2012 ASSESSMENT OF THE ISO NEW ENGLAND ELECTRICITY MARKETS

Prepared by:

David B. Patton, PhD
Pallas LeeVanSchaick, PhD
Jie Chen, PhD

EXTERNAL MARKET MONITOR FOR ISO-NE

MAY 2013

Table of Contents

List of Figures

List of Tables

Guide to Abbreviations

APR	Alternative Price Rule
ASM	Ancillary Services Market
CONE	Cost of New Entry
CT DPUC	Connecticut Department of Public Utility Control
EMM	External Market Monitor
FCA	Forward Capacity Auction
FCM	Forward Capacity Market
FTR	Financial Transmission Rights
GW	Gigawatt (1 GW = 1,000 MW)
HHI	Herfindahl-Hirschman Index, a measure of market concentration
ISO	Independent System Operator
ISO-NE	ISO New England Inc.
LFRM	Locational Forward Reserve Market
LMP	Locational Marginal Price
LOC	Lost Opportunity Cost, a component of the regulation price
LSR	Local Sourcing Requirement
MMbtu	Million British Thermal Units, a measure of energy content
IMM	Internal Market Monitor
MW	Megawatt
MWh	Energy associated with producing 1 MW for one hour
NCPC	Net Commitment Period Compensation
NEMA	North East Massachusetts
NERC	North American Electric Reliability Corporation
NPCC	Northeast Power Coordinating Council, Inc.
NYISO	New York ISO
PER	Peak Energy Rent
RA	Reconfiguration Auction
RAA	Reserve Adequacy Assessment
RCP	Regulation Clearing Price
RCPF	Reserve Constraint Penalty Factors
RMR	Reliability Must-Run
RTO	Regional Transmission Organization
SEMA	South East Massachusetts
SCR	Special Constraint Resources
SMD	Standard Market Design
TMNSR	Ten-minute non-spinning reserves
TMOR	Thirty-minute operating reserves
TMSR	Ten-minute spinning reserves
UDS	Real-time dispatch software

Preface

Potomac Economics serves as the External Market Monitor for ISO-NE. In this role, we are responsible for evaluating the competitive performance, design, and operation of the wholesale electricity markets operated by ISO-NE.[1] In this assessment, we provide our annual evaluation of the ISO's markets for 2012 and our recommendations for future improvements. This report complements the Annual Markets Report, which provides the Internal Market Monitor's evaluation of the market outcomes in 2012.

We wish to express our appreciation to the Internal Market Monitor and other staff of the ISO for providing the data and information necessary to produce this report.

1 The functions of the External Market Monitor are listed in Appendix III.A.2.2 of "Market Rule 1."

Executive Summary

This report assesses the efficiency and competitiveness of New England's wholesale electricity markets in 2012. Since ISO-NE began operations in 1999, it has made significant enhancements to the energy market and introduced markets for other products that have improved overall efficiency. ISO-NE's markets currently include:

- Day-ahead and real-time energy, which coordinate commitment and production from the region's generation and demand resources, and facilitate wholesale energy trading;
- Financial Transmission Rights (FTRs), which allow participants to hedge the congestion costs associated with delivering power to a location that is constrained by the limits of the transmission network;
- Forward and real-time operating reserves, which are intended to ensure that sufficient resources are available to satisfy demand when generation outages or other contingencies occur;
- Regulation, which allows the ISO to instruct specific generators to increase or decrease output moment-by-moment to keep system supply and demand in balance; and
- Forward Capacity Market (FCM), which is intended to provide efficient long-term market signals to govern decisions to invest in new generation and demand resources and to maintain existing resources.

These markets provide substantial benefits to the region by ensuring that the lowest-cost supplies are used to satisfy demand in the short-term and by establishing transparent, efficient wholesale price signals that govern investment and retirement decisions in the long-term. The markets achieve the short-term benefits by coordinating the commitment and dispatch of the region's resources, which is essential due to the physical characteristics of electricity and the transmission network used to deliver it to customers. This coordination affects not only the prices and production costs of electricity, but also the level of reliability with which it is delivered.

A. Introduction and Summary of Findings

In addition to providing a summary of market outcomes in 2012, this report includes findings in two primary areas: the competitive performance of the markets and the operational efficiency of the markets. The broad findings in each of these areas are discussed below.

1. Competitive Performance of the Markets

Based on our evaluation of the markets in New England (in both constrained areas and the broader market), we find that the markets performed competitively in 2012. Although structural analyses indicate potential market power under certain conditions in some areas, our assessment raised no significant competitive concerns associated with suppliers' market conduct. In addition, the ISO automated its market power mitigation process, which improved the effectiveness of the mitigation in preventing the exercise of market power under conditions when a supplier may face limited competition.

