating factor to achieve a similar impact and capture resources marginal reliability contributions. ISO-NE differs from other capacity markets in that it operates on an ICAP basis. PJM, MISO, and NYISO all rely upon an estimate of equivalent forced outage rate demand (EFORd) to account for unplanned outages to qualify UCAP for traditional resources. While the underlying principles are similar across these three markets (*i.e.*, UCAP is the function of a resource's ICAP rating and EFORd), the details in the calculation of EFORd differ slightly. We provide below a short summary of the methodology applied in each of these jurisdictions.

PJM uses an annual EFORd value for conventional resources calculated based on forced outage data from October through September of the previous year.[27] The EFORd is finalized for all resources at least one month prior to the third incremental auction in PJM (roughly three months before the start of the delivery period).[28] If a resource has less than twelve months of available service data, a class average EFORd is applied for that resource. This EFORd calculation includes outages that are deemed outside management control events, including events related to transmission/distribution, acts of nature, fuel quality, and unforeseen regulatory action among others.[29]

NYISO uses a seasonal EFORd, calculated separately for its summer and winter auctions based on a rolling annual average of resource availability. For the winter capability period, the EFORd considers the average outages over the twelve-month periods ending in January, February, March, April, May, and June from the prior year (*i.e.*, the average of those six twelve-month periods). The summer EFORd is calculated similarly, considering the six twelve-month periods ending July, August, September, October, and November.[30] For new generating resources, NYISO relies upon NERC class averages (if NERC averages are unavailable, NYISO estimates EFORd based on the class average of same type resources). Outages that are considered outside of management control events are counted as forced outages, similar to PJM.[31]

MISO relies on the three-year average EFORd to calculate UCAP.[32] Unlike PJM and NYISO, outages that are considered outside of management control are excluded from MISO's derating calculations, which is referred to in MISO as the XEFORd. In instances when the resource has fewer than three years of available outage data, MISO will use all the data that is available unless a resource has less than twelve months of available data in which case, they will use a class average XEFORd based on fuel type and size.

---

27 See PJM, PJM Manual 18: PJM Capacity Market, Section 4.2.5 (p. 68), January 1, 2019.

28 See PJM, Reliability Assurance Agreement among Load Serving Entities in the PJM Region, Schedule 5, Section B, September 17, 2010. For overview of PJM incremental auction schedule, see PJM, RPM 101 Overview of Reliability Pricing Model, slide 53, April 18. 2017.

29 Prior to the introduction of capacity performance in PJM during the 2018/2019 delivery year, the EFORd calculation excluded such events. For a more complete description of outside management control events, see PJM eGADS OMC ("Outside Management Control") Guidelines.

30 NYISO, Manual 4 Installed Capacity Manual, pp. 50-51, March 2019.

31 *Ibid*, p. 58.

32 MISO, Business Practices Manual 11: Resource Adequacy, Appendix H, pp. 131-139, February 20, 2019. Available from: https://www.misoenergy.org/legal/business-practice-manuals/.

AESO created its own resource qualification methodology to best meet system needs during periods of system stress.[33] For conventional resources, AESO uses an availability factor, which captures the availability of a resource during the tightest 250 supply cushion hours in each year, over a five-year period, for a total of 1,250 hours. This availability factor is applied to a resource's capacity to calculate an availability factor UCAP, or "AF UCAP."[34] Unlike in the other markets surveyed above, all outages, included planned outages and physical delists, count against availability.[35] When a resource has fewer than 300 hours of available historical data, AESO uses a class average to fill in the remaining hours.[36]

Ireland specifies derating percentages according to technology class (*i.e.*, gas turbine, hydro, solar) in its capacity market accounting for planned and unplanned outages.[37] For some resource types, there are curves that specify a derating factor based on the nameplate capacity of the resource.[38] The individual percentages and curves are derived using a model that repeatedly simulates the probability of scarcity in the I-SEM and Great Britain market.[39] All resources of that type use the same derate no matter the age or condition of the resource. If a new resource joins the market for which there is not a specified derate, that resource uses a system-wide curve.

Great Britain has a very similar scheme to Ireland. In Great Britain, generators are derated based on derating factors for each resource type.[40] The derating factors are calculated based on fleet availability during the seven preceding Core Winter Periods.[41]

ISO-NE differs from the markets described above in that it relies on an ICAP based capacity rating approach. This means that ISO-NE does not consider a resource's unplanned or planned outages in its procurement of capacity. Intermittent resources are adjusted to account for their intermittent nature. Additionally, a resource's capacity commitment can be reduced to account for failure to perform in prior years.[42]

---

33 Note that the summaries regarding AESO reflect the most recent proposal before the capacity market design was cancelled.

34 AESO, Alberta Electric System Operator Application for Approval of the First Set of ISO Rules to Establish and Operate the Capacity Market, p. 77, January 31, 2019. Available via the AUC eFiling System: http://www.auc.ab.ca/pages/apply-or-access-applications.aspx.

35 *Ibid,* p. 79.

36 *Ibid,* p. 79.

37 SEM-O, Capacity Market: The Quick Guide to Understanding Qualification, April 3, 2019, pp. 3–4.

38 SEM-O, Capacity Market – Final Auction Information Pack, August 3, 2018, pp. 14–16.

39 Capacity Requirement and De-Rating Factor Methodology Detailed Design: Decision Paper, SEM-16-082, December 8, 2016, p. 21.

40 National Grid, Capacity Market Auction Guidelines, July 19, 2018, pp. 5–6.

41 Informal Consolidated Version of the Capacity Market Rules, Rule 2.3.5, July 26, 2019.

42 CRA, A Case Study in Capacity Market Design and Considerations for Alberta, March 30, 2017, Page 83.

ADVANTAGES OF THE QCAP APPROACH

We recommend adopting a QCAP approach. A QCAP approach provides three primary advantages over ICAP, including: (1) *uniformity and interchangeability,* with 1 MW of QCAP contributing the same expected reliability value regardless of resource type, age, or other characteristics, which provides greater assurance of meeting reliability objectives especially if procured resources differ from those assumed in the LOLH study; (2) *fairness,* in that suppliers are rewarded in proportion to expected reliability value; and (3) *more economic asset selection,* in that the auction will be more likely to procure the resources that provide the most reliability value at the lowest price, and this also has the beneficial effect of incentivizing and rewarding reliability improvements in the fleet.

The advantages of the QCAP approach are particularly pronounced when (1) actual resources' reliability characteristics differ from those modeled in the LOLH study; and (2) their characteristics differ from each other. For example, suppose the market clears a set of resources that have an average outage rate that is higher than that assumed in conducting the LOLH study. In this example, the ICAP approach may not achieve the desired level of reliability, as it clears the same amount of ICAP capacity as required by the LOLH study, but the capacity that it cleared is more prone to outages than the capacity assumed in the LOLH study. Additionally, under an ICAP approach, the auction may be more likely to clear capacity that has these outage risks since that might be the cheapest available capacity. These concerns are mitigated under a QCAP approach, where the reserve margin is set to reflect the necessary capacity to achieve the LOLH target, and the auction offers reflect the marginal capacity value of the resource. This way, the QCAP approach procures exactly the needed amount of capacity from the auction, without any potential for under- or over-procurement.

This key advantage of the QCAP approach is illustrated in the example in Table 10. In this example, a reliability study determined it would need a 130% ICAP reserve margin or a 117% QCAP reserve margin to achieve the target reliability of 3 LOLH given a 10% average fleet outage rate. If the 10% outage rate assumption was correct in the study and reflects the cleared capacity, using an ICAP or QCAP approach both yield the reliability standard. However, if the assumed outage rate is incorrect, the two approaches yield very different results. Under the ICAP approach, the auction still procures capacity to achieve a 130% ICAP reserve margin although the performance of the capacity procured differ from what was expected. This means that if the outage rate is much higher than anticipated, the system will be less reliable, evidenced by the 20% outage rate which yields a LOLH of 10 hours. Alternatively, if the outage rate is lower, the market over procures capacity, and consumers overpay for a level of reliability that might not be necessary. By contrast, the QCAP approach is able to procure the desired level of capacity as determined by the reliability study, since a QCAP approach procures capacity that reflects its actual marginal reliability contribution.

Table 10: ICAP versus QCAP Performance with Variable Outage Assumptions

	Cleared Resource Actual Outage Rate	QCAP (% of Peak Load)	ICAP (% of Peak Load)	LOLH	
ICAP Methodology	10%	117%	130%	3	Uncertainty yields variable reliability under ICAP method; increases cost if higher ICAP needed to avoid exceeding 3 LOLH
	20%	104%	130%	10	
	0%	130%	130%	~0	
	Cleared Resource Actual Outage Rate	**QCAP** *(% of Peak Load)*	**ICAP** *(% of Peak Load)*	**LOLH**	
QCAP Methodology	10%	117%	130%	3	Uncertainty yields consistent reliability under QCAP methodology
	20%	117%	146%	3	
	0%	117%	117%	3	

This example highlights how a QCAP approach yields consistent reliability no matter the underlying makeup of the fleet. It also shows how an ICAP approach can lead to variable reliability, which leads to higher costs if it clears additional capacity above the target reliability level, or breaching the reliability standard, when the resources cleared in the capacity market differ substantially from those expected and modeled in the LOLH study.

Finally, the QCAP approach achieves the most economic asset selection and ensures that resources clear in the fairest manner. Since capacity bids under a QCAP approach reflect each resource's marginal reliability value, it is possible to clearly and efficiently rank resources by their ability to contribute to reliability. This might not be true under the ICAP approach, where resources with a higher outage probability may be able to offer at a lower price and, therefore outcompete resources that could provide more reliable capacity at lower cost. This is captured in the 20% outage scenario above, where resources that were less reliable than expected cleared and caused a lower reliability outcome. Table 11 summarizes the pros and cons of each approach.

Table 11: Advantages and Disadvantages of QCAP and ICAP Systems

	ICAP	QCAP
Advantages	• Simpler to calculate resources' capacity ratings (but availability ratings and calculations may still be required for penalty assessments, so administrative savings are minimal)	• Uniformity among resources provides better assurance that the capacity auction will achieve the desired reliability level • Most level playing field among resource types, with payments proportional to reliability value (which incentivizes and rewards cost-effective reliability enhancements) • More likely to procure better-performing resources • More compatible with future treatment of intermittent and energy-limited resources
Disadvantages	• Adverse selection of resources that under-spends on maintenance and fuel security and thus perform poorly	• Different from historical approach • Potentially increased administrative effort

	• No protection from procuring a fleet with higher outage rates than assumed in the LOLH model, resulting in poor reliability • If playing field is levelized through higher penalties for non-performance, the magnitude of these greater payment adjustments would be less transparent to the market than a simple QCAP-based price • Favors traditional resources over intermittent and energy-limited resources that presumably will be derated for their unavailability during peaks	

On the above considerations, we recommend adopting a QCAP approach, rating resources according to their expected reliability value (in terms of MW-equivalents of always-performing capacity) and defining the reliability requirement accordingly. This provides better assurance that the FCM achieves the desired reliability level, creates a level playing field across resource types based on their reliability value, rewards better performing resources appropriately, and ensures fair and equitable treatment.

Recommendation
Capacity Rating Approach • Adopt a "qualified capacity" (QCAP) approach that accounts for the marginal reliability value of each resource type

B. Minimum Size for Participation

In general, the guiding principle should be to enable the smallest minimum participant size possible that is operationally feasible for EMA to qualify. This is especially important to enable participation by technology types that may be relatively disaggregated, like storage and demand response. Thus, the threshold for the minimum size of a supply resource allowed to participate in the FCM must balance two key factors:

- The value of enabling broad participation from all resource types that may be able to provide cost-effective capacity (which tends to imply a lower participation threshold); versus
- The administrative cost of qualifying very small resources that provide relatively low value to the system (which tends to imply a higher participation threshold).

As described in Table 12, other markets have weighed these factors differently, resulting in a range of minimum participation size between 0.1 MW and 2 MW.

Table 12: Minimum Participation Size in Other Capacity Markets

Jurisdiction	Resource Type	Minimum Participation Size	Aggregation Allowed to Reach Minimum Size?
PJM	Generation Metered	0.1 MW	Only in same location
	DR	0.1 MW	Yes
ISO-NE	Generation Metered	1 MW	No
	DR	0.1 MW	Yes
NYISO	Generation Metered	1 MW	Only in same location and w/ same owner
	DR	0.1 MW	Yes
UK	All	2 MW	Unclear
AESO	All	1 MW	Unclear

After consultation with EMA, we recommend a minimum participation size of 1 MW (in ICAP terms, not QCAP) for all resource types, with aggregation allowed to reach this minimum size. We believe this provides an appropriate balance of the factors described above in the Singapore context.

C. Traditional Generation

Traditional fossil generators such as CCGTs, OCGTs, and oil-fired steam plants can be dispatched to serve load when needed, as long as they are online and available. Thus, their marginal reliability value is similar to an ideal, always-performing resource but must be adjusted for expected unavailability to serve during possible shortages. As noted below, the relevant measure of unavailability in Singapore should include planned outages as well as unplanned since shortages can occur any time in the year, including when generators plan their maintenance outages. Response times matter too, since supply shortages occur by surprise when large generators trip, so we propose a maximum response time that is consistent with resource capabilities and system needs.

QCAP Equation and Rationale

We propose that each traditional generator's QCAP be given by its ICAP discounted by its expected planned outage rate (POR) and historical unplanned outage rate (UOR) as described in the equation below:

$$QCAP_{thermal} = ICAP_{thermal} \times (1 - POR) \times (1 - UOR)$$

where: *POR* = declared planned outages for the delivery year in days / total number of days in a year[43]

UOR = one-year historical unplanned outages in days / (total number of days in a year – historical planned outages in days)

[43] The use of "day" in defining the POR and UOR is transitional due to current data availability limitations. EMA intends to transition to half-hourly data as higher granularity data becomes available.

We propose that this approach is applied only to existing conventional dispatchable resources with sufficient historical operational data to estimate the necessary parameters. Non-dispatchable resources, new resources, and other resources without sufficient historical operational data will require different approaches, described below.

The QCAP rating should be adjusted over time to reflect the recent performance of the resource. For example, if a resource with a high capacity rating significantly underperformed one year, it should not be relied upon to provide that same level of capacity again until it has proven that it has addressed the underlying issues and can perform reliably again.

As part of the resource qualification process for each auction, PSO will require that all existing traditional generators submit historical performance records, demonstrating the ability of a resource to maintain at least 90% loading of its installed capacity for a minimum of two hours. Similarly, during the delivery year, traditional generators' holding a CSO will also be required to continue demonstrating their ability to maintain at least 90% loading of its registered capacity for a minimum of two hours, in each calendar quarter. In addition, PSO may also conduct surprise tests on generators to further assess their capabilities.

In addition, the QCAP rating for traditional generators that do not possess sufficient historical operational data (*i.e.*, new resources), will be based on class average performance metrics—that is, the unplanned outage rate of new CCGTs will be based on the average unplanned outage rate of all existing CCGTs (possibly differentiated by CCGT technology class).

ACCOUNTING FOR UNPLANNED AND PLANNED OUTAGES IN DETERMINING QCAP

Under the QCAP approach, we recommend accounting for all unplanned and planned outages to accurately capture a resource's marginal reliability value. While most markets simply rely on an UOR to derate from ICAP, we recommend that the Singapore market also account for resources' POR to capture the full range of outages that may impact a resource's ability to provide capacity when needed. Most international capacity markets have more seasonal demand variation and are able to plan all of their maintenance in shoulder months. However, in Singapore, additional maintenance scheduled in any month can impact reliability due to Singapore's relatively consistent peak load profile throughout the year. By accounting for all types of outages, the QCAP approach incentivizes resources to maximize their overall resource adequacy value by optimally managing maintenance decisions while also not jeopardizing overall system reliability.

We recommend using an annual UOR and POR to determine QCAP in Singapore, as opposed to focusing on estimating availability for one particular time period during the year, because shortages in Singapore are equally likely to occur in any time of the year. This follows from the fact that the annual load duration curve in Singapore is relatively flat compared to other jurisdictions, and that the primary drivers of resource shortages are unplanned or forced outages, not peak load.

The POR will be defined as the share of hours across the delivery year during which the resource declares to be unavailable due to planned maintenance outages, based on maintenance schedules. The UOR will be based on historical data for the past one year, aligned with the probability that the resource was not available due to unplanned outages (including forced outages, outages for ad hoc, urgent repair and/or unplanned derates). For both outage types, we propose that the planned/unplanned outage duration ends when a unit is back in operation, *i.e.*, connected to the power system and running at or above Minimum Stable Loading level for a minimum of four periods (two hours).[44] Furthermore, for both outage types, the POR and UOR will be updated in the forward period as new information becomes available. In particular, the planned outage rate will be updated to reflect maintenance schedule changes closer to the delivery year. The unplanned outage rate will be updated to reflect more recent historical outage patterns. QCAP revisions in the forward period resulting from updated POR and/or UOR may enable a resource to offer additional capacity into a rebalancing auction (when QCAP increases above cleared capacity) or require a resource to buy-out part of their CSO in a rebalancing auction (when updated QCAP falls below cleared capacity).

Estimating the UOR requires a precise definition of what constitutes an unplanned outage or unplanned derate. The North American Electric Reliability Council (NERC) defines an unplanned, or "forced," outage as the removal from service availability of a generation for emergency reasons or being unavailable due an unanticipated failure.[45] Different jurisdictions have different views on what types of outages are considered unanticipated. To increase the likelihood that QCAP capacity is available, we recommend a conservative approach that includes all unplanned outages. This consistent with the PJM and NYISO definition, that includes outages that are deemed outside management control events, including events related to transmission/distribution, acts of nature, fuel quality, and unforeseen regulatory action.

---

44 As time is required to update the rules/manuals, in the first Compressed Auction, the planned and unplanned outages (in days) of existing units shall be based on the duration indicated in the approved Generating Unit Outage Request Form, that market participants submit to PSO for maintenance, repair, upgrading or commissioning, as stipulated in the latest version of PSO's System Operation Manual.

45 North American Electric Reliability Council, "Glossary of Terms used in NERC Reliability Standards," Updated May 13, 2019. Available at: www.nerc.com/files/glossary_of_terms.pdf.

D. Solar

Standalone solar resources are intermittent in nature; they can only contribute to reliability to the extent they provide energy when supplies become scarce and load might otherwise be shed. Hence, for stability of the power system, EMA proposes for solar generation which participates in the FCM auction to mitigate intermittency, through the Intermittency Pricing Mechanism and/or coupling with their own solution(s).[46] Nevertheless, even if they have less value per installed nameplate MW than always-available dispatchable resources, they should be recognized in the capacity market with a derated value. Ideally, the capacity rating for solar resources should be set such that each MW of QCAP from solar resources provides the same marginal improvement to reliability metrics (*i.e.*, LOLH) as a MW of QCAP from dispatchable resources. To accurately and simply capture the reliability contribution of solar resources, we recommend setting the capacity rating of solar resources at their expected average capacity factor during on-peak periods (to be revisited as penetration increases).

Solar QCAP Rating Approach

Similar to dispatchable resources, the capacity rating for a given solar resource must consider the overall nameplate capacity of the resource as well as the expectation for how often the resource will be unavailable during shortage conditions. This general capacity rating can be captured simply in the formula below:

$$QCAP_{solar} = Nameplate\ Capacity \times Derated\ Performance\ Factor$$

There are a few potential approaches to set the derated performance factor. The most simple would be to set the derated performance factor equal to the simple average of the hourly capacity factors (where "hourly capacity factor" describes the amount of output from the solar resource in a given hour divided by its nameplate capacity):

$$Simple\ Performance\ Factor = \frac{1}{N}\sum_{i=1}^{N} g_i$$

where: N is the number of periods (in one year, with half-hourly periods, $N = 17{,}520$)
g_i is the metered generation observed in period i for 1 MW nameplate resource

This approach does not consider when the solar resource generates and whether or not those times coincide with when the system is actually at risk of shortage. From our analysis of the provided data on standalone solar, this approach yields a factor of 17.9%. Therefore, a 10 MW solar resource would have a QCAP of 1.79 MW.

46 EMA intends to separately, via the Intermittency Pricing Mechanism announced in October 2018, allocate to intermittent generation sources their fair share of reserves costs. See EMA, "Intermittency Pricing Mechanism for Intermittent Generation Sources in the National Electricity Market of Singapore," October 30, 2018. Available at: https://www.ema.gov.sg/cmsmedia/Final%20Determination%20Paper%20-%20Intermittency%20Pricing%20Mechanism%20vf.pdf.

