Prepared For:

NextEra Energy Resources

Juno Beach, Florida

Performance Incentives in ISO New England's Forward Capacity Market

Prepared By:

Robert Stoddard

Charles River Associates

200 Clarendon Street

Boston, Massachusetts 02116

Date: 2 July 2013

CRA Project No. D18053-01

Disclaimer

Table of contents

1. Overview and Summary

ISO New England (ISO) is currently conducting a stakeholder process to redefine the product it buys through the Forward Capacity Market (FCM). In particular, ISO has proposed to replace the current penalty structure for failure of a Capacity Resource to be available during a Shortage Event with a Performance Incentive package. ISO's explicit goal is to provide an increased incentive for suppliers to be available during critical periods and to provide clear financial incentives for capacity suppliers to make economic investments to improve performance, whether through investment in dual-fuel capability, purchase of firm fuel transportation, or other creative approaches to enhancing system reliability.

NextEra Energy Resources (NextEra) has retained Charles River Associates' Senior Consultant Robert Stoddard to consider the ISO proposal and independently assess the range of possible improvements in the FCM design to address the reliability concerns raised by ISO. The opinions expressed in this paper are solely those of Mr. Stoddard and do not necessarily represent those of NextEra.

1.1. Executive Summary

To ensure their efficacy into the future, the ISO energy and capacity markets need to be updated. While generators and suppliers are concerned that the removal of the FCM price floor limits the opportunity to recover their costs, the ISO's overriding concern is unit performance during periods of scarcity and firm fuel. For the most part, ISO's Performance Incentives proposal addresses each of these issues. But, with several modifications, as briefly described below, Performance Incentives can be significantly improved and can more efficiently and effectively achieve the ISO's desired goals.

The implementation of more stringent environmental standards and a general reduction in the cost of natural gas as a fuel have produced a significant increase in the proportion of natural gas generating resources as a percentage of the New England supply mix. ISO has recognized that during certain system conditions the limits on gas infrastructure can raise reliability and performance concerns for those same gas resources as they compete with residential and commercial customers for fuel. Understanding this reality, ISO has designed its Performance Incentives proposal as a way of incentivizing greater levels of resource availability and flexibility, particularly during periods of scarcity. Performance Incentives should directly translate into pricing that better provides generators and suppliers with the opportunity to recover their costs—specifically, providing an upfront revenue stream to finance unit upgrades, firm fuel arrangements, secondary fuels, or other investments to improve unit availability. With a few important modifications, as suggested below, the ISO's Performance Incentives proposal can be successful in providing the right incentives for performance.

In brief, the Performance Incentives proposal would pay an extra Performance Payment Rate (PPR), tentatively set at $5,000/MWH, to all resources providing energy or reserves during specified scarcity conditions. In contrast to Locational Marginal Price (LMP)-based energy charges, which ISO collects from loads, ISO will collect the Performance Incentives charges from capacity suppliers in proportion to their Capacity Supply Obligation (CSO). In a particular scarcity event, a capacity supplier may have enough dispatched resources to earn back these Performance Incentives charges; nonetheless, the supplier earns lower net Performance Incentives revenues than it would have had it not taken on a CSO. Consequently, under the Performance Incentives proposal, all resources will have a strong incentive to perform during scarcity conditions, and capacity suppliers will have a direct financial incentive to take steps to assure that their physical assets will be available to hedge their financial obligation to pay the Performance Incentives charges.

While only capacity suppliers pay the PPR payments, there necessarily will be additional costs to the market through higher capacity payments. The Performance Incentives proposal properly adds incentives and risks to encourage performance, and the additional risks to capacity suppliers will be reflected in capacity offers and auction results. As ISO has appropriately stated, capacity suppliers cannot and should not accept these costs and the associated risks without compensation. By taking a CSO, a resource is giving up the right to earn the PPR on its full output during scarcity conditions; instead, it will only earn (or pay) the difference between its actual output and its prorated share of system requirements. This extra financial cost (and risk) of being a capacity supplier likely and appropriately will add to FCM offers an amount reflecting the value of a generator's calculated risk of loss to equilibrium capacity costs, depending on expected system conditions, the PPR level, the level of perceived risk, and the risk tolerance of capacity suppliers.

In this, the goal of the Performance Incentives proposal is sound: provide a sustained pricing incentive to suppliers so that they will undertake investments in existing or new resources not only to meet the *planning reserve* requirement, but also to provide ISO the resources to ensure real-time operational reliability. The question, though, is whether the Performance Incentives program, as currently proposed, is well calibrated to achieve this goal in a cost-effective way.

This Charles River Associates White Paper reviews ISO's Performance Incentives proposal and concludes that, without the risk mitigants discussed below and without the integration of a downward-sloping demand curve and other market design changes, the level of risk is likely higher than necessary to provide appropriate incentives to invest in reliability enhancements. These high risks may lead some resources, especially those with long notice periods that are not typically committed except during peak periods, to withdraw from the capacity market or retire altogether. Such an outcome is not necessarily a sound or desirable goal—certainly not in the near term, given ISO's Winter Operations Summary memorandum. Instead of simply ramping up risk to push out certain resources, the goal should be to create a fully functioning market that reasonably incents resources to be available and reasonably punishes those resources that are not available when they should be. Unfortunately, Performance Incentives blurs the distinction between *planning* reserves and *operating* reserves. Generators that provide valuable fuel diversity and are frequently dispatched during high load periods (consistent with their role as planning reserves) may not be available quickly during shoulder periods—a job for operating reserves. Asking the entire supply stack to invest in preparation of offering operating reserves is not cost efficient and will require the markets to provide significantly higher revenue. The investments required cannot be recouped in a single year, and the reformed FCM needs to provide stability in addition to higher revenues.

The following modifications will enhance the ISO's Performance Incentives proposal:

1. **Retain stop-loss provisions.** The current FCM structure of nested limits on daily, monthly, and annual loss limits should be retained, with modifications. Reasonable stop-loss limits have a very small effect, prospectively, on expected performance penalties and, therefore, on incentives to undertake cost-effective reliability investments. By truncating the potential for extreme losses, stop-loss provisions will lower the risk premium suppliers will include in their offer price, as well as their costs in posting financial assurance.

2. **Retain a limited set of excused unavailability.** ISO's proposed "no-excuses" or "no-fault" approach to measuring availability may have the virtue of simplicity, but it also can inflict substantial costs for events outside the capacity supplier's control. Such costs have no incentive effect; they are simply an additional cost and risk that will be priced into capacity offers. This paper outlines a proposal for excused unavailability that is much more restrictive than the current exemptions and, therefore, will reduce risks and costs without loss of meaningful incentive effects.

3. **Enhance Supplemental Availability Bilaterals.** Capacity suppliers are used to facing and managing risk, but it is essential that ISO provide them risk management tools tailored for the task. With suppliers exposed to high performance charges under a broader range of conditions, the Supplemental Availability Bilaterals should be updated to reflect their more extensive function. The stop-loss provisions must be crafted to allow transparent price formation around a common, fungible product, and ISO should expand its bulletin board clearinghouse to facilitate liquid trading.

4. **Implement a capacity demand curve.** As we have seen in the Forward Capacity Auctions (FCAs) to date, the outcomes are polar; each auction has cleared either at the price floor or cap. With the elimination of the price floor, capacity prices could fall to levels that will not support the ongoing investment in reliability improvement measures sought by ISO, even if the capacity supply balance is fairly tight. ISO, its Board of Directors, and the Internal Market Monitor (IMM) have all expressed support for integrating a demand curve construct within FCM. And the ISO's External Market Monitor (EMM), Dr. David Patton, recently reiterated his call for integrating a demand curve within FCM in his "2012 Assessment of the ISO New England Electricity Markets." Implementing a variable resource requirement in the FCA will provide greater durability of the capacity price signal for investment in new and existing resources, reducing risks to suppliers and costs to consumers.

The Performance Incentives program also creates substantial challenges for the IMM. Competitive capacity offers will include a new component: the expected cost of paying the Performance Incentives to performing units, and the associated risk. Suppliers also will need to add in expected Performance Incentive receipts when weighing whether they can cover their going forward costs (for existing resources) or all-in costs (for new resources). The IMM has made solid proposals so far; this White Paper outlines additional considerations:

- Additional costs of accepting a CSO, including costs of posting financial assurance, must be allowed in dynamic and static delist bids;
- Substantial deference to a supplier's business judgment about future Performance Incentive costs, with mitigation only when a capacity supplier possesses structural market power and fails to support its offer price;
- Capacity suppliers should be allowed to include a risk premium in their supply offers above the expected value of future Performance Incentive charges;
- Different segments of Capacity Resources may face different risk profiles; therefore, as in the energy market, potential Capacity Resources should be allowed to offer several price/quantity pairs to reflect different levels of risk at higher output levels; and
- If the actual frequency of scarcity events exceeds the frequency used in offer mitigation, the PPR should be scaled down to limit the downside risk to capacity suppliers.

The ISO is looking for resources to improve their performance during times of scarcity. Owners of these assets will only make the necessary investments if they see a reasonable opportunity to earn a positive return on the invested capital. That confidence will come only from stable revenues, unlike those witnessed by the binary capacity market design currently employed. A demand curve would provide the stability in revenues needed to support investments in reliability. Moreover, capacity demand curves reduce rate volatility for customers and reduce total consumer costs by reducing the frequency of shortage conditions.

It is important that a vigorous discussion of these ideas occurs holistically in the stakeholder process to ensure that a robust and commercially reasonable Performance Incentives program can be put into place that balances the risks and incentives needed to ensure long-term reliability of the New England bulk power system.

1.2. Rationale for Capacity Markets

A functional market does more than merely set the spot prices at which products are bought and sold; it also supports long-term capital investment in an efficient mix of production capacity to meet current and future needs, as well as informs the long-term purchase decisions of customers and, therefore, their investment decisions. Electricity markets should be no exception to this general rule.

Some have argued that "energy only" markets should be the gold standard. In such a market, real-time prices can rise to very high levels during scarcity conditions, inducing marginal resources to generate additional power or, more likely, customers who place a relatively low value on electricity to curtail their usage. The potential for high-scarcity pricing then would ripple through various forward prices. The day-ahead price would reflect the market's expectations for potential scarcity in real time, and monthly, seasonal, or annual contract pricing also would reflect the hedge value against potential scarcity (as well as the fundamental cost of producing electricity). Under this theory, power plant owners invest or retire assets in anticipation of these higher future energy prices.

Electricity differs in important ways from nearly all other commodities. It follows, therefore, that the way we pay for a reliable bulk power system should differ from the standard spot markets for cocoa beans or strawberries. This point of view is controversial. Some leading thinkers advocate strongly for an "energy only" market, which pays only for power delivered—potentially at very high prices during times of scarcity. Even an "energy only" market diverges from the classic commodity spot market, however, because it pays units for providing operating reserves—that is, some resources are paid not for providing the commodity, but for being willing and able to do so quickly in the event of a system contingency. Simply put, capacity markets are the logical extension of paying for operating reserves. We could equally well call a "capacity market" a "planning reserves market." Just as we pay resources to remain on standby to provide short-term operating reserves that cover unexpected short-term contingencies, the capacity market pays resources to remain available to provide longer-term planning reserves that cover unexpected longer-term contingencies, including higher-than-expected load growth, extreme peak loads, and outages of generation and transmission. Just as the operating reserves market prices the value of having standby generation available to meet an in-day reliability standard, the "planning reserves market" (i.e., capacity market) prices the value of meeting a planning standard.

This formulation is slightly different from the "missing money" argument first articulated by Cramton and Stoft.[1] The existence of a capacity market should not be a substitute for sound pricing in the underlying energy and operating reserves markets. No proponent of "energy only" markets has demonstrated that establishing an efficient scarcity pricing mechanism fully closes the "missing money" gap—nor could it, because the planning reserve requirement is not set based on economics, but on engineering and political judgment. Regulators mandate minimum installed reserve margins that largely remove periods of scarcity. In economic terms, regulators have an implied high value on planning reserves that is not included in any price in the energy or operating reserve markets. By mandating that supply exceed expected demand (including demand for operating reserves) by 10% or more, regulators—in this case, ISO in its role as the security coordinator for New England—drive down energy prices and, consequently, margins. In any other industry, low margins would spur the exit of capacity and cut off new investment. In electricity markets, however, such disinvestment may be forestalled by ISO to ensure sufficient resources. It is more challenging, though, to mandate new investment when margins are poor. Capacity markets bridge this gap between the planning reserve margin set by regulators and the margin that naturally would arise from an "energy only" market design.

[1] Cramton, Peter and Steven Stoft, "The Convergence of Market Designs for Adequate Generating Capacity," White Paper for the Electricity Oversight Board of California, 25 April 2006, available at http://stoft.com/metaPage/lib/Cramton-Stoft-EOB-2006-04-ICAP-energy-convergence.pdf.

Capacity markets also can help address further implementation challenges to a purely bid-based “energy only” market:

- First, absent some other mechanism, very high prices during scarcity conditions can only be achieved with very high offers from suppliers. Most organized markets, including ISO’s, sharply constrain the ability of a supplier to offer at much above marginal cost. Allowing suppliers to increase their offer prices above costs invites the abuse of market power. It also exposes suppliers to after-the-fact litigation for pricing practices, which may make scarcity *pricing* even rarer than scarcity *events*. By including automatic premiums during scarcity events, a capacity market design can improve the operation of the real-time markets.
- Second, the cost of financing major capital projects depends greatly on the risk of the project. This risk profile is dominated by the uncertainty of future revenues. Replacing some of the future price risk inherent in an energy-only market with a relatively more stable stream of capacity revenues can allow new investment to pay lower financing costs, a savings that is passed on to customers through competitive markets. While in theory customers can achieve this same savings in an energy-only market through long-term contracts with suppliers, in New England such contracting is unlikely. Most New England electric distribution companies have little or no role in long-term contracting, and competitive retail suppliers have made only limited inroads. Thus there are few opportunities in the region for long-term contracting that would simultaneously allow end-use customers to manage their power costs and support financing of new power plant construction.
- Third, a capacity market can recognize the distinction between the time frames in which resources are available. The characteristics of resources that make good planning reserves are not necessarily the same as those that make good operating reserves. Operating reserves are needed to respond to contingencies that need to be resolved in a matter of seconds or minutes, whereas planning reserves are intended to address contingencies that develop over the course of hours or days, such as when alternative supplies are unavailable (for example, because of limited gas pipeline capacity or extended unit outages) or when loads exceed planning levels perhaps because of strong macroeconomic growth or temperature extremes. Thus, units that have low fixed costs to retain on the system, such as older steam units, can provide value on a planning basis even if, in a typical operating year, they are rarely needed in dispatch.

To address these issues, ISO and other Regional Transmission Operators (RTOs) have adopted capacity markets. These markets effectively pay a “standby” price to qualified capacity suppliers on a time scale consistent with planning reserve requirements. These capacity resources may include demand-side resources as well as active generation. Capacity markets can therefore provide the financial returns necessary to fund new and existing supply resources sufficient to meet planning reserve requirements.