Energy prices fell 22 percent from 2011 to 2012, due primarily to the reduction in natural gas prices (the dominant fuel in New England), which fell 21 percent on average from 2011.[2] In a competitive market, suppliers have strong incentives to offer their supply at prices close to their short-run marginal costs of production.[3] Because fuel costs constitute the vast majority of the marginal costs of most generation, lower fuel costs translate into lower offer prices and market clearing prices in a well-functioning, competitive market. The correspondence of fuel prices and offer prices in New England is an indication of the competitiveness of ISO-NE's markets.

Other variations in supply and demand also contributed to the decrease in energy prices:

- On the demand side, average load fell in 2012 by 1 percent across all hours and by 6 percent in the first quarter because of milder winter weather. The summer peak load fell 7 percent from 2011 to 2012, although the average load in the summer months rose modestly in 2012.

- On the supply side, a new 620 MW gas-fired combined cycle unit in Connecticut entered the market in mid-2011. In addition, nuclear generation and imports from neighboring areas rose by an average of 570 MW combined in 2012. Consequently, electricity prices fell as lower demand and increased low-cost supply led high-cost generating resources to operate less frequently.

2 Natural gas prices are based on the day-ahead prices reported by Platts for the Algonquin pipeline for the City Gate Rate.

3 Short-run marginal costs are the incremental costs of producing additional output in a timeframe short enough to preclude expanding, retiring or converting the assets to another use. These costs include any foregone opportunity costs of producing such output. For convenience, we will refer to these costs as "marginal costs". The incentive to submit offers at prices close to marginal cost is affected by the design of the market. This incentive exists in markets that establish clearing prices paid to all sellers, as is the case in the ISO-NE markets. Markets that make payments to suppliers based on the supplier's offer (i.e., pay-as-offer markets) create incentives for suppliers to raise their offers above their marginal costs.

2. Operational Efficiency of the Markets

Efficient real-time prices are critically important because they:

- Provide incentives for market participants to operate in a manner that maintains reliability at the lowest overall cost;
- Facilitate efficient day-ahead scheduling, resource commitments, and the arrangement of reliable fuel supplies for those resources; and
- Contribute to efficient investment in supply and demand response resources with flexible operating characteristics in the long term.

We find that both the day-ahead and real-time markets operated relatively efficiently in 2012 as prices appropriately reflected the effects of lower fuel prices and load levels. However, we also find that real-time prices often do not fully reflect the cost of satisfying demand and maintaining reliability during tight market conditions, particularly when fast-start resources or demand response resources are deployed in the real-time market. We make several recommendations in this report to address the efficiency of real-time prices.

Two significant factors have led to noteworthy changes in market operations in recent years. First, following the completion of transmission upgrades in Boston, Connecticut, and Southeast Massachusetts by 2009, the ISO has needed to commit far less capacity for local reliability. Hence, the total uplift charges from NCPC payments fell from $387 million in 2008 to $87 million in 2012.

Second, natural gas system limitations that prevent generators from responding to the ISO's commitment instructions have become more frequent, especially during the winter months. This has led the ISO to commit additional capacity that is not gas-dependent for system reliability. Such reliability commitments can often lead to significant surplus capacity in real time, which tends to depress energy and ancillary services prices in the real-time market.

The ISO is proposing changes to increase the incentive for suppliers to make fuel arrangements to increase their availability in real time, and improve real-time price signals by bringing the market requirements into better alignment with its reliability requirements.

3. Recommendations

Overall, we conclude that the markets performed competitively in 2012 and were operated well by the ISO. Based on the results of our assessment, however, we offer ten recommendations to further improve the performance of the New England markets. Eight of the ten were also recommended in our *2011 Annual Assessment*. This is expected since many of the recommendations require substantial resources and must be prioritized with the ISO's other projects and initiatives. Most of these recommendations are either currently being evaluated by the ISO or have been included in the Wholesale Markets Plan for implementation over the next five years. Two of the recommendations made in the 2011 report are not included in this report because the ISO has nearly completed market changes to address the recommendations.[4] A table of recommendations can be found at the end of this Executive Summary.

B. Energy Prices and Congestion

Average real-time energy prices decreased 22 percent, from approximately $49 per MWh in 2011 to $38 per MWh in 2012.[5] This was due primarily to substantially lower natural gas prices, which fell 21 percent from 2011. This is important because natural gas-fired resources are most frequently on the margin in New England. Several other changes in supply and demand also contributed to the reduction in energy prices, including:

- Lower load levels – Average load decreased 1 percent overall and 6 percent in the first quarter from 2011 to 2012 because of mild winter weather. The annual peak load fell 7 percent from 2011 to 2012, leading to fewer demand response activations during the summer, although there were more hours in 2012 when load was at moderately high levels (e.g., 20 GW).

- Increases in net imports – Average net imports from neighboring areas, particularly Hydro Quebec and Upstate New York, increased by an average of 310 MW in 2012. The increase was largest during the coldest winter months (December, January, February) when net imports increased by an average of 400 MW over the previous year.