The second approach similarly takes an average of hourly capacity factors, but it weights these values by the "Probability of Lost Load" (POLL) in that hour, where the hourly POLL is an output of EMA's reliability modeling:

$$POLL\ Weighted\ Performance\ Factor = \frac{\sum_{i=1}^{N} w_i \cdot g_i}{\sum_{i=1}^{N} w_i}$$

where: N is the number of periods (in one year, with half-hourly periods, $N = 17{,}520$)
g_i is the metered generation observed in period i for 1 MW nameplate resource
w_i is the system POLL in period i

By accounting for the probability of shortages in each hour, this approach captures the contribution of the output from a given solar resource to address the system's reliability needs. It places more value on the daytime and evening hours that are at higher risk of experiencing a shortage, while placing almost no value on the overnight and early morning hours that present almost no risk of experiencing a shortage.

Because of the positive correlation between solar output and POLL, this approach yields a factor of 31.9%. A 10 MW solar resource would accordingly have a QCAP of 3.19 MW. This is nearly twice the value suggested by the simple capacity factor after taking into account the pro-cyclical pattern between solar output and POLL. However, this approach is more complicated and is sensitive to assumptions in the POLL modeling.

The third approach takes the simple average of solar output across on-peak periods defined as those hours with non-zero and positive POLL:

$$On\ Peak\ Performance\ Factor = \frac{1}{N_P}\sum_{i=1}^{N_P} g_i$$

where: N_P is the number of on-peak periods (*e.g.*, if defined as 9:00am to 10:00pm, with half-hourly periods, $N_P = 9{,}490$)
g_i is the metered generation observed in period i for 1 MW nameplate resource

This approach combines the benefits from each of the first two approaches. It is simple and yet able to capture the capacity value of solar when shortages are most likely. In fact, the factor from this approach is 31.5%, almost equal to the estimate from the more robust POLL-weighted average approach, as seen in Table 13.

Table 13: Performance Factors by Approach

Simple Performance Factor	17.9%
POLL-Weighted Performance Factor	31.9%
On-Peak Performance Factor	31.5%

Given the relatively low levels of solar penetration currently in Singapore, the on-peak average capacity factor is workable. However, it may be less accurate in the future as more solar enters the system and reliability needs shift towards decreasing the POLL in the hours when solar generates and increasing POLL in the evening when solar is not generating as much. This shift will reduce the marginal reliability value of solar overall, and especially during the peak period.

In this high-solar future, using the broad on-peak period average would likely mischaracterize the contribution of solar resources to helping meet the reliability standard. With high enough solar penetration, it will be necessary to embrace a more precise rating approach.

The on-peak average performance factor for an existing solar resource can be calculated based on the actual historical generation of the solar resource. Relying on historical data from a representative period, would be sufficient to inform an expectation of future performance. Existing solar resources will be qualified based on their metered historical MWh generation data. New solar resources that do not have such historical data would need to rely on class averages that reflect expected performance. EMA should determine and publish the on-peak periods in advance of resource qualification for each auction.

Recommendations

Capacity Rating Approach

- QCAP for a solar resource will reflect its specific on-peak average capacity factor
- Existing resources qualified using one-year of historical metered generation data
- New resources qualified using class averages

E. Demand Response

Other jurisdictions have proven that Demand Response (DR) can provide a major source of reliable and flexible capacity that can help ensure resource adequacy cost-effectively. Similar to other sources of capacity, DR can help maintain adequate supply to meet demand even under adverse conditions of high load combined with multiple generator outages. It allows loads with flexibility to contribute to meeting reliability standard.

Demand response may be perceived as riskier than generation because it is not "steel in the ground," but it can be managed to be just as reliable. Even though some individual end-user participants may not always materialize or respond, DR aggregators assemble portfolios conservatively to solidify their ability to meet their obligations. The experience in other jurisdictions has been that DR performs at least as well as traditional generation when called upon.

While Singapore already has DR and interruptible load (IL) programs in the energy and/or ancillary services markets, a new DR capacity product calls for a wholly different design. The qualification requirements for DR should reflect the needs of the system in Singapore while recognizing the characteristics of the resource.

We have developed an approach to qualify DR based on lessons learned from other jurisdictions. The proposed design meets EMA's design criteria of encouraging participation, and accurately representing the contribution of the resource to system reliability without over-complicating the market design and implementation. In short, our recommendation is to qualify DR based on plans (similar to other new resources), subject to certain performance requirements: that resources have to be available during the hours that they are qualified for when determining QCAP, and to offer into the real-time energy and/or ancillary services markets, and they have to be able to reduce load for at least four hours.

The Process for Qualifying DR Resources

Other jurisdictions that have successfully developed and relied on DR resources have shown that most DR is provided by aggregators. These aggregators have highly specialized skills in recruiting customers and harnessing their operational flexibility to shed load or dispatch behind-the-meter generation when called. A reality of their business model that must be accounted for in the qualification process is that they take time to recruit their customers and then typically sign contracts for only a few years. They may not have all of their customers committed four years in advance at the time of the base auction, particularly not in the first several years of growing their customer bases. Therefore, they have to be able to qualify capacity based on plans to attract and retain customers, not just firm contracts.

Aggregators must submit a customer acquisition and retention plan to qualify capacity for the auction, just as a planned generator would have to submit a project development plan, subject to review by EMA. Plans must include a marketing strategy, target new/existing customers or market segments, assume share of each segment the aggregator can recruit, and estimates of how much net load reduction is realistic with each type of customer, corresponding to their sizes and operational characteristics. If any end-use customers have already been contracted, sites of the load should be detailed in the plan.

DR aggregators must include development milestones in their plans, similar to construction and testing milestones for generation resources. This includes marketing and customer acquisition milestones, as well as a timeline for securing necessary permits, financing, and equipment orders. Finally, the plan should specify testing milestones and an expected commercial online date, including a final testing milestone several weeks before the commitment period begins.

EMA would review the proposed resource using all of the customer and timeline information provided in the business plan. EMA would also review all DR providers' plans in aggregate, to identify potential overlap of end-use customer targets and derate the portfolios if necessary. Each aggregator's plan must be approved prior to the auction, in order to place an offer and earn a capacity obligation, and secured with the same financial assurance instruments as other new resources, as discussed in Section V below. After the forward auction, EMA would monitor the achievement of milestones, where certain progress failures may result in requiring the aggregator to acquire replacement capacity (and transfer financial assurance obligations) or pay penalties.

Technical Requirements Affecting Qualification and Ratings

Demand response resources are subject to technical requirements that affect their qualification and capacity rating. These technical requirements should be rooted in the specific reliability needs of the system, as well as the capabilities DR aggregators and participating end-use customers.

Required Hours of Availability. To provide resource adequacy value, DR has to be available during the hours when shortages are most likely. In Singapore, this generally corresponds to non-holiday weekday peak hours, aligning with the times when businesses are using energy and could provide load reductions. Similar to the approach for solar, EMA should determine and publish the required hours of availability in advance of the qualification period for each auction and should define these hours to reflect the hours with greatest shortage risk (*i.e.*, on-peak periods). Providers may also be allowed to nominate a more restricted set of hours, within the required hours of availability, for a reduced capacity rating. For example, if a DR provider is only available from 12:00pm to 6:00pm (6 out of, for example, 13 on-peak hours), its QCAP would be $^{6}/_{13}$ of its self-nominated DR capacity.[47] This accounts for the resource's ability to contribute during hours of peak reliability need.

Maximum Notification Time. Like generation resources, DR is subject to start-up time limitations. We recommend specifying a single maximum notification time for all DR. While some other jurisdictions enable longer or shorter notification times, shortages in Singapore are primarily driven by unexpected forced outages rather than forecastable high peak load conditions. This situation lends itself to shorter notification time requirements to react to sudden generator outages.

Duration Requirements. This duration could be set equal to the average shortage duration, the 90th percentile shortage duration, or another measure of the expected duration of shortage events. EMA has determined that four hours would cover most possible shortage events and should therefore be the requirement.

Allowable Interruptions. The qualification rules must clearly define any maximum limit on the total number of interruptions during a delivery year or specify that there is no interruption limit. We recommend specifying "no limit" to avoid having to discount the QCAP relative to other resources, and at no great cost to the DR provider.

Measurement and Verification

There are two main concepts for DR to provide capacity, and each is best suited to different types of end-use customers. We recommend accommodating both. The firm service level (FSL) approach is suitable for customers with baseload essential loads and varying non-essential loads that they are willing to shed. An FSL asset agrees to reduce its load to a specified firm level in the event it is called upon. It is compensated based on the difference between a forecast baseline load and a target FSL of load to which it must reduce consumption when called upon. The forecast baseline should be established as the average historical consumption from the prior year corresponding to the hours when the resource would need to be available (see above). For example, a customer with a 10 MW forecast baseline load and a 4 MW FSL would have to reduce to 4 MW when called. It would be credited for providing 6 MW of capacity, even if it was consuming at 9 MW or 11 MW just before being called. This reflects the fact that the forecast baseline load (reflecting the customer's peak load contribution and how much capacity it pays for over the course of the year) is based on annual characteristics, and would not be updated to reflect behavior on an hourly basis. Thus, there is no reason the customer would have to "maintain" its consumption at its forecast baseline in order to be fully compliant.

[47] DR providers that can only provide load reductions during hours that do not fall within the required hours of availability, would effectively have a QCAP of 0 MW.

By contrast, the guaranteed load drop (GLD) approach is more suitable for customers whose consumption might vary but whose load reduction ability is constant: those with backup generation or a specific fixed load they can interrupt on short notice. A GLD asset's capacity value is equal to the amount of load it can shed from a running baseline when called upon. The running baseline would be dynamic, based on consumption in the trailing non-event hours from prior days, with adjustments for the level of consumption in the day of, in the intervals just prior. For example, a resource with a 4 MW CSO that is consuming 10 MW when called would be required to reduce net load to 6 MW. The resource would be credited for providing 4 MW of capacity. As with FSL, there is no pre-set baseline of consumption that the customer must maintain in order to comply, as that would be unnecessary and economically inefficient, except for the CSO itself. Consumption below the CSO obviously precludes the customer from providing sufficient load drop to meet their CSO, and they would be penalized accordingly as discussed in Section X.

Both GLD and FSL assets must prove their availability and performance, as all capacity resources must. But because DR is like other rarely-run resources, monitoring relies primarily on a few actual events, complemented by in-year testing requirements. We recommend that EMA require that the DR resources respond to surprise tests initiated by PSO throughout the delivery year, rather than allowing providers to self-schedule tests, to yield more realistic and accurate results. Testing output must match or exceed an aggregator's committed capacity.

Operational and Other Market Issues

DR differs from traditional generation resources in how it is deployed. DR capacity is required to provide load reductions only during shortage or emergency conditions. This results in only rare deployment, which is attractive to the vast majority of customers providing DR who are willing to provide an option but do not want their business operations impacted frequently.

Here is how deployment would work: when shortage conditions are anticipated based on conditions in the pre-dispatch process ahead of real-time, all DR resources are required to offer into the real-time energy and/or ancillary services markets for scheduling based on the prevailing DR and IL framework. Their offer price can be up to the relevant market price cap and scheduled based on real-time clearing outcomes. Real-time price formation protocols would allow them to set the price at their offer, rather than depressing prices if required to offer at zero price. In addition, PSO can still, in accordance with the real-time market rules, activate DR resources (subject to their hours of availability and maximum notification time) during an actual scarcity period, even if they had offered but were not scheduled for that period. Failure to comply will constitute a breach of the CSO.

DR resources can also participate in real-time energy and/or ancillary services markets, as long as it is still able to meet its capacity requirements when called upon, as described above. Details of interactions with the existing DR and Interruptible Load (IL) framework are currently being reviewed by EMA.

Jurisdictional Review

Brattle has conducted a review of DR capacity products in other jurisdictions, summarized in Table 14 below. These jurisdictions have significant DR participation throughout their wholesale markets and utilize qualification and operational requirements tailored to the needs and capabilities of their system.

Table 14 Summary of Demand Response in Other Jurisdictions.

	PJM	ISO-NE	AESO	ERCOT ERS
Qualifying Market	Capacity	Capacity	Capacity	ERS is a DR-only capacity product within ERCOT's "energy-only" market
Qualification Criteria	Unlimited interruptions 30-min lead time (can apply for 60- and 120-min if necessary) Qualified based off of customer acquisition plan	Unlimited interruptions 10- and 30-min lead time Qualified based off of customer acquisition plan	Based on customer acquisition plan If DR is not able to produce >75% of its stated UCAP by second rebalancing auction, it must buy out of the difference between tested production and UCAP	10- and 30-min lead time Can qualify as weather sensitive or non-weather sensitive Qualify for 3-4 hour time blocks across three seasons
Measurement Approach	Both firm service level and guaranteed load drop	Only firm service level	Both firm service level and guaranteed load drop	Both firm service level and guaranteed load drop
DR Operational Process	Called when all non-emergency resources are exhausted. Longer lead-time DR called first. Dispatched according to energy offer or strike price. Can set prices in RT at strike price	Called during shortage conditions. Dispatched according to energy offer. Can set prices in RT, at offer price		Called in emergency conditions. 30-min reserves can be called if reserves are under 2,300 MW; 10-min reserves can be called if under 1,750 MW. Special provisions to avoid RT price reversal

Summary of Recommendations for Qualifying DR

The recommended framework for DR resources is summarized in the table below.

As discussed in Section II, EMA is concerned with having significantly more DR resources providing dispatchable capacity in the Singapore power system, compared to today (with only 11.8 MW of registered IL capacity). EMA therefore intends to impose a 200 MW cap on the amount of cleared capacity provided by DR, for the first compressed auction in Q3 2021 for delivery period Q4 2023. This maximum cap will be reviewed prior to each auction by EMA, taking into account the track record and operational experience with more DR resources in the Singapore system. Likewise, the QCAP rating approach for DR could also be reviewed after more operational experience is gained.

Recommendations and Next Steps

Basis for Qualification

- New and existing DR resources in the forward auction should be qualified based on their business plan to acquire and retain customers
- Plans will be monitored for progress on achieving development milestones, with forfeiture of financial assurance for not meeting the milestones
- A 200 MW cap on the amount of cleared capacity provided by DR for the first compressed auction in Q3 2021 for delivery period Q4 2023. This maximum cap will be reviewed prior to each auction

Capacity Rating Approach

- Allow new DR resources to submit their own capacity ratings in their business plan
- Accurate self-rating is enforced via measurement and verification (M&V) at the individual asset level and aggregated portfolio level, with penalties similar to other capacity resources

Firm Service Level (FSL) or Guaranteed Load Drop (GLD) Approaches

- Allow both to accommodate different types of customers
- FSL uses a static historical baseline corresponding to all hours of availability, its DR contribution is calculated as the baseline minus its FSL, and compliance involves reducing net load to FSL
- GLD uses a running baseline, and its compliance is based on reducing net load by GLD

Performance Requirements

- **Required hours of availability**: DR resources should provide capacity during the defined required hours of availability. EMA will consider allowing resources to nominate more restricted hours and have their QCAP rating derated as a fraction of the required hours of availability
- **Notification time**: EMA to establish a reasonable maximum response time consistent with resource capabilities and system needs
- **Duration**: The dispatch duration should cover the length of typical shortage events, which EMA has determined to be at least four hours
- **Interruptions**: There should not be a limit on the number of interruptions for which DR assets are responsible for responding

Dispatch protocols:

- **Deployment**: In the pre-dispatch process, during anticipated shortage conditions, DR resources shall offer into the real-time energy and/or ancillary services markets for scheduling based on the prevailing DR and IL framework. During an actual scarcity period, PSO may still activate DR resources that had offered but were not scheduled for load curtailment in that period

Next Steps

- Solicit input from stakeholders to determine key parameters, such as allowable notification times and interruption durations

F. Storage

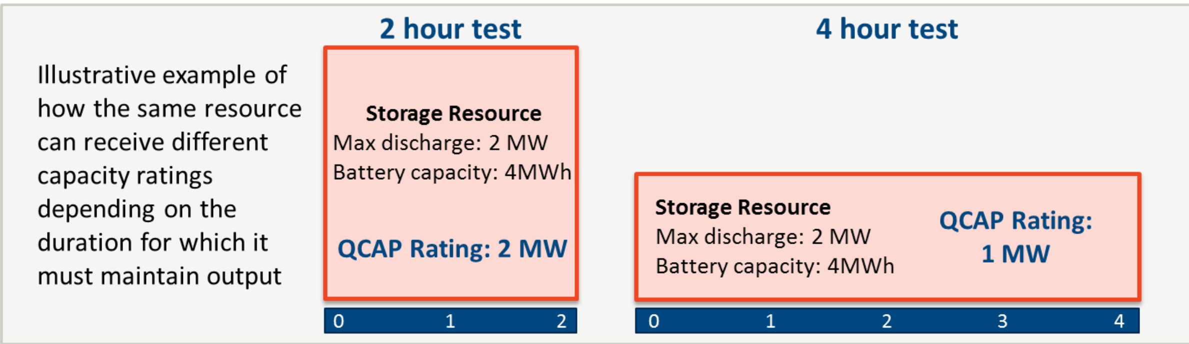
Storage resources provide a valuable source of capacity as they offer dispatchable, highly flexible and reliable capacity to the system. However, they can be limited in their ability to perform for extended periods of time which may prevent them from providing capacity throughout the entirety of a shortage event. This constraint should be properly accounted for in its capacity rating. We recommend rating a storage resource based on the maximum output it could capably sustain during an average outage in Singapore.

Rating storage resources by their maximum output that is capably sustained over a specific duration is widely accepted practice across other North American capacity markets. However, across these jurisdictions there is no consensus on the specific duration. ISO-NE uses two hours; NYISO, MISO and AESO use four hours; PJM uses ten hours.[48]

We recommend that the duration used to evaluate storage resources cover most shortage events, which EMA has determined to be four hours.[49] This allows for storage resources to be evaluated based on the most likely duration that they would be expected to perform. A longer duration requirement would be more conservative but would lower the rating for any resources constrained by their storage capacity. This approach could be attractive if there are concerns that storage resources might not be able to produce according to their manufacturer specified maximum discharge and capacity because they are not fully charged at the start of a shortage event.

A four-hour duration requirement would not preclude the participation of shorter-duration batteries, but would allow them to participate at a reduced QCAP value corresponding to their duration as a fraction of the requirement. This is captured in the illustrative example in Figure 13.

Figure 13: Illustrative Example of Storage Resource Rating



The same storage resource, with a 2 MW maximum discharge and a 4 MWh storage capacity, is given different capacity ratings under different performance tests. If the resource is expected to perform for only two hours, it is able to be rated based upon its maximum discharge since it would be able to sustain that output over the entire period. However, over the course of four hours, the resource is only able to provide 1 MW.

Obligations for storage resources in the delivery year should be similar to those for other dispatchable resources. First, unless storage resources self-nominate a restricted availability

48 The ten-hour requirement in PJM was based upon a practice where traditional generators were tested for their ratings over ten hours. However, this requirement has been criticized by industry which says it is unnecessarily stringent, making it difficult for storage to compete. FERC has asked PJM to review this rule.

49 This is based on the average duration for a conventional CCGT in warm state to ramp up to full load and relieve shortage conditions.

window (which would yield a lower QCAP rating), they should be available to provide energy or ancillary services during scarcity periods. This may occur more than once per day; storage resources should evaluate the risks of failing to perform during scarcity periods due to lower than anticipated state-of-charge when determining their capacity offers. Finally, storage resources should be subject to similar in-year monitoring as other dispatchable resources.

As discussed in Section II and similarly to DR, EMA intends to set a 200 MW maximum on the amount of cleared capacity provided by storage resources (that provide energy and not regulation nor reserves), for the first compressed auction in Q3 2021 for delivery period Q4 2023. This maximum cap will be reviewed prior to each auction by EMA, taking into account the track record and operational experience with more storage resources in the Singapore system. Likewise, the QCAP rating approach for storage could also be reviewed after more operational experience is gained.

Recommendations

Capacity Rating Approach

- QCAP for a storage resource should reflect the maximum output it can capably sustain over four hours, corresponding to the average shortage duration for Singapore
- A 200 MW maximum cap on the amount of cleared capacity provided by storage resources (that provide energy and not regulation nor reserves) for the first compressed auction in Q3 2021 for delivery period Q4 2023. This maximum cap will be reviewed prior to each auction

G. Solar + Storage

Potential entrants may choose to install both solar and storage generation, which are connected to the grid as a single joint unit. Such solar plus storage (S+S) facilities can be valuable resources to meeting the system reliability objectives, leveraging cheap abundant energy from the solar resource and the flexibility and control from the storage resource. While these facilities offer benefits from both solar and storage resources, they also carry with them similar challenges from both resources. It is not always clear when and for how long a S+S will be able to provide capacity. These challenges manifest themselves differently in AC-coupled and DC-coupled S+S facilities, which necessitates different ratings for each type.[50] We recommend rating AC-coupled S+S facilities as the sum of their individual solar and storage resource ratings. For new DC-coupled, we recommend allowing developers to propose their own rating which can be no more than either the size of their solar capacity or inverter, while existing DC-coupled will be rated based upon historical performance.