1.3. Performance Incentives in Capacity Market Designs

The simplest form of a performance incentive is to derate the quantity of capacity that a supplier can offer based on its historic performance. The New York market still relies exclusively on this Unforced Capacity (UCAP) construct, which derates capacity by its Average Effective Forced Outage Rate under Demand (AEFORd). The PJM Interconnection (PJM) also includes a similar mechanism as one of the performance incentives in its Reliability Pricing Model (RPM). In this design, a forced outage today reduces the quantity of capacity you can sell going forward.

This simple UCAP approach has several merits, though it is not intended to be a proxy for scarcity pricing. One merit of the UCAP design is certainty of revenues and charges. The financial penalties associated with a forced outage are incurred over one or more years, minimizing the potentially disruptive effect of a large bump in cash flows to a generator. Moreover, because penalties cannot reduce capacity payments below zero, there is no need for the RTO to have financial assurance to cover a potential penalty payment. Another benefit is that the RTO purchases a set of Capacity Resources that are collectively more likely to be available. By paying more to units that have higher availability, this form of capacity market selects higher-availability resources. On the other hand, average availability metrics are just that: averages. Although the EFORd statistic includes some statistical weighting, there is little in the metric that distinguishes a forced outage when there are ample alternatives from an outage when the system is stressed. As implemented by PJM, the UCAP concept also excludes outages deemed outside management control, implicitly recognizing the difference between serving as a planning resource and operating reserve.[2]

Two of the current capacity market designs include mechanisms that attempt to improve on this broad UCAP concept.

- PJM not only derates capacity using the EFORd metric, discussed above, but it also assesses financial charges and credits based on a measure of availability during a pre-defined set of Peak Demand-Hour Periods of about 500 hours, about 300 in the summer months and another 200 in the winter.[3] If a Capacity Resource is on partial or full-forced outage during a Peak-Hour Period when its cost-based offer for energy would have been less than the applicable Locational Marginal Price, or when the resource would have been called to provide operating reserves, then its peak-period availability (measured as the Equivalent Forced Outage Rate-Peak or EFORp) is

[2] These exceptions are defined in accordance with the standards and guidelines of NERC, as defined in the Generating Availability Data System, Data Reporting Instructions.

[3] PJM Tariff, Attachment DD, Section 10.

reduced.[4] Each unit's "Peak Capacity" is computed as its ICAP times (1-EFORp) and compared to the unit's average UCAP. On a portfolio basis, each resource owner is assessed a penalty equal to the capacity clearing price times the shortfall of its Peak Capacity below its UCAP. These penalty payments are used to reward capacity suppliers whose resources' Peak Capacity was above UCAP.[5]

- ISO measures availability during periods of actual operating reserves shortages of at least 30 contiguous minutes. A resource with availability below its CSO is assessed a stiff penalty of five percent of the annual capacity payment. For example, if the capacity clearing price is $4/kW-month, each MW unavailable at a resource during a 30-minute Shortage Event would be penalized $4,800 ($2,400/MW-hour).

The PJM approach has been reasonably successful. PJM has not had any absolute generation scarcity situation (recognizing that the extended outages after Superstorm Sandy were primarily transmission related and, in any case, well outside the planning envelope). Equivalent Availability Factors (a standard measure of overall unit availability, including forced, maintenance, and planned outages) has fallen slightly since the start of the RPM design, from an average of 85.2% from 2007 to 2009 to an average of 83.9% during 2011 and 2012. This success has been relatively modest, though, in part because measuring performance over 500 hours dilutes the incentive from the RPM market design to invest in performance enhancement. Furthermore, the PJM IMM also has criticized the exemption of outages deemed outside of the resource's management control (OMC Forced Outages), particularly the exemption for "lack of fuel," which constituted 37% of the OMC Forced Outages and 4.6% of all Forced Outages. In the 2012 State of the Market Report for PJM, the IMM notes that "[l]ack of fuel is especially noteworthy because the lack of fuel reasons are arguably not outside the control of management. Virtually any issue with fuel supply can be addressed by additional expenditures."[6]

The ISO approach has either worked brilliantly or poorly, depending upon the yardstick one applies. As the goal of the severe penalties was to avoid Shortage Events and the attendant reliability risks, FCM has done superbly. There have been no Shortage Events since the start of FCM, which could reasonably be credited, at least in part, to the strong incentives to

4 There are three notable exceptions to this rule: First, a resource that is already paying a penalty for non-delivery of capacity (such as a test failure) is not subject to double penalties. Second, "for single-fueled, natural gas-fired units, a failure to perform during the winter Peak-Hour Period shall be excused … if the Capacity Market Seller … can demonstrate … that the failure was due to non-availability of gas to supply the unit." Attachment DD, Section 10(e). While this clause would clearly exempt a unit if the gas pipeline failed to deliver firm gas, it is unclear whether the failure of the supplier to secure firm gas supply and firm gas transportation would be excused. Third, resources with fewer than 50 service hours during Peak-Hour Periods in a year are evaluated at the lower of their EFORp or EFORd.

5 Incentive payments are capped at the capacity clearing price; any surplus revenue is shared *pro rata* among loads.

6 Monitoring Analytics, *2012 State of the Market Report for PJM*, p. 164.

ensure resources are available when a Shortage Event might occur.[7] If enough resources make themselves available, then no Shortage Event need occur if ISO manages those resources properly. ISO, however, sees the performance differently: it notes that

> Empirical analyses of generating unit performance indicate that, at times of high system stress, a significant share of the region's generating fleet fails to respond to ISO dispatch instructions according to their offered capabilities. … [O]lder units that are relied upon for peaking service, ramping, or reserves are not performing within their offered parameters. These shortcomings became manifest in operational events on June 24, 2010, September 2, 2010, and January 24, 2011 (including a NERC violation related to inadequate generation contingency response on September 2). More generally, an examination of dispatch response performance following the 36 largest system contingency events over the last three years indicates that, on average, New England's non-hydro generating fleet delivered less than 60% of the additional power requested of these resources by the ISO. In sum, at times of greatest need, many resources are delivering far below the performance ability represented in their supply offers.[8]

ISO also promotes its Performance Incentives proposal as a means to address two other substantive issues. First, the region is relying increasingly on natural gas as the primary fuel for the region's generation fleet. The ISO concludes: "As the region's reliance on natural gas expands, greater private investment in hardware, fuel arrangements, or other supplier-selected solutions to ensure resource availability is essential. Changes to the FCM can improve participants' incentives to undertake these investments."[9] Second, the growth of intermittent renewable generation places a premium on the availability and flexibility of the remaining resources able to be dispatched. FERC has addressed the need for "pay for performance" in most operating reserve markets, and changes in design of the capacity market should not be considered a substitute for continued improvements in the operating reserves markets. Nonetheless, because both of these issues involve capital investment (as well as short-term operational changes), there is a rationale for including these incentives within the FCM.

These incentives come at a cost. Although Capacity Resource owners fund the Performance Incentive payments in the first instance, they require a higher capacity payment to offset the expected cost. As the IMM has acknowledged, capacity offers must be allowed to rise by this expected Performance Incentives payout, and implementing this change in bid mitigation rules is essential to the viability of the ISO's proposal. ISO has commissioned The Analysis Group to estimate what this cost increase might be, and their findings are expected shortly

7 The Settlement Agreement provides for loss of capacity compensation for Capacity Resources that fail to perform (subject to certain well-defined excuses) during these Shortage Events. On any critical day, a resource can have its compensation reduced up to 10% of annual FCA Payment if it is not available and, in any month, a resource can lose up to two and a half months of its annual FCA Payment. The availability provisions in the FCM also have a variety of other mechanisms designed to balance, in the specific context of the settlement, the desire for strong economic incentives for Capacity Resources to be available when needed without creating unnecessary payment risk to resources.

8 ISO, "FCM Performance Incentives, October 2012", pp.1, 2., footnotes omitted (henceforth "ISO White Paper").

9 ISO White Paper, p. 2.

but are likely to range between $1 billion and $4 billion in annual capacity cost increases. This White Paper proposes ways of moderating that cost without losing the reliability benefits sought by ISO.

1.4. Summary

This White Paper explores potential changes to the FCM design to address the concerns raised by ISO. ISO's proposal can and should, with limited modifications, bring into greater balance risk and costs. If implemented as currently proposed, the FCM Performance Incentives likely would:

- Provide insufficient incentive for investment in flexibility;
- Create unintended incentives in the Real-Time Market;
- Increase costs and emissions through excessive self-commitment;
- Increase the cost of new entry; and
- Increase costs to consumers needlessly.

As a starting point, though, the ISO's proposal has several merits. Consequently, rather than starting from whole cloth, this White Paper suggests several ways to tailor the proposal to improve its long-run operation. In summary, this White Paper makes the following proposals:

- Retention of daily, monthly, and annual stop-loss levels;
- Stop-loss implementation must support a fully tradable CSO;
- Support for more flexible Supplemental Availability Bilaterals;
- Measurement of portfolio performance;
- Limited penalty exemptions;
- Multi-part capacity supply offers from resources; and
- Implementation of a downward-sloping demand curve for capacity.

2. Current FCM Design and the ISO Proposal

This section reviews the existing performance incentives and the proposed changes, providing our assessment of the limitations of the new proposal in reaching the intended goals.

2.1. Current Performance Incentives

The current ISO market design provides a three-tiered incentive for flexible generation to be available for commitment and dispatch to ISO.

First, prices in the energy market increase during periods when supply is tight; thus, resources that are not generating power or providing reserves miss the opportunity to earn net margins. As noted in the introduction, this energy-market incentive should, in theory, be the core driver of unit operations and capital investment; as the EMM states, the incentive for resources to be available in the operating horizon "is most efficiently provided by the energy and operating reserve prices themselves."[10] Higher prices follow high loads naturally in the LMP markets, because ISO will need to dispatch resources with higher incremental production costs to meet the higher levels of demand, thus raising the price paid to all energy and reserve resources. To augment this natural price increase, ISO has adopted Reserve Constraint Penalty Factors (RCPFs) that are automatically added to reserve prices when the quantity of a reserve product falls below the target level:

- System Thirty-Minute Non-Spinning Reserves RCPF = $850/MWh
- System Ten-Minute Operating Reserves RCPF = $500/MWh
- Local Ten-Minute Operating Reserves RCPF = $250/MWh

Because ISO co-optimizes the reserve and energy markets, these RCPFs also increase energy LMPs. Thus, during extreme conditions, LMPs can top $2,350/MWh.[11] Day-Ahead prices also can rise markedly, reflecting the market's views that Real-Time LMPs can be driven up by scarcity.

The second factor driving high performance in the current market design are the obligations imposed on resources that have been committed in the Locational Forward Reserve Market. Such resources face penalties if they are not available to provide operating reserves (or energy). This incentive structure, however, only applies to the small portion of resources cleared in the Locational Forward Reserve Market.

The third factor is the penalty structure during Shortage Events defined in the current FCM. ISO assesses capacity suppliers penalties or credits for performance during defined Shortage Events, which is any period of 30 or more contiguous minutes of system-wide operating

10 Potomac Economics, "EMM Responses to Questions on Performance Incentives Proposal," February 19, 2013, p. 3.

11 The $1,000 offer cap plus the RCPFs; LMPs can exceed this level because of congestion costs.

reserves shortage or, for an import-constrained capacity zone, when Operating Procedure 7 or Actions 6, 12, or 13 of Operating Procedure 4 are declared based on adequacy and not security.[12] Each Capacity Resource is assigned a score equal to the time-weighted MWs available during the Shortage Event, where the available MWs are defined as the resource's Economic Maximum Limit under three circumstances:

- A resource that is online and following dispatch instructions;
- For an off-line resource following dispatch instructions and with a cold notification time plus cold startup time of 30 minutes or less;
- For an off-line resource following dispatch instructions and (a) a cold notification time plus cold startup time less than or equal to 12 hours and (b) the resource was competitively offered into the energy market but was not committed by the ISO and was consequently unavailable within 30 minutes.

A resource not following dispatch is deemed available at the resource's metered output. Otherwise, the resource's available MWs are set to zero unless it meets one of these exceptions:

- The resource was not committed due to an outage or derate of transmission equipment within the New England Control Area other than equipment owned by the resource owner;[13]
- The resource was denied a self-schedule request by ISO and therefore was not available in the real-time market; or
- A new generating capacity resource has completed construction but is waiting for planned transmission to be put into service.

Availability may be adjusted upward if either:

- The resource owner entered into a Supplemental Availability Bilateral and the contracted supplemental resource had available MWs above the resource's own CSO; or
- The resource was on a partial or full planned outage approved in the ISO's annual maintenance scheduling process, in which case the outage MWs are deemed available.

Additional rules apply for import capacity resources, intermittent power resources, demand resources, and settlement-only resources.

[12] ISO Tariff, Section III.13.7.1.1.1(a) and (b). ISO-NE has proposed and the Participants Committee has approved (over the objections of Generators) that the definition of "Shortage Event" be changed to "whenever the 30-minute RCPF is triggered and remains in effect for 5 minutes." This rule change has not yet been filed with or accepted by FERC.

[13] Except of Wyman 4, Stony Brook, or Maine Independence, where the radial interconnector is deemed part of the transmission system.

Resources then are subject to penalties on unavailable MWs equal to five percent of the resource's annual FCA payment per Shortage Event. Penalties for Shortage Events extending beyond five hours increase by one percentage point per hour. There are three levels of caps on availability penalties:

- In any operating day, penalties cannot exceed 10% of a resource's annual FCA payment;
- In any obligation month, penalties cannot exceed 2.5/12 (approximately 21%) of a resource's annual FCA payment; and
- Over the course of an annual capacity commitment period, penalties cannot exceed FCA revenues less Peak Energy Rent adjustments. Consequently, although it is possible to have a net negative capacity payment in any month, over the course of the year, a Capacity Resource cannot end up owing ISO for having cleared the FCA but may not have recouped any of the costs it incurred to supply capacity during the year.

As the foregoing description suggests, the current Shortage Event mechanism was the product of a lengthy negotiation among ISO, suppliers, and loads. The complexity of the rule was designed to balance risks, costs, and performance.

Suppliers should *not* be expected to second-guess the ISO's commitment or dispatch. As the security coordinator for New England, ISO has the responsibility to ensure system reliability. Furthermore, ISO's security-constrained unit commitment and dispatch are designed to provide the least-cost, reliable solution to a complex problem—a problem that no other entity has sufficient information to compute. The current Shortage Event rules, therefore, do not penalize resources that were offered to ISO but not committed by ISO, as doing otherwise would encourage suppliers to self-commit resources that ISO had not determined to be part of the least-cost solution. Unnecessary commitments increase fuel consumption, air emissions, and costs. Self-commitment also can exacerbate reliability issues by using up scarce fuel-oil supplies to keep resources online, thus reducing the future optionality of ISO to use these resources later should gas supplies become constrained. Of course, ISO's commitment decisions are made under the assumption that committed resources will perform according to their bid operating parameters, so the current Shortage Event rules penalize units not following dispatch.[14]

2.2. ISO's Performance Incentives Proposal

ISO has proposed a plan that replaces the Shortage Event penalties with a high performance payment or penalty, tentatively set at $5,000/MWH, during any five-minute interval in the real-time market in which 30-minute operating reserves (TMOR) or 10-minute non-spinning reserves (TMNSR) are deficient. All resources providing energy or reserves receive this PPR

[14] The rules only penalize *under*-generation, which in a reserve shortage situation is nearly always the problem. One could construct scenarios, though, in which *over*-generation of energy and under-provision of reserves was impairing reliability.

for the MWh or reserves actually delivered; however, all resources with a CSO would be charged or credited this PPR for each MW of their CSOs, scaled down based on actual load and reserve requirements by a Balancing Ratio. Functionally, therefore, the new Performance Incentives would operate similarly to the Peak Energy Rent (PER) deduction, but is necessarily distinct from the PER because the PPR is not included in the posted LMP.