---

4 Recommendations #8 and #9 from the 2011 Annual Assessment are set for implementation by the eighth Forward Capacity Auction in February 2014. These include implementing buyer-side and supplier-side mitigation measures and modeling all eight load zones in the auction.

5 The price at the New England Hub, which is representative of the New England market, is reported here. The average electricity price is weighted by the New England load level in each hour.

- Increases in nuclear generation – The average output from nuclear units increased by 260 MW from 2011 to 2012 because they experienced fewer outages in 2012.

- Increased internal supply – A new 620 MW gas-fired combined cycle unit in Connecticut entered the market in mid-2011.

1. Congestion and Financial Transmission Rights

New England has experienced very little congestion into historically-constrained areas, such as Boston, Connecticut, and Lower Southeast Massachusetts, since transmission upgrades were completed in 2009. In 2012, most of the price separation between net exporting regions and net importing regions was due to transmission losses, rather than to transmission congestion. Reductions in congestion-related Locational Marginal Price (LMP) differences result in less overall congestion revenue being collected in the day-ahead and real-time markets.

Total day-ahead congestion revenues totaled $30 million in 2012, up from $18 million in 2011. The increase in congestion revenue was due to several factors:

- Peak load conditions occurred more frequently in 2012, leading to more frequent congestion into import-constrained areas in the summer months;
- Congestion increased in areas where planned transmission outages substantially affected the network capability; and
- Natural gas prices rose substantially in November and December 2012, increasing redispatch costs and associated congestion-related price differences.

Nonetheless, the recent levels of congestion revenue are far below the historic levels that prevailed before transmission upgrades were completed in 2009 (e.g., congestion revenue averaged $138 million between 2006 and 2008). Likewise, the recent levels of congestion revenue are far lower than levels seen in other LMP markets in 2012 (e.g., NYISO exceeded $300 million, MISO exceeded $700 million, and PJM exceeded $500 million).[6] Given the relatively small congestion price differences between net-importing regions and net-exporting regions, future investment in new resources is most likely to occur in areas where it is less costly to build and operate resources until generation retirements and/or load growth change the pattern of network flows.

---

6 These markets are larger than ISO-NE. NYISO serves almost 50 percent more load and the other markets are four to five times larger. However, these markets exhibited 10 to 25 times more congestion revenue in 2012.

The ISO uses most of the congestion revenues to fund the economic property rights to the transmission system in the form of FTRs.[7] The ISO operates annual and monthly markets for FTRs, which allow participants to hedge the congestion and associated basis risk between any two locations on the network. Since FTR auctions are forward financial markets, efficient FTR prices should reflect the expectations of market participants regarding congestion in the day-ahead market.

Our analysis of FTR prices indicates:

- In 2012, annual FTR prices generally over-estimated the congestion that prevailed in the energy market.
- Monthly FTR prices were more consistent with congestion patterns, which is to be expected due to additional information that becomes available regarding system conditions.
- The consistency of FTR prices and congestion improved substantially overall in 2011 and 2012 from prior years.
- Overall, we conclude that the FTR markets performed reasonably well in 2012.

Congestion revenue is used to fund the FTRs sold by ISO-NE. The congestion revenue of $30 million collected by the ISO in 2012 was sufficient to fully fund the target value of the FTRs.

2. Day-Ahead to Real-Time Price Convergence

When prices in the day-ahead market converge well with the real-time market, it indicates that the day-ahead market accurately represents expected real-time market conditions. This is important because most supply and demand settlements occur in the day-ahead market and FTRs settle against day-ahead congestion prices. Additionally, most generation is committed through the day-ahead market, so good price convergence leads to a more economic commitment of resources and the arrangement of fuel supplies at lower cost.

We evaluated price convergence at the New England Hub, which is broadly representative of prices in most areas of New England.

7 FTRs entitle the holder to the congestion price difference between the FTR's sink and source in the day-ahead market (i.e., the congestion price at the sink minus the congestion price at the source).

We found that price convergence between the day-ahead and real-time markets was not optimal in 2012. Average real-time prices have been persistently higher than average day-ahead prices in the past few years, which is unusual since electricity markets typically exhibit slightly higher day-ahead prices. We do not believe this result is efficient because small day-ahead premiums generally lead to a more efficient commitment of the system's resources.

Section V shows that real-time energy prices frequently do not reflect the full costs of the marginal source of supply. For example, when high-cost peaking resources are committed to satisfy the real-time demand, real-time prices generally do not reflect the full costs of such resources. Because the real-time prices are understated in these cases, day-ahead prices would have to be slightly higher than the actual real-time prices in order to efficiently facilitate a day-ahead commitment of resources to fully satisfy the real-time system needs.

One reason for the pattern of real-time premiums in the past few years is that the average allocation of NCPC charges to virtual load has increased after May 2010 (which would otherwise have a strong incentive to buy at the lower day-ahead price and sell at the higher real-time price). Hence, the allocation of NCPC charges has likely inhibited the natural market response to the sustained real-time price premiums. This is discussed in the next sub-section.