---

[50] In an AC-coupled S+S facility, PV output flows through one inverter to the grid. For PV to be stored and then injected, power flows through the DC-AC inverters twice to store and then once more to inject. Battery injection passes through the AC-DC inverter once.

In a DC-coupled S+S facility, PV output flows through one inverter to the grid. For PV to be stored and then injected, power flows through the DC-DC inverter twice and a DC-AC inverter once. Battery injection passes through the DC-DC and AC-DC inverter once each.

AC-Coupled QCAP Approach

In an AC-coupled system, both the solar and storage components have inverters that connect them to the grid, allowing both resources to be independently metered. Because the operator has sight of each of the component resources, it is possible to give each resource its own capacity rating. When qualifying for the capacity market, the capacity rating of the overall AC-coupled S+S system should be the sum of two underlying solar and storage resources, reflecting each of their values to the grid. This is accepted practice in other jurisdictions.

DC-Coupled QCAP Approach

DC-coupled S+S systems pose additional challenges because the solar and storage components are tied behind a single inverter. Without sight of the individual resources, the system must be assigned a single capacity value. Determining this capacity value is complex because the specific relationships between solar size, battery size, inverter size and control system greatly impact their possible capacity contributions. For example, a facility that has lots of solar and storage capacity but a very small inverter may produce at near-constant, but low, levels throughout the day, while a facility with a larger inverter and less storage may have more volatile and less controllable production.

Best practices have not been established in other jurisdictions. PJM does not allow DC-coupled capacity supply resources that are not separately metered (and if they are, they would be qualified similarly to AC-coupled); ISO-NE qualifies these resources based on their production during a pre-specified peak period. Other jurisdictions have not developed approaches to address this unique resource.

We recommend a flexible approach. Existing DC-coupled S+S systems should receive capacity ratings based on their historical metered supply into the power grid. We recommend that new DC-coupled S+S nominate their own capacity value, limited by size of the solar array and AC-inverter, before they produce sufficient historical metered data. We believe that abuse of this rating approach by new resources will be limited, due to the potential for significant penalties if they underperform relative to their capacity rating.

Recommendations

Capacity Rating Approach

- AC-coupled S+S facilities should be rated based on the sum of their solar and storage resource capacity ratings
- New DC-coupled S+S facilities should be rated based on determination by the provider, while new DC-coupled S+S facilities should be rated based on historical metered supply into the power grid

H. Imports

Imports can be cost-effective sources of incremental capacity supply for Singapore. We understand that EMA is considering the participation of imports in the FCM. Specifically, 100 MW of imports by 2021 through a trial and another 100 MW of imports around 2022/23.

Beyond that, EMA will study the scope for additional electricity import capacity, taking into account demand growth and new plantings.

There are two primary ways of enabling capacity from imports: the first is to qualify capacity from a specific external resource (plant to grid), similar to how local resources are qualified. The second is to qualify generic external capacity (grid to grid), which is not directly tied to an individual resource, but it still able to reliably provide its rated level of capacity.

Similar to local resources, imports must be accurately qualified and rated. For specific external resources, the qualification process should parallel that of local resources, with a few additional requirements to confirm that the capacity can be reliably imported when needed. In most markets, there are two essential requirements for importers to demonstrate:

- **Non-recallability** is a guarantee from the host region that the resources producing energy for export are not committed to provide capacity in their region, and that the host region will not curtail the associated exports even under emergency conditions.
- **Deliverability** is a requirement that external generation must be supported by firm transmission capability and rights sufficient to support the transfer of the capacity when needed (both within the host region and across the interface into the region qualifying the capacity). To ensure the deliverability of capacity from imports, capacity markets require importing resources to provide proof of deliverability across two segments of a contiguous transmission path: (1) from the resource to a landing point at the border with the importing market, and (2) into the importing market such that the capacity is "deliverable" to the system from a capacity injection perspective.

In addition to these requirements, imports should also face the same performance requirements and obligations as local resources. Imports should also be subject to the same performance and availability penalties as local resources, described in Section X. The QCAP ratings they receive should be aligned with the resource-specific methodologies used for local resources described in the preceding sections, with an additional de-rating factor used to account for expected interconnector failure/outages, further reducing the QCAP of imported capacity.

While we recognize that capacity markets are generally designed to be technology neutral, we recommend that imports be restricted to coming from generation capacity, including traditional thermal generators, hydro, solar, storage, wind and S+S. We recommend excluding demand resources such as DR, energy efficiency and aggregated distributed energy resources, as is common practice in ISO-NE, PJM, MISO and NYISO. There are significant challenges associated with enabling cross-border demand resources. The standards for qualification, verification, availability, dispatch, and controllability for non-traditional resources can vary to a large degree between jurisdictions. The reliability value contributed by these resources can be significantly affected by when and how the resources are dispatched in the energy market and within emergency conditions, factors that are largely infeasible to influence in another neighboring jurisdiction.

An alternative approach for achieving fully resource-neutral trade is to qualify generic capacity that is certified by the host system, regardless of the underlying resource type. The net commitment to ship capacity across the border in emergency conditions would then be fulfilled on a system-to-system basis (rather than resource-to-system). This approach has the potential to increase efficiency, reduce complexity of qualifying external resources, and enable more fungible capacity trade. Though a generic UCAP approach has not previously been implemented in other markets, MISO has previously considered this option as a means of enhancing capacity trade at the PJM seam.

Importantly, the EMA must agree with neighboring regions to allow external resources to promise firm transmission rights and non-recallability. Efficiently enabling imports will benefit from an integrated approach in the capacity auction, energy market, and transmission rights framework within the EMA and with neighboring regions.

Recommendations and Next Steps

Imports Approach

- Qualified Capacity for imports shall be based on self-declaration with additional non-recallability and deliverability requirements. Importers that fail to meet their self-declared levels will be penalized
- Allow for both qualification of specific external resources and generic capacity

Next Steps

- EMA to work with external systems to solidify plans to offer non-recallability

I. Qualification Timeline

It is necessary that all new and existing resources are properly qualified before the start of the auction period to ensure those that clear will be able to fulfill their capacity supply obligation. The qualification process for new and refurbished resources needs to establish that the resource fully intends on becoming operational and is able to do so by the delivery year if it does clear in the capacity auction. For existing resources, the qualification process is less involved and is largely to provide updates to capacity ratings and process requests for mothballing or retirement.

For both new and existing resources, the qualification timeline should have clearly defined deadlines so that resources can understand exactly what is required at each step throughout the process. The qualification timeline should be published far enough before the auction period to allow resources sufficient time to collect and prepare the necessary materials and allow the qualifying body ample time to review and respond to the provided information. We recommend that the qualification timeline should initially be conservatively long as EMA first qualifies resources and should be adjusted as necessary as EMA develops efficient practices and solidifies their understanding of the timing to review qualification materials. In other jurisdictions, the qualification timeline for new resources varies substantially due to requirements to align with other ongoing processes and timelines. For example, qualification begins eight to nine months before the auction in ISO-NE, six months before the auction in the UK, and only a few months in PJM.

While markets employ different qualification timelines for new resources based on their specific requirements and capabilities, there are some overarching similarities throughout their processes that we believe could shape the qualification process in Singapore:

- **Submission of Interest:** The earliest period throughout the qualification process, the submission of interest (SOI) window allows for resources to submit preliminary materials describing their proposed project. These materials can include necessary project information including the location, proposed MW and type of resource, as well as an initiated interconnection request in markets where the duration of interconnection studies may be a limiting time constraint. A resource may also have to prove that they have secured control of the site during this time.
- **New Qualification Package:** Following approval from EMA regarding the materials submitted during the SOI window, new resources must submit the full package of information necessary to be qualified. This package includes the necessary information to be given a QCAP rating, including intermittent data for intermittent resources. At this time, or after approval of the new qualification package, resources must post financial assurance requirements providing credit to solidify their promise to provide capacity and to support the market administrator's ability to collect replacement costs and penalties in the event the resources fail to materialize, as described in Section V. Proposed new resources must also submit a critical path schedule (CPS) outlining the timing of major construction and financial milestones they plan to achieve as they progress towards becoming operation. EMA should review the CPS to ensure that the proposed timeline is feasible and offer feedback if necessary.
- **Critical Path Schedule Monitoring:** Upon clearing the capacity auction, resources should progress along their approved CPS, providing updates and evidence to EMA throughout. In the case that new resources fall behind on their schedule, EMA should communicate with the resource to understand if they are at risk of not becoming operational by the delivery year. In the event that the resource appears unable to materialize nor transfer their cleared MWs, they are subject to losing their capacity supply obligation and any financial assurance requirements.

In addition to the SOI window and new qualification package deadlines, EMA may want to include dispute periods as a way to formally address any complaints from resources that feel as if they were inaccurately rated or improperly rejected. This period will allow resources to collect any additional information that might have been missing from their initial materials that could change the outcome of their decision. Similar to the recommendations provided above, the duration and timing of these dispute periods should initially be long enough to allow EMA time to learn how to efficiently process such disputes.

Recommendations and Next Steps

Approach

- Qualification timeline should initially be long enough to allow for EMA to learn to efficiently process resources, then be adjusted as necessary
- Qualification process can include initial submission of interest window, followed by a new qualification package deadline with potential for formal dispute windows. After the auction, it will be necessary to monitor the progress of a resource along its critical path schedule

Next Steps

- EMA to internally determine feasible qualification timeline

V. Financial Assurance Requirements

Forward capacity markets require financial assurances (FAs) from new resources, both to solidify their promise to provide capacity and to support the market administrator's ability to collect replacement costs and penalties in the event the resources fail to materialize. When designing FA policies—especially the size and the punitive forfeiture terms—it is necessary to weigh the tradeoffs of ensuring that participants are properly incentivized to deliver on its capacity obligation versus creating barriers to entry.

Generally, these FAs are collected from suppliers during the qualification period before the auction, then updated based on the results of the auction. Once the resource demonstrates commercial operation, its FA should be returned. However, if the resource fails to reach commercial operation by the delivery year, it generally forfeits its FA, except in jurisdictions that allow suppliers to transfer their FAs if/when they buy out of their capacity obligations bilaterally or through rebalancing auctions.

Additional credit requirements may apply to cover exposures to performance and availability penalties during the delivery year, for new resources as well as existing ones.

This section draws upon the experience of PJM, ISO-NE, AESO, UK and Ireland to provide recommendations for EMA on how to structure the FA requirement for new resources and how the FA requirement may be updated throughout the qualification process.

A. Structure of Financial Assurance Requirement

The FA requirement should be structured to allow the market operator to procure replacement capacity in time for the delivery year should the new resource fail to materialize, as well as provide a financial incentive for new resources to reach commercial operation. The markets that we surveyed offer varying approaches to accomplish these goals.

PJM: The FA requirement in PJM is initially very high compared to other markets, although it does decrease before the delivery period. This high initial level provides a very strong incentive for resources to only offer into the capacity market if they truly intend to become operational.

Although this high FA requirement would likely provide enough capital to procure replacement capacity if necessary, it does not explicitly have structural characteristics that ensure that the market operator would be able to procure replacement capacity. For example, if the FA requirement was based off of the clearing price of the auction, rather than Net CONE, it would reflect the revealed cost to replace capacity.

The FA requirement in PJM is the product of the capacity (MW) submitted or committed and the auction credit rate. The value of the auction credit rate, expressed in $/MW-day, can differ before and after the auction in PJM.[51]

$$Before\ Auction\text{: Max}\{\mathbf{50\%\ Net\ CONE}, \text{USD } \$20/\text{MWday}\}$$

$$After\ Auction\text{: Max}\{20\%\ \text{Clearing Price}, \$20/\text{MWday}, \text{Min}(\mathbf{50\%\ Net\ CONE}, 1.5\ \text{Net CONE} - \text{Clearing Price})\}$$

The terms shown in **bold** express the most likely ones to be in effect, based on typical values for Net CONE and auction clearing prices. Before the auction, the auction credit rate is likely determined by the 50% Net CONE term since Net CONE is around USD$300/MW-day, making it much greater than the floor of USD$20/MW-day. Similarly, after the auction, the 50% Net CONE is likely to set the auction credit rate since the clearing price has been less than Net CONE in recent years in PJM. If the clearing price is very high, however, then the auction credit rate could increase with the 20% clearing price term. A high clearing price would warrant a higher financial assurance value since the short market conditions that cause high capacity prices would also indicate that it would be very expensive if a planned resource would not be able to fulfill their capacity supply obligation. The higher financial assurance payment offers an even stronger incentive for participants to fulfill their obligations in these short market conditions.

After the auction, a participant's financial assurance commitment can decrease if it remains on schedule as the delivery date approaches. Upon hitting certain milestones, participants may provide evidence and apply for reductions in FA requirements.[52] Thus, the FA declines over time as milestones are met, indicating a reduced risk of ultimate deficiency.

Figure 14: PJM Reduction in Financial Assurance Schedule

Credit Reductions (%) for Planned Generation Capacity Resources and Planned External Generation Capacity Resources	
Credit Milestones Certified by an Independent Engineer (where applicable)	**Incremental Credit Reduction from Initial Credit Requirements**
Effective date of fully executed ISA for Planned Generation Capacity Resources or agreement equivalent to ISA for Planned External Generation Capacity Resources	50%
Financial Close	15%
Full Notice to Proceed and Commencement of construction (e.g. footers or foundation poured)	5%
Main power generating equipment delivered	5%
Commencement of Interconnection Service	25%

51 PJM, Credit Overview and Supplement to the PJM Credit Policy, October 21, 2019, pp. 16-17.

52 PJM, Credit Overview and Supplement to the PJM Credit Policy, October 21, 2019, pp. 14-15.

As it passes certain milestones, the risk of it not coming to market by the delivery period decreases, allowing for a corresponding decrease in collateral, until the entire FA requirement is returned upon reaching commercial operation.

ISO-NE: The FA requirement for ISO-NE is generally smaller than that of PJM. In ISO-NE, the FA requirement is equal to Net CONE times the committed capacity times a multiplier that *increases* over time (unlike PJM, where the multiplier decreases over time).[53]

$$\mathit{Current\ Collateral}\text{: Net CONE} \times \text{Committed Capacity} \times \text{Increasing Multiplier}$$

The multiplier in the collateral calculation begins at one (*i.e.*, one month, with the clearing price expressed in \$/kW-month) and increases by one prior to each subsequent reconfiguration auction, ending with a value of three before the delivery period.[54] These changes in the multiplier mean that the FA is initially one twelfth of the annual revenue a resource can expect from the capacity market, then one sixth after the first reconfiguration auction, then one quarter in between the last reconfiguration auction and the delivery period. Thus, the credit requirement rises to a quarter of the auction price, which varies, but might be a quarter of Net CONE in long-term expectations, compared to half of Net CONE in most PJM auctions, so it is roughly half the size.

Employing this increasing multiplier is the opposite of the declining schedule approach in PJM. Rather than reflecting the declining risk of participants not fulfilling their capacity supply obligation, the increasing multiplier reflects the increasing cost of procuring replacement capacity if the resource failed to materialize. If a resource does not become commercial by the start of the delivery period, the multiplier for its FA will increase by one every six months until the capacity supply obligation is fulfilled or terminated.[55] If a resource never becomes commercial, it forfeits its collateral.[56]

ISO-NE's relatively small FA has raised concerns among stakeholders that it enables financial participation (rather than physical) when market participants expect reconfiguration auction

---

[53] ISO-NE recently switched to using Net CONE instead of the capacity clearing price to define its FA requirement believing that Net CONE offers a number of improvements over using the clearing price. First it provides a more stable level of collateral that is not exposed to the volatility of the market clearing prices. Second, it offers a consistent collateral before and after the auction, whereas when based on the capacity price participants had to produce pre-auction collateral based upon the much higher starting price, then post-auction collateral based upon the lower eventual clearing price. Not only does this simplify the process for participants, it allows them to not have to secure collateral based on the auction price cap (the starting price in ISO-NE's descending clock auction). See ISO-NE, ISO New England Inc., Docket No. ER20-_____-000; Changes to ISO New England Financial Assurance Policy: Net CONE, November 15, 2019, PDF 4-5.

[54] ISO-NE, Exhibit IA – ISO New England Financial Assurance Policy, September 17, 2019, PDF 56.

[55] ISO-NE, Exhibit IA – ISO New England Financial Assurance Policy, September 17, 2019, PDF 56-57.

[56] Specifically, the ISO draws down all of the existing collateral and issues an invoice for any remaining shortfall, if necessary. See ISO-NE, Exhibit IA – ISO New England Financial Assurance Policy, September 17, 2019, PDF 57-58.

prices to fall, or for other reasons. The situation was most acute in the first seven delivery periods, when excess capacity combined with a price floor on the forward auction nearly guaranteed that prices would fall in the reconfiguration auctions. Speculators could offer phony capacity and collect the difference, mitigated only by not returning the FA (even when buying out of one's CSO). Now that the forward price floor has been eliminated, this threat is lessened, although likely-phony capacity did enter a recent auction, probably because it needed to clear in order to gain a state siting permit that might have value in the future.[57] This depressed the auction clearing price and angered many physical suppliers, who called for a higher FA, or other measures, to solidify the commitment of physical resources.

To address this concern, ISO-NE was recently was approved by FERC to add a component to its FA calculation equal to the profit to be made from a participant selling its position.[58] If a resource sheds its capacity obligation, it would not only forfeit its original collateral, but also this adder which is equal to the profit that it would make. This can raise the cost of offering non-physical supply into the market.

AESO: AESO did propose general FA rules or "security requirement," policies (although these policies were never accepted or implemented since the capacity market was discontinued). The proposal is more similar to PJM, with a schedule allowing the FA requirement to decrease as the delivery date approached.

AESO proposed a FA requirement to be a function of Gross CONE, a capital recovery factor for a new resource and a 5% factor representing bond performance.[59]

$$FA\ Requirement = \left(\mathrm{CONE} * \frac{1}{\mathrm{CRF_{New}}}\right) * 5\%, \text{ where}$$

$$Capital\ Recovery\ Factor_{New} = \frac{\mathrm{i}(1+\mathrm{i})^{\mathrm{n}}}{(1+\mathrm{i})^{\mathrm{n}} - 1}$$

i = discount rate for Gross CONE determination

n = 20 year plant life

The capital recovery factor is supposed to capture how much of the project investment is recovered annually by delineating the number of years over which a project investment can be recovered.[60]

---

57 In Connecticut, a new combined cycle cleared the capacity market, a key hurdle in the process for securing a permit with the local siting council. Some speculate that its bid was artificially low to ensure that it would clear the auction to get the permit, understanding that it might be able to later sell its position and end with a potential profit or at least low net cost. See Pilon, Matt, Proposed $700M Killingly power plant clears key hurdle, Hartford Business, February 7, 2019.

58 ISO-NE, ISO New England Inc. and New England Power Pool, Docket No. ER20-____-000; Changes to ISO New England Financial, PDF 9.

59 AESO, Proposal: 2. Supply Participation, p. 5.

60 AESO, Rational: 2. Supply Participation, p. 8.

Similar to PJM, AESO proposed a declining FA requirement. In its proposal, participants would have their FA requirement decrease linearly as the resource passes milestones by certain auction thresholds.[61] Similar to PJM, this declining requirement reflects the decreased risk that the resource is not going to provide capacity on time.

United Kingdom: The United Kingdom Capacity Market Rules do not have FA requirements as defined in the markets above, but do feature fixed penalties if resources do not materialize. The UK is allowed to terminate capacity contracts for failure to meet a number of milestones and guidelines. Under the market rules, the UK assesses a "termination fee" at one of five rates based on which rule the capacity provider violated. Figure 15 shows the termination charge schedule which features five tiers of termination fees before and after the 2016 Market Rules Amendment came into force. After the 2016 Capacity Market Amendments, the termination fees were increased and a higher final termination fee was added.