An important difference between the Performance Incentives and the current Shortage Events design is the proposed "no excuses" approach. Both ISO and the EMM defend this approach as consistent with the incentives provided by real-time energy pricing during shortages. Under an idealized energy-only market, units get paid precisely for what they deliver. While one cannot argue with this consistency, we discuss below whether this consistency gives the best outcomes in all cases. We argue that retaining a small subset of the current performance exemptions will reduce the risk and cost of the Performance Incentives program without any effect on incentives for capacity suppliers to invest in reliability.

Another important difference is that the Performance Incentives proposal directly adds costs to the decision to become a Capacity Resource. The current market design is a closed system: penalties during Shortage Events are paid only to other Capacity Resources. This approach, however, artificially values energy from Capacity Resources during a Shortage Event more highly than energy from other resources. ISO's proposal corrects this design flaw, but in so doing makes the Performance Incentives payments assessed to Capacity Resources a truly incremental cost. This cost is not trivial. Although we await The Analysis Group's study of this issue, ISO has released estimates that suggest that this extra financial cost (and risk) of being a capacity supplier will likely add between $2 and $10/kW-month to equilibrium capacity costs, depending on expected system conditions, the PPR level, and the level of perceived risk and the risk tolerance of capacity suppliers.[15] This translates into increased capacity payments (above levels that would otherwise have occurred under the current design) between $800 million and $4 billion. This increase, however, would be moderated by decreases in energy costs due to increased unit self-commitment, lower reliance on expensive intra-day gas, and fewer shortage events. Moreover, new entrants setting the capacity clearing price may perceive an opportunity to earn extra revenues from Performance Incentives payments, thereby causing the long-run equilibrium price for capacity to increase by less than this full cost.

ISO's PI proposal has many strong elements. In particular, the Performance Incentives mechanism should markedly increase the effective LMP during scarcity conditions by setting the effective payment rate for energy to the sum of the LMP and the PPR.[16] In so doing, ISO

[15] This range is calculated from the range of "annual expected hours of system operating reserve deficiencies at criteria" shown in Table 1 of the ISO New England memo, "Operating Reserve Deficiency Information – At Criteria", dated May 29, 2013, and assuming a $5,000/MWh PPR and an average 75% Balancing Ratio during system operating reserve deficiencies.

[16] Even if the Performance Incentives proposal causes some suppression of the LMP due to self-commitment and self-supply, this effect will be more than offset by the addition of the PPR during scarcity conditions.

harnesses market forces to encourage increased production during the scarcity condition and, more importantly, provides a clear financial incentive for capacity suppliers to invest in new technologies or improvements to existing resources so that they will be available to address scarcity conditions. ISO's proposed implementation of this scarcity pricing has two positive elements. First, by incorporating the expected scarcity earnings into the capacity price, it reduces the volatility of future cash streams to suppliers that should, in principle, support cost-effective incremental investments in reliability. Second, it sidesteps a potentially serious problem with how the Peak Energy Rent deduction would play out if the scarcity price were included in LMPs. ISO's Performance Incentives proposal also should have the effect of increasing the slope of the capacity supply curve, which should contribute to reduced year-to-year volatility in capacity prices. (On this point, however, we suggest that a sloped demand curve would further enhance price stability.)

ISO has made a bold proposal to treat all Capacity Resources equally, regardless of whether they are fossil-fueled, renewable, or demand resources. We strongly support ISO on this point. To quote from the labor movement, "equal pay for equal work" is a sound principle.

The IMM also has carefully considered the ramifications of this new Performance Incentives proposal on the level of competitive offers into the FCA and reached generally sound conclusions. The Performance Incentives proposal imposes substantial costs on capacity suppliers, and it follows that these costs must be allowed into their supply offers. The IMM's analytic framework is generally sound and provides the right first-order answer. We propose certain refinements, though, to better align the offer mitigation levels to what a competitive market would bring forward.

For these reasons, we generally support the ISO's Performance Incentive proposal. After careful consideration, though, we offer some observations of potential weaknesses in the proposal and matching remedies.

3. Issues with Performance Incentives Proposal

ISO's proposal raises a series of issues, both by what it includes and what it omits. ISO hews closely to a purist view of the allocation of risks without squarely addressing whether placing (literally) unlimited liability on capacity suppliers is reasonable. Imposing risk on suppliers in a competitive market is not free; every business requires a higher rate of return to carry a higher level of risk. There are many ways in which ISO's proposal imposes risks on suppliers that are unlikely to be a cost-effective means to achieving higher reliability, especially when we recognize that "reliability at any price" is not the goal of sound market design.

Many of the issues with the Performance Incentives arise because the proposal muddies the distinction between *planning* reserves and *operating* reserves. ISO's proposed standard blurs the horizon on which reserves are secured. NERC requires a higher level of reserves in the planning horizon than it does in the operating horizon precisely because system operators prudently plan for a different, broader set of contingencies in the two horizons. The operating reserve margin includes sufficient reserves to respond to operating contingencies on the system, particularly the loss of a major generation or transmission facility. The planning reserve margin, by contrast, needs to be higher to provide a greater likelihood that ISO will have sufficient available resources to commit to meet the energy and reserve requirements of the system, recognizing that at any given time, there will be some level of generation and transmission outages that make some resources in the planning reserve margin unavailable.

Measuring *actual* deployment of a resource to meet operating reserves may materially understate its cost-effective contribution to meeting the planning reserve margin. For example, a higher-cost oil-fired steam unit may be an efficient part of the planning reserve margin, even though in normal operation it would be committed only during periods of peak loads. In the event of a gas interruption or a cold snap (when gas supplies might be short), however, this oil-fired unit could be committed and provide important operating reserves. Indeed, ISO used at least one of the large oil-fired units in this way during January 2013. But because of its high commitment and dispatch cost, this oil-fired unit would generally not be committed and, therefore, would incur substantial availability penalties—even when the unit is available and has ample fuel. Anticipation of these penalties could result in retirement. By blurring the lines between planning reserves and operating reserves, ISO may push valuable resources out of the market, even when their going forward costs are fairly modest and their contribution to fuel diversity and overall system reliability are high. Pushing these resources out of the market is likely to materially increase costs to ratepayers, who ultimately bear the cost of constructing new supply.

Many of the modifications proposed in this paper are intended to address this issue.

3.1. Stop-Loss Provisions

The current Shortage Event penalties are capped on a daily, monthly, and annual basis. Annually, a resource cannot be penalized more than its FCM annual payment, net of PER charges. This approach provided a clear limit to the liability a supplier was taking on a

Capacity Resource, while still providing substantial economic penalties for failure to perform. As noted above, a supplier could lose five percent of the *annual* payment in one hour, 10% in one day, and 21% in one month.

Stop-loss provisions are helpful primarily because the risk of high losses will be priced into capacity supply offers. As discussed in Section 4.3, suppliers will price the potential for a large negative outcome into their offers disproportionately to their probability, owing to the burden and disruption that such negative cash flows have on a business including, but not limited to triggering of debt covenants. Furthermore, a stop-loss eliminates or reduces the need for a Financial Assurance requirement from existing capacity suppliers. Posting credit is costly, and therefore limiting the amount of credit required lowers offer prices in the FCA. Also, a stop-loss provides assurance to new resources that their capacity obligation could not become a liability, which helps in arranging financing on reasonable terms. In our experience, lenders extend terms based on a worst plausible case basis, and the potential for negative FCA revenues will have a sharply negative effect on the ability of entrants to secure project financing.

There are two important downsides, however, to including an annual cap. First, limiting the penalty risk on an annual basis only means that the stiff per-event penalties, which start at five percent of the annual payment, could be exhausted by a string of events in one or two months. For example, poor performance in the summer could leave a Capacity Resource with no additional incentive to arrange for backup fuel supplies in the winter. To address this concern, the current FCM design also includes daily and monthly limits, with the goal of spreading out the limited pot of penalty money over the full year. Second, capping penalties at the net capacity payment could create a "free option." Suppliers might accept a CSO for a resource with very poor availability, hoping to hit a year with few Shortage Events and earn capacity payments even if it rarely was available to support system reliability. To address this issue, the FCM includes a rule specifically to deal with the poorly performing units, eventually flushing them out of the market altogether.[17]

In its October 2012 White Paper, ISO included no cap, or stop-loss, on PPR payments. It justified this omission by reference again to the idealized energy-only market. As stated before, though, there are important differences between an idealized energy-only market and the mechanism ISO has proposed. In an energy-only market, a unit with no availability is paid nothing; its down-side risk is limited to its operating costs. That option exists within the ISO Performance Incentive framework: simply decline to take a CSO. Presumably, however, ISO wants and expects that most supply resources will take a CSO, and so most resources will end up in a position where they have a very limited upside, constrained by the PER Deduction and the net PPR payments, but an almost unlimited downside.[18]

---

[17] ISO Tariff III.13.7.1.1.5.

[18] In the extreme, but very unlikely case that a resource was unavailable at all during a year and the system was perpetually in a TMNSR shortage condition, the PPR and PER charges to a Capacity Resource would top $40.7 million per MW of CSO.

Realistically, however, corporate risk managers will not accept an unlimited liability obligation at any price. Even though most generating assets are held in station-specific Limited Liability Corporations (LLCs) designed to cap the corporate parent's exposure to PPR payments, the prospect that a major equipment failure could bankrupt the LLC because of its CSOs will be a non-starter for many market participants. It seems unlikely that ISO will be calling in the repo man to seize the physical assets, so ISO will need to establish Financial Assurance requirements for capacity suppliers that will, in effect, establish a stop-loss rate for the PPR payments independent of the value of the underlying Capacity Resource.

Since it is clear that a stop-loss will exist, either by design or default (literally), we should give attention to this important risk mitigation tool up front to ensure that the overall design works with the stop-loss.

3.1.1. Annual Stop-Loss

In designing a stop-loss, there are three major questions to address:

- To what metric is the annual stop-loss linked?
- What multiplier is applied to that metric?
- How is the stop-loss implemented?

Reference Price for Stop-Loss

There are two points of view as to what the appropriate reference point for the annual stop-loss metric should be. While this debate is worthwhile, it is more important that an effective stop-loss mechanism be implemented than that it start at one number or another.

If the decision were to be based entirely on economic theory, the capacity clearing price would be an appropriate reference point for the stop-loss. The capacity clearing price was set where the supply curve intersects the (currently vertical) demand curve. Thus, at least at the time of the FCA, the cost of covering a capacity shortfall would have been some, probably small, increment above the FCA clearing price. ISO had the option, at this price, to increase reliability by securing additional Capacity Resources, yet did not do so, indicating that this price is the revealed value of incremental reliability. This logic suggests that the FCA clearing price should be the foundation for the stop-loss, rather than the FCA starting price or some measure of the Cost of New Entry (CONE). Doing so ties the penalties for reduced capacity to both the opportunity cost of capacity and the revealed value of incremental capacity.

ISO and some parties favor the use of some fixed stop-loss value, principally because (a) it allows suppliers to build a known risk level into their offers, and (b) it ensures that the penalty will be high enough to provide the desired performance incentives. Both of these issues can be addressed by other means, discussed below. While a fixed value has the apparent virtue of simplicity, it lacks any basis in economics and creates a material inequity between suppliers. The starting price of the descending clock auction is entirely arbitrary—in principle, it could start at any high number. The current value of $15/kW-month was chosen in the Settlement Agreement as twice the deemed value of CONE, but ISO intentionally unlinked this starting value from any economic meaning in a subsequent tariff update. To vest that $15/kW-month with economic meaning now reverses that unlinking.

More importantly, though, using a single fixed number creates a substantial inequity among Capacity Resources that clear at different price points. Consider the outcomes of the most recent FCA, in which new Capacity Resources in NEMA cleared at $14.999/kW-month, existing Capacity Resources in NEMA will be paid $6.661/kW-month, Connecticut Capacity Resources will be paid $2.883/kW-month, and all other Capacity Resources will be paid $2.774/kW-month.[19] These resources, therefore, face very different upside earnings potential, so it would be logical to assume that they would face symmetric downside potential, as well. Furthermore, the high prices in NEMA reflect resource scarcity in that capacity zone, so underperformance would be likely to have a more serious reliability impact. Both of these considerations suggest that the stop-loss rate should be linked to the payment rate, rather than choosing some arbitrary fixed value.

That said, a potential weakness of linking the stop-loss limits to the capacity payment rate is that there is a potential for that rate to be very low and so, in turn, the incentive for Capacity Resources to perform would be small. This can be remedied simply by inserting a floor value for the stop-loss rate. In previous filings, ISO has defended $1.00/kW-month as a reasonable estimate of the marginal going-forward costs of existing resources. To reflect the need for some amount of at-risk compensation, we propose a floor, for example, of 150% of this value, or $1.50/kW-month. As ISO has correctly noted, however, capacity clearing prices should be much more robust than this cost-based value to reflect foregone Performance Incentives earnings, so in practice, this floor is unlikely to bind.

Proponents of a single fixed price support this view for two reasons: simplicity for bidders and sufficiency of the penalty in practice. Setting a floor on the stop-loss limit addresses the second of these concerns. Moreover, if a sloped demand curve for capacity is adopted, then a low capacity price occurs only when there is a higher-than-required planning reserve margin. Scarcity conditions should, therefore, be infrequent and the need for suppliers to take expensive steps to secure high availability is appropriately reduced. As to the issue of simplicity, bidding into the FCA requires a fairly high level of sophistication already, and the Performance Incentives proposal further adds to the complexity. Capacity suppliers are sophisticated bidders with extensive analytic capability, however, so adding one more element into their decision-making process is not unreasonable. In the descending clock auction format, suppliers can update their offer price based on the current auction price and, therefore, maximum penalty rate. Static Delist bids can be parameterized based on the current auction price, as well, so the IMM can approve a bid structure based on the current FCA round's price, rather than a fixed value. While this adds some minor complexity, it creates substantial value by resolving the potential inequities of using the same penalty cap for suppliers receiving markedly different payment rates.

19 ISO New England, "Forward Capacity Market (FCA 7) Result Report," Table 3, available at http://www.iso-ne.com/markets/othrmkts_data/fcm/cal_results/ccp17/fca17/fca_7_result_report.pdf.