3. Virtual Trading and Uplift Allocation

Virtual trading plays an important role in overall market efficiency by improving price convergence between day-ahead and real-time markets, thereby promoting efficient commitment and scheduling of resources in the day-ahead market. Virtual trading in the day-ahead market consists of purchases or sales of energy that are not associated with physical demand or physical generating resources. Since no physical energy will be supplied or consumed in real time, virtual transactions scheduled in the day-ahead market are settled against real-time energy prices and are only profitable when they contribute to price convergence between the two markets.

ISO New England allows virtual traders to schedule transactions at every pricing location. This includes individual nodes and more aggregated locations, such as the New England Hub and load zones. Virtual transaction quantities at individual nodes decreased sharply in May 2010 and remained relatively low throughout 2011 and 2012. This was due primarily to the correction of a day-ahead modeling inconsistency that allowed virtual transactions to earn sustained profits at a small number of nodes.

The reduction in virtual trading volumes after May 2010 caused NCPC costs to be allocated to a smaller quantity of real-time deviations (which is explained below), thereby increasing the average NCPC charge rate to virtual transactions. The allocation of Economic NCPC to virtual transactions increased significantly from an average of $0.68 per MWh in 2009 to $2.10 in 2010, $1.98 in 2011, and $2.11 in 2012. This increased allocation to virtual transactions has placed downward pressure on virtual trading volumes and likely hindered the day-ahead market's natural response to transitory price differences between the day-ahead and real-time market.

Most NCPC charges result from supplemental commitments for system-wide needs (known as Economic NCPC) and are allocated to "real-time deviations" between day-ahead and real-time schedules.[8] In reality, some deviations are "harming" and tend to increase NCPC, while others are "helping" and reduce NCPC. For example, underscheduling physical load in the day-ahead market can cause the ISO to commit additional units in real-time, which are likely to increase NCPC—this is a "harming" deviation. Conversely, "helping" deviations, such as over-scheduling load (including virtual load), generally result in higher levels of resource commitments in the day-ahead market and, therefore, usually decrease the ISO's need to make additional commitments, thereby avoiding NCPC. The current allocation does not distinguish between helping and harming deviations and is, therefore, not consistent with cost causation. Hence, this allocation assigns NCPC costs to transactions that actually tend to *reduce* the need for supplemental commitments, including virtual load.

Additionally, NCPC charges are caused by many factors other than real-time deviations, such as when peaking resources are dispatched but do not set LMPs or when supplemental commitments are made for forecasted needs that do not materialize. We find that the current allocation scheme allocates costs to helping deviations, which likely reduce NCPC charges, and over-allocates costs to harming deviations relative to the portion of the NCPC they likely cause.

---

8 Real-Time Deviations include Real-Time Load Obligation Deviations, which are positive or negative differences between day-ahead scheduled load and actual real-time load, uninstructed generation deviations from day-ahead schedules, virtual load schedules and virtual supply schedules.

→ *We recommend that the ISO modify the allocation of Economic NCPC charges to participants that cause the NCPC, which would generally involve not allocating NCPC costs to virtual load and other real-time deviations that cannot reasonably be argued to cause real-time economic NCPC.*

This recommendation is not consistent with the IMM's recommendation on this issue. However, we will continue to work with the ISO and its IMM to develop changes to the NCPC allocation that would address this issue and improve the incentives for efficient day-ahead scheduling by market participants.

C. Reserve and Regulation Markets

The ISO operates a forward reserve market where reserves are procured in seasonal auctions, a real-time regulation market, and a real-time reserve market where reserves are scheduled with local requirements and co-optimized with the real-time energy market. These markets provide mechanisms for the wholesale market to meet the reliability needs of the system, thereby reducing the need for out-of-market actions by the operators.

1. Real-Time Reserve Market Results

Overall, the clearing prices for operating reserves increased from in 2012. Outside the local constrained areas:

- The average 10-minute spinning reserve ("TMSR") clearing price increased from $1.04 per MWh in 2011 to $1.65 per MWh in 2012;
- The average 10-minute non-synchronous reserve ("TMNSR") price increased from $0.39 per MWh to $0.98 per MWh; and
- The average 30-minute reserve TMOR price rose from $0.25 per MWh to $0.97 per MWh.

These increases resulted from higher prices in the second half of 2012, which were primarily due to the following two significant market changes:

- The RCPF for system-level 30 minute reserves was increased from $100 to $500 per MWh in June 2012. This led the real-time market to set much higher clearing prices and incur higher re-dispatch costs during tight operating conditions.

- ✓ After June 1, the clearing price averaged $174 per MWh when the 30-minute reserve constraint was binding.
- ✓ Previously, the real-time model would have simply been short of reserves in these intervals and set the clearing price at $100 per MWh, or the operators would have had to maintain adequate reserve levels through out-of-market actions.