Figure 15: UK Capacity Contract Termination Charges

Timeline	Tier	Termination Fee (£/MW)
Before the 2016 Amendment	Termination Fee 1	5,000
	Termination Fee 2	25,000
After the 2016 Amendment	Termination Fee 3	10,000
	Termination Fee 4	15,000
	Termination Fee 5	35,000

The market rules provide around thirty provisions for disqualification. Based on the provision violated, capacity providers are assessed the corresponding termination fee. The violations that qualify a capacity provider for termination under the lower tiers (Termination Fee 1 and Termination Fee 3) are mostly administrative. Capacity providers would be eligible for Termination Fees 1 and 3 if they failed to provide evidence of an accepted grid connection offer, failed to comply with metering standards, or make a transfer sale without complying with market conditions.[62] Capacity providers become eligible for the tier two termination fees (Termination Fee 2 and Termination Fee 4) if they fail to achieve financial or developmental milestones. Capacity providers only qualify for the highest tier Termination Fee 5 if they fail to achieve the minimum completion requirement, or alter, do not receive, or cease to have a grid connection agreement.

Unlike the ISO-NE or PJM, the UK penalties are not intended to be time-varying, but are designed to penalize specific violations more severely than others. These "fixed-cost" penalties are much simpler than the FA requirements that we have observed in other markets since they do not change based on market conditions and do not require resources to provide credit in advance. However, there are considerable drawbacks. Although the penalties do incentivize performance, they do not ensure that the market operator will be able to procure replacement

[61] AESO, Rational: 2. Supply Participation, p. 9.

[62] These examples are a subset of the conditions that would make a provider eligible for termination under the EMR market rules.

capacity based on the actual current market conditions, as is possible in markets that determine the FA requirements based on the results of auctions or estimates of Net CONE. Additionally, because participants do not provide advance credit, the market operator risks not collecting the full penalty upon determining deficiency from the participant.

SEM (Ireland): Similar to the UK, the Irish Capacity Market, operated by SEM, levies fixed termination fees for participants that fail to achieve commercial operation. The SEM charges a termination payment for projects that fail to meet developmental milestones that *increases* over time, similar to the increasing FA requirements seen in ISO-NE. The termination fee starts at a low level and reaches its highest level before the last routine event through which capacity could be procured. The termination payments are calculated yearly and intended to capture the cost to consumers of undelivered capacity, the level of delayed liquidated damages available from a typical EPC contract, and the level of penalties for undelivered capacity to which an existing unit would be exposed.

The 2022/23 SEM Capacity Auction set these termination payments to begin at €10,000/MW (about 12% of Net CONE)[63]. This charge escalates to its highest level at €40,000/MW (46% of Net CONE) by start of the Capacity Year.

Figure 16: SEM Capacity Contract Termination Charges

Date / Event	Termination Charge Rate (€/MW)
More than 13 months prior to the beginning of the Capacity Year	10,000
From 13 months to the beginning of the Capacity Year	30,000
From the beginning of the Capacity Year	40,000

Source: SEM, "Capacity Remuneration Mechanism, Capacity Auction Parameters," 2019:6.

The decision by the SEM to use a levelized and increasing termination charge is intended to encourage early termination by projects (in recognition of the fact that later termination costs customers more) and to penalize larger projects proportional to their risk to customers.

Recommendation for Singapore: To provide certainty to suppliers of new resources, we recommend that the FA requirement in Singapore be defined based on the Net CONE value (similar to the approaches in PJM and ISO-NE, and related to the Gross CONE approach in AESO). In contrast to the approach in PJM, it will not depend on the outcomes of the auction. The FA requirement rate should be aligned with the maximum penalty for a completely non-performing resource, or 0.3 times the Net CONE.[64] This penalty provides a financial incentive for the new resources to reach commercial operation and should prove sufficient, or even more than sufficient, to procure replacement capacity should the resource fail to do so. The formula below describes this proposed FA requirement rate:

---

[63] The lowest SEM Net Cone was €86.0/kW de-rated for Capacity Year 2022/23. SEM, "Best New Entrant Net Cost of New Entrant Consultation Paper," 2018:24.

[64] See Section X for more discussion of potential penalties.

$$FA\ Requirement\ Rate = 0.3 \times \text{Net CONE}$$

Some capacity markets, such as PJM and AESO, feature FA requirements that decrease as the delivery period approaches reflecting declining risk of the resource not materializing, while others, such as ISO-NE and Ireland, increase over time reflecting the increased cost of procuring replacement capacity. We do not believe that the decreasing risk of the resource not achieving commercial operation needs to be reflected in the FA requirements, as suppliers may already benefit from declining risk as construction milestones are met through more favourable rates for FA granted by the bank or other credit lending body responsible for posting suppliers' FA requirements.

B. Return of Financial Assurance Requirement

While the overall structure drives much of the FA requirement, other considerations, such as how it is returned, are important to the effectiveness of the FA requirement. In all markets, the FA requirement for new resources is returned when the resource achieves commercial operation. Some markets, such as PJM, also allow for the return of the FA requirement if the resource is able to shed its capacity supply obligation through bilateral transactions or in rebalancing auctions.[65] Other markets, such as ISO-NE, do not allow for such transfers and will recover the full FA requirement unless the cleared resource comes to market itself.[66]

We recommend allowing for the return of FA requirements for a resource that sheds or transfers its CSO as doing so provides certain efficiencies. If it is more expensive for the resource to fulfill its CSO than it would be for a resource who does not yet have a CSO, then it is efficient to allow these participants to trade and not build in an additional cost by forcing the first resource to lose its FA. However, this approach works best if the demand forecast is not biased. If the demand forecast is predictably high (*i.e.*, based on a high forecast, then brought down closer to the delivery year), then financial players will potentially be able to make a profit without needing to provide capacity by securing positions in the initial auction and buying out in later rebalancing auctions as the demand forecast and price decrease.

C. Continued Financial Assurances Once Operational

The FA requirements described until this point have only been relevant to new resources, however, all new and operational resources may potentially face penalties in excess of their revenues which EMA must ensure that they are able to collect. If a resource underperforms its capacity supply obligation, it could be liable to pay up to 30% of its total expected revenue in net penalties, as described in Section X. To ensure that this penalty can be collected if necessary,

---

65 PJM, Credit Overview and Supplement to the PJM Credit Policy, October 21, 2019, p. 18.

66 ISO-NE, Exhibit IA – ISO New England Financial Assurance Policy, September 17, 2019, PDF 57-58.

we recommend that resources be required to post credit requirements aligned with their maximum potential exposure throughout their obligation period, and to allow leniency, an additional factor of 25% is applied. The formula below describes this proposed credit requirement rate:

$$Credit\ Requirement\ Rate = 0.25 \times 0.3 \times \text{Clearing Price Received}$$

To incentivise resource performance, EMA could waive credit requirements for select resources that have demonstrated sufficient reliability and/or have a low risk of failing to meet their obligations.

Recommendations

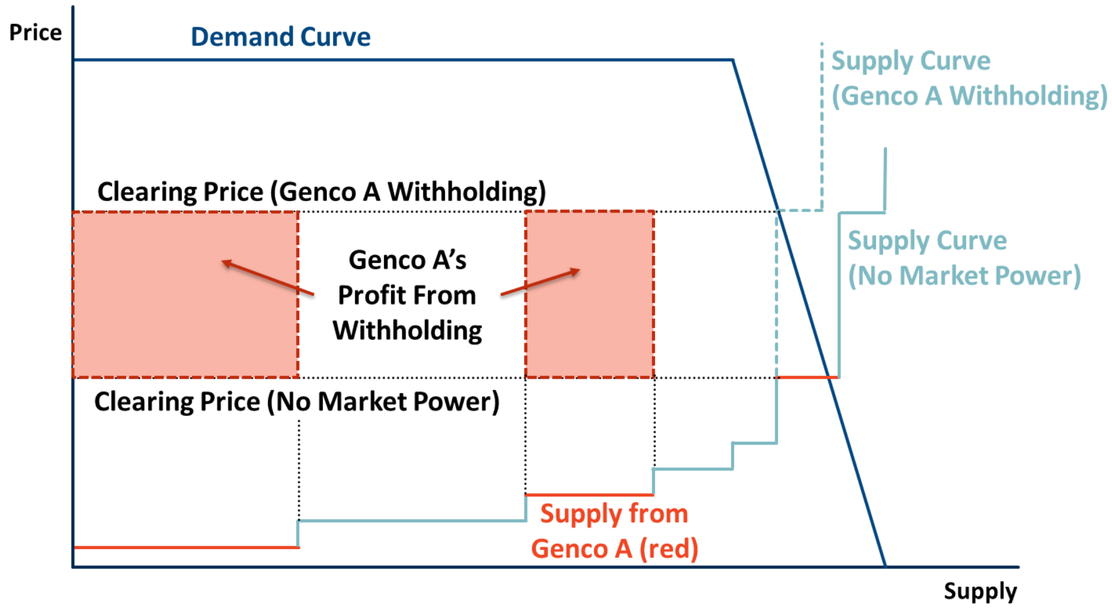
Financial Assurance Requirements

- FA requirement should support the market operator to procure replacement capacity in time for the delivery year should the new resource fail to materialize. More importantly, this provides a financial incentive for all new types of resources to reach commercial operation
- FA rate should be 0.3x Net CONE
- Allow for FA requirements to be returned as soon as when the new resource has met all construction and delivery milestones, or if participant successfully sheds or transfers its CSO
- Require new and existing resources to post credit requirements covering potential performance penalties

VI.Capacity Market Power Monitoring and Mitigation

All capacity markets are considered structurally uncompetitive at least some of the time because residual supply tends to be small (with little excess beyond peak load plus reserve margin) relative to the size of some suppliers. Singapore is no different, with several suppliers being large enough to be pivotal or become pivotal as excess capacity diminishes. Although supplier shares cannot increase above 25% due to the existing capacity market share cap, 25% is a large share. Such large participants could have the incentive and ability to increase the price by inefficiently withholding capacity, as illustrated in Figure 17.

Figure 17: Illustration of Incentive and Ability to Raise Market Prices in the FCM



Withholding could occur physically, by not offering or prematurely retiring a resource. Withholding could also occur economically, by offering a resource at a price above the cost of providing capacity, with the intention of not clearing the auction. The FCM can protect against both physical and economic withholding, through capacity market power monitoring and mitigation by the market monitor.

To address physical withholding, resources (including aggregated resources) that are 10MW and above (in ICAP terms) should be required to offer their full QCAP.[67] Resources that wish not to participate in the auction and rather retire or mothball for the delivery year need to receive a must-offer requirement exemption prior to the auction in order to do so and would not be allowed to participate in the energy market in the delivery year. The decision to grant a must-offer requirement exemption would be contingent on the resource demonstrating that they would be unable to meet a capacity supply obligation.

For clarity:

- New resources will not be subject to the must-offer requirement, given that it is difficult to force suppliers to enter and it is better to discipline behavior with competition from other potential new resources[68];
- Existing capacity suppliers will be subject to the must-offer requirement;
- Embedded (*i.e.*, behind-the-meter) generation will not be subject to the must-offer requirement, as their capacity is primarily intended to serve on-site load; and

---

67 Existing resources that have an ICAP of less than 10MW will not be subject to this requirement.

68 New resources that are in the process of construction and have yet to reach commercial operation should decide on an auction to participate based on their confidence of being available in the respective delivery year.

- Demand response, even that associated with assets that have been participating and appear "existing" would not be subject to the must-offer requirement because their ongoing provision of capacity requires ongoing renewal of contracts securing end-users' willingness to be curtailed or utilize behind-the-meter backup generation (and end-users do not have incentive to participate in schemes to withhold capacity and increase prices given that they are net consumers of energy and capacity).

Table 15: Summary of Must-Offer Requirements for Resources with ICAP of 10 MW or above

<table>
<tr><th>Resource Type</th><th>Operational</th><th>Mothballing/ Mothballed</th><th>Retirement</th><th>New</th></tr>
<tr><td>Thermal Resources</td><td>Yes</td><td>Yes</td><td rowspan="5">Requirement to seek EMA's approval is based on license conditions</td><td>Depends</td></tr>
<tr><td>Imports</td><td>Yes</td><td>Yes</td><td>Depends</td></tr>
<tr><td>Solar</td><td>Yes</td><td>Yes</td><td>Depends</td></tr>
<tr><td>Storage</td><td>Yes</td><td>Yes</td><td>Depends</td></tr>
<tr><td>Embedded Generation</td><td>No</td><td>No</td><td>No</td></tr>
<tr><td>Demand Response</td><td>No</td><td>No</td><td>No</td><td>No</td></tr>
</table>

To prevent economic withholding, the market monitor will cap ("mitigate") the offer prices for existing non-DR capacity[69] of market participants that are deemed likely to have both the *incentive* and *ability* to exercise market power. To determine which capacity suppliers should have their offers capped, the market monitor will employ a market power screen to test each supplier. There are many different types of market power screens used in other jurisdictions, such as the three-pivotal supplier test, the single-pivotal supplier test, the conduct and impact test, or an incentive test.[70] We evaluated the options based on three criteria: (1) the ability to

[69] New resources and DR would be exempt for the same reasons as noted above.

[70] PJM uses a Market Structure test based on a three-pivotal supplier test. If the required capacity cannot be met with the output of the two largest suppliers, plus the output of the supplier being tested, then all three are jointly pivotal. These three suppliers would be able to manipulate prices by jointly withholding output. See PJM Tariff Attachment DD: Reliability Pricing Model, Section 6.3. We have previously raised the concern that this test is too stringent as it would mitigate even very small suppliers; see Reitzes *et al.*, "Review of PJM's Market Power Mitigation Practices in Comparison to Other Organized Electricity Markets," September 2007.

NYISO uses similar monitoring and mitigation measures, based on a single pivotal supplier test. Of particular interest are several measures that are specifically applied only to market-internal import-constrained capacity zones, particularly New York City which has a high concentration of both supply and demand. These factors tend to increase the risk and impact of market power exercise relative to larger and more structurally-competitive capacity zones. See NYISO Tariff Attachment H: Market Power Mitigation Measures, Section 23.2.1.

MISO's monitoring and mitigation measures are quite different from those in PJM and NYISO, partly because of the region's traditionally-regulated market structure in which the vast majority of supply and demand are represented by vertically-integrated, cost-of-service-regulated utilities that avoid over-mitigation, which is ultimately inefficient and discourages participation in the market; (2) the ability to avoid under-mitigation (either excusing a resource from mitigation, or mitigating to a lower price that is still higher than the competitive level); and (3) the complexity and controversy of design and implementation.

EMA selected the one-pivotal supplier test (1PST), and we support this decision. This test is conducted by first calculating the residual supply index (RSI) as follows:

$$RSI = \frac{Total\ Supply - Supply\ from\ Supplier\ being\ Tested}{Total\ Demand}$$

If the RSI (in QCAP terms) is less than or equal to 1 for a supplier, then there is not enough supply in the market without the supply from that supplier, so the supplier fails the test.

The 1PST test is unlikely to over-mitigate compared to some other tests, such as the 3PST (where the RSI is calculated using the joint supply for the supplier being tested and the two largest suppliers). It could under-mitigate small, non-pivotal suppliers who nevertheless have some ability to affect the price, but with limited consequences given their size and competition from new resources. And it is much simpler to implement than the various types of incentive tests or conduct and impact tests.

Only suppliers whose portfolios are large enough to fail the 1PST would be subject to mitigation, but that does not mean that all their offers would be mitigated. Offers would be mitigated only if above pre-defined thresholds. Defining a "no-review" threshold can reduce the administrative burden of mitigation and can limit the risk of over-mitigating. In principle, such thresholds should reflect the expected competitive offer level, that is, the net avoidable going-forward costs[71] of providing capacity, either generically or by resource type. However, during an initial transition period, EMA intends to define a single threshold for all resources using the vesting contract parameters because the data is available and reliable, and because this approach is suitable for the applicable fleet in Singapore that is comprised largely of CCGTs. Further, EMA is opting to use fixed annual running cost without deducting net E&AS revenues, to avoid potential over-mitigation that could result from estimation error.[72]

have balanced positions and so have little incentive to manipulate capacity auction prices. In that context, and to minimize its interference in the auction, MISO imposes mitigation measures only if it determines that exercise of market power could increase auction clearing prices by an impact threshold of at least 10% of the Cost of New Entry (CONE). In that case, must-offer or offer-cap mitigation measures may be applied. See MISO Tariff Module D, Section 64.2.1(e).

71 Competitive offers at net avoidable going-forward costs may consider: (a) capital and fixed costs incurred in the immediate year, minus (b) energy and ancillary services margins expected in the immediate year, minus (c) future net margins expected for the remainder of the asset life. If the capacity obligation exposes suppliers to non-performance risk, the rational offer price would not drop below the expected penalty size.

72 Based on EMA's "Review of the Long Run Marginal Cost Parameters for Setting the Vesting Contract Price for 2019 and 2020", the fixed annual running cost is about S$55/kW-year for a 432.2 MW F-Class CCGT on an installed capacity (ICAP) basis. As the vesting parameters are reviewed biennially (or when deemed necessary), with a mid-term review of the capital cost parameters, EMA intends

Resources that fail the market power screen and exceed the no-review threshold would be subject to possible offer mitigation. The identities of capacity suppliers whose offer prices have been mitigated would remain confidential. To enforce that their offers are competitive and reasonably reflect net avoidable going-forward costs, the market monitor (*i.e.*, EMA) would provide the resources with two options:

1. Submit a pre-determined default offer cap equal to the no-review threshold; or
2. Request a resource-specific offer cap consistent with the net avoidable going-forward cost of that resource.

In individual resource-specific offer reviews, suppliers would have to present their costs and their projected net revenue estimates, and their plans in case they do not clear. Potential net revenue should be considered (as they are in other jurisdictions) because it reduces the net avoidable going-forward cost, thus affecting the resource-specific competitive offer level. Their alternative plans are important because they determine the net avoidable going-forward cost if the resource does not take on a CSO (versus taking on a CSO). Net avoidable costs are generally highest if the resource would retire (and promises it would do so if it does not clear), since in that case all ongoing fixed costs would be avoided and net E&AS revenues foregone. For plants that would mothball, net avoidable going-forward costs may be lower, since in that case some of the ongoing fixed costs might not be avoided, such as property taxes and certain insurance. For a resource that plans to stay online and sell energy and ancillary services, the avoidable cost is the going-forward incremental fixed costs to meet its CSO. EMA views that the cost of expected penalties arising from the inability a resource to meet its CSO should not be included in its offer, as doing so would not support reliability objectives.

Auction results should also be reviewed ex-post to detect anti-competitive behavior. This review should comprise a thorough analysis of supplier bidding behavior and market outcomes. If market power is found to still be a concern after the ex-ante mitigation, the EMA may wish to adjust the market power screens and/or pursue administrative action against offending parties.

On a preliminary basis, we recommend that EMA use the following timeline for market power mitigation in advance of the four-year forward auctions beginning in 2022 (a compressed timeline may be needed for earlier auctions):

- Six months in advance of auction: EMA to publish the no-review threshold.
- Four months in advance of auction: last day for suppliers to request must-offer exemption; last day to request resource-specific offer cap (and provide all information to support such a request).
- Three months in advance of auction: EMA to notify suppliers of determination on proposed must-offer exemptions and resource-specific offer caps.
- Two and a half months in advance of auction: last day for suppliers to notify EMA of disagreement with determination and submit additional information justifying must-offer obligation exemption request and resource-specific offer cap request.

---

to use the latest updated values at the time of the relevant auction, to determine the no-review threshold.

- One month in advance of auction: EMA to notify suppliers of final determination on must offer exemption and resource specific offer caps.

EMA has introduced since 2016 a generation capacity market share cap of 25% to prevent structural increase in market concentration in the generation market. In line with this regulatory policy and also to provide a forward signal in the FCM to enforce this, EMA intends to impose a 25% market share cap on each supplier based on the total CSOs procured in the FCM for each delivery year (*i.e.*, a 25% FCM Cap). In line with the principle of no forced divestment should the total CSOs procured in the FCM exceed the 25% FCM Cap due to the conduct of any rebalancing auction held in the forward period, such as with an updated demand curve reflecting lower demand in the delivery year and/or due to the actions of other suppliers, any supplier which had been cleared in the base auction will not be required to unwind its CSO. However, any bilateral transactions or offers into the rebalancing auction (*e.g.*, incremental sell offers) that result in the supplier further exceeding the 25% FCM Cap in the delivery year will not be allowed.

Recommendations

Market Power Mitigation

- All existing resources (including aggregated resources) that are 10MW and above are required to offer their full QCAP, unless exempted
- EMA to use a single pivotal supplier test to identify suppliers with market power
- EMA to develop offer price cap for suppliers based on vesting contract parameters, in the first instance
- Suppliers who wish to have a higher resource-specific offer cap must prove their avoidable net going-forward costs of providing capacity are above the no-review threshold and petition EMA for a higher resource-specific cap
- A 25% FCM Cap on each supplier, based on the total CSOs procured for each delivery year, to mitigate structural increase in market concentration

VII. Forward Capacity Auction

We recommend that the FCM have a four-year forward auction. The auction should have a uniform clearing price paid to all resources, conducted as a single-round auction, with sealed bids. This auction structure can maximize reliability at the lowest possible societal cost and has a strong performance record in other capacity market contexts.