Multiplier for Annual Stop-Loss

The next question is, what multiplier should be applied to the reference price to set the annual stop loss-level? Too low a level leaves the "free option" problem largely unaddressed. Even if a capacity supplier has little expectation of actually performing during scarcity hours, if there is no premium atop the capacity price, then there is relatively little risk for that low-quality resource to taking on a CSO. Too high a level imposes a great deal of risk with little added incentive for change. Where is the "Goldilocks" solution?

In addressing this question, it is important to bear in mind the ultimate goal of the FCM: to secure reliable capacity, supporting cost-effective investment in new and existing resources. Imposing large penalties is a means to an end, not the end itself. The goal of the program should be to foster investment in improved reliability, not to issue massive penalties. Of course, without penalties, there will be no incentives. The question is, how much risk is needed to provide that incentive?

Given information published so far by ISO, when the system is at criteria (i.e., the installed capacity margin equals the planning reserve requirement) it appears that capacity prices will need to increase by approximately $5/kW-month ($60/kW-year) to cover the expected Performance Incentive lost opportunity costs.[20] For a typical 550-MW combined cycle plant, that's approximately $30 million in annual capacity payments.[21] To put this amount in perspective, we sought to compare this "at risk" amount to the operating profits earned by typical independent power producers. As a benchmark, we reviewed a number of independent power producers' (IPPs') financial data and selected that of GenOn Energy, which was recently acquired by NRG Energy, as a reasonable proxy for other IPPs. Unlike many other IPPs, GenOn was a "pure play" generation company with substantial assets in the Northeast markets, concentrated primarily on fossil-fueled plants. GenOn Energy's fleet-wide US operating income was $209 million in 2011, and the company had an operating loss of $136 million in 2012, on 4,680 MW. Scaling these figures to a 550 MW plant gives $25 million of operating profits in 2011 and $16 million of operating losses in 2012. In this context, putting $30 million of capacity revenues at risk is already a major swing in the operating income, i.e. $5/kW-month.

Given this large fraction of income at risk, the FCM should have an annual stop-loss above the annual FCA capacity payments. PJM caps penalties for non-delivery at 120% of the capacity rate, in response to stakeholder feedback that a 200% cap created too much risk. There has been some sentiment that this cap is too low to adequately address the "free option" issue, but the PJM RPM design includes many excused outages, and capacity suppliers are not foregoing peak energy rents and scarcity payments, as well. In light of

[20] A $5/kW-month increase is consistent with 16 scarcity hours at an average Balancing Ratio of 75%, which is generally consistent with numbers published by ISO to date. See, e.g. footnote 15.

[21] Taking the extreme case where, but for the Performance Incentives, the capacity price would have been approximately zero.

these differences, an annual stop-loss rate of 120% of a Capacity Resource's annual capacity payment rate (or $18/kW-year, whichever is higher) is a reasonable balance of interests.

Implementation of Stop-Loss

ISO has also raised the possibility of including PPR *payments* as part of the funds subject to clawback through PPR penalties. This proposal creates perverse incentives. First, as noted above, carrying forward positive balances creates unacceptable accounting issues for corporations. Second, it has the curious property of increasing the maximum penalty for resources that have actually performed. Third, it markedly complicates the ability to choose a rational delist offer price, because the definitive stop-loss cap is replaced by one that depends on the timing and mix of performance during the year.

The stop-loss cap should be set against the capacity price alone, not the sum of capacity price and PPR payments made. The capacity price is, in effect, a prepayment of the expected level of PPR payments during scarcity conditions. Capacity suppliers face a risk that the number of scarcity hours will exceed the expected value (H) imbedded in the capacity price—when this occurs, a normally-performing resource would have been better off not accepting a CSO. ISO's proposal of adding actual PPR payments to the capacity price to determine the penalty cap exacerbates this risk. When there are many scarcity hours, not only are Capacity Resources losing (relative to not having taken a CSO), but also they are exposed to higher penalty caps. Compounding the risk in this way will lead to higher offers in the FCA and higher capacity charges for consumers.

3.1.2. The Need for Daily and Monthly Stop-Loss

Although ISO has acknowledged the need for some annual stop-loss in its Performance Incentives proposal, it has resisted nested stop-loss limits, such as those included in the current FCM design that caps daily and monthly losses. ISO's concern is misplaced. The Performance Incentives should be focused on providing *forward-looking* incentives to enhance availability across the whole year. This goal is not accomplished by inflicting massive penalties in a short time. Instead, potential losses should be calibrated to the costs of economically efficient actions that could reasonably be taken to improve performance.

As noted above, a concern raised by introducing an annual stop-loss is that severe conditions in the summer could burn through the stop-loss for some critical units, which would subsequently have muted incentives to perform in real time. This is not to say that there will be no incentives, or that the FCM Performance Incentives will have failed. First, resources have strong incentives to earn high LMPs during shortage periods, independent of any PPR payments. Second, anticipating the potential for high PPR payments, the unit owner may have invested to make the unit more reliable by, for example, securing firm gas, installing dual-fuel capability, or making engineering changes to the unit to provide fast start times or more flexibility in ramping. Nonetheless, if a unit has run through to its stop-less level in the summer, it will have a lower incentive at the margin to perform in the winter than it would have without the annual stop-loss.

One way to address this problem is to follow the lead of the current Shortage Event penalties and set sub-annual stop-loss limits. A daily stop-loss is appropriate, inasmuch as there is only so much that can be done intra-day to solve a performance problem—charging a unit on an ISO-approved maintenance outage, for example, with a massive PPR penalty has no practical incentive effect that a substantial but limited charge would not accomplish. Burning through a large portion of the annual stop-loss in a single day, however, does risk "running out of bullets" down the line. It is better, therefore, to limit the daily PPR charge at the current level of 5% to 10% of the annual FCA payment.

Many of ISO's expressed reliability concerns are about winter performance, but the capacity year starts in June. To ensure that poor performance during an unusual summer does not exhaust the annual stop-loss, there needs to be some form of monthly or seasonal stop-loss. The current Shortage Event penalty is capped monthly at 2.5 times the monthly FCA payment, a level that was thoughtfully chosen by ISO, the IMM, and market participants to balance the in-month performance incentives against the goal of reserving some portion of the annual cap for winter and shoulder seasons.

This monthly 2.5x cap provides strong incentives to undertake appropriate investments. Referring back to the hypothetical 550-MW combined cycle plant, its monthly cap would be about $6.5 million, placing $13 million at risk just in January and February, the two winter months with the greatest gas delivery issues historically. Fuel burn at such a unit, with a typical 50% capacity factor, would be about 1.4 million MMBtu. The capacity payment at risk, therefore, provides almost $10/MMBtu of incentives to secure gas supplies, which is well in excess of the firm gas transportation rate historically. Thus, the 2.5x monthly cap provides plenty of headroom for incentives without adding undue and costly risk.

3.1.3. Empirical Analysis of Stop-Loss Limits

To explore the practical effect of different stop-loss levels, we examined how representative units from four different classes of generating resources might be affected by the Performance Incentive proposal to assess how different stop-loss levels would affect the risk to a set of typical New England units, on their incentive to invest in performance enhancements, and the impact on the competitive offer price (which informs how much more consumers would pay were one of these units on the margin of the FCA).

In these simulations, we performed a Monte Carlo simulation of unit availability and scarcity conditions. The monthly probabilities were set, for unit availability, from historical operating data for certain NextEra units, which are broadly representative of each of the four major unit types in New England: baseload (including nuclear and other resources typically on dispatch because of low incremental costs), short-notice mid-merit resources (including non-baseloaded combined cycle units), long-notice mid-merit resources (including gas-oil steam), and quick-start peakers. For scarcity conditions, we used the frequency data from ISO's Performance Incentives presentations to examine the impact on a typical summer, winter, and shoulder month. Scarcity conditions and unit outages are modeled as independent events, which tends to understate the risk, especially for large units in import-constrained areas.

These simulations show that, on a prospective basis, a monthly stop-loss has a very modest dampening of the expected penalty and, consequently, a very modest change in the incentive to avoid penalties. The following table summarizes (risk-neutral) expected value of net Performance Incentive payments by unit type, assuming a $5/kW-month capacity price. Note, however, that rational capacity suppliers might be expected to hedge their exposure during planned outages or shoulder seasons, during which their Capacity Resources are not available should a scarcity event occur. The hedge premiums likely would increase the expected cost for all categories of units, but particularly baseload plants, such as nuclear units, that have long scheduled maintenance outages.

Table 1: Expected Value of Simulated Penalty/Bonus Payments under Monthly Stop-Loss

	1x	2x	3x	No Stop-Loss
Long-Notice Mid-Merit	($1.66)	($2.31)	($2.61)	($2.82)
Short Notice Mid-Merit	($0.80)	($1.06)	($1.11)	($1.12)
Peaker	$0.13	($0.02)	($0.06)	($0.08)
Baseload	$0.47	$0.45	$0.45	$0.45

As Table 1 indicates, a 1x monthly stop-loss does markedly reduce the expected penalty rate for nearly all classes of units. A monthly stop-loss between 2x and 3x, however, yields an expected penalty very close to that without any stop-loss. What the stop-loss does, however, is three-fold. First, it truncates the very infrequent extreme outliers, thereby reducing risk of bankruptcy and cost of high financial assurance. Second, it preserves more of the annual cap to provide incentives in subsequent months. Third, it lowers the equilibrium bid and, consequently, the expected clearing price of the FCA.

3.1.4. Finality of Stop-Loss Monthly Settlements

ISO has shown some willingness to consider stop-loss, but in a form that is unworkable from a commercial perspective. What ISO has discussed is a cumulative cap: Penalties in June would be capped at some multiple of the June payments—including both capacity payments and PPR earnings. July penalties would be capped at a multiple of capacity payments and PPR earnings in *both* June and July, less any penalties paid in June, and so on. The result would be that all payments under the FCM tariff—including both capacity payments and PPR earnings—would be at risk until May 31 of the delivery year.

This rolling approach is unworkable for two reasons:

- **Corporate accounting.** Under generally accepted accounting principles, the capacity supplier cannot recognize any of the revenue it earns, either from capacity payments or PPR, until the end of the delivery year. Moreover, auditors may require not only that all of these earnings be reserved, but further that the potential penalty multiplier against those earnings be reserved. Given that capacity revenues range anywhere from one-third to nearly all of some resource's revenues, the consequences of an open-ended clawback of revenues would be seriously disruptive to suppliers.

- **Fungibility of CSOs**. The market is best served by a robust secondary market for trading CSOs within the delivery year. Such a market encourages capacity suppliers who anticipate that a resource may be unavailable (for example, during scheduled maintenance or refueling) to enter into a Supplemental Availability Bilateral. ISO's rolling approach to penalties, however, disrupts the sound functioning of a secondary market. Such a market should operate to replace units that are *prospectively* most likely to be unavailable. Instead, the suppliers who will value these contracts most are those who had the best performance *retrospectively* (and hence have the highest amount at risk). Ironically, resources with a very poor track record would have relatively less incentive to find replacement resources, even if they anticipated continued poor performance. Each CSO, therefore, accumulates its own idiosyncratic value, destroying the price transparency and fungibility essential to a well-functioning secondary market.

It is essential, therefore, that each monthly settlement be final. This finality allows suppliers to recognize revenues (or penalties) in an orderly way, to access an economically sound secondary market for replacement capacity, and to ensure equitable payments to composite resources.

3.2. Limited Excused Unavailability

The current Shortage Hour penalty design thoughtfully considered a broad range of reasons why a resource might not be producing energy or providing reserves in a given moment, and attempted to weigh the question of whether each of these situations should be subject to financial penalties. The thought process that went into these decisions focused on the key question: What cost-effective actions could the resource owner have undertaken to be available? Imposing penalties in other situations adds more risk (and, therefore, costs to consumers) but either limited or overly costly actions to improve reliability.

ISO's proposal simply sweeps this question aside, placing all risks on the income statements of capacity suppliers. Some risks clearly belong there: failure to procure fuel supplies, forced outages, unscheduled maintenance outages, labor strikes, etc. There is another end of the spectrum, however, where there is no conceivable action by a supplier and, consequently, where the penalty is simply a cost with no incentive benefit. ISO clearly believes that the evidence over the first years of operation under FCM justifies redrawing the balance of risk, but redrawing the line is not the same as ISO proposes. FCM is designed to procure *planning reserves* on a different horizon than *operating reserves*; the reason the planning reserve margin is higher than the expected peak load plus required reserves is precisely to accommodate a normal level of resource outages. The high level of the PPR proposed by ISO, however, is only supportable by an appeal to a Value of Lost Load. There is a disconnect here when this charge is applied to units unavailable for precisely the sort of reasons we hold a higher planning reserve margin.

At the top of that list are transmission outages. Expecting suppliers to reinforce the high-voltage transmission system is highly cost-ineffective. In PJM, about one percent of forced outages were related to transmission outages (and, therefore, are excused under the RPM rules).[22] Transmission constraints caused by transmission outages or deratings likely also contribute to failure to commit a unit or to dispatch it fully. It is neither the business nor the responsibility of capacity suppliers to maintain the transmission grid, so it is unclear what reliability benefit charging penalties related to transmission outages creates. Moreover, transmission outages will disproportionately affect resources in import-constrained load pockets; penalizing these units for transmission outages may create the perverse incentive for strategically located Capacity Resources to be pushed out of the capacity market by other resources not facing this locational risk, even though the Capacity Resources inside the load pocket are more effective at supporting system and local reliability.

Also in this exempt category should be hurricanes, floods, tornados, and other natural disasters. These accounted for six percent of forced outages in PJM during 2012, which by all accounts, was one of the worst years on record, with Superstorm Sandy as the centerpiece. While disaster preparedness is an important reliability step, the devastation and expense from a major natural disaster are bad enough without subsequently imposing massive fines for non-performance.

A third area to examine is the treatment of a resource that was fully available but not committed by ISO. ISO has stated that generators should simply self-commit going into days where a scarcity condition may arise, but this is a weak claim. ISO is in a unique position to judge the likely need for resources, based on the information that only ISO has regarding the operating status of each resource, the condition of the transmission system, integrated load forecasts, and scheduled imports and exports. Capacity suppliers are in a poor position to second-guess ISO's commitment decisions. Moreover, ISO's security constrained commitment and dispatch are intended to be the least-cost means to meet load reliably; by definition, then, self-commitment raises costs for uncertain benefits. Self-commitment wastefully burns fuel, which may be in scarce supply in the winter, thus increasing the risk of a system-level scarcity condition. Self-commitment can cause its own reliability issues by creating over-generation in off-peak periods; consequently, ISO approval is required to self-commit. If a unit sought self-commitment, was denied, and was then unavailable during a scarcity condition, what incentive is the resulting penalty intended to provide? Therefore, at a minimum, units denied self-commitment and, as a result, are unavailable during a scarcity condition should be exempt.

The current Shortage Event definition exempts a broader class of uncommitted units, namely those with notification and cold start times of 12 hours or less, but which were not committed by ISO for economics. The argument for exempting these units is that they did precisely what

[22] Monitoring Analytics, *2012 State of the Market Report for PJM*, Table 4-13, p. 164.