- The system-level 30-minute reserve requirement rose in July 2012 consistent with the 25 percent increase in the system-level 10-minute requirement. This contributed to more frequent binding constraints and higher clearing prices.

The ISO also has local reserve zones in Boston, Southwest Connecticut, and Connecticut, but real-time reserve prices were comparable to prices outside the local areas, reflecting that local reserve constraints have rarely been binding. This has generally been the case since the completion of transmission upgrades in Connecticut and Boston in mid-2009.

The ISO plans to increase the 30-minute operating reserve requirement to procure additional “replacement reserves”. We believe this is a valuable change because it will:

- Allow the ISO’s true reliability needs to be more fully specified and priced. This has become increasingly important over the past two years as concerns regarding the availability of fuel and the performance of generation; and
- Improve suppliers’ incentive to be available and perform in real time.

However, this change would be even more effective if the ISO had the ability to vary the replacement reserve quantity as reliability dictates. Under cold weather conditions when the ISO’s concern regarding fuel availability is heightened, it may be reasonable to procure a larger quantity of replacement reserves. Under mild conditions when fuel uncertainty and load forecast uncertainty are both minimal, it may be reasonable not to procure replacement reserves. Allowing the ISO to determine the quantity of replacement reserves it needs will help maintain consistency between the market outcomes and the ISO’s reliability requirements.

→ *We recommend that the ISO have the capability to vary the quantity of replacement reserves during the operating day in order to improve consistency between the market outcomes and the ISO’s reliability needs.*

The ISO has recently revised its procedures for auditing the 10-minute and 30-minute reserve capabilities of off-line and on-line resources to improve their accuracy. Such efforts are beneficial because they will help ensure that the real-time market procures a sufficient quantity of operating reserves and that real-time prices more accurately reflect the cost of maintaining reliability.

2. Operating Reserves in the Day-Ahead Market

The day-ahead market coordinates the procurement of an efficient set of resources to satisfy the needs of the system over the operating day, while respecting transmission constraints and other limitations. Doing this in the day-ahead timeframe enables resources to be committed, considering start-up costs, minimum run times, and other operational inflexibilities. Currently, the ISO procures only energy through the day-ahead market, although forward reserve providers have certain day-ahead obligations.

Procuring operating reserves in the day-ahead market would allow the ISO to procure the amount of reserves it needs for the following day and to set clearing prices that reflect the costs of satisfying the operating reserve obligations. Such markets would also likely help address the ISO’s concerns regarding unit availability. The day-ahead reserve schedules would be established in a timeframe in which suppliers can make arrangements for fuel and staffing to allow them to be available in real time and respond to reserve deployments.

→ *We recommend that the ISO consider introducing day-ahead operating reserve markets that are co-optimized with the day-ahead energy market.*

3. Forward Reserve Market Results

The Locational Forward Reserve Market (LFRM) is a seasonal auction held twice a year where suppliers sell reserves which they are then obligated to provide in real-time. LFRM obligations must be provided from an online resource with unused capacity or an offline resource capable of starting quickly (i.e., fast-start generators). The auction procures operating reserves for All of New England, Boston, Connecticut, and Southwest Connecticut. [9] This report evaluates the results of recent forward reserve auctions and examines how suppliers satisfied their obligations in the real-time market.

9 The ISO used to procure forward reserves for Rest-of-System. However, the Rest of System 30-Minute Operating Reserves purchase requirement was eliminated before the Summer 2011 Procurement Period.

The TMOR prices in the three local areas cleared at the same levels as the system TMOR prices in both the 2011/2012 and the 2012/13 Capability Periods because none of the local requirements were binding. Outside of the local reserve areas, the system TMOR requirement was the only binding constraint. Therefore, the clearing price has been the same for all forward reserve products in all locations for these Capability Periods.

Prices fell roughly 24 percent from an average of $4.40 per kW-month in the 2011/12 Capability Period to $3.35 per kW-month in the 2012/13 Capability Period. Most of the decrease was related to the reduction in the Forward Capacity Market clearing prices. We also found that 99 percent of the resources assigned to satisfy forward reserve obligations in 2012 were fast-start resources capable of providing offline reserves. This is consistent with our expectations because these resources can satisfy their forward reserve obligations at a very low cost.

The ISO may wish to consider the long-term viability of the forward reserve market for several reasons. First, it has not achieved one of its primary objectives, which was to lower NCPC by purchasing forward reserves from high-cost units frequently committed for reliability. Second, the Locational Forward Reserve Market is largely redundant with the locational requirement in the Forward Capacity Market. Third, the forward reserve requirements are determined seasonally and the obligations of forward reserve suppliers are not consistent with the day-to-day operational needs of the system. In fact, the forward procurements do not ensure that sufficient reserves will be available during the operating day.