A. Auction Design

We recommend a single-round, sealed-bid, uniform clearing price auction. This is the auction structure that is most likely to achieve efficiency and deliver the targeted reliability at the lowest cost. In this section we review the alternatives and provide justifications for the recommended approach to each design element.

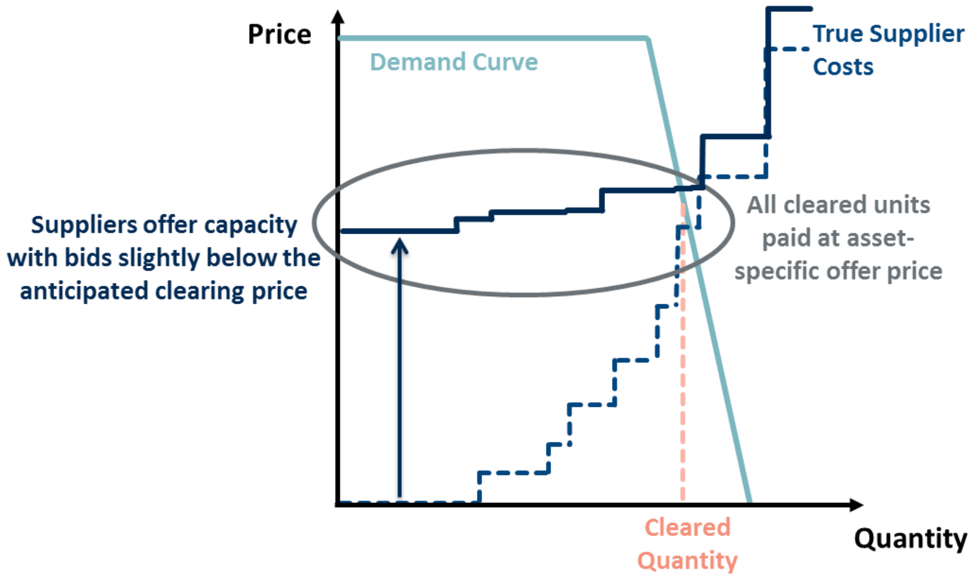
Pay all Cleared Resources a Uniform Price

In a uniform price auction construct, all cleared suppliers receive the same price per MW of capacity. This incents suppliers to offer at cost—the absolute minimum price they are willing to accept to provide capacity, that is, at their net avoidable going-forward costs—except in cases of market power. As a result, the clearing price in the auction should reflect the marginal cost of capacity, which is most likely to ensure least-cost procurement of capacity and to provide efficient long-term signals for investment and consumption.

There are two primary alternatives to this approach: "pay-as-bid," in which cleared resources are paid according to the price they offered capacity into the market, and differentiation between new and existing resources, which allows prices to separate between new and existing resources if the marginal new resource has a higher offer price than the marginal existing resource. We explore each of these alternatives below.

Pay-as-bid. In an alternative pay-as-bid approach, all cleared suppliers are paid their bid price. This approach is not used in any centralized wholesale energy or capacity market, but it is used in other contexts, most commonly in decentralized commodity markets. We see no substantial benefits over uniform price auctions in any centralized market context.

Figure 18: Supplier Offer Behavior in Pay-As-Bid Auction



Under a pay-as-bid construct, suppliers have the incentive to bid at the price of the most expensive offer they expect to be accepted (that is, at the expected clearing price), as illustrated in Figure 18. Thus, the auction does not elicit information about suppliers' marginal costs (as in a uniform price auction), but rather about suppliers' expectations of the clearing price.

Theoretically, these two approaches could produce the same prices if suppliers accurately estimate the marginal cost of capacity. However, in practice, the pay-as-bid construct will likely not achieve the efficient price signals achieved by uniform pricing. The pay-as-bid construct invites sellers to offer above their costs, and uncertainty over the clearing price is likely to result in inefficient results. If low-cost resources offer too high (due to incorrect beliefs about the auction clearing price) and fail to clear, they may exit or fail to enter while higher-cost resources with lower offers enter instead. This issue of distorted merit order is illustrated in Figure 19. In addition, suppliers with a larger generation portfolio are likely to have more information about the potential clearing price, and would be at an advantage compared to smaller suppliers who risk guessing the clearing price wrong and inefficiently fail to clear their resource. Finally, monitoring for the abuse of market power is inherently difficult in this construct, where offers reflect participant beliefs rather than private costs.

Figure 19: Example of Distorted Merit Order in Pay-as-Bid Auctions

Differentiated payments for new and existing capacity. In a second alternative to uniform pricing, the clearing price could be differentiated between new and existing resources to reflect the marginal offer under each resource type. It is often believed that differentiating payments will save consumers money overall, based on the implicit assumption that existing suppliers have lower net avoidable going-forward costs than new resources, and therefore do not "need" the same high capacity payments necessary to attract new resources.

However, these arguments are flawed. There are three key reasons that uniform pricing for new and existing resources is best:

- With any market-oriented approach, all-in prices are expected to reflect long-run marginal costs in the long run. Thus, differentiation would achieve no net gain for consumers in the long run.
- Paying all resources the same price for the same product, regardless of how it is produced, is consistent with current principles and practice in the energy market and best practices in other commodity markets, including all other capacity markets.
- Uniform clearing will minimize societal costs, which minimizes consumer costs in the long run. Allowing the market to express demand for the capacity product, and treating all suppliers of that product the same, will allow the market to find the least cost resource mix. If the price for existing resources is not allowed to rise to the same prices facing new resources, the market will not accommodate efficient upgrades to existing resources or provide incentives for efficient retirement decisions. From a societal perspective, price differentiation is inefficient because it reduces competition, can

induce inefficient high-cost investments in new resources, and often leads to inefficient retirement of lower-cost existing resources.

We conducted quantitative analysis comparing auction outcomes and consumer costs between a uniform-price FCM and one that differentiates payments between new and existing resources.[73] This quantitative analysis suggests that consumer costs are similar under either method over a 20-year horizon. Initial consumer savings (a wealth transfer from existing generators to consumers) are likely offset by long-term higher prices. This reflects the fact that new entrants must offer at very high prices to recover capital costs rapidly in initial years (the period when they are still considered "new," assumed to be five years in our analysis).

USE A SINGLE-ROUND, SEALED-BID AUCTION

Single-Round Auction. We recommend that the FCM auction be conducted in a single round. Multi-round auctions are used to allow resources to amend offers during the auction clearing process. However, such auctions can be more complex to administer and increase the risk of participants engaging in gaming behavior. The advantages and disadvantages of each approach are outlined below in Table 16.

73 Our model assumes that under the alternative approach where new and existing resources clear separately, new entrants must recover most of their capital costs during the years they are considered to be "new" resources in the auction. Thereafter, lower "existing" capacity payments are assumed to cover only their ongoing fixed O&M costs, and their only capital recovery will derive from net E&AS revenues. This affects how new suppliers offer into the auction because it compresses the amount of time during which they can recover capital costs. In our base case, we assume that resources offer at prices consistent with recovering capital costs over just five years instead of over the 25-year economic life of the plant.

Supply offers in our analysis are assumed to reflect avoidable going-forward fixed costs, net of expected E&AS revenues. We assume that new resources are efficient enough to earn net revenues in the energy market. Solar resources' net energy revenues are calculated based on an assumed capacity factor over the year, while the net revenue for new CCGTs is an input assumption to the model. Existing thermal resources are assumed to be on the margin, making a negligible profit in the energy market. Demand increases annually according to EMA projections and at a consistent rate after the projections end, and the demand curve has been simplified to be a vertical curve.

Table 16: Comparison of Single-Round and Multi-Round Auction Formats

Format	Advantages	Disadvantages
Single Round	• Simplicity helps prevent design flaws. • Less exposure to the exercise of gaming, market power, and collusion. • Lower implementation, transaction, and overhead costs (both for the market administrator and market participants). • Easier to implement in a zonal framework (N/A in Singapore) or any other structure that would add complexity to the types of constraints reflected in the auction (*e.g.*, flexibility requirements, seasonal requirements, dynamic effective load carrying capability ratings that depend on penetration levels, or clean energy requirements). So far these are N/A in Singapore but that could change in the future.	• *Theoretically:* No price discovery during the auction. • *In practice:* We do not view price discovery as particularly valuable or important in capacity auctions. Price discovery is useful in other contexts (such as leases on an oil reservoir) when the true value of the contract is the same across all bidders, but the bidders all have different information on the "common value" (*e.g.*, the amount of oil in the ground). Thus, a more accurate price is achieved by allowing bidders to pool information via multiple rounds (this avoids under-bidding to avoid "the winner's curse"). In the capacity market, this logic does not apply since there is no "common value" aspect of the capacity market contract. • *In theory:* The other theoretical benefit of multi-round auctions is related to products that have a "contingent value" such as spectrum auctions, where the value of a radio broadcast right in one area is higher if also receiving the same spectrum in a neighboring area. • *In practice:* This benefit also does not apply in capacity auction contexts since there is only one product being cleared.
Multi-Round ("Descending Clock")	• Price discovery in early rounds may help marginal suppliers decide what to bid. This is likely to be minor or not applicable in the context of a capacity market, where price discovery is not likely to be efficiency improving (as most bidders' costs are private costs and little information about private value can be gleaned from other bidders' behavior, as discussed above). • Similar result as single-round auction in higher price ranges where bids must be pre-approved by the market monitor, at least for existing resources. • Better clearing with multi-product auctions (though not as efficient as if those multi-product auctions can be simultaneously co-optimized).	• More complicated if using zonal capacity product (N/A in Singapore). • More exposure to the exercise of market power, gaming, and collusion (as price discovery may allow participants to infer when they are pivotal and change their offers accordingly).

We believe it is best to use the standard single round approach for two reasons:

- This approach helps limit exercise of market power. The descending clock approach could allow participants to infer when they are pivotal and exercise market power when they would not have risked doing so in the single round format.
- The primary theoretical advantage of descending clock auctions, which is that bidders can learn information about other participant's cost of providing the good being auctioned, and that "crowdsourcing" this information may lead to better bids, is largely

> not applicable in capacity markets. In capacity markets, most of the costs each participant would incur to provide the capacity product are privately known and not highly correlated across participants.

Across other capacity markets, only ISO-NE and UK use a multi-round, descending clock approach. We are of the view that the benefits of a multi-round design in these capacity auctions are overstated. The theoretical benefits of multi-round auctions are much more applicable in other types of products, as described above. Further, the way that ISO-NE implements the multi-round auction makes the outcome similar to a single-round auction in any case (as any medium- or high-price offers for existing resources are capped in advance, market participants are not able to change their offer prices as information is gleaned over rounds of the auction).

Sealed-Bid Auction. Market participants in the FCM will submit sealed bids. In a sealed bid auction, the offers of the participants are not revealed to the other participants during the auction. The additional information made available to participants via open bidding may introduce greater opportunities for gaming. This shortcoming has led all existing capacity auctions to use the sealed bid approach.

B. Offer Format and Auction Clearing

We recommend that EMA enable resources to represent their offers using up to ten offer segments, each defined by a price and quantity. All segments must be divisible (can partially clear), except the first segment, which the supplier can specify as divisible or non-divisible ("lumpy"). Higher-priced segments will not clear unless lower-priced segments clear first. Lumpy offer segments can be guaranteed all-or-nothing clearing. Allowing multiple offer segments will allow suppliers to represent a range of potential underlying cost structures of their supply resources. Suppliers may be able to offer additional capacity at a higher marginal cost, for example via inlet chilling or refurbishment, at incremental cost, or adding higher-cost demand response to a portfolio. It is also consistent with best practices in other jurisdictions. Allowing both divisible and non-divisible offers allows efficient clearing of discrete units of capacity that may affect how suppliers make investment and/or operations and maintenance decisions in the forward and delivery periods. Offer segments can be no smaller than 0.1 MW of QCAP, aligned with practices in other jurisdictions.

The auction will be cleared by maximizing the objective function of social surplus (consumer plus producer surplus), subject to all constraints. EMA is proposing to implement specific rules for treatment of marginal, non-divisible offers. If the marginal offer is non-divisible, the auction will minimize consumer cost subject to procuring at least the minimum acceptable reliability level. The auction clearing price will be set at the higher of (1) the value of the demand curve at the cleared quantity or (2) the offer price of the marginal offer.

Key elements of this recommendation include (1) the use of multi-block offers; (2) enabling both divisible and non-divisible offers; (3) the auction clearing objective function; and (4) how auction clearing prices are determined. We discuss the rationale for the recommended design choices for each element below.

Multi- vs. Single-Block Offers. Key considerations for whether to allow resources to submit multiple price-quantity offer tranches are summarized in the table below.

Table 17: Advantages and Disadvantages of Enabling Multi-Block Offers

Advantages	Disadvantages
• Allows resources to more accurately reflect the incremental cost structure of capacity • Resources may have marginal costs increasing with quantity of capacity obligation (*e.g.*, due to environmental retrofit) • Reduces the lumpiness of the supply curves, and improves the efficiency of the market outcome	• May be easier to exercise market power, as it is somewhat riskier for a resource to economically withhold its entire QCAP than a portion of it in its attempt to raise the market clearing price

All other jurisdictions with a capacity market (except UK) allow offers with five to ten blocks as this maximizes economic efficiency. We view the benefit of having accurate representation of costs will outweigh any potential costs of incremental market power. Additionally, the incremental market power should be mitigated through the capacity market power mitigation measures described in Section VI.

Divisible vs. Non-Divisible Offers. Enabling both divisible and non-divisible offers affects whether a marginal price-quantity offer can be cleared partially. Providing flexibility enables more resources, particularly new entrants or refurbished units, to participate in the FCM, as generation investments are inherently lumpy. All other jurisdictions (excepting the UK) allow both divisible and non-divisible offers.[74] Key considerations for making the decision are summarized below.

Table 18: Advantages and Disadvantages of Enabling Non-Divisible Offers

Advantages	Disadvantages
• Avoids clearing existing resources at the level below their economic minimum, *i.e.*, the levels below which the capacity cannot be provided at reasonable cost • Prevent new resources that are in the development stage from having to re-size if the offer is partially cleared, which could disincentivize new resources from participating	• Somewhat more complex market clearing algorithm required

We recommend allowing the first offer segment of each supplier offer to be non-divisible, with all other segments divisible. This enables efficient market outcomes and allows suppliers more flexibility to accurately reflect the cost of providing capacity in their bid offers. Both existing and new resources would have the opportunity to designate their economic minimum quantity, where applicable, as non-divisible in the first offer tranche.

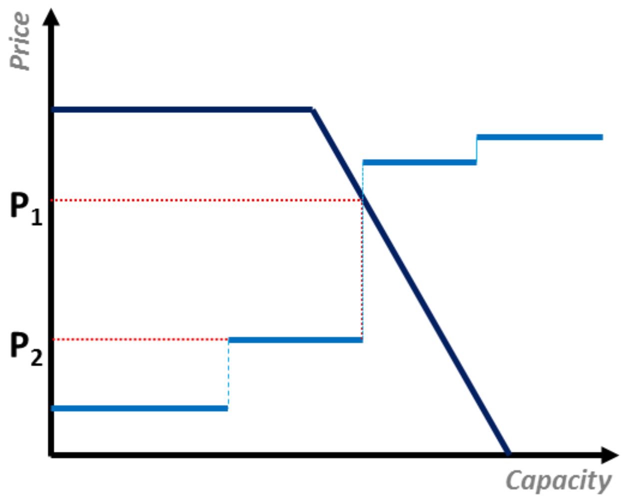
74 PJM allows suppliers to specify a minimum level below which their offer cannot be divided, and provides make-whole payments to any suppliers whose offers are cleared below this minimum level.

Auction Clearing Objective Function. When allowing non-divisible offers, it is possible that the marginal offer would be non-divisible. In this case, EMA proposes that the auction is cleared to choose the set of supply offers that minimize consumer cost, only if doing so procures at least the minimum acceptable reliability level. To do so, the clearing engine will first determine the clearing price and volume of: (i) clearing the non-divisible offer; (ii) not clearing the non-divisible offer; or (iii) skipping to the next higher priced divisible offer, thereafter (a) prioritize clearing the option that results in a procurement volume of at least the minimum acceptable reliability level, otherwise (b) compare the outcomes and clear the option that results in least procurement costs.

While this differs from other markets, such as PJM and ISO-NE, that maximize social surplus through minimizing deadweight loss, it is relatively aligned with the UK approach, that has a different objective function, which is to maximize social surplus when doing so does not cause a net loss in consumer surplus, *i.e.*, loss in consumer surplus is not greater than reduction in deadweight loss.

Determining the Auction Clearing Price. There could be instances where the entire cleared supply curve lies below the demand curve. We illustrate this below. In such instances, a choice has to be made as to whether the market clearing price should be set by the demand curve at P_1 ("incremental value" approach) or by the supply curve at P_2 ("incremental cost" approach).

Figure 20: Illustration of Case When All Cleared Supply Is Below Demand



All markets except for the UK and Ireland have adopted the incremental value approach. Though it may yield somewhat higher procurement costs in a particular auction, this approach more accurately reflects the value of incremental capacity, providing an enhanced price signal for incremental capacity to enter in subsequent auctions and potentially reducing long-run costs. It also mitigates year-to-year capacity price volatility. We recommend the incremental value approach.

Tie-Breaks. There might be rare instances in which there is a tie between two or more offers of the same price, that can clear the auction at the same outcome prescribed by the auction clearing objective function. In such circumstances, EMA reserves the right to determine which offer(s) will clear the auction in discussion with the relevant resources.[75] Some potential considerations can include clearing divisible offers proportionately, and if the offers are non-divisible, the prioritization of resources willing to accept a shorter commitment term where eligible (see Section VII.C below on multi-year commitments).

C. Commitment Term

The default commitment term for all resources will be a single year (after one shorter delivery period in Q4 2023). It may be beneficial, however, to allow some resources to “lock-in” the clearing price for multiple years after they are initially cleared in the auction. The potential advantage to this approach is that it reduces revenue uncertainty and may help reduce financing costs for capital-intensive new resources. This may lower FCM clearing prices needed to attract new resources and/or reduce the risk of failing to attract sufficient new resources when needed. In addition, it may facilitate greater competition by attracting investments that would not have otherwise participated in the capacity market.

However, even if improving revenue certainty can allow suppliers to offer lower prices, this does not necessarily translate to lower overall costs for consumers. Most importantly, providing price certainty does not eliminate risk; it merely shifts risk from suppliers to consumers. This risk manifests as the potential for paying above-market prices to locked-in capacity in subsequent years. Whether the efficiency of market clearing outcomes is improved depends on whether suppliers or buyers are better able to manage or absorb the risk. Other potential disadvantages are described in the table below.

[75] The choice of offer which will clear the auction will not affect the auction clearing price, as the offers would be of the same price.

Table 19: Advantages and Disadvantages of Multi-Year Price Lock-Ins

Advantages	Disadvantages
• Revenue certainty may be beneficial to lower financing costs, thereby lowering supply offers and market clearing prices. • Encourages greater participation in the FCM, which can help to improve resource adequacy and mitigate against under-supply conditions when significant new capacity is needed.	• Risk of locking-in expensive supply, increasing costs in subsequent years when capacity prices are lower. • Discriminates against existing resources and may distort the incentives for generation owners. With the option for a lock-in on new resources, generation owners may have less incentive to invest in maintaining existing resources and more incentive to build new resources, even if maintaining existing resources is the lower-cost option for providing capacity to the market. • Special provisions to incentivize new investment could be distortionary if they reduce investors' incentives to carefully assess future market conditions. In particular, lock-ins diminish the importance of future market conditions (supply and demand outlook, technology costs, *etc.*) and increase the importance of current market conditions for suppliers making investment decisions.

Several other jurisdictions with capacity markets allow new and refurbished resources to lock-in prices in this way, as described in Table 20.