ISO requested of them. On the other hand, ISO's commitment decisions are influenced by unit characteristics, and those characteristics are under the control of the plant owner—perhaps only to a small degree for existing resources, but to a large degree for new resources. For this reason, the exception for uncommitted units in the current design is probably overbroad to achieve the long-run transformation of the New England resource mix.

Rather than having to draw a bright line, a possible approach to address this issue would be to ratchet up the PPR over the course of a scarcity condition created by a contingency. In the first intervals of such a scarcity condition, there have been no NERC criteria violations. A relatively low PPR is therefore appropriate. If the scarcity condition continues beyond 30 minutes, the PPR should step up, and again at each half hour until the full PPR value is reached. (If instead the PPR was variable, ramping up as scarcity conditions developed, the potential for a reserve shortage would already have been visible in the PPR, except for very large contingencies, thus addressing the issue of providing advance notice for faster-starting units to commit.) Combining this approach with a straight exemption for units unavailable or on reduced dispatch due to transmission outages or derations, acts of God, or denied self-commitment, would produce a balanced approach to risk and incentives.

3.3. Tools for Supplier Risk Mitigation

Under ISO's proposal, capacity suppliers will bear far more risk than they do under the current FCM design. Risk is a two-edged sword: on the one hand, increasing risk on suppliers may elicit changes in investment and operations that will ultimately provide greater reliability at lower cost; on the other hand, suppliers will rationally seek to recover costs of risk management (for those risks they can manage) or enhanced returns to compensate (for those risks that they cannot manage). Consequently, simply dumping more risks on capacity suppliers is not always a good thing: shifting risks to capacity suppliers that they cannot manage does not increase system reliability, but it will increase capacity costs paid by customers.

By hewing closely to an economically pure design, ISO's proposal lacks important features to adapt sound theory into reasonable commercial practice. ISO relies on the concept that in an energy-only market with very high scarcity pricing, suppliers would not be insulated from any risk. As a practical matter, however, there are commercial issues raised by the particular circumstances of the region. No one supposes, in an energy-only market, that all transactions occur in the spot market. Instead, when faced with the potential for high and volatile pricing, both loads and suppliers will seek to enter into contracts that provide a mutually beneficial level of risk mitigation. In New England, however, state law largely forbids long-term contracting by investor-owned utilities, even though these utilities serve the majority of customers. The year-to-year, or short multi-year, contracts that these utilities use to serve default customers simply do not offer the same risk management tools that are typical in more developed retail markets, such as ERCOT or England.

Given this fact of incomplete contracting in New England, more risk mitigation tools should be included in the Performance Incentive design. Carefully chosen, these tools will reduce the total level of risk that will be borne by capacity suppliers (and, consequently, priced into their capacity offers) under the Performance Incentives program with minimal reduction in the incentives to provide flexible, reliable capacity. ISO should (1) enhance the functionality and flexibility of Supplemental Availability Bilaterals, (2) measure a capacity supplier's performance across the whole of its portfolio of Capacity Resources, and (3) facilitate trading and transparent pricing of Supplemental Availability Bilaterals through a bulletin board. These measures are discussed in the following three subsections.

3.3.1. Enhanced Supplemental Availability Bilaterals

One of the most significant changes proposed by ISO is the elimination of all exemptions from the PPR charges or credits. For example, generation resources are currently excused from availability charges during scheduled annual maintenance. The logic supporting this exemption is that system reliability is well served by generation units taking reasonable maintenance (and, for nuclear units, refueling) outages, and that annual maintenance scheduled with ISO approval is not a risk to reliability, because sound maintenance scheduling ensures that there are enough other Capacity Resources to meet expected peak loads and reserves even absent the units under maintenance.

Although the risk of scarcity conditions should be low during scheduled maintenance periods, few capacity suppliers would want to be fully exposed to PPR payments while their generator is out of service. A rational approach to hedge this risk is through a Supplemental Availability Bilateral with another Capacity Resource. Certain rules in the current ISO Tariff, however, limit the flexibility of the Supplemental Availability Bilateral as a hedging vehicle, given the broader responsibility that ISO is placing on capacity suppliers.

Foremost, ISO should abandon the limitation that the Supplemental Capacity Resource be in the same Reserve Zone as the Supplemented Capacity Resource, or be nested within that Reserve Zone. Tariff Section III.13.5.3.2. If ISO is handing responsibility for managing availability over to capacity suppliers and eliminating all exemptions, including those when ISO makes errors in load forecasting and unit commitment, it cannot simultaneously retain authority to second-guess a capacity supplier entering into a capacity hedge. However, in settling the bilateral, ISO should recognize that the Supplemental Capacity Resource may be an approximate (or, to use a term of art, a "dirty") hedge for the Supplemented Capacity Resource. For example, if a unit in Maine is supplementing the capacity of a unit in Northeast Massachusetts (NEMA) that is on a maintenance outage, and then if there is a scarcity condition in NEMA only, then the NEMA unit is exposed to the full PPR charge because the Maine unit will not have earned any PPR credits during that scarcity condition. This exposure will lead most resource owners to seek first to contract with nearby supplemental resources, but in Resource Zones with a very high ownership concentration (like NEMA), that perfect hedge product may not be available at a competitive price. Allowing hedges from other Reserve Zones creates competitive pricing pressure in concentrated, import-constrained zones. While ISO would forego the must-offer requirement in the energy market from an in-zone resource, ISO doesn't have that today during scheduled maintenance or other excused outages: resources don't arrange Supplemental Availability Bilaterals in those cases.

ISO should also reconcile the difference in the designation period between Section III.13.5.3.1.2 (specifying that the contracts “shall be in Operating Day increments, no less than one Operating Day) and Section III.13.5.3.2.2 (specifying “the term of the transaction, which shall be in hourly increments coinciding with hourly boundaries, no less than one hour”) in favor of the more flexible, hourly designation. Generators taking a weekend maintenance outage may rationally want to start the outage on Friday evening and restart the unit on Monday morning and, therefore, would prefer a hedge contract to match.

Finally, ISO should eliminate the hourly review of Section III.13.5.3.2.3. In line with the new philosophy of “pay for performance,” the Supplemental Capacity Resource will be paid according to its actual performance, but that performance will be credited to the Supplemented Capacity Resource instead. There is, therefore, no need for an hourly review of unit status, conformance with dispatch instructions, or other matters that do not affect the PPR payments.

These changes to the Supplemental Availability Bilaterals are important because the parties cannot match the same contract without ISO. A side-contract between two capacity suppliers for one to cover the PPR risk of the other will necessarily be a financial contract. The PPR payments, however, are subject to *pro rata* reductions in the event that sufficient revenues are not collected to fund those payments. By contrast, a Supplement Availability Bilateral allows the performance of the Supplemental Capacity Resource to be transferred to the Supplementing Capacity Resource (provided, as noted above, that the Supplemental Capacity Resource was able to respond to the scarcity condition), thereby removing the risk of underfunding.

3.3.2. Measurement of Portfolio Performance

Another important risk management tool in today’s FCM design is the evaluation of a capacity supplier on a portfolio basis. This method of evaluating performance during Shortage Events is critical under the current plan because of the lack of symmetry between penalties and rewards: although penalties are collected only from under-performing units, payments are spread among units actually generating as well as units exempted from performing. As a consequence, the payment per MW would be less than the penalty per MW.

Under the Performance Incentive plan proposed by ISO, this problem would still be present. Although in theory there is a balance of payments on a MWh basis, in practice there could be a disconnect under certain conditions. These conditions are precisely the same as those that would motivate a supplier to arrange a Supplemental Availability Bilateral. For example, the Balancing Ratio should be capped at 1.0, i.e. capacity suppliers should not be expected to deliver more capacity than their CSO. If this limit is binding because of high internal loads or high exports, the energy and reserves from performing resources will necessarily exceed the under-deliveries from capacity suppliers. Because the net settlement must equal zero, the price per MWh paid to performing resources will be less than the per-MWh charge to suppliers.[23]

To help manage this risk for capacity suppliers, ISO should continue its current practice of determining MWh shortfalls on a portfolio basis. That is, a MWh of over performance at one resource should offset a MWh of underperformance at another in the same portfolio. This, in effect, would create an automatic Supplemental Availability Bilateral among each of a supplier's resources, side-stepping the fiction of having a market participant contracting with itself.

3.3.3. Facilitating Supplemental Availability Bilaterals

With greater risks created by the Performance Incentives, we expect to see greater activity in the Supplemental Availability Bilateral market. For example, units on maintenance outages will almost surely want to cover their position, and other resources should be in a position to hedge that risk because the Balancing Ratio is likely to be low during non-peak months when maintenance is typically scheduled. Supporting a robust secondary market for these bilaterals is in ISO's interest, as it enhances the incentives of resources that will be available for dispatch to perform.

One tangible way that ISO can help support this market is to expand upon its bulletin board for posting offers and bids for Supplemental Availability. Compared to the current market design, we expect there to be a large number of potential buyers and sellers. Matching needs and availability will be challenging for many market participants, particularly those without sophisticated trading operations. A bulletin board is a simple, low-cost way of facilitating these transactions. The current bulletin board could serve as a useful starting point for a more complete trading platform.

[23] Capping the Balancing Ratio at 1.0 is consistent with the implementation of a parallel concept, the Peak Energy Rent Scaling Factor, as defined in the current Tariff, Section III.13.7.2.7.1.1.1(a). The alternative, allowing the Balancing Ration to exceed 1.0, would lead to the bizarre outcome that a Capacity Supplier is paying an availability penalty even when it has fully met its CSO. Allowing the BR to exceed 1.0 fundamentally changes the unit of capacity that the ISO is procuring in the FCA from a fixed MW amount to a "slice of system," but without any change to the compensation to the capacity supplier.

4. Capacity Offer Monitoring and Mitigation

4.1. Profit-Maximizing Offer Under Risk Neutrality

Market monitoring and offer mitigation will be critical elements in the Performance Incentive design, either as proposed by ISO or as modified along the lines discussed in Section 3. In a short-run capacity market with few in-period availability penalties, such as the New York market, market monitors have used cash going forward costs as the basis for offer mitigation. Under the Performance Incentive plan, however, these going forward costs are only one component of the costs of taking on a CSO. A more substantive cost—and one that is much harder to quantify before the fact—is the foregone opportunity to earn high Performance Payments during scarcity conditions had the resource not taken a CSO.

ISO and the IMM have acknowledged this shift and have proposed to modify the Dynamic Delist threshold upward from $1/kW-month to a value reflecting the expected break-even offer price.[24] ISO presents an elegant derivation of this break-even offer price, showing that this price is independent of a unit's expected performance, but equal to the product of (a) the PPR, (b) the expected average Balancing Ratio during scarcity conditions (Br), and (c) the expected hours of scarcity conditions (H).[25] This product, PPR×Br×H, is the profit-maximizing offer for a risk-neutral capacity supplier whose going forward costs are less than its expected market earnings.[26]

It is initially surprising that this profit-maximizing offer is independent of the resource's expected performance. To understand why this is true, it is useful to think of the Performance Incentive as three distinct cash streams:

1. A stream of Performance Incentives payments, where ISO pays performing resources at the PPR based on the resource's actual performance during scarcity conditions;
2. A stream of Performance Incentives charges, which ISO collects from Capacity Resources in proportion to their CSO and the Balancing Ratio; and
3. A stream of capacity payments paid to Capacity Resources.

Note that *all* resources receive the first stream of payments (depending, of course, on their actual performance). Neither the second nor the third streams depend on the measured performance of the Capacity Resource. Thus, the capacity payment has to cover the stream of Performance Incentives charges, which is independent of the resource's performance.

---

[24] Capacity sellers offering resources at or below the Dynamic Delist threshold are exempt from offer review or mitigation.

[25] ISO, Presentation No. 4, April 10, 2013, slides 8–18.

[26] This derivation is true only if there are no additional direct costs of taking on a CSO. This assumption is unlikely to be true; at a minimum, capacity suppliers will likely need to post some sort of financial assurance in light of their obligation to pay Performance Incentive charges in the future, the cost of which would appropriately be included in the profit-maximizing capacity offer.

A supplier's evaluation of whether it will cover its going forward costs, however, does depend on its resource's expected future performance. For a risk-neutral existing supplier, the profit-maximizing capacity offer is the sum of the product shown above (PPR×Br×H) plus the difference between the going forward costs and the expected actual market earnings of the resource, which is the sum of the net margins in the energy and reserves markets plus net Performance Incentive payments PPR×A×H, where "A" is the expected performance of the resource during scarcity conditions, expressed in MW.[27] Note that this last product, PPR×A×H, may be negative if "A" is less than "Br". Thus, a unit with low availability will need to increase its capacity offer price to reflect the potential cost of availability penalties as well as its net going forward costs.

This consideration of unit performance exists today, of course: even a "deep in the money" resource like a nuclear station could fail to cover its going forward costs if the plant were unable to operate because of, say, a relicensing dispute. There is an extra challenge added by the addition of the Performance Incentives, however, because the performance in a relatively small number of hours can make an enormous difference in the resource's earnings. It cannot even be assured that earnings will be higher than they would be absent the Performance Incentives: to the extent that other resources perform better than they would have, or self-commit and depress energy prices, it is possible to earn less.

4.2. Risk-Aversion in Capacity Suppliers

The derivations above hinge on the assumption that suppliers are risk-neutral. While this is the typical assumption in industrial organization textbooks, the positive risk-return relation derived from financial portfolio theory is applicable here.[28] Here, "risk" is simply the broad term reflecting future uncertainty of financial outcomes. Modern portfolio theory holds that, when comparing two courses of action, the one with higher risk requires a higher expected return. That is, risk carries a price, contrary to the risk neutrality assumption.

Mathematical variance, however, is not the only factor at play. The concept of risk and return in modern portfolio theory is entirely driven by expected values, which is, for our purposes, flawed in two important ways.

First, power plant operators have a more structured understanding of their operational risks than just a set of probabilities:

> "Options theory and modern portfolio theory have at least one important conceptual difference from the probabilistic risk assessment [PRA] done by nuclear power [plants]. A PRA is what economists would call a structural model. The components of a system and their relationships are modeled in Monte Carlo simulations. If valve X fails, it causes a loss of back pressure on pump Y, causing a drop in flow to vessel Z, and so on.

[27] In ISO's presentation, "going forward costs" already excludes net margins. Other costs related to participation in the capacity market would also be included; see footnote 5.

[28] *See*, e.g. Markowitz, H.M. (March 1952). "Portfolio Selection". *The Journal of Finance* **7** (1): 77–91.

> But in the Black–Scholes equation and modern portfolio theory, there is no attempt to explain an underlying structure to price changes. Various outcomes are simply given probabilities. And, unlike the PRA, if there is no history of a particular system-level event like a liquidity crisis, there is no way to compute the odds of it. If nuclear engineers ran risk management this way, they would never be able to compute the odds of a meltdown at a particular plant until several similar events occurred in the same reactor design."[29]

In this probability risk assessment framework, power plant operators will model outage scenarios that could result in very long outages, rather than assuming a neat distribution around a mean. Furthermore, generators will take into account the correlation of the unavailability of their plant and the risk of scarcity conditions. Especially for very large stations and large units, the risk of correlation between one's outage and scarcity conditions may be substantial. It is not sufficient, therefore, merely to base an offer on a unit's historical operations and a historical measure of scarcity conditions. Failing to consider the correlation of these terms will lead to an overly optimistic view on future outcomes.