4. Regulation Market

The regulation market performed competitively in 2012, with an average of approximately 815 MW of available supply competing to serve an average of 60 MW of regulation demand.[10] The significant excess supply generally limited competitive concerns in the regulation market. However, regulation supply was sometimes tight in low-demand periods when many regulation-capable resources were offline, leading to transitory periods of high regulation prices.[11]

10 The average available supply is the average of offered regulation capabilities from committed resources in each hour.

11 These types of transitory high regulation prices are normal market outcomes and generally do not raise competitive concerns.

Regulation market expenses fell modestly from $13.3 million in 2011 to $11.6 million in 2012. This reduction was due in part to the reduction in natural gas prices over the same period.

In October 2011, FERC issued Order 755 on Frequency Regulation Compensation, which requires ISO-NE and other ISOs to operate regulation markets that compensate generators for "actual service provided, including a capacity payment that includes the marginal unit's opportunity costs and a payment for performance that reflects the quantity of frequency regulation service provided."[12] The ISO submitted its latest proposals for complying with the Order on February 6, 2013. If accepted, they will become effective on or after January 1, 2015. It is likely that the regulation market will continue to perform competitively given the large amount of supply in the market relative to the demand.

D. External Interface Scheduling

Efficient scheduling of the interfaces between New England and its neighbors can have a significant effect on the ISO-NE market outcomes. Hence, we evaluate transaction scheduling between New England and the three adjacent regions: Quebec, New Brunswick, and New York.

1. Quebec and New Brunswick Interfaces

Power is usually imported from Quebec and New Brunswick -- net imports averaged 1,640 MW during peak hours and 1,480 MW during off-peak hours in 2012. This is characteristic of the efficient management of hydroelectric resources, whereby the largest imports are made in periods with the highest prices. Most of these imports are from Hydro Quebec, which exported the most power to New England in the summer months and in periods with high natural gas prices (i.e., typically the winter months).

2. New York Interface

New England and New York are connected by one large interface between northern New England and eastern upstate New York, and by two small interfaces between Connecticut and Long Island. Exports are consistently scheduled from Connecticut to Long Island over the

---

12 *Frequency Regulation Compensation in the Organized Wholesale Power Markets,* Order No. 755, 76 Federal Register 67260 (Oct. 31, 2011), FERC Stats. & Regs. ¶ 31,324 (2011) (Order 755).

smaller interfaces (averaging roughly 390 MW during peak hours in 2012), while participants schedule power flows that can alternate directions on the larger interface depending on the relative prices. On average, New England imported roughly 235 MW from New York across the larger interface in 2012.

The spread in natural gas prices between New England and New York is an important driver of the variations in interchange between the two markets. The spread averaged $0.70 per MMbtu in four months (January, February, November, and December) of 2012, significantly higher than the average of $0.10 per MMbtu in the other eight months. Accordingly, New England imported an average of 625 MW from New York during peak hours in these four winter months, compared to only 45 MW in other eight months.

On an hourly basis, market participants should arbitrage the prices between New York and New England by scheduling power from the low-priced market to the high-priced market. However, uncertainty and long scheduling lead times have prevented participants from fully utilizing the interfaces. This has caused large real-time price differences to frequently occur between the two markets, even when the interfaces are not fully utilized. We found that power was scheduled in the inefficient direction (*from* the high-priced market *to* the low priced-market) 48 percent of the time across the primary interface and in 42 percent of the time over the two smaller interfaces in 2012. This resulted in substantial inefficiencies and higher costs in both areas. It also degrades reliability because the interchange will not adjust predictably to changes in supply or demand changes in New England.

To address this issue, ISO-NE and NYISO are developing a new scheduling process intended to improve the efficiency of the interchange between the two control areas. The Coordinated Transaction Scheduling (CTS) process is under development to allow intra-hour changes in the interchange between control areas and is scheduled to be effective in 2015. Under CTS, the ISOs will schedule interchange based on short-term forecasts of market conditions and new bidding procedures that allow firms to submit bids that are jointly evaluated by the ISOs.

→ *We recommend that the ISO continue to place a high priority on the implementation of CTS.*

ISO-NE and NYISO are also considering market-to-market congestion management coordination, which are procedures for enabling one ISO to redispatch its internal resources to relieve congestion in the other control area when it is efficient to do so. The estimated benefits of this initiative are substantially lower than the benefits of the CTS initiative given the current low levels of congestion in New England, so we continue to place a much higher priority on implementing CTS.