Table 20: Price Lock-ins in Other Jurisdictions

Jurisdiction	Eligibility	Term	Rationale
ISO-NE	New & refurbished	Up to 7 years	Smaller markets with lower investor confidence and/or shorter history of capacity market; deemed necessary to provide revenue stability to attract sufficient investment
Great Britain	New & refurbished	Up to 15 years for new; up to 3 for refurbished	
Ireland	New & refurbished	Up to 10 years	
IESO	New & refurbished*	Unclear	
PJM	No Lock-In*	-	Significant investor confidence. Not deemed necessary to attract new investments
NYISO	No Lock-In	-	Most investments supported by long-term contracting (by traditional utilities); not necessary to attract new investments
MISO	No Lock-In	-	

Notes: Ontario suspended design and implementation of the FCM so details of the multi-year lock-in were not finalized. PJM does have a very narrowly defined price lock-in for the purpose of supporting prices in small, transmission constrained zones where a large new resource could suppress capacity prices for a sustained period. It is almost never triggered and sufficiently different in scope and design that we do not consider it here.

Notably, in other jurisdictions that have a multi-year lock-in (or multi-year commitment) provision, existing resources are never included. These resources have already entered the market and made large, irreversible investments. A price lock-in is not needed to retain them, as a single-year term is sufficient to recover net avoidable going-forward fixed costs if they clear the capacity market. Furthermore, allowing existing resources to lock in prices for multiple years may artificially delay economic retirements and hinder investment in new resources. Finally, it could substantially reduce the liquidity in subsequent auctions, increasing market power concerns.

EMA proposes to offer a 10-year multi-year commitment (MYC) for new/repowered CCGTs with an economic lifespan of at least 25 years that meet the proposed heat rate standard for power generation (see Annex C of the Third Consultation Paper for EMA's consultation on heat rate for power generation). Qualifying new units would be able to lock in for ten years the price from the first auction that they clear. For the delivery years after the end of the MYC, the CCGT will be considered an existing unit and will not be eligible for MYC in the auctions for those delivery years.

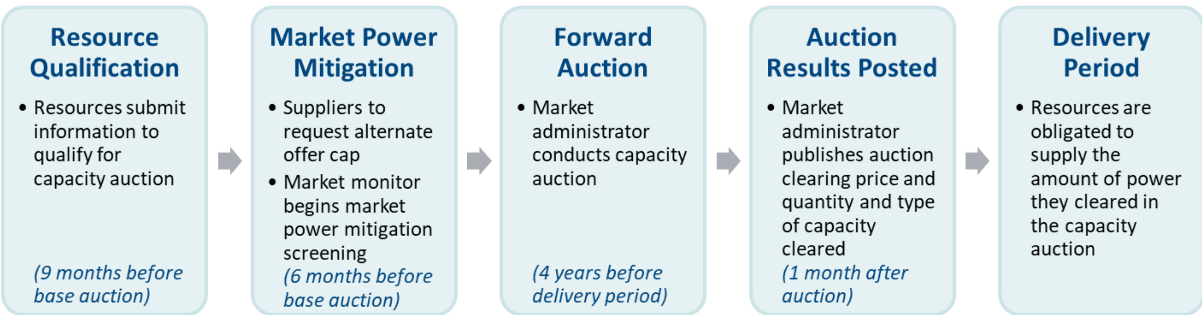
This proposal reflects EMA's belief that gas-fired CCGTs will continue to be the main generation technology to meet Singapore's baseload electricity demand efficiently. They are also proven frequency responsive resources that provide online reserves which are essential for maintaining power system security. With growing electricity demand and significant CCGT capacity reaching end of life, EMA believes there is a need to facilitate the adoption of more efficient CCGTs to meet baseload demand as well as provide reliable online reserves so that the overall energy efficiency of the power generation sector can also be improved.

EMA proposes this provision for CCGTs entering during the first decade of the FCM. EMA will review and calibrate the duration and eligibility threshold for MYC over time, as new technologies evolve and mature.

D. Auction Timelines

The FCM market rules will establish the timing of events leading up to the auction, immediately after the auction, and for the period between the auction and the delivery year. These procedures are illustrated in Figure 21.

Figure 21: Preliminary Timeline for Base Forward Auction



Pre-auction: During the pre-auction period, the EMA will need time to qualify resources, and to implement market power mitigation procedures (see Section V on market power mitigation). Other jurisdictions begin these processes five to nine months before the auction. This will require an assessment of how much time is required to conduct these functions and establish the timelines appropriately.

Post-auction: After each auction, the results should be published in a timely manner, usually within a few weeks. The published auction results should, at a minimum, include information on the clearing price, how much capacity cleared, and what types of resources cleared. The lag time allows the EMA to assess auction performance to check for ex-post signs of market power abuse or other inefficiencies, then to publish the results of that assessment.[76] On longer time scales, the overall performance of the FCM should also be assessed after every few years (perhaps more frequently at the beginning of FCM implementation).

Forward period: The forward period refers to the time between the auction and the start of the delivery year. We recommend a four-year forward period, consistent with EMA's historical records of the lead-time needed to incorporate planned new CCGTs in Singapore. A shorter forward period may limit the types of resources that could make their development contingent on clearing the FCM. A longer forward period would increase the uncertainties that exist between the base auction and delivery of capacity. This would increase risks for suppliers by introducing more uncertainty regarding the status of their resource so far in the future; it would increase risks for consumers of over-procuring capacity based on larger errors in such long-term forecasts.

E. Recommendations for Singapore

Recommendations

Capacity Auction Design

- Adopt a uniform-price, single-round, sealed-bid auction design
- Allow offers to have multiple segments
- Allow the first segment of each offer to be non-divisible; all other segments must be divisible
- Clear the auction to maximize social surplus, with specific rules for treatment of non-divisible offers
- Determine prices as the higher of the value of the demand curve at the cleared quantity and the offer price of the marginal offer

Commitment Term

- Adopt a default one-year commitment term (delivery year)
- Allow a 10-year multi-year commitment for new/repowered CCGTs with an economic lifespan of at least 25 years and which meet the proposed heat rate standard for power generation

VIII. Rebalancing Auctions

Rebalancing auctions are designed to address changes in market conditions (both demand-side and supply-side) between when the base auction occurs and when the delivery year starts. They enable:

---

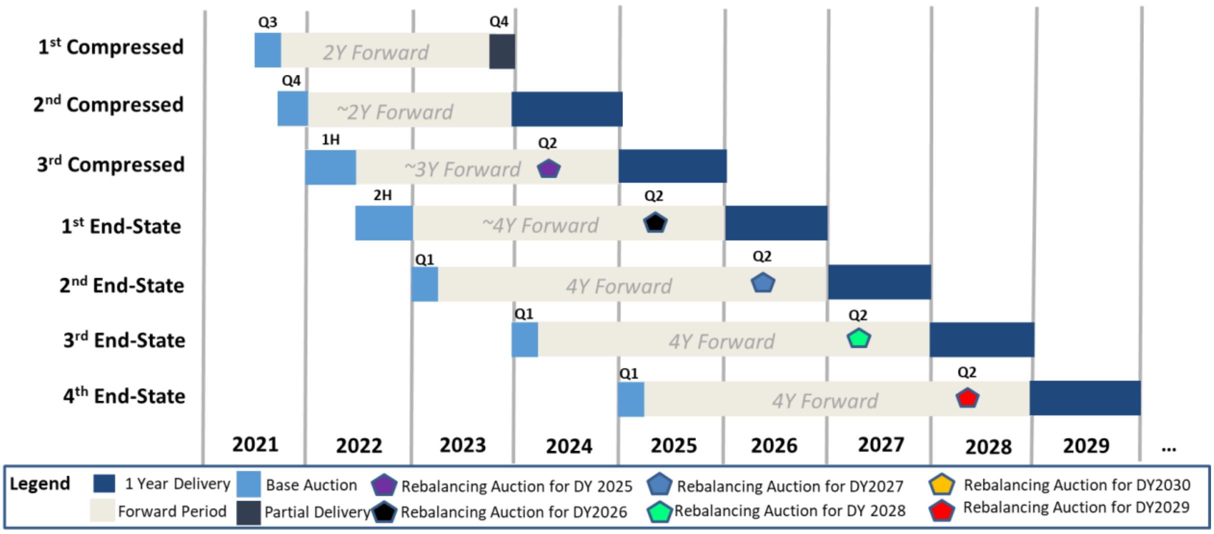
76 The market monitor (*i.e.*, EMA) may reject the auction results prior to publication, if it appears that there are grounds to suspect irregularities in relation to the auction and/or auction results. Thereafter, the market monitor should notify the auction participants of the annulment, and may also instruct the market administrator (*i.e.*, EMC) to re-run the relevant auction. This approach is consistent with that employed in other markets, including Ireland.

- The procurement of additional supply if the load forecast increases prior to the delivery year;
- The release of excess capacity and recovery of some costs for consumers if the load forecast decreases prior to the delivery year;
- New resources to enter closer to the delivery year; and
- Resources with a forward commitment to buy out if their QCAP has decreased, or they are otherwise unable to deliver, or if delivery is no longer cost-effective.

EMA intends to conduct one scheduled rebalancing auction during the forward period, held about 8 to 9 months before the delivery year, to facilitate opportunities for new resources to enter the FCM as well as for changes in forecasted demand. The timing of the rebalancing auctions and base auctions for different delivery years can be staggered to prevent having to conduct multiple auctions in a short time period. EMA could decide to hold more ad hoc rebalancing auctions during the forward period if the need arises.

Figure 22 shows the proposed timeline for the rebalancing auctions, with the pentagon shapes representing the rebalancing auctions for the respective delivery year (*e.g.*, the rebalancing auction for the delivery year 2025 will be held in Q2 2024).

Figure 22: Preliminary Timeline for Base Forward Auction



Prior to each rebalancing auction, EMA will conduct qualification for any new resources eligible for the relevant delivery year. In addition, the QCAP of all previously-qualified resources (including previously-cleared resources) should be updated to reflect updated information on development plans for new resources and on planned and unplanned outage rates for existing resources. Previously-cleared resources whose updated QCAP falls below their CSO will be required to buy out of the excess obligation, as discussed below.

The rebalancing auctions can be conducted in a manner similar to the base auctions:

Auction Format and Demand Curve: The same auction format should apply as in the base auctions. In addition, while auction parameters (primarily the load forecast) may be updated, the demand curve shape in the rebalancing auction will otherwise be unchanged from the forward auction. This is to prevent any systematic discrepancies in auction format or curve shape and position, which have the potential to create incentives for suppliers to arbitrage between these auctions to capture the value differential between these curves.

Market Power Mitigation: Existing qualified capacity should have an obligation to offer into the rebalancing auction and suppliers will remain capped at a 25% market share of the total CSO procured, just as in the base auction, but we recommend no other market power mitigation measures be enforced. This is because the potential benefits of mitigating any market power abuse are likely to be outweighed by the costs, for two reasons:

- First, no suppliers are anticipated to have significant market power in the rebalancing auction, as they have sold off significant long positions in the base auction, with relatively little incremental capacity exposed to the rebalancing auction price; and
- Second, even if some suppliers could exercise market power in the rebalancing auction, the higher prices would only apply to the small quantities of demand transacted in the rebalancing auction, so this would have a negligible effect on total consumer costs.

Auction Clearing Mechanism: The auction should be cleared on a gross basis, with all supply and demand in the market represented in the auction. The demand curve shape is the same as in the base auction, providing for a clear way of seeing the effect of updated auction parameters on the administrative demand curve. Settlements will be on a net basis—that is, only the incremental or decremental cleared quantities would be settled at the rebalancing auction price. This allows market participants that do not wish to change their position to be unaffected by the rebalancing price.

Supply Resources Offers and Bids: During the rebalancing auctions, market participants may want (or need) to change their capacity commitments because of changes in resources' QCAP. To allow for these types of adjustments, market participants should be allowed to submit the following types of offers and bids:

- <u>Incremental Sell Offers</u>: Enable suppliers to offer uncommitted capacity that has been made available, not cleared in the previous auction(s), or new capacity that requires a shorter lead time (for example, demand response and imports);
- <u>Repricing and Buy-Out Bids</u>: Repricing Bids enable suppliers to buy out of their committed positions for financial reasons. Buy-Out Bids can be used for capacity that is physically unable to deliver on a prior capacity commitment, and they must submit a non-price QCAP adjustment offer. Buy-Out Bids can be used for (1) resources wishing to guarantee a reduction of their capacity commitments regardless of capacity clearing price; (2) new resources that have not achieved development milestones and that are required to buy out of their capacity obligations; and (3) existing or new resources whose qualified QCAP at the time of the rebalancing auction is below the previously committed QCAP quantity; or
- <u>Do Nothing</u>: Enable capacity suppliers who do not wish to change their supply to participate as price takers on the supply side during the rebalancing auctions. This will not incur any settlement as a result of the auction; the capacity price of their previously committed positions will remain unchanged and this typically forms the bulk of market participants' submissions in the rebalancing auction.

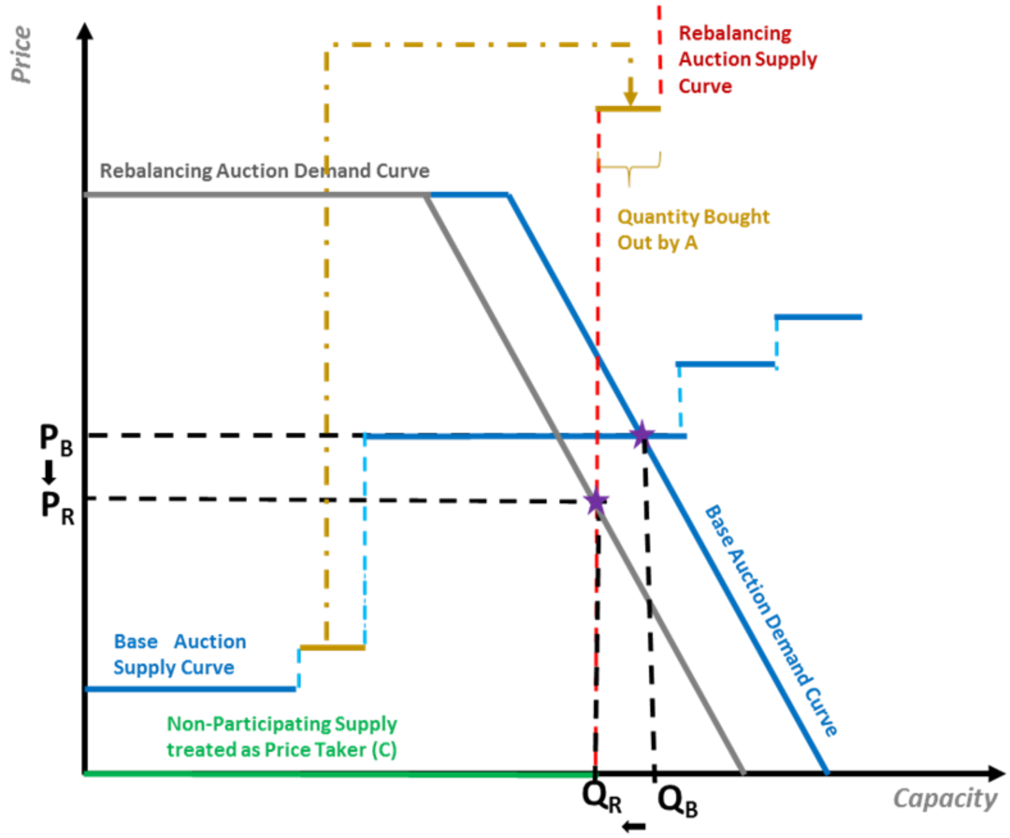
Settlements would occur as follows:

- Any uncommitted capacity, whether new or existing, clearing the rebalancing auction would receive the rebalancing auction clearing price multiplied by the quantity cleared. The quantity cleared becomes its CSO for the delivery year.
- If any previously committed capacity successfully buys out with a re-pricing or buy-out bid, its CSO is reduced accordingly. The net financial settlement works as follows: the resource is paid only the retained CSO quantity at the base auction price, and it is paid for the shed quantity at the base auction price minus the rebalancing auction price (or pays if negative). This is equivalent to typical settlement of forward markets, whereby the seller is still paid the forward price on the original forward position but must buy out of the quantity it will not provide by paying the subsequent market clearing price on that quantity (in this case, the rebalancing auction price).
- There is no net financial settlement for resources whose CSOs don't change, including those with base auction commitments who participated as price takers in the rebalancing auctions. As their CSO remains they will be paid based on the base auction clearing price and quantity.

There are many combinations of scenarios that can arise in the rebalancing auctions. Below we present a few examples.

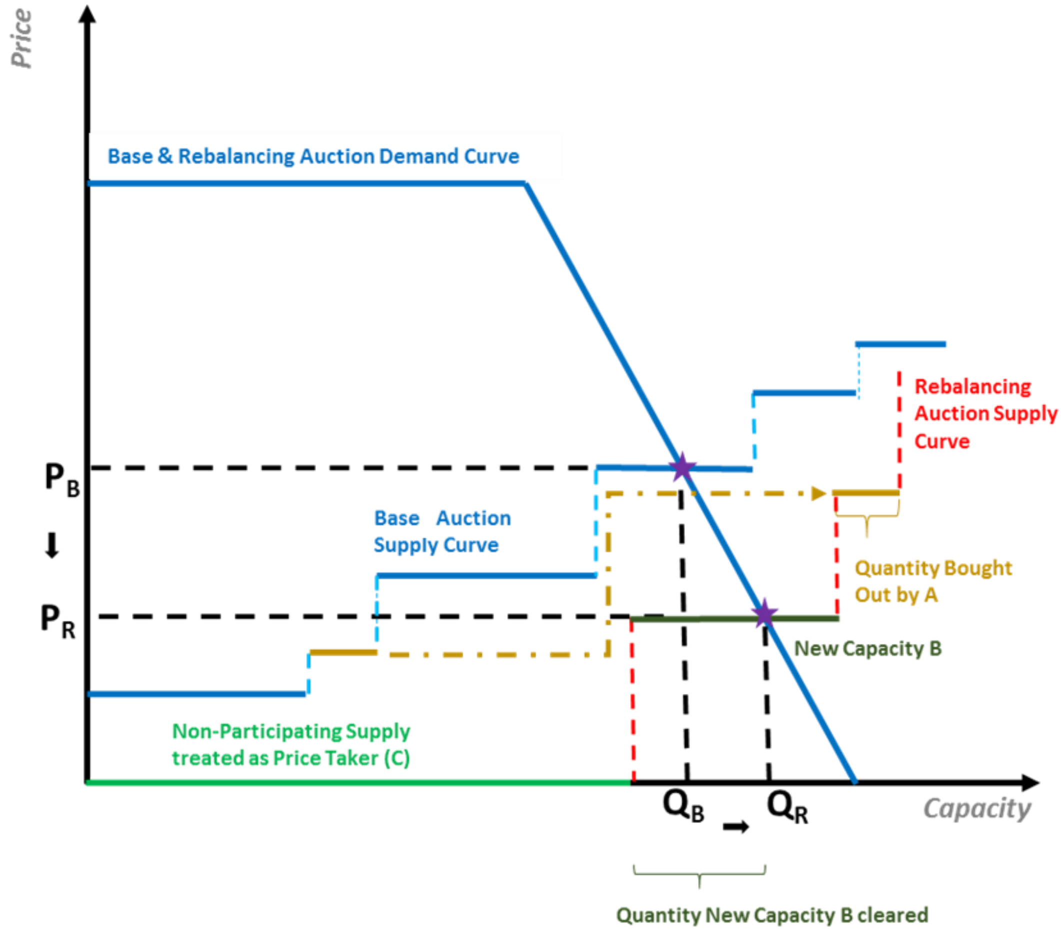
- <u>Scenario 1: Load forecast decreases, a committed resource submits a QCAP reduction bid and its previously procured capacity is released.</u> In the example illustrated below, there is a decrease in load forecast in the rebalancing auction relative to the base auction. This shifts the rebalancing auction demand curve to the left (in grey). A supplier with an existing CSO (A) submits a QCAP reduction bid above the price cap, indicating its willingness to buy-out of its CSO at any price. The rebalancing auction clears at P_R, below the base auction price P_B. Supplier (A) pays the rebalancing auction clearing price P_R. As $P_R < P_B$, this results in a net revenue to A of ($P_B - P_R$) multiplied by the quantity bought out ($Q_B - Q_R$). Remaining non-participating supply (C) remains unaffected.