The second flaw with a simple risk-return model is that doing so treats outcomes symmetrically. Empirical studies of corporate behavior suggests otherwise.

> Studies of managers indicate the performance and aspiration constructs found in the behavioral theory of the firm are central to managers' concepts of risk. Mao (1970) found executives characterized risk in terms of failure to meet a target rather than in terms of variance. March and Shapira (1987) reported that 80 percent of the executives they surveyed considered only negative outcomes when thinking about risk. Baird and Thomas's (1990) survey results indicated financial analysts specializing in six different industries considered the size and probability of a loss the most important of seven risk definitions. Hence, managerial surveys suggest that downside concepts of risk—those specified in terms of failure to perform at an aspired-to level—are much more relevant to practicing managers than performance variability, which includes both upside and downside outcomes.[30]

Perhaps most succinctly, Robert Porter stated, "[r]isk is a function of how poorly a strategy will perform if the 'wrong' scenario occurs."[31]

---

[29] Douglas W. Hubbard, "The Failure of Risk Management", p. 67, John Wiley & Sons, 2009.

[30] Miller, Kent D. and Michael J. Leiblein (February 1996), "Corporate Risk-Return Relations: Returns Variability vs. Downside Risk", *Acad. Management Journal,* 39(1): 91–122., *citing* Mao J. C. T. (1970), "Survey of capital budgeting: Theory and practice", *Journal of Finance*, 25: 349–360; March J. G., Shapira Z. (1987), "Managerial perspectives on risk and risk taking", *Management Science*, 33: 1404–1418; and Baird I. S., Thomas H. (1990), "What is risk anyway?: Using and measuring risk in strategic management" in Bettis R. A., Thomas H. (Eds.), *Risk, strategy, and management*: 21–52. Greenwich, CT: JAI Press.

[31] Porter M. E. (1985), *Competitive advantage: Creating and sustaining superior performance*. New York: Free Press.

4.3. Offers from Risk-Averse Suppliers

Viewed through the lens of a financial product, ISO's proposal pays a fixed amount to capacity suppliers, while capacity suppliers take on the obligation to pay the $PPM \times Br_t \times CSO$ (where Br_t is the Balancing Ratio during the scarcity condition at time *t*) for an unknown length of scarcity conditions. At first glance, this proposal looks to add risk to a supplier. That first impression is wrong, at least for some suppliers, because we have left out an important piece of the ISO proposal: to pay each performing supplier's $PPM \times A_{it}$ (where A_{it} is the actual performance of resource *i* at time *t*). Thus, a supplier that takes on a CSO and exactly matches its resource output to the system Balancing Ratio is perfectly hedged. It is merely exchanging the Performance Incentive net settlement for the capacity revenues. Consequently, taking a CSO will be variance-reducing for a wide range of capacity suppliers.

Let's work through some simple examples to see how risk changes for different types of suppliers. We'll assume PPR=$5,000/MWh, H=15, and Br=75%, yielding an expected level of incentive payments of $56,250/MW-year. The risk-neutral, profit-maximizing capacity offer would then be $4.688/kW-year (plus any FA costs).

Suppose we now model a range of potential actual outcomes. One fact is that scarcity conditions will occur with a skewed distribution: there cannot be fewer than zero hours of scarcity conditions, but unexpected system conditions can force scarcity hours up to very high levels. The table below shows a skewed distribution that has an expected value of 15, and then computes the financial outcomes for a resource with performance that exactly matches the system Balancing Ratio:

Outcomes with A=0.75									
N	5	10	15	20	25	30	35		
Probability	20%	26%	20%	15%	9%	7%	3%		
PI Payment	$ 18,750	$ 37,500	$ 56,250	$ 75,000	$ 93,750	$ 112,500	$ 131,250		
PI Charge	$ 18,750	$ 37,500	$ 56,250	$ 75,000	$ 93,750	$ 112,500	$ 131,250		
CapRevenue	$ 56,250	$ 56,250	$ 56,250	$ 56,250	$ 56,250	$ 56,250	$ 56,250		
								Average	Risk (std dev)
Net w/o CSO	$ 18,750	$ 37,500	$ 56,250	$ 75,000	$ 93,750	$ 112,500	$ 131,250	$ 56,250	$ 30,695
Net w/CSO	$ 56,250	$ 56,250	$ 56,250	$ 56,250	$ 56,250	$ 56,250	$ 56,250	$ 56,250	$ -

Note that, because the CSO was priced with correct assumptions about both H and Br, the expected return is the same regardless of whether one takes the CSO. The actual financial outcome, of course, depends critically on N (the number of scarcity hours). As this table shows, for a resource with availability that exactly matches the system Br, taking the CSO provides a perfect hedge for the Performance Incentive payment, lowering the standard deviation of returns (on the capacity side of the income statement) from $30,695/MW to zero. Note that there is no other good means to hedge this risk. Because the Performance Incentive is outside of the energy markets, the usual route of hedging risk through an energy bilateral will be ineffective.

A different generator may have a very different expected outcome. Consider a baseload plant with 90% availability. Leaving aside any concern about possible correlation between plant outages and scarcity conditions, the simple risk assessment would be:

Outcomes with A=0.9									
N	5	10	15	20	25	30	35		
Probability	20%	26%	20%	15%	9%	7%	3%		
PI Payment	$ 22,500	$ 45,000	$ 67,500	$ 90,000	$ 112,500	$ 135,000	$ 157,500		
PI Charge	$ 18,750	$ 37,500	$ 56,250	$ 75,000	$ 93,750	$ 112,500	$ 131,250		
CapRevenue	$ 56,250	$ 56,250	$ 56,250	$ 56,250	$ 56,250	$ 56,250	$ 56,250		
								Average	Risk (std dev)
Net w/o CSO	$ 22,500	$ 45,000	$ 67,500	$ 90,000	$ 112,500	$ 135,000	$ 157,500	$ 67,500	$ 36,834
Net w/CSO	$ 60,000	$ 63,750	$ 67,500	$ 71,250	$ 75,000	$ 78,750	$ 82,500	$ 67,500	$ 6,139

Here again, the CSO reduces the total risk by giving away upside (when the actual number of shortage hours is higher than expected) but reducing the downside risk (when there are few shortage hours).

What about resources that are rarely used, even during scarcity periods? Consider a unit that only hits 25% of scarcity hours—perhaps because it has a long start-up time, or because the scarcity conditions occurred in short blocks of time and the unit, although a quick-start resource was not usually dispatched in economic merit. Here is how the basic analysis looks for this resource:

Outcomes with A=0.25									
N	5	10	15	20	25	30	35		
Probability	20%	26%	20%	15%	9%	7%	3%		
PI Payment	$ 6,250	$ 12,500	$ 18,750	$ 25,000	$ 31,250	$ 37,500	$ 43,750		
PI Charge	$ 18,750	$ 37,500	$ 56,250	$ 75,000	$ 93,750	$ 112,500	$ 131,250		
CapRevenue	$ 56,250	$ 56,250	$ 56,250	$ 56,250	$ 56,250	$ 56,250	$ 56,250		
								Average	Risk (std dev)
Net w/o CSO	$ 6,250	$ 12,500	$ 18,750	$ 25,000	$ 31,250	$ 37,500	$ 43,750	$ 18,750	$ 10,232
Net w/CSO	$ 43,750	$ 31,250	$ 18,750	$ 6,250	$ (6,250)	$ (18,750)	$ (31,250)	$ 18,750	$ 20,463

For this resource, taking a CSO has doubled the riskiness of its FCM payments.

At the simplest level, using modern portfolio theory alone, we can conclude that allowing resources to include a risk factor would move FCA outcomes in the direction that, presumably, ISO is seeking. Capacity Resources with average to high availability during scarcity conditions will find that taking a CSO lowers the variance of its payments, while resources with low availability during scarcity will want to include a premium in their capacity offers to reflect the higher risk.

A more sophisticated analysis shows that even resources with high availability may find that taking a CSO under the new Performance Incentives adds risk. Consider a generator in a small Reserve Zone that has generally excellent availability, but when it is off-line for maintenance, its absence sharply increases the risk of a scarcity condition in the Reserve Zone. In the example below, this is modeled as the unit having A=90% for the first 15 scarcity hours, but an assumed full outage for hours above 15 (triggered by the unit's unavailability):

Outcomes with A=0.9 for first 15 hours, 0% thereafter									
N	5	10	15	20	25	30	35		
Probability	20%	26%	20%	15%	9%	7%	3%		
PI Payment	$ 22,500	$ 45,000	$ 67,500	$ 67,500	$ 67,500	$ 67,500	$ 67,500		
PI Charge	$ 18,750	$ 37,500	$ 56,250	$ 75,000	$ 93,750	$ 112,500	$ 131,250		
CapRevenue	$ 56,250	$ 56,250	$ 56,250	$ 56,250	$ 56,250	$ 56,250	$ 56,250		
								Average	Risk (std dev)
Net w/o CSO	$ 22,500	$ 45,000	$ 67,500	$ 67,500	$ 67,500	$ 67,500	$ 67,500	$ 52,650	$ 17,779
Net w/CSO	$ 60,000	$ 63,750	$ 67,500	$ 48,750	$ 30,000	$ 11,250	$ (7,500)	$ 52,650	$ 19,077

Here again, taking the CSO has increased the risk of the generation owner. In this case, however, the ISO presumably wants this resource to take the CSO, as otherwise it will face a planning reserve deficiency in the Reserve Zone. Paying the incumbent a small premium to compensate for its added risk is likely to be far cheaper than securing new replacement resources. Allowing the large generator to bid a risk premium also reflects genuine risk to the system. From a reliability perspective, the FCA should be designed to select several small resources rather than one large resource, all else being equal.

A final element needs to be layered into this discussion, namely the aversion to strongly negative outcomes. The tables above have all stopped at 35 scarcity hours, but there is no physical reason why there could not be many more. In Porter's "wrong' scenario" for a capacity supplier, it finds itself on an extended outage during a hot spell or a cold snap, quickly racking up massive losses. This could be a "career-limiting" outcome for the resource's manager, and potentially lead to the default and bankruptcy of the entity (which, as discussed above, will probably be a facility-specific LLC). The most obvious way of managing this risk is to implement stop-losses, discussed in Section 3.1. Absent such a stop-loss, suppliers will be seeking a premium to take unlimited liability and, at a minimum, high Financial Assurance costs will need to be factored into capacity offers.

To estimate the size of this asymmetric risk element, we return to the empirical study used to support the 2.5x monthly stop-loss, described above in Section. 3.1.3. To simulate a risk-averse capacity supplier, we weighted any net losses in a month by 1.5 and recomputed the expected gains or losses over the year, assuming no monthly stop-loss. Table 2 summarizes the findings, assuming a monthly stop-loss of 2.5x, of the premium over the capacity price without the PI program (i.e., set either to be the avoidable going forward cost of existing resources or the all-in cost of new resources):

Table 2: Risk-Averse and Risk-Neutral Bid Adders by Unit Class ($/kW-month)

	Risk Neutral Valuation	Risk Averse Valuation	Implied Risk Premium
Long-Notice Mid-Merit	($2.82)	($4.20)	$1.38
Short Notice Mid-Merit	($1.12)	($1.79)	$0.67
Peaker	($0.08)	($0.40)	$0.32
Baseload	$0.45	$0.33	$0.12

As expected, in each case the possibility of negative outcomes increases the competitive offer from a risk-averse supplier. Unsurprisingly, resource types with the lowest average availability (because they are not generally committed except during high peak loads, and they have slow start times) have the greatest risk premium. Note that these risk premiums are a pure cost to consumers. Consequently, including stop-losses to limit extreme downside risks is likely to result in lower capacity prices.

4.4. Partial Delist Bids

The current market design allows partial delist bids only for Dynamic Delist bids; going into the auction, a resource must be offered up to its full qualified level or withdrawn in total. This limitation is reasonable under current rules because the Shortage Event rules also treat the resource as a block. As long as a resource is following ISO dispatch and meets other criteria discussed above, its entire EcoMax rating is deemed available.

With a change to the performance incentives, ISO should allow partial delist bids from resources, i.e. allow bidding a curve similar to what is used in the energy markets or for Dynamic Delist bids. As the foregoing discussion concluded, even though the risk-neutral offer price will be the same for all blocks of a unit (indeed, for all blocks of all units, aside from going-forward costs considerations), the a risk-averse supplier may want to offer these blocks at different prices reflecting the different risk profiles across the asset's blocks.

For example, consider a typical gas-fired generator. Its Summer Claimed Capability will be set by testing at ambient conditions not above 90 degrees F. On a very hot day, however, that resource's output will likely be reduced by several percentage points. Those same days are also more likely to have scarcity conditions. Therefore, the supplier's estimate of the availability of that top block of the generator (the parameter A above) is likely to be much lower than its estimate of A for the lower block of the unit. As we saw in the previous section, that difference in A creates a different risk profile for the top block, and therefore justifies a different risk adder.

The IMM has acknowledged the need for offer curves from individual resources; we look forward to future discussions to develop this element of the Performance Incentive design.

4.5. Offer Mitigation

In light of the foregoing discussion, offer price mitigation will be far more challenging than it is under any other capacity market design, including the current FCM. A competitive offer in the FCM now requires the supplier to assess (and the IMM to review) a series of forecasts:

1. Avoidable costs. For existing resources, these are the going forward costs plus recovery on and of capital needed to meet environmental regulations or to enhance reliability pursuant to the Performance Incentives. Historically, ISO's IMM has been reluctant to acknowledge the recovery of such capital expenditures; PJM's Avoidable Project Investment framework should be adopted by ISO to ensure that such expenditures can be reflected in capacity offers. For new resources, its direct costs are relevant only for calculating the floor price used in supply-side market power mitigation (the Alternative Price Rule).
2. Future net margins in the energy and reserves markets. This mitigation is little changed from today's approach, but there is the potential for a transition period in which historical margins may overstate future margins, under the assumption that the

Performance Incentives work as designed and reduce the frequency of tight system operations, such as we saw in January 2013. The probabilistic cost of the PER adjustment must be backed out from the expected net margins.

3. Net payments under the Performance Incentive plan. This new item will be the most challenging. Sophisticated market participants will want to study not only historical patterns, but also project the frequency, duration, and severity of scarcity conditions based on expected future market conditions. They will also want to study the potential for correlation and potential causation between unavailability of their resources and scarcity conditions. This component of the capacity offer price will then rationally include not only the expected foregone Performance Incentive payments, but also an appropriate risk premium or discount (discussed above).

4.5.1. Setting the Dynamic Delist Offer Threshold

The IMM will only need to review offers from existing resources that offer above the Dynamic Delist offer threshold. ISO has currently proposed that this be set equal to the PPM×Br×H, as discussed in Section 4.1. This raises the question of how ISO will estimate Br and H. History may be a poor guide, especially given the substantial capacity surplus the region has enjoyed over most of the past decade. One Reserve Zone, NEMA, has already crossed over to needing new capacity, and as margins there and poolwide tighten, the frequency and severity of scarcity conditions should rise.