E. Real-Time Pricing and Market Performance

The goal of the real-time market is to coordinate the use of resources to efficiently satisfy the reliability needs of the system. To the extent that reliability needs are not fully satisfied by the market, the ISO must procure needed resources outside of the market. However, these out-of-market actions tend to undermine the market prices because the prices will not fully reflect the reliability needs of the system. Efficient real-time prices are important because they encourage competitive scheduling by suppliers, participation by demand response, and investment in new resources when and where needed. We evaluated five aspects of the real-time market related to pricing and dispatch in 2012 and make the following conclusions and recommendations:

1. Real-Time Pricing of Fast-Start Resources

Fast-start generators are routinely deployed economically, but the resulting costs are often not fully reflected in real-time prices. In 2012, 68 percent of the fast-start capacity that was started in the real-time market did not recoup its offer. This leads fast-start resources with flexible characteristics to be substantially under-valued in the real-time market, despite the fact that they provide significant economic and reliability benefits. If the average total offers of these units were fully reflected in the energy price, the average real-time LMP would increase approximately $2.20 per MWh in 2012. If these price increases were reflected in the calculation of NCPC uplift charges, we estimate that they would have been $6.3 million lower in 2012.

→ *We recommend that the ISO evaluate potential changes in the pricing methodology that would allow the deployment costs of fast-start generators to be more fully reflected in the real-time market prices.*

2. Real-Time Pricing in Forecasted and Actual Operating Reserve Shortages

The ISO replaced the $100 RCPF for system-level 30-minute reserves with the $500 RCPF on June 1, 2012. Before the increase, the ISO often had to curtail exports or take other manual actions outside the market in order to maintain adequate reserves. This led to inefficiently low real-time prices that did not properly reflect the cost of maintaining reliability. The new RCPF level provides market participants better incentives to schedule in the day-ahead market and schedule net imports from external areas that will lower the costs of maintaining reliability.

3. Real-Time Pricing During Demand Response Activations

A total of nearly 2,800 MW of demand resources were enrolled by the end of 2012. Participation by demand response in the market has been beneficial in many ways. Demand response contributes to reliable system operations, long-term resource adequacy, lower costs, decreased price volatility, and reduced supplier market power. Even modest reductions in consumption by end-users during high-price periods can significantly reduce the costs of committing and dispatching generation to satisfy the needs of the system. These benefits underscore the value of designing wholesale markets that provide transparent economic signals and market processes that facilitate demand response.

However, the inflexibility of demand response resources presents significant challenges for efficient real-time pricing. When demand response resources are on the margin, prices should reflect the marginal cost of the foregone consumption by the demand response resources. Because they are generally not dispatchable, they do not set real-time energy prices and tend to lower prices by reducing the apparent demand in the market. In 2012, there were no capacity deficiencies that required activation of emergency demand response resources, so the market outcomes were not affected by these pricing issues. Nonetheless, the activation of demand response is likely to occur more frequently in the future, which makes it important to address this pricing issue as the region's reliance on demand response resources grows.

- → *We recommend that the ISO develop rules for allowing the costs of activating non-dispatchable demand response resources to be reflected in clearing prices when there would have been a shortage without the activation of demand response resources.*

4. Ex Ante and Ex Post Pricing

ISO-NE re-calculates prices after each interval (i.e., "ex post pricing") rather than using the "ex ante" prices produced by the real-time dispatch model. Our evaluation of ISO-NE's ex post pricing results indicates that it (i) creates a small upward bias in real-time prices in most areas, and (ii) sometimes distorts the value of congestion into constrained areas.

→ *We recommend that the ISO consider modifying the inputs from UDS to the ex post pricing model to improve the consistency of the ex post and ex ante prices.*

5. Price Corrections

We find that price corrections were very infrequent in 2012, which reduces uncertainty for market participants transacting in the ISO-NE wholesale market. Furthermore, a large share of the price corrections that did occur affected a very small number of pricing nodes.

F. System Operations

The wholesale market should provide efficient incentives for participants to make resources available to meet the ISO's reliability requirements. When the wholesale market does not meet all of these requirements, the ISO will commit additional generation or take other actions to maintain reliability. In addition to the NCPC costs that result from these actions, these commitments result in surplus supply that lowers real-time prices and reduces scheduling incentives in the day-ahead market. Hence, such actions should be undertaken only when necessary. In this section, we evaluate several aspects of the ISO's operations and processes for satisfying reliability requirements in 2012.

1. Accuracy of Load Forecasting

The day-ahead load forecast is important because market participants may use it and other available information to inform their decisions regarding fuel procurement, management of energy limitations, formulation of day-ahead bids and offers, and outage scheduling. In addition, the ISO uses the forecast to estimate the amount of resources that will be needed to satisfy the load and reserve requirements of the system. Based on our analysis of ISO-NE's daily peak load forecasts, we found that the average day-ahead load forecast was slightly higher than the average real-time load in the peak load hour of each day in 2012. Overall, load forecasting was relatively accurate and generally superior to load forecasting in other RTO markets.

2. Supplemental Commitment for Local Reliability

Supplemental commitment for local reliability has been low since mid-2009 when significant transmission upgrades in historically import-constrained areas were completed. These upgrades have allowed additional imports to these areas, reducing the amount of online and quick start capacity that must be available internally. In 2012, the amount of capacity committed for most local reliability issues averaged 155 MW, down considerably from 1,000 MW in 2008.