Figure 23: Rebalancing Auction Clearing Example 1



- <u>Scenario 2. No change in load forecast, a new resource offers at a low price and displaces (*i.e.*, takes on the CSO in place of) a forward-committed resource that submits a higher re-pricing bid; the low new offer results in more capacity even if that resource only partially clears.</u> In this example, there is no change to the load forecast, so the demand curve is the same as in the base auction. A supplier with an existing CSO (A) submits a re-pricing bid for its full capacity commitment from the base auction. A new supplier with capacity (B) offers incremental capacity and partially clears, setting the rebalancing auction clearing price P_R below the base auction clearing price P_B. When the auction clears, supplier (A) pays the rebalancing auction price P_R to buy out of its obligation. As it receives the base auction price P_B, and $P_B > P_R$, this results in a net revenue to A of ($P_B - P_R$) multiplied by the quantity bought out. The market operator procures an additional capacity ($Q_R - Q_B$) at the rebalancing auction clearing price P_R. The new capacity receives the rebalancing auction clearing price P_R multiplied by its cleared capacity. Remaining non-participating supply (C) remains unaffected.

Figure 24: Rebalancing Auction Clearing Example 2



Recommendations
Rebalancing Auction Design and Format • For each delivery year, conduct one scheduled rebalancing auction after the base auction, with EMA having the discretion for more if the need arises • Maintain auction format and parameters from base auction, including Net CONE, required reserve margin, auction clearing, offer format, *etc.* Peak load forecast can be adjusted to reflect most recent forecasts • Clear auction with all supply and demand represented (on a "gross basis") and settle only incremental cleared quantities at the rebalancing auction clearing price (*i.e.*, on a "net basis")

IX. Bilateral Transactions

Gencos, other market participants, and independent retailers may wish to transact CSOs for a variety of reasons outside of the centralized capacity auctions, both during the forward period, and potentially during the delivery year. These transactions may be used to hedge capacity costs or to assign a CSO to another qualified supplier in cases of unexpected inability to provide capacity during the delivery year. The market design should enable these bilateral transactions.

We recommend that EMC, as the wholesale market operator and administrator, develop a mechanism to track the bilateral exchange of CSOs from the auction. For simplicity, we do not advise facilitating and tracking financial bilateral transactions. Market participants can transact financially on their own outside the FCM. What the market administrator must do is track when one entity assumes the physical obligations of another.

The requirements and limitations for bilateral transactions are few:

- CSOs can only be exchanged between resources that have been qualified and rated as per Section IV. This requires the provision of sufficient financial assurance for new resources, as any financial assurance requirements transfer with the obligation;
- CSOs can be exchanged at any time in the (a) forward period, except for a short period surrounding each base or rebalancing auction to ensure that all committed capacity can be accounted for in the auctions, and (b) delivery year to enable participants to efficiently manage their capacity obligations;
- CSOs can be exchanged for a full delivery year or portions of a delivery year. This provides an appropriate balance between providing flexibility to capacity suppliers to efficiently transact their CSOs while limiting administrative burden;
- CSOs can be exchanged in increments of 0.1 MW, but in no case will a supplier be allowed to hold a CSO of less than 1.0 MW for any single supply resource; and
- All bilateral transactions are subject to EMA's final approval.

Outside of the capacity auction, large loads or retailers may also want to enter into bilateral hedges with capacity suppliers to lock in capacity prices even before the base auction. These transactions are purely financial; there is no exchange of physical CSO that must be tracked.

X. Supply Obligations and Performance Penalties

Suppliers receiving a CSO will be subject to obligations that require them to participate in the real-time market for energy and/or ancillary services. In addition, they may have other obligations such as participating in performance testing and data collection activities necessary to calculate qualified capacity levels.

Performance assessments measure compliance with obligations, and associated penalties determine how compliance will be incentivized. The combined incentives from energy market prices and potential capacity market penalties encourage efficient operations and investment.

A. Obligations on Capacity Resources

Obligations on the capacity product procured during the capacity auction have to be clearly defined. As a starting point, best practices in other jurisdictions with a day-ahead energy market is to enforce a must-offer requirement to ensure the full available capacity of committed resources. These obligations accomplish two objectives:

- Ensure availability during shortage conditions; and
- Mitigate the potential for exercise of market power.

As several stakeholders have emphasized, in the SWEM, the absence of a day-ahead market precludes strict must-offer requirements, as resources are self-committed. As a result, they may not be available in the short timeframes required by the real-time market, for example if they have longer start-up times (some conventional generation) or notice periods (some demand response). Instead, we recommend accomplishing the two objectives listed above through alternative mechanisms. To ensure availability during shortage conditions, we recommend obligating all resources that are available to offer in the real-time market. Resources that are available but not scheduled in the real-time market are still liable to be activated for emergency, out-of-market commitment by the system operator, during high risk and emergency operating states. As this is an out-of-market commitment for emergency, EMA does not intend to offer any compensation for such activations and resources are expected to calibrate their capacity offer prices accordingly. Additionally, EMA may conduct periodic ex-post reviews of suppliers' operational behavior to ensure their pattern of self-commitment is consistent with competitive behavior.

B. Penalties for Resource Unavailability

Incentives for resource performance during shortage conditions can come both from the energy market and from the capacity market. We recommend real-time energy market prices reflect marginal system costs, including scarcity and the costs of administrative actions during shortage conditions, up to the energy market price cap.

However, one of the key lessons of the first 15 years of experience with capacity markets internationally is the need for additional incentives to solidify performance against capacity supply obligations. Several jurisdictions have accordingly established mechanisms to measure and incentivize suppliers' availability during pre-defined hours of the year and/or shortage conditions. The purpose is to reward sellers for maintaining availability for dispatch to the system operator, especially during times when the resource is most likely to be needed for supply adequacy. Key principles for the design of these incentives are that they should be strong, consistent with the implied value of reliability, and with all revenues at stake (and potential net penalties in cases of severe under-performance); and focused on when capacity is needed.

In alignment with practices in other jurisdictions, we have designed a penalty framework that we recommend for Singapore. Key elements include the penalty structure, the measured average capacity delivered, treatment of scarcity periods, penalties rates and revenues at stake, and settlements.

Penalty Structure. Penalties will be assessed for any under-performance on an annual basis. Underperformance due to unplanned or planned outages will be treated similarly. Under-performance will be evaluated by comparing "measured average delivered capacity" (defined below) against CSO. For example, if ICAP of thermal resource is 100 MW and CSO is 90 MW (with 10% planned + unplanned outages) but realized outages are 20%, a penalty would be assessed for 10 MW of underperformance.

Measured Average Capacity Delivered reflects the average amount of capacity actually provided, accounting for realized planned and unplanned outages.

Treatment of Scarcity Periods. The proposed definition of a scarcity period is a dispatch period when (1) there is any system energy, reserve and/or regulation shortfall, and (2) when PSO has to intervene by activating resources to avert any physical shortfall or restore the power system to normalcy.[77] Penalties should be higher during these periods of system stress, shortage, or near-shortage when the operator requires full delivery from all capacity resources. This penalizes relatively inflexible resources that struggle to deliver during critical periods in favor of flexible and well-performing resources. Thus, when calculating measured average capacity delivered, all scarcity periods should be weighed more heavily by a factor of 100. This rate reflects internal Brattle and EMA analysis of the additional scarcity period multiplier needed to allow total incentives during scarcity periods to reflect the value of lost load (VOLL). Given that the EMA can and will stringently assess planned maintenance requests, we do not find it necessary to apply the scarcity period multiplier to planned outages. While doing so would send the efficient signals to reduce planned maintenance outages, it would also unnecessarily and unfairly penalize supply resources that were granted permission for the applicable outages.

Revenues at Stake. For each MW of underperformance, the penalty rate will be the maximum of (1) 130% of the clearing price received for the CSO; (2) 100 percent of the latest rebalancing auction price; or (3) 20 percent of the auction price cap.[78] Element (1) ensures that total size of potential penalties can *exceed* capacity revenues in event of non-delivery to incentivize compliance with the capacity supply obligation. Element (2) ensures that deficient suppliers have an incentive to procure replacement capacity. Element (3) ensures penalties are not too low in circumstances when capacity market prices are low.

The determination of the exact penalty rate is qualitative and somewhat subjective. Other jurisdictions have not used quantitative approaches to answer this question. Important factors that must be evaluated qualitatively include the following:

- The penalty rate must be high enough that suppliers are not incentivized to over-state their QCAP or petition for a higher QCAP than they believe they will be able to provide.
- The penalty rate must be high enough that suppliers are not indifferent to outages. They should have sufficient incentive to maximize their availability in the delivery year, and potentially make costly operational/maintenance decisions to be able to meet their stated QCAP.

---

77 The EMC currently issues market advisories to spot market participants. These advisory notices pertain to the incidence and extent of any projected energy, reserve and regulation shortfalls for dispatch periods included in the relevant short-term and pre-dispatch schedules. This would provide a sufficient signal for resources to be available to provide energy and/or ancillary services. See EMA, "Singapore Electricity Market Rules Chapter 6 Market Operation," January, 2020. Available at: https://www.emcsg.com/f283,7862/Chapter_6_Market_Operation_1Jan20.pdf

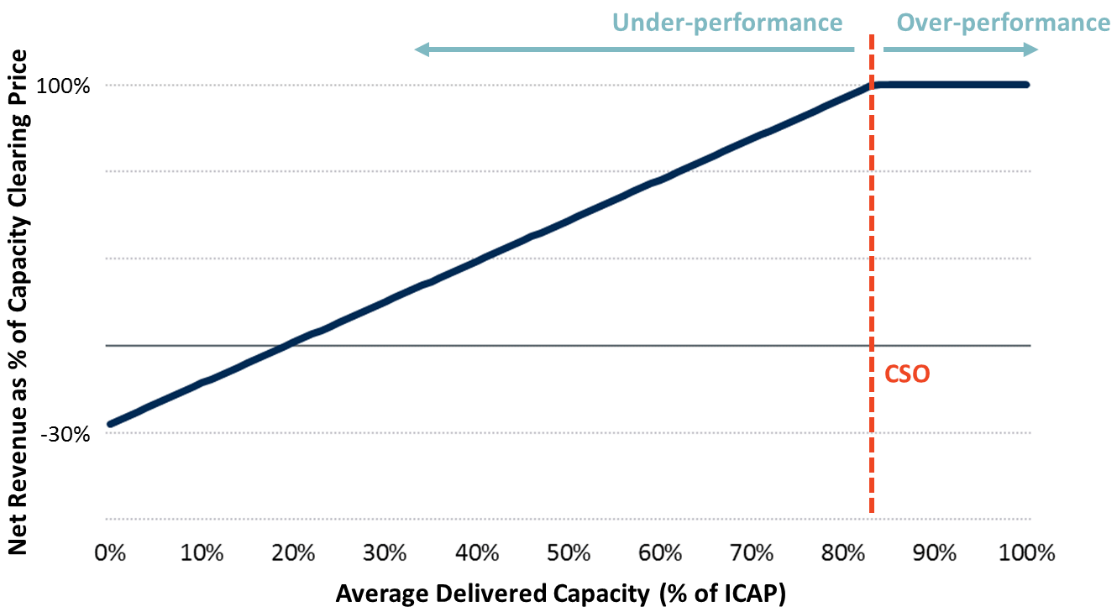
78 These penalties are defined relative to the expected capacity revenue. For example, a supplier being assessed a (gross) penalty of 130% of the clearing price received would net a penalty of 30% after the capacity revenue is accounted for. In extreme cases, this penalty structure allows for total claw back of expected revenues plus a 30% additional punitive penalty rate.

- The penalty rate should be low enough to avoid introducing undue year-to-year revenue risk.
- The penalty rate cannot be too high because of the asymmetric nature of the penalty structure combined with natural uncertainty and year-to-year variation in availability due to random outages. That is, suppliers can lose money for under-performance, and this risk is not compensated by additional revenues for over-performance. Thus, higher penalty rates incentivize suppliers to maximize expected revenues by understating their QCAP, which will result in both higher costs per MW for procured resources and costly over-procurement system-wide.

Settlements. Penalties will be estimated on a rolling basis each month based on all available year-to-date information, with penalties assessed as a deduction against capacity payments. A suppliers' over-performance in certain months would decrease penalties for underperformance in other months.

Figure 25 below illustrates how penalties assessed as a function of average delivered capacity impact a suppliers' total revenues from the capacity market.

Figure 25: Illustration of How Total Capacity Revenues Reflect Average Delivered Capacity



Below we provide examples of how measured average capacity delivered and penalties would be calculated for various resource types in a way that reflects the standards on which they were qualified. Note that for non-dispatchable resources, non-performance during any scarcity periods occurring outside of the peak period used for qualification would not impact resources' average available capacity calculation. The penalty mechanism is not the appropriate mechanism to adjust revenues for limited availability windows, as this introduces unnecessary revenue risk. Rather, limited availability windows should be accounted for in resource qualification and the QCAP methodology.

- **Thermal.** Thermal resources are qualified to reflect their expected availability during scarcity periods. This is calculated to reflect the QCAP calculation used during qualification. Thus, measured average delivered capacity is calculated based on the actual delivery year POR and UOR data.

- **Solar.** Solar resources are qualified based on their average capacity factor during the defined on-peak periods (say 9:00am to 10:00pm). Correspondingly, measured average delivered capacity is calculated as the weighted average generation actually observed in these hours, with the weights chosen as described above to account for scarcity hours. Suppliers take the risk of a year being less sunny than expected (particularly if occurring during scarcity periods).
- **Demand Response.** Demand responses resources are qualified based on their claimed capability during the periods for which they are qualified for delivery. The average claimed capacity provided in the delivery period is calculated as the annual weighted average during expected hours of delivery:
  - 0 MW, when outside of hours of nominated availability; and
  - ICAP MW, when during hours of nominated availability.

  The measured average capacity delivered is calculated as the average claimed capacity multiplied by the "realized performance rate" during the delivery year, which is the average number of MW delivered during hours when dispatched as a fraction of QCAP MW. For DR that is rarely dispatched (*e.g.*, less than once in each quarter of the delivery year), the realized performance rate would include performance during surprise testing events.
- **Storage.** Storage resources are qualified at their maximum sustained discharge for the required duration of four hours, that is aligned with the average shortage duration in Singapore as determined by EMA to be four hours. The average claimed capacity is calculated similar to above for DR, and the measured average delivered capacity is calculated as the average claimed capacity multiplied by the realized performance rate as defined above.

Refer to the attached spreadsheet for illustrative examples on penalties for resource unavailability.

Illustrative Examples
on Penalties for Reso

Recommendations

Penalties for Resource Unavailability

- Penalize suppliers whose measured average delivered capacity is below their QCAP rating
- Penalties to be no lower than 1.3 times the price received for capacity on a $/MW-year basis, such that severe under-performance can yield substantial penalties
- When calculating measured average delivered capacity, weight actual realized scarcity hours higher by a factor of 100. Methodology varies by resource type to reflect resource characteristics

XI. Settlements and Cost Allocation

The costs of procuring capacity in the FCM should be allocated to consumers in a manner that sends fair and efficient price signals for them to reduce load and mitigate the need for capacity to maintain reliability.

Accordingly, we recommend for the capacity costs be allocated to consumers or retailers that serve end-users in proportion to actual MWh consumption during the different period-types, where higher capacity costs will be incurred for consumption during peak (and potentially also mid-peak) periods. We elaborate on and justify this recommendation below.

A. Principles and Best Practices

A key principle of cost allocation that underlies our recommendation is that the allocation of costs should be aligned with the drivers of those costs. This ensures that the market can send accurate price signals so consumers can respond efficiently. The capacity market is intended to ensure sufficient resources are available to serve load during shortage or near-shortage conditions, which generally align with times that the load on the system is the greatest.[79] Thus, costs should be allocated in a way that reflects consumption during those peak periods: consumers that consume more during those periods and contribute more to the peak demand level should contribute more to capacity cost recovery. By aligning the price signal with the peak period, consumers have an incentive to reduce consumption during the system's peak, which should allow the market to efficiently reduce the need for additional capacity in the future and, in turn, reduce the overall capacity cost of the system.

The precise definition of the peak period used to determine capacity cost allocation varies across jurisdictions, generally reflecting underlying characteristics of each market:

- In markets where the annual load factor is high (*i.e.*, a non-peaky load profile where the average load is close to the annual peak load), a **wide peak period approach** is most efficient. Under this approach, the peak period can be pre-defined to include a wide range of hours throughout the year. This "ex-ante" approach allows consumers to plan to reduce their consumption broadly over many peak hours. This incentive aligns with

[79] The alignment is not perfect due to planned and unplanned maintenance outages, variation in output from variable renewable energy, and other factors.

value since such markets are vulnerable over a broad range of hours when loads are close to their maximum, and random generator outages can cause shortages.

- In markets where annual load factor is lower, it is likely more efficient to adopt a **narrow peak period approach**, where consumers are charged based on their consumption during only a few of the highest realized load hours during the year. Since it is only possible to determine when these peak load hours occurred after the fact, it is necessary to select these hours on an ex-post basis, and they are used to allocate capacity costs for the *following* delivery year. With the price signal concentrated in a few high-consumption hours, consumers have a strong incentive to anticipate when these highest load hours will occur and reduce their consumption during those periods. These strong price signals will likely lead to stronger demand reductions in those hours, efficiently reducing capacity costs.

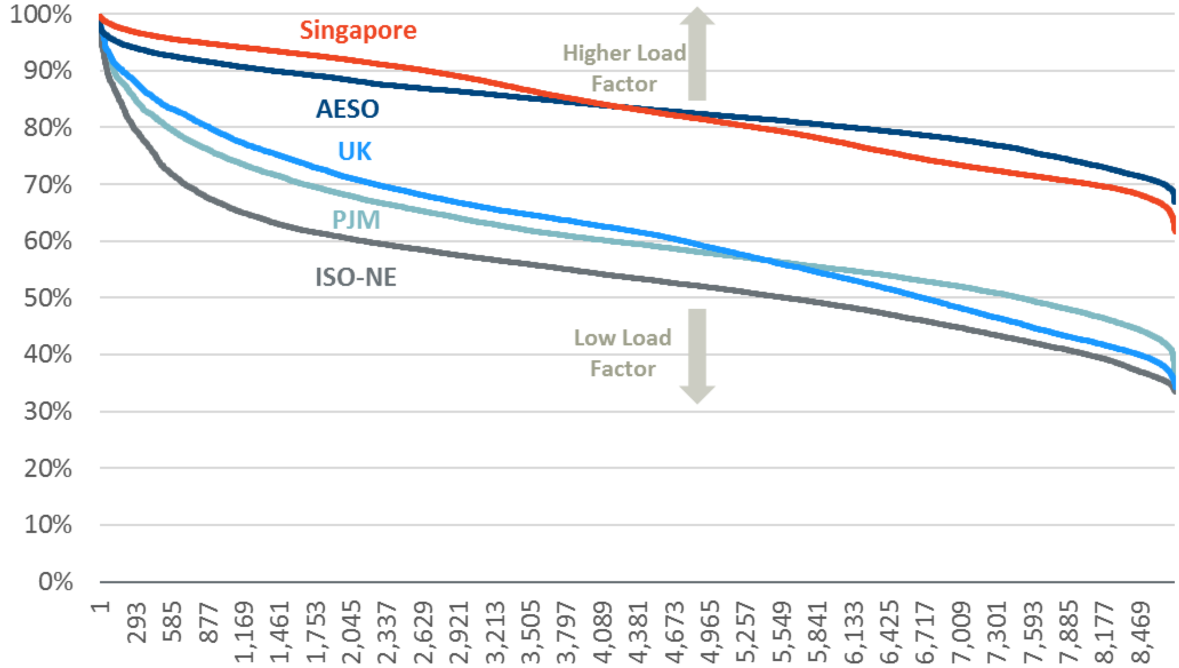
Selecting the peak period for cost allocation purposes can have significant impacts on market efficiency. For example, consider a market with a relatively flat annual load profile but where capacity costs are allocated on an ex-post basis across only the highest few peak hours in the year. Consumers would aim to reduce their consumption in just those highest hours but might not considerably reduce the need for capacity with the absence of an incentive to reduce consumption during other high consumption periods. Conversely, if a market had very pronounced peak loads but costs were allocated according to consumption over many hours, then capacity costs would likely stay relatively unchanged, as consumers would broadly reduce their consumption, but not focus those efforts in the most important, highest load hours. It is important that the cost allocation approach reflects the actual drivers of the costs so that price signals incentivize efficient behavior from the consumers.

None of the cost allocation approaches presented here runs any risk of under-recovery due to ex-post allocation of capacity costs. In each delivery year, *actual* capacity costs are fully allocated according to the methods described. Thus, even if consumers reduce their consumption during the relevant peaks, the total costs allocated will still be sufficient. In the medium and long run, load reduction during the peaks will reduce capacity needs and system costs for all consumers.