Another challenge will be the modeling of demand response. It matters not only to demand resources but to all resources how demand resources are modeled and treated under the Performance Incentive plan. As Enernoc has pointed out, demand resources are currently not able to provide reserves. This fact raises a series of what-ifs:

- *Will demand resources that have not been activated be counted as TMOR for purposes of computing scarcity conditions?* Presumably not.
- *Will a demand resource that is a Capacity Resource and has not been activated pay the PPR? Or will it receive PPR payments?* Presumably the demand resource will pay the PPR in all scarcity intervals like any other Capacity Resource, and it will only be paid PPR rates when qualified as (negative) energy or reserves.
- *Will demand resource be qualified to provide reserves prior to FCA9?* Assuming our answers to the first two questions are correct, the answer to this is very important. Demand resources represents approximately 10% of the forecast peak load; consequently, the frequency of scarcity conditions will vary dramatically depending on whether demand resources are considered reserve capable. For the same reason, if demand resources are not reserve-qualified by FCA9, the cost and risk to demand response aggregators of participating in the FCM will rise dramatically, and therefore fewer demand resources would be expected to stay in the market, again changing the calculus about the future frequency and duration of scarcity conditions.

In addition to addressing these issues, the ISO should consider adding a margin above this base level of PPM×Br×H. Most market mitigation standards allow some variance above or below, as applicable, recognizing that the IMM does not have perfect information and that there is a cost to over-mitigation of competitive offers. The Dynamic Delist offer threshold should be set 10% above PPM×Br×H, the same margin allowed in Reference Prices in the energy markets. This buffer should sharply lower the number of offers that the IMM will need to review, allowing it to concentrate its limited resources on the offers most likely to have material impacts on the FCA clearing price. (Below we propose a further limit on liability for Performance Incentives payments if the Br×H used by the IMM to mitigate offers diverges sharply from actual experience.

4.5.2. Review of Above-Threshold Offers

Even with a (slightly) higher Dynamic Delist threshold, the IMM is likely to receive a substantial number of Static Delist bids. The Tariff already provides clear guidance to the IMM on the evaluation of risk-adjusted going forward costs for the existing components of the FCM. ISO will need to adopt new rules governing the evaluation of offers relying, in part, on the potential for higher-than-average costs associated with the Performance Incentive.

Individual market participants are likely to develop their independent views about critical parameters to forecast foregone Performance Incentive payments. In particular, the number of hours of scarcity conditions (H) presents a real challenge to forecasters because it is such a knife-edge system condition. Adopting a variable PPR that increases in value gradually as the system approaches scarcity conditions likely would result in a narrower range of forecasts. It's easier to estimate that tomorrow will be cooler than average than it is to estimate whether the low temperature will fall below 32 degrees. Furthermore, because the PPR would be continuously decreasing across a range of system, a difference in view about the frequency of absolute reserve scarcity conditions is likely to have less effect on the expected Performance Incentives payments. A full development of such a variable PPR is outside the scope of this white paper.

The Performance Incentive structure will only be viable if the IMM allows reasonable offers into the market, even if those offers are based on expectations of future outcomes different from the ISO's or the IMM's. The challenge for the IMM is distinguishing competitive offers from economic withholding. In reviewing these above-threshold offers, the IMM should consider:

1. The ability of the supplier to profit from economic withholding. As in the energy market, cost-based bids should be used only as a fallback in situations when a supplier has structural market power. The current FCA offer mitigation omits this important feature because, in the original design, incumbent suppliers had substantial latitude to price under the Dynamic Delist process, which extended up to 80% of CONE in the original market design. In the subsequent tightening of the Dynamic Delist threshold, there should have been a compensating structural test. The need for a structural test prior to offer mitigation is even more pronounced under the Performance Incentives plan because of the potentially wide range of competitive offers, depending on the risk aversion and forecasts of the frequency of scarcity events. These parameters are not easily reviewed, and therefore substantial

deference should be shown to the as-submitted offers. Capacity suppliers that are not pivotal and have small market shares should, therefore, be exempt from mitigation.

2. The process by which the market participant arrived at its offer. Substantial deference should be given to a supplier has undertaken a reasoned analysis of future market conditions, supported by empirical analysis.

3. Consistency of analysis within a portfolio. A hallmark of economic withholding is that marginal resources are bid up to establish a higher price for the rest of the portfolio. Key parameters of the Performance Incentive cost, however, are common across the portfolio. Using a uniformly high estimate of H, for example, raises all offer prices within the portfolio, contrary to the hypothesis of economic withholding.

4. Consistency of conclusions across the market. Absent overt (and illegal) collusion among market participants, only the IMM will see the range of estimates for key, market-wide parameters H and Br. The IMM should use this "wisdom of the markets" to identify outliers for more detailed review. If, however, there is a rough consensus about a parameter, it should be allowed to stand even if that consensus differs from the official view of ISO or the IMM.

To the extent that the IMM replaces the judgment of capacity suppliers with its own judgment about key parameters, it is forcing the supplier to take on risks at a price that, in the suppliers' business judgment, is not compensatory. Such decisions should therefore, at a minimum, be reviewable by FERC prior to the auction. This review process is integral to the current FCA timetable and should, therefore, be extended to resources mitigated because of Performance Incentives parameters as well as unit costs.

4.5.3. Liability Caps Based on Dynamic Delist Threshold

The Dynamic Delist Threshold is calculated using two key assumptions: the Balancing Ratio (Br) and event frequency (H). Presumably, supply offer that incorporate higher values for either assumption will be subjected to review and (potentially) restated to use ISO's assumptions instead. This would be a substantial departure from normal offer-price mitigation, where key parameters are most often measurable engineering parameters (such as a generator's heat rate), auditable accounting costs (such as payroll and taxes), or market-tested values (such as forward gas prices). In this new program, though, the key drivers will necessarily be based on projections and studies.

In order for this market design change to be robust, there must be accountability for the reasonableness of these studies. ISO will face tremendous political pressure to low-ball its estimates of Br and H: preliminary studies by ISO of the potential number of scarcity hours ranges from 6 hours to 32 hours, which would correspond to a potential range of Dynamic Delist Thresholds of about $2/kW-month to $10/kW-month. The corresponding range of implied costs to consumers is vast, ranging from about $800 million to over $4 billion in incremental capacity costs. It is critical to understand that these are *real costs* for capacity suppliers: they will be charged the *actual* PPM×Br×H that occurs during the capacity delivery year.[32] As long as the projected PPM×Br×H used for mitigation purposes has the same expected value, before the fact, as the actual PPM×Br×H, the market is fair.[33] But if the mitigation values understate expected value, the offer mitigation process will systematically shift potentially billions of dollars from capacity suppliers that otherwise would and should be paid by load-serving entities. The prospect of excessive mitigation is a serious potential risk.

Naturally there will be variations in the actual outcomes from year to year: some years may have very few scarcity events, while others have more. The risk, however, is not symmetric: H is a small number relative to the 8,760 hours in the year, and so the potential upside for capacity suppliers in a year with few events is far smaller than the potential downside in a year with many events.

One approach to reflecting the stakeholders' collective judgment about the value of the Performance Incentives – as imbedded in the filed PPR, Br, and H – would be to sharply scale down the PPR rate (prospectively) as the actual value of H approaches and then surpasses the forecast value. For example, Appendix A discusses a system in which the PPR is scaled down, beginning when Br×H reaches 75% of its target value. The change is only made prospectively, so each day has a known value of the PPR going in, and settlement charges and credits are final at the end of each month.

In this example, where target H is 12 and target Br is 0.75, the impact on incentives and payments when these forecasts are accurate is very small: at H=12, payments from Capacity Resources are 94% of the payment under the unmodified system. If H is large, however, the modification keeps the total cost to suppliers under check: at H=50, payments are 197% of the forecast level, compared to 417% under the unmodified system.

This approach limits the downside risk associated with taking on a CSO without dulling Performance Incentives. First, it must be recognized that a reduction in risk will reduce the offer price of suppliers, lowering consumer costs. Second, if ISO's estimates of Br and H are generally accurate, then this approach has little effect at all; it is only when ISO estimated a much lower incidence of events than actually occurred that there is a material effect. Third, this scaled PPR still places substantial risks on capacity suppliers, both overall and on the margin. The marginal incentive to perform is relatively strong, with the PPR after 50 hours equal to $900/MWh, roughly double the RCPF today.[34] If this reduction in PPR level is deemed as providing insufficient incentives to perform at the margin, an alternative would be to assess loads the difference between the modified PPR rate collected from capacity suppliers and the $5,000/MWh target. With this change, the system is effectively a risk-sharing approach, with capacity suppliers taking on the entire cost of the Performance Incentives up 80% of the target Br×H and a decreasing portion thereafter.

---

32 Although it is true that the majority of these charges will be paid back to these same resources, they would receive these payments regardless of whether they were Capacity Resources. As ISO has correctly noted, the *entire* amount of the performance charges is appropriately classified as costs to capacity suppliers.

33 Although a risk premium is still justified.

34 For simplicity, I assume that Br is correctly estimated and only adjust outcomes in H.

5. A Demand Curve Will Improve FCM Effectiveness

An important goal of ISO supporting the introduction of revised Performance Incentives is to foster cost-effective investment that will lead greater system flexibility and resilience. Suppliers are willing to make investments, but only if they see a reasonable opportunity to earn an appropriate return on capital. The greater the systematic risk in the market, the higher that return must be. To the extent that ISO can redesign the FCM to reduce *systematic* risk without blunting the *idiosyncratic* risks that spur good investments, ISO can lower the required return and, therefore, cost to consumers.

5.1. Rationale for a Demand Curve

The most substantial risk-reduction ISO could achieve is to adopt a capacity demand curve. Although a vertical demand curve was a central design feature of the FCA, there was never any economic rationale supporting a vertical demand curve. One point of view supporting the vertical demand curve is that it is cheaper for consumers because it doesn't overbuy or overpay for resources. While this may be true in a single period marked by surplus, it is not true—indeed, cannot be true—in the long run. Research by Prof. Benjamin Hobbs and his colleagues at Johns Hopkins University shows that a downward-sloping demand curve lowers total costs to consumers because some surplus of supply reduces periods of high energy prices.[35] Moreover, a downward-sloping demand curve produces a more reliable system over time, with more years where the planning reserve requirement is met or exceeded.

Some mechanism to moderate swings of the capacity price is sorely needed. Outcomes to date in the FCAs have been binary: either there was surplus capacity and the capacity clearing price was set at the floor, or there was a need for new capacity and the capacity clearing price was set at the administrative cap. Static Delist bids that might have set the clearing price were not allowed to set the market clearing price, but instead taken off the table by declaring the resource needed for local reliability. Dynamic Delist bids have not been a factor because of the high levels of over-supply; moreover, recent rule changes have gutted the intended role of Dynamic Delist bids of moderating price swings. As Dr. Patton, the ISO's EMM notes, "[w]hen the floor is eliminated beginning in FCA 8, the clearing price will likely fall significantly due to the level of existing capacity and the vertical demand curve implicit in the FCM design."[36] In short, none of the price stabilization mechanisms that we included in the original FCM design have had any durable success.

---

[35] "Affidavit of Benjamin F. Hobbs on Behalf of PJM Interconnection, L.L.C." FERC Docket Nos. EL05-148-000 and ER05-1410-000 (August 31, 2005); and Hobbs, Benjamin F., Javier G. Inon, Ming-Che Hu, and Steven E. Stoft, "Capacity Markets: Review and a Dynamic Assessment of Demand Curve Approaches," IEEE Power Engineering Society General Meeting, (1) 514-522, June 2005

[36] David B. Patton, PhD, "Highlights of the 2012 Annual Report on ISO New England Markets," presentation to the NEPOOL Participants Committee, June 27, 2013, ("EMM 2012 Highlights") p.26.

These wild swings in the capacity price are not conducive to new investment, either in existing or new resources. The sponsors of the Salem Harbor Repowering project have testified that the lack of confidence about future capacity prices will make project finance challenging, even with the project's $15/kW-month five-year capacity price.[37] Pool-wide capacity prices are widely seen to be below $1/kW-month until new capacity is needed, perhaps around 2020, providing little support for new investment in existing resources to meet the reliability challenges facing ISO. This conclusion is supported by the EMM's recent report, which states that "the current FCM design is not likely [to] facilitate the efficient entry and exit of resources in New England. Most of the new investment in generation under FCM has been motivated by out-of-market payments related to [Requests for Proposals] of the Connecticut [Department of Utility Control]. A large share of capacity that has attempted to go out-of-service by delisting has been unable to do so for reliability reasons."[38]

If the FCM is to support meaningful investment, it needs some new mechanism to reduce the volatility of capacity prices. The tool for this job with the best track record is an administrative demand curve or, as PJM refers to it, a variable resource requirement. Dr. Patton concurs: "We believe it is critical to introduce market reforms to address these issues … and recommend that the ISO adopt a sloped demand curve that recognizes the benefits of installed capacity beyond what is necessary to satisfy planning reserve requirements."[39] The economic logic for such a curve is compelling. A vertical demand curve implicitly values any capacity above the installed reserve margin (IRM) as having no value, and any capacity below IRM as having infinite value, albeit arrested by an administrative price cap. As with operating reserves, discussed above, this cannot be correct from first principles. First, the exact level of the IRM is not some physical constant dictated by the laws of physics: it is not like a rocket that, if it fails to achieve 11.3 km/sec, will not escape the earth's gravity well. Instead, there is a range of installed capacity that gives varying levels of reliability. More capacity provides more reliability, and this incremental reliability has some non-zero incremental value.

Adding a demand curve adds value to consumers in two additional ways. First, as noted by Prof. Hobbs, supporting additional resources is likely to result in lower energy prices. This effect will spill over into the FCA prices under the proposed Performance Incentives, too, because extra resources will reduce the expected frequency of scarcity conditions (the parameter H) and, therefore, lower the competitive offer price for capacity.[40] Second, the demand curve will result in less volatility in capacity prices and, consequently, lower volatility

---

[37] Initial Comments of Footprint Power LLC, Investigation by the Department of Public Utilities on its own Motion into the Need for Additional Capacity in NEMA/Boston within the Next Ten Years, Pursuant to Chapter 209, Section 40 of the Acts of 2012 "An Act Relative to Competitively Priced Electricity in the Commonwealth" and Pursuant to G.L. c. 164, § 76, D.P.U. 12-77.

[38] EMM 2012 Highlights, p.26.

[39] Ibid.

[40] There is a consistency problem raised by this interaction, discussed below.

of expected returns to energy market investments. Because of the positive risk-return relationship discussed above, this risk reduction will lower the required return on that investment, and this lower return requirement will be incorporated in lower competitive offers for new resources.