Reduced commitment for local reliability has also contributed to a decline in the amount of daily surplus capacity (i.e., the amount of online reserves and fast-start reserves minus the real-time reserve requirement in the peak load hour) from an average of nearly 1,700 MW prior to the transmission upgrades in mid-2009 to 1,200 MW in 2012 during daily peak load hours. This decline in surplus online capacity has affected the market in a number of ways that are discussed throughout the report.

3. Supplemental Commitment for System-Wide Reliability

Given the effect of surplus capacity on prices, it is important to evaluate the supplemental commitments made by the ISO and self-commitments made by market participants after the day-ahead market. Both types of commitments can depress real-time prices inefficiently, while supplemental commitments by the ISO also lead to increased uplift costs.

Although supplemental commitment to satisfy local reliability requirements has fallen in recent years, the ISO has increasingly needed to make supplemental commitments to satisfy New England’s system-wide reliability requirements. Our evaluation indicates that supplemental commitments to meet the system-wide capacity needs increased from under 100 MW before 2009 to an average of 335 MW in 2012.

After reviewing the supplemental commitments and the surplus capacity levels that resulted from real-time operating conditions, we found that roughly 44 percent of the capacity that was supplementally committed in 2012 was actually needed to maintain system level reserves in retrospect.[13] This is not surprising because resource commitments are "lumpy" (i.e., the market cannot commit exactly the quantity it needs) and commitment decisions are often made well in advance when there is significant uncertainty regarding the necessity of the supplemental commitments.

It is also important to recognize that New England has a limited quantity of fast-start resources, which help ensure that sufficient capacity will be available when unexpected conditions arise. The lack of fast-start resources leads the ISO in some cases to rely on slower-starting units that must be notified well in advance of the operating hour when uncertainty regarding load, imports, and generator availability is high. Most of the commitments of slow-starting units are made overnight, more than 12 hours before the forecasted peak. Furthermore, ISO-NE is heavily reliant on gas-fired generating capacity, which can suddenly become unavailable due to the limitations of the natural gas network. Consequently, the ISO may supplementally commit oil-fired and/or dual-fueled capacity in order to protect the system in the event that some generators are unavailable due to limited gas supplies.

4. Uplift Charges

Uplift charges increased from $76 million in 2011 to $99 million in 2012. Several factors contributed to the increase. First, out-of-market capacity payments (i.e., FCM reliability credits, which are payments to rejected delist capacity under FCM) increased from $1.4 million in 2011 to $11.4 million in 2012. In 2011, reliability credits were paid to units in Connecticut because their de-list requests were rejected in the first Forward Capacity Commitment Period (i.e., June 2010 to May 2011) for reliability reasons. Likewise, reliability credits were paid to units in Boston in 2012 because their de-list requests were rejected in the third Forward Capacity

13 This is a simple evaluation that treats any surplus capacity (i.e., the amount of online and available offline capacity in excess of the demand for energy and reserve requirements) as "not needed" for the system. This simple evaluation tends to understate the necessity of supplemental commitments because: 1) the evaluation is based on hourly integrated peak rather than the higher instantaneous peak, and 2) the ISO cannot commit just a portion of a unit. For example, suppose the ISO needs an additional 200 MW of capacity to satisfy system reliability needs and commits the most economic unit with a capacity of 300 MW. In this evaluation, 100 MW of capacity would be deemed as "not needed".

Commitment Period (i.e., June 2012 to May 2013) for reliability reasons.[14] The design issues in the FCM that caused this inconsistency between the ISO's reliability needs and its FCM procurements have been resolved. However, the FCM reliability credits will not be eliminated until June 2016.

Second, the uplift payments for voltage support rose from $6 million in 2011 to $15 million in 2012, which was due primarily to increased reliability commitments for voltage support in Western Central Massachusetts. The increased need for voltage support was attributable to several planned transmission outages, some of which were required to incorporate transmission upgrades to the area.

Third, the "Economic" category of uplift payments associated with non fast-start resources rose from $47 million in 2011 to $54 million in 2012. The increase in supplemental commitment for system-wide reliability needs has led to concomitant increases in the NCPC payments to such resources.

However, these increases were partly offset by lower natural gas prices in 2012, which led to reduced commitment costs for reliability units. Absent the change in fuel prices, the increase in uplift payments would have been more significant.

5. Conclusions: Market Operations

Our assessment of system operations indicates that the ISO's operations to maintain adequate reserve levels in 2012 were reasonably accurate and consistent with the ISO's procedures. Our analyses show that market clearing prices are heavily dependent on the amount of surplus capacity that is available in the real-time market, especially under relatively tight operating conditions. Hence, factors that lead to artificially high levels of surplus capacity tend to:

- Reduce the incentive for units to be available in real time;
- Dampen economic signals to invest in better performance and availability for