Approaches in Other Jurisdictions

As discussed above, markets with higher load factors should find it beneficial to allocate costs using a broader, ex-ante peak period definition, whereas markets with a lower load factors and more pronounced peaks should adopt a narrower, ex-post peak period definition. This pattern is generally confirmed in our review of other jurisdictions. Alberta (designated as "AESO" in the graph) relies on ex-ante, wide peak period definitions corresponding to its high load factor, as shown in Figure 26. PJM and ISO-NE rely on ex-post, narrow peak period definitions and have the lowest load factors. The UK has a load factor more similar to PJM but still determined that a wide peak period definition was most appropriate for reasons we have not been able to confirm.

Figure 26: Load Duration Curves across Markets



Sources and Notes: The Singapore, AESO, PJM, and ISO-NE load duration curves reflect 2018 load. UK load duration curve reflects 2015 load (latest we could find publicly available). AESO, PJM, ISO-NE data from Ventyx Velocity Suite. UK load data from European Network of Transmission System Operators for Electricity. Singapore load data provided by EMA.

Alberta's (AESO's) proposed approach allocates costs according to MWh consumption during a broad range of hours in peak and mid-peak blocks.[80] These blocks were determined by an analysis of the distribution of expected unserved energy (EUE) as follows:

- The peak blocks represent the very highest load hours throughout the year in August through October, for hours ending 16:00 – 18:00 (HE16–HE18) and November through February HE18–HE19; these hours receive the highest cost allocation (on a per-MWh basis); and
- The mid-peak block represents the other hours with non-negligible EUE potential throughout the year (HE8–HE23 excluding the already designated peak hours);[81] these hours receive a lower cost allocation.

By choosing to allocate costs on a wide range of hours, the AESO would be able to incentivize load reduction during peak times throughout the year, which is valuable given their very flat load duration curve.

---

80 Although the AESO capacity market was recently cancelled, their proposal materials offer another legitimate point of reference. AESO, *Tariff Design for Capacity Market and Bulk and Regional Transmission Cost Allocation*, March 2019.

81 AESO considered hours with unserved energy contribution greater than 0.0007% per hour across months. The period they chose had a lower bound of about 10 observations of a given month-hour exceeding the 0.0007% EUE cutoff.

The UK also relies on a broad peak period definition, even though its load duration curve is not quite as flat as that of Alberta.[82] The UK defines its peak periods, or "periods of high demand," from hour ending 17:00 – 19:00 on any workday between November and February. Consumers are charged based on their volumetric (MWh) consumption during these periods, similar to Alberta's approach.

PJM, which has a relatively much lower load factor, and thus higher peak demand, relies on a narrow peak period definition, where costs are allocated based on *actual* consumption during the five highest coincident peak hours in the year.[83] The effect of this approach is that capacity costs are allocated according to consumption during very few (five) hours in each year. Each consumer's peak load contribution (PLC) is calculated as its consumption during the (five) coincident peaks and each retailer who serves load has an obligation that reflects the sum of the PLCs across all its consumers. Because these periods are determined after the delivery year, they are used to allocate capacity costs for the *next* delivery year. This necessitates an additional step of tracking consumers as they can potentially switch between retailers from one year to the next.

ISO-NE allocates capacity costs using a single coincident peak methodology.[84] Similar to PJM, ISO-NE has relatively very high peak loads such that it seeks to focus capacity price signals on just the highest load hour of the year. Relying on this narrow peak methodology, capacity cost allocations are based upon consumption during the annual system-wide coincident peak load for the prior year.

B. Recommendations for Singapore

We recommend that the costs of Singapore's FCM be allocated to consumers or retailers in proportion to actual consumption during different period-types, where higher capacity costs will be incurred for consumption during peak (and potentially also mid-peak) periods. In this section we step through each component of this recommendation.

Costs Allocated in Proportion to Actual Consumption

As outlined above, capacity costs should be allocated in a way that reflects the cost drivers. Since the load duration curve in Singapore is relatively flat, as shown in Figure 26, we propose to adopt an ex-ante, "wide peak" approach where the costs are allocated across a broad set of pre-established hours. This has the advantage of spreading the cost allocation across many hours that contribute to incurring capacity costs and gives consumers a defined set of hours during which they receive an incentive to reduce their load. To reflect prevailing supply/demand conditions, some hours (*e.g.*, peak periods) could be allocated higher capacity costs per MWh.

Allocating costs to a very narrow set of hours defined after the delivery year, as in PJM and ISO-NE, would not be appropriate for Singapore. This coincident-peak allocation approach

---

82 EMR Settlement Limited, *G12 – Supplier Capacity Market Demand Forecast*, June 2018.

83 PJM, *Manual 18: PJM Capacity Market*, Section 7 (pp. 149-155), January 2019.

84 ISO-NE, Demand-Side Settlement – FCM Charges, October 2018.

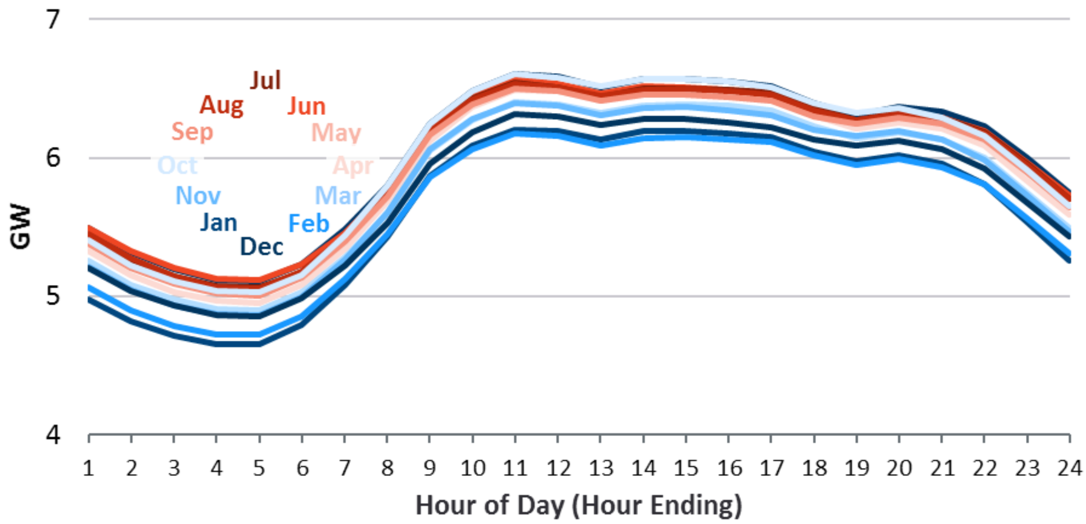
would send a price signal that is too concentrated given the flat load duration curve of the Singapore system.

Regarding the mechanics of cost allocation, we recommend establishing a volumetric rate (S$/MWh) that applies to all consumption during the peak period, described below. The rate would be calculated to recover the appropriate capacity costs over the expected volume of consumption. As wholesale electricity charges in the Singapore market are set on an ex-ante basis, the rate could also be implemented in the same manner. However, this means that monthly (or quarterly) true-ups could be used to continually adjust the rate going forward if there is slight under-collection or over-collection in preceding months.

Costs Allocated to Peak Hours

We propose to use a constant set of hours throughout the year to define the peak period. As shown in Figure 27, the daily load profile is almost identical across months such that the highest load hours remain fairly constant. Additionally, using a consistent set of hours will help to keep the peak period definition simple, although it may not perfectly capture intra-day granularity such as the dip during the midday lunch hour.

Figure 27: Representative Daily Load Profile across Months



Sources and Notes: Representative daily load profiles reflect average monthly-hourly load during 2014-2018. Singapore load data provided by EMA. Note axis does not begin at zero.

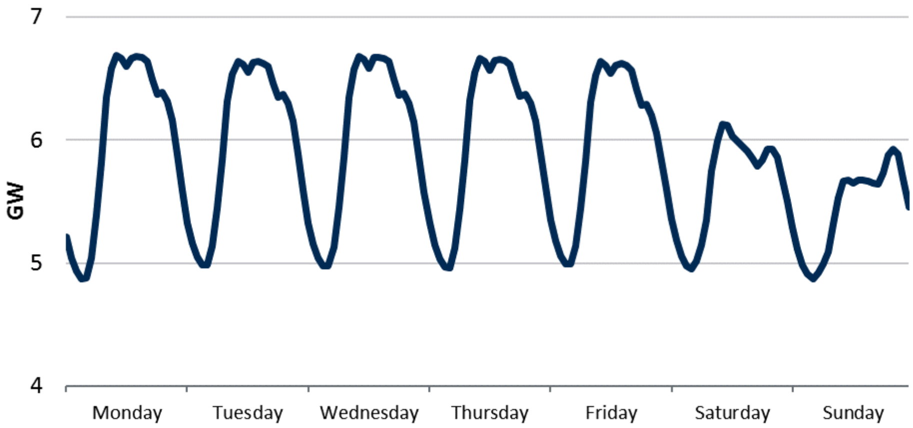
The threshold to determine the exact definition of peak hours within the day is somewhat subjective but should reflect the marginal reliability cost associated with incremental electricity usage, or inversely, the reliability value to the system of conserving a marginal MW. This value is proportional to the hourly probability of loss of load (POLL) given by internal EMA reliability modeling. In today's system, this probability is highest in the late morning through evening, when average system load is highest across all days. EMA should determine and publish the definition of peak period before each auction.

An alternative approach is to define a "peak" and "mid-peak" period, as proposed in Alberta; the peak period would have higher per-MWh costs allocated to it to reflect the higher value of consumption/conservation during those hours. The disadvantage of such a solution is that it likely only marginally improves the efficiency of the price signals, while somewhat adding to the complexity of both the cost allocation design and the price patterns to which consumers would be expected to respond.

ALLOCATION OF CAPACITY COSTS

We propose to allocate capacity costs in proportion to actual MWh consumption during the different period-types, where higher capacity costs will be incurred for consumption during peak (and potentially also mid-peak) periods.[85] The data shows that load on weekends is much lower and does not present significant risk of shortage events. As shown in Figure 28, the weekends have much lower average loads than weekdays.

Figure 28: Representative Weekly Load Profile



Sources and Notes: Representative weekly load profile reflects average hourly load during 2014-2018 across days of the week. Note axis does not begin at zero. Data provided by EMA.

Weekends have considerably lower average load as well as daily peak load. This is further captured in Figure 29 and Figure 30, where we observe that Saturday and Sunday have significantly lower load throughout the peak period and do not contain a single observation in the top five percent of load throughout the year. Thus, we conclude that consumption on weekends is very unlikely to contribute to potential shortage conditions, and thus should not be allocated the bulk of the capacity cost.

Figure 29: Average Load in Each Hour and Day of Week (GW)

	Hour Ending																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
Monday	5.2	5.0	4.9	4.9	4.9	5.0	5.4	5.8	6.3	6.6	6.7	6.7	6.6	6.7	6.7	6.7	6.6	6.5	6.4	6.4	6.3	6.2	5.9	5.6
Tuesday	5.3	5.2	5.0	5.0	5.0	5.1	5.5	5.8	6.3	6.5	6.6	6.6	6.5	6.6	6.6	6.6	6.6	6.5	6.4	6.4	6.3	6.2	5.9	5.6
Wednesday	5.3	5.2	5.0	5.0	5.0	5.1	5.5	5.9	6.4	6.6	6.7	6.7	6.6	6.7	6.7	6.7	6.6	6.5	6.4	6.4	6.3	6.1	5.9	5.6
Thursday	5.3	5.1	5.0	5.0	5.0	5.1	5.4	5.8	6.3	6.6	6.7	6.6	6.6	6.6	6.7	6.6	6.6	6.5	6.4	6.4	6.3	6.1	5.9	5.6
Friday	5.4	5.2	5.1	5.0	5.0	5.1	5.4	5.8	6.3	6.5	6.6	6.6	6.5	6.6	6.6	6.6	6.6	6.4	6.3	6.3	6.2	6.1	5.8	5.6
Saturday	5.4	5.2	5.1	5.0	5.0	5.0	5.1	5.4	5.7	6.0	6.1	6.1	6.0	6.0	5.9	5.9	5.9	5.8	5.8	5.9	5.9	5.9	5.7	5.5
Sunday	5.3	5.1	5.0	4.9	4.9	4.9	5.0	5.1	5.3	5.5	5.7	5.7	5.6	5.7	5.7	5.7	5.6	5.6	5.7	5.9	5.9	5.9	5.7	5.5

Sources and Notes: Table reports average load in 2018 across each day-of-week and hour. Darker red shading indicates higher load. Data provided by EMA.

85 We have therefore only analyzed weekday vs. weekend load, as load on holidays in Singapore follows a similar pattern to that observed on weekends (as we have observed in other markets), and as such should also be excluded from the definition of peak hours for cost allocation purposes.

Figure 30: Distribution of Hours in Top 5% of Load

	Hour Ending																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
Monday	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	4%	3%	2%	3%	3%	3%	3%	1%	0%	0%	0%	0%	0%	0%
Tuesday	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	3%	3%	2%	3%	3%	3%	3%	1%	0%	0%	0%	0%	0%	0%
Wednesday	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	3%	2%	1%	2%	3%	3%	2%	0%	0%	0%	0%	0%	0%	0%
Thursday	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	2%	2%	1%	3%	3%	3%	2%	0%	0%	0%	0%	0%	0%	0%
Friday	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	3%	2%	1%	2%	3%	3%	2%	0%	0%	0%	0%	0%	0%	0%
Saturday	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Sunday	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

Sources and Notes: For each day-of-week and hour in 2018, we report the fraction of half-hour observations that fall in that period and are in the top 5% of highest system load observations. Data provided by EMA.

Costs Allocated in Every Month of the Year

Singapore experiences relatively small variations in load patterns across the year. As a result, shortage events could occur in any month and, therefore, load in all months drives capacity costs. This relatively consistent monthly load pattern is captured in Figure 27.

Since load is relatively even across months, the supply cushion, which represents the difference between load and the available capacity to serve that load, is very similar during peak hours of each day throughout the year. As a result, we expect that the effect of a marginal unit of consumption on reliability during on-peak, weekday hours should be relatively similar throughout the year. Following the principle of cost causation, this implies that costs should be equally distributed across all months of the year.

Allocation of Costs to Self-Suppliers With Embedded Generation

Some market participants have both load facilities and embedded generation (EG) facilities on the same site (EG Consumers), with the generation primarily intended to serve on-site electricity, heat, steam, and other needs (rather than generation for the market). The EG could include thermal generation units, solar and storage. We recommend treating these participants in a way that respects their unique characteristics *and* interacts appropriately with the FCM.

Thus, we recommend that EMA enable EG Consumers to choose whether to pay capacity charges based on a declared maximum withdrawal (DMW) from the grid, or on a gross basis. If the EG Consumer does not nominate their choice, capacity charges will be allocated based on gross treatment.

DMW treatment. The EG Consumer who opts for this treatment will be required to project its peak demand to be served from the grid (PD) four years ahead of each delivery year. For such a consumer, EMA will procure sufficient generation capacity from the FCM base auction to support its PD projection in the delivery year, and the EG Consumer will be required to install a load limiting device (LLD).[86] During the delivery year, the capacity charges will be allocated in the following manner:

[86] A LLD is a device installed at the grid connection point *i.e.*, between the grid and a consumer's premises including his load facilities therein. The LLD will be automatically triggered (when certain operational set points are breached) to discontinue electricity supply to the consumer.

- If the EG Consumer's consumption via the grid is within a ±5% tolerance of its PD projection, it will pay prevailing capacity charges based on its consumption;
- If the EG Consumer's consumption is less than 95% of its PD projection, it will still be required to pay capacity charges based on 95% of its PD projection; and
- If the EG Consumer's consumption is more than 105% of its PD projection, it will pay the prevailing capacity charge for 105% of its PD projection and twice the prevailing capacity charge for consumption in excess of 105% of its PD projection, for the first two half-hourly occurrences in a given delivery month. The LLD will be triggered from the third half-hourly occurrence onwards, and the counter would be reset for each delivery month.

The DMW treatment is consistent with EMA's proposed framework to facilitate the entry of large discrete loads (LDLs).[87]

Gross treatment. The EG Consumer will be required to pay capacity charges in the delivery year based on their half-hourly gross load. The gross load is defined as the total of (1) electricity drawn from the grid, and (2) electricity generated from the EG and consumed on-site. For such a consumer, EMA will procure sufficient generation capacity from the FCM auction based on its gross load.

Participation of EG Consumers in the FCM. EG Consumers would have the option to utilize any excess capacity from their EG (*i.e.*, only the capacity for injecting excess electricity into the power grid) for participation in the FCM on a voluntary basis; choosing not to participate in the FCM would not preclude the EG Consumer from participating in the real-time energy and/or ancillary services market.

EG Consumers choosing to participate in the FCM must meet all requirements for participation as any other supply resource. For example, they will be qualified in the same manner, and will be subject to the same obligations and penalties:

- Must meet the minimum threshold of 1 MW for participation;
- Must qualify the EG resource(s);
- In the delivery year, cleared EG capacity will receive capacity payment for meeting its capacity supply obligation (CSO); and
- Penalties apply for failure to meet its CSO.

C. Settlement Framework

As the market operator and administrator (of both the spot market and FCM), EMC will be the appropriate party to levy and collect the capacity charge from all energy market participants in respect of their half-hourly energy purchase under the electricity market rules, similar to existing wholesale market charges. EMA is developing the detailed operational process to effect

87 See EMA, "Framework for Serving Electricity Demand of Large Discrete Loads Consultation Paper," March 6, 2020. Available at: https://www.ema.gov.sg/cmsmedia/LDL%20-%20Consultation%20Paper.pdf.

the above cost allocation method and settlement framework, and will separately consult retailers, the Market Support Services Licensee and EMC in due course.

Recommendations and Next Steps
Cost Allocation Approach • Define the peak (and potentially also mid-peak) periods as certain hours of the year • Determine S$/MWh rate to allocate capacity costs volumetrically to energy consumed during defined peak (and potentially also mid-peak) periods • Enable self-suppliers to participate on a gross basis, or an "exemption" basis based on a self-nominated DMW
Next Steps • Finalize peak period definition • Consult stakeholders on the settlement framework separately

XII. Reforms to Energy/Ancillary Services

The introduction of a FCM can be complemented by changes to the existing energy and ancillary services markets to ensure the combined markets function efficiently.

Given that the FCM provides for the recovery of fixed costs, offers in the energy market should be mitigated to their short-run marginal costs. EMA had previously implemented an energy market power mitigation in the form of vesting contracts, that has since been phased out owing to lower market power concern with the short-term overcapacity situation in the SWEM. However, with implementation of the FCM expected to increase the correlation of supply and demand, coupled with EMA's expectation of the reserve margin tightening in the coming years, suppliers will have increasing incentive and ability to exercise market power in the real-time energy market.

Thus, to emulate a perfectly competitive market and allow the real-time energy market to always clear the resources with the lowest costs, EMA intends to implement a one-pivotal supplier test (1PST) for the energy market in 2023. This is consistent with practices in other markets, and for the purpose of familiarity to market participants is similar to the capacity market power mitigation mechanism for the FCM as proposed in Section VI. Under this mechanism, depending on total supply offers and demand in each real-time energy market dispatch interval, only suppliers who fail the 1PST will be subject to mitigation to a reference level. EMA proposes that the reference level be set at 3x the short-run marginal cost of a CCGT to provide a reasonable buffer for suppliers to recover their variable costs.[88] Further, EMA

88 Based on EMA's "Review of the Long Run Marginal Cost Parameters for Setting the Vesting Contract Price for 2019 and 2020", the short-run marginal cost of a F-Class CCGT (including carbon price, variable non-fuel cost and the fuel component) is about S$118/MWh, *i.e.*, the representative mitigation level would be approximately $355/MWh. In the transitional period EMA intends to adopt these variable cost components, subject to updates as part of the vesting contract procedures.

intends to maintain the energy market price cap of $4,500/MWh to continue enabling efficient pricing signals.

XIII. Conclusion

We have worked with EMA to design a FCM for Singapore to meet the objectives described in Section I—to maintain resource adequacy, through price signals that reflect needs and can attract and retain sufficient resources, and by harnessing competition to maximize economic efficiency to minimize long-run costs to consumers.

The proposed FCM design outlined in this document builds on lessons learned in other jurisdictions with FCMs that have successfully met similar objectives, while recognizing the unique characteristics and requirements of Singapore's power system and market.

The proposed design works as a package in which the various design elements complement each other, such that changing one element would have consequences for others. We envision that this design can be continually reviewed and refined after implementation, to ensure that the market evolves coherently to address longer-term trends and fundamental technology-shifts while meeting Singapore's needs into the future.

BOSTON	WASHINGTON	MADRID
NEW YORK	TORONTO	ROME
SAN FRANCISCO	LONDON	SYDNEY

THE Brattle GROUP