5.2. Design of the Capacity Demand Curve

Economics strongly suggests that the demand curve for capacity should be downward sloping, and that this slope should become flatter as capacity increases relative to the installed reserve margin (IRM). Demand should be downward sloping for two reasons: Adding extra capacity beyond the administratively determined IRM increases the reliability of the system and decreases the expected energy price, both of which have non-zero value to consumers. Conversely, falling one megawatt short of the IRM should not imply that consumers are willing to pay a very high price to cover that shortfall; that small shortage only increases the reliability risk by some small fraction. Thus, the current vertical demand “curve” does not reflect the economic fundamentals of buying planning reserves.

Economics also indicates that the slope should decrease (become flatter) along the curve.[41] With such a curve, at capacity levels below target capacity, the capacity price rises quickly to ensure that the market does not go too far short even if new capacity costs more than had been estimated—running short of capacity has serious economic consequences. Above the target capacity, the price tapers off more gradually, consistent with the idea that extra capacity has decreasing incremental value. There is also a maximum price to ensure that consumers are not unexpectedly exposed to prices outside of the range of reasonableness.

The shape of the demand curve should also reflect other key elements of the market design; in particular, the forward commitment period. The functional goal of adding slope to the demand curve is to reduce price volatility, and this volatility depends both on the curvature of the demand *and* supply curves. With a short forward commitment period, the supply curve will be very inelastic, and so the demand curve needs to have more slope to damp volatility; by contrast, a multi-year forward design, such as FCM, creates some extra slope in the supply curve by allowing timely entry and exit of resources; thus the demand curve can be relatively steeper than needed with a short forward commitment period. The curves ISO proposed in 2006 were calibrated for a market that cleared shortly before the start of the capacity capability year. In a prompt market design, there is greater risk that capacity will vary from year to year because new entry must anticipate high prices in the future and build, prospectively, to meet as-yet unrevealed demand. One of the innovations of the FCM was to operate more than three years before the delivery date, thereby allowing time for planned resources to clear the market and come on line, and for the orderly retirement of existing resources.

In a three-year forward design, it will be possible to set the target level of capacity closer to the minimum IRM level. The target capacity needs to be above the minimum IRM because,

41 I.e., that the curve should be convex to the origin.

in order to ensure that the *average* capacity revenue over time is sufficient to support new entry, and the market must spend some time short of the target capacity used to build the demand curve. If the target is set exactly at the minimum requirement, the pool will be capacity deficient approximately half the time in order to support new capacity, which is not only unacceptable from a reliability point of view, but also unduly expensive to consumers, who end up paying more in energy scarcity premiums than they would with a more generously drawn curve. In PJM, the target capacity exceeds the minimum requirement by one percentage point, but it should be noted that even some of the smaller Locational Deliverability Areas of PJM are larger than ISO's load. ISO's target, therefore, should be above one percentage point, but below the eight percentage points proposed in the Locational Installed Capacity (LICAP) market design. The exact level should be determined by simulation analysis similar to that performed by Prof. Hobbs.

Aside from this target capacity point with a clearing price at the net cost of new entry, the Variable Resource Requirement curves implemented in PJM have a weak foundation in economic theory. Properly drawn, the curves should be convex to the origin, i.e. more steeply sloped in tight capacity conditions and becoming flatter with increasing levels of capacity reserves. This pattern reflects a decreasing marginal value of incremental capacity. Both ISO and PJM proposed such curves in their original capacity market filings; it was only through a settlement process that PJM's curves became concave to the origin (standing the idea that incremental capacity is most valuable when capacity is scarce on its head) and that ISO's curves became vertical lines (implicitly denying the incremental value of capacity, and placing nearly unbounded value on achieving the target IRM).

In developing a new capacity demand curve for New England, ISO should return to the empirical analysis it did in the LICAP filing, informed as well by the analysis done by Prof. Hobbs for PJM. Prof. Hobbs' analysis assumed that all new entry was through combustion turbines, however, which allows for finer granularity of new entry than is likely to occur in fact (especially given the smaller gap between the capital cost of a combined cycle plant and a combustion turbine in today's market). This simplifying assumption was reasonable in a pool as large as PJM, but in the smaller ISO control area and, in particular, the smaller Reserve Zones, the curves will need to be set to provide sufficient buffer if new capacity needs are met by the addition of an efficiently scaled combined cycle plant.[42]

[42] This paragraph is concerned with the *slope* of the demand curve, which must be great enough to accommodate normal levels of supply variation without large price swings. The *height* of the demand curve is based on Net CONE, typically determined with reference to a simple-cycle combustion turbine. As both ISO witness Stoft and Capacity Suppliers witness Stoddard explained in the LICAP case, in an balanced market, Net CONE does not depend on the underlying technology: all efficient new technologies will have the same Net CONE, thereby equalizing a developer's expected return on equity. Net CONE is usually estimated based on combustion turbines because CTs have the lowest energy market earnings, making the estimate less subject to predictions of future energy market outcomes.

5.3. Bidding Into the FCA

A critical issue raised by adding a demand curve is that the reservation price for capacity suppliers is a function of the quantity and composition of the Capacity Resources. This is not a new issue—expectations of future energy margins and PER deductions depend on what else is on the system—but the addition of the potentially high PPR earnings brings this issue to the forefront. Estimates of the parameters H and Br are highly sensitive to the mix and overall quantity of resources. The linkage with quantity is obvious: if there is an 18% reserve margin, scarcity conditions will occur less off than if there is a 14% reserve margin. The composition of resources is also important:

- A system with a higher mix of intermittent renewables is likely to have more scarcity conditions (given that New England is geographically small and, therefore, has only a modest degree of wind and solar diversity compared to a larger region like MISO).
- Until demand resources are able to provide reserves (and system operators actually rely on them for reserves, rather than triggering OP4 actions), a higher proportion of demand response in the resource mix will be associated with more scarcity conditions.

Given the likely levels of Performance Incentive payments and their impact on competitive capacity offers, this linkage is too important to ignore. At the indicative values of a PPR of $5,000/MWh, 15 hours of expected scarcity conditions adds about $4.70/kW-month to capacity offers. Halving or doubling the number of expected hours creates a very wide range of equilibrium capacity offers.

This consideration raises issues for the auction design and market monitoring.

Auction Design Considerations

If ISO were to use a sealed-bid auction design, such as is used in PJM or NYISO, market participants would need to have the ability to submit parameterized offers. While possible, this would require a level of complexity and sophistication on the part of all capacity auction participants that seems unwieldy, requiring in principle capacity suppliers to develop, and the IMM to approve, in advance of the auction a complete path of strategies taking into account potentially binding constraints, levels of surplus capacity at various prices, and other posted information (for example, about the expected level of demand resource activation).

As it turns out, the current Descending Clock Auction (DCA) format is well suited for this purpose. At the end of each round, the auction manager currently provides the total surplus in each capacity zone. Additional information, such as the aggregate MWs of intermittent resources and, separately, of demand resources, could also be provided to market participants to help inform their revised reserve prices for the next round.

Change will be required in the current DCA to implement the variable resource requirement (demand curve). This change is straightforward, though, as it merely requires the auction manager to use a different decision rule as to when the auction closes, increasing the quantity it will purchase as the price falls.

Market Monitoring Considerations

The linkage between a variable reserve margin and the competitive capacity offer price raises more substantive issues for market monitoring.

First, the Dynamic Delist threshold should (weakly) decrease round to round, as the target installed capacity margin increases. So, for example, if the target at the opening price is to secure 97% of the nominal Installed Reserve Margin, the value of H used in setting the Dynamic Delist threshold will be high. As the clock ticks down and the capacity price falls, the target capacity procurement level will also rise and, with it, H will decline. This adjustment will only be relevant, of course, once the DCA price is below the Dynamic Delist threshold.

Static Delist bids are also more complex to monitor. Because forecasts of H and Br will vary through the course of the FCA, the Static Delist bids will need to be parameterized, based on information that will be posted during the course of the auction.

6. Conclusions

ISO has proposed a bold Performance Incentive proposal that would integrate a strong—perhaps overly strong—incentive to perform during scarcity conditions compared to the current FCM design. This additional incentive to deliver power when it is needed most is commendable, but the ISO proposal does not give due weight to the question of whether consumers are truly well-served by this approach—especially in light of the incomplete retail access and limited availability of long-term contracts in New England.

If ISO intends the Performance Incentive program to bring forth investment better tailored to the long-term reliability needs of the region, it needs to focus on the risk profile created by the market design. ISO's proposal relies on levying potentially very high penalties on Capacity Resources that provide less energy or reserves than its *pro rata* share of resource requirements during scarcity conditions. While this performance payment creates a strong incentive to take steps prospectively to reduce the risk, the possibility of severe penalties creates substantial risk for investors, who in turn will require a commensurately high rate of return on capital, raising prices in the long run for consumers.

This white paper has laid out a series of revisions to the ISO Performance Incentive program designed to retain the substantial incentives for providing energy and reserves when they are needed most, but tempering the risk by adding several elements: a demand curve for capacity, the addition of stop-loss provisions, limited excused unavailability for capacity suppliers, portfolio availability assessment, and enhanced ability to use bilateral contracts to hedge risks.

The success of the Performance Incentive program, either as proposed by ISO or as outlined herein, depends on the willingness of the IMM to allow competitive offers reflecting the added opportunity costs and changed risk profile taken on by capacity suppliers. Under the current design, the IMM could focus on accounting costs, with some limited provision for Peak Energy Rents. Under the Performance Incentive program, however, capacity offers will be priced primarily based on the expected level of foregone incentive payments. This critical role of competitive, unmitigated offers requires a deeper rethink of the mitigation program, including the addition of a structural test, as well as inclusion of a dynamic PPR or similar risk mitigation approach to ensure that the IMM does not unreasonably burden capacity suppliers with uncompensated risk by unnecessary or poorly calibrated mitigation.

With the amendments proposed in this white paper, we believe that the ISO's Forward Capacity Market will function more effectively in shaping the investment in new and existing resources to secure the long-term reliability of the New England system at reasonable prices to consumers.

Appendix A: Numeric Example of Declining PPR

ISO has proposed that the PPR should be a fixed value throughout the capacity delivery year. As discussed in Section 4.5.3 above, however, this creates a very large risk for capacity suppliers if ISO's forecast of the number of scarcity hours (H) and the balancing ratio in those hours (Br) is biased downward, and if ISO has imposed its forecasts of H and Br through offer mitigation.

Section 4.5.3 refers to a particular numerical example, discussed in more detail here. In this example, we assume that ISO stakeholders have agreed that, for offer mitigation purposes, PPR=$5,000/MWh, H=12 hours, and Br=0.75. These values set not only the Dynamic Delist threshold but also, presumably, would serve as a guide for mitigating Static Delist bids. In this example, we assume IRR=34,000 MW. The level of capacity prices should rise by PPM×Br×H, which is $3.75/kW-month or $1.53 billion annually. We refer to this cost as the "target," inasmuch as it represents a prepayment by load of the expected level of Performance Incentive charges.

The declining PPM discussed above retains the base PPM rate of $5,000/MWh. Once the actual number of shortage hours reaches 75% of the target H, the PPM is scaled by the ratio of 75% × target(Br×H) / actual(Br×H). The change is only made prospectively, so each day has a known value of the PPR going in, and settlement charges and credits are final at the end of each month.

In the table below we calculate how this declining PPM affects the total payments as the actual H increases and compares that payment to the target $1.53 billion. For the sake of expositional clarity, we assume that Br is correctly forecast at 0.75.

This approach has several desirable traits. When actual H and forecast H are close, the actual payments are very close to target: 94%. Even if ISO severely under forecasts H, the PPR remains at a robust level: at 50 hours (more than four times the forecast), the PPR is still at $900/MWh, which layers on top of the RCPF and energy prices to yield effective prices around $3,000/MWh. The downside risk to capacity suppliers is still considerable, with more than double the cumulative cost at 50 hours than at the target 12 hours. The extreme risk is greatly reduced; payments without the modified PPR would be 417% of the target at 50 hours. This risk reduction should translate directly into lower-priced offers in the FCA.

H	Br	H x Br	PPR ($/MWh)	Cumulative PI Charges	% of Target	Modified PPR ($/MWh)	Modified Cumulative PI Charges	% of Target
2	0.75	1.5	$ 5,000	$ 255,000,000	17%	$ 5,000	$ 255,000,000	17%
4	0.75	3.0	$ 5,000	$ 510,000,000	33%	$ 5,000	$ 510,000,000	33%
6	0.75	4.5	$ 5,000	$ 765,000,000	50%	$ 5,000	$ 765,000,000	50%
8	0.75	6.0	$ 5,000	$ 1,020,000,000	67%	$ 5,000	$ 1,020,000,000	67%
10	0.75	7.5	$ 5,000	$ 1,275,000,000	83%	$ 4,500	$ 1,249,500,000	82%
12	0.75	9.0	$ 5,000	$ 1,530,000,000	100%	$ 3,750	$ 1,440,750,000	94%
14	0.75	10.5	$ 5,000	$ 1,785,000,000	117%	$ 3,214	$ 1,604,678,571	105%
16	0.75	12.0	$ 5,000	$ 2,040,000,000	133%	$ 2,813	$ 1,748,116,071	114%
18	0.75	13.5	$ 5,000	$ 2,295,000,000	150%	$ 2,500	$ 1,875,616,071	123%
20	0.75	15.0	$ 5,000	$ 2,550,000,000	167%	$ 2,250	$ 1,990,366,071	130%
22	0.75	16.5	$ 5,000	$ 2,805,000,000	183%	$ 2,045	$ 2,094,684,253	137%
24	0.75	18.0	$ 5,000	$ 3,060,000,000	200%	$ 1,875	$ 2,190,309,253	143%
26	0.75	19.5	$ 5,000	$ 3,315,000,000	217%	$ 1,731	$ 2,278,578,484	149%
28	0.75	21.0	$ 5,000	$ 3,570,000,000	233%	$ 1,607	$ 2,360,542,770	154%
30	0.75	22.5	$ 5,000	$ 3,825,000,000	250%	$ 1,500	$ 2,437,042,770	159%
32	0.75	24.0	$ 5,000	$ 4,080,000,000	267%	$ 1,406	$ 2,508,761,520	164%
34	0.75	25.5	$ 5,000	$ 4,335,000,000	283%	$ 1,324	$ 2,576,261,520	168%
36	0.75	27.0	$ 5,000	$ 4,590,000,000	300%	$ 1,250	$ 2,640,011,520	173%
38	0.75	28.5	$ 5,000	$ 4,845,000,000	317%	$ 1,184	$ 2,700,406,257	176%
40	0.75	30.0	$ 5,000	$ 5,100,000,000	333%	$ 1,125	$ 2,757,781,257	180%
42	0.75	31.5	$ 5,000	$ 5,355,000,000	350%	$ 1,071	$ 2,812,424,114	184%
44	0.75	33.0	$ 5,000	$ 5,610,000,000	367%	$ 1,023	$ 2,864,583,205	187%
46	0.75	34.5	$ 5,000	$ 5,865,000,000	383%	$ 978	$ 2,914,474,509	190%
48	0.75	36.0	$ 5,000	$ 6,120,000,000	400%	$ 938	$ 2,962,287,009	194%
50	0.75	37.5	$ 5,000	$ 6,375,000,000	417%	$ 900	$ 3,008,187,009	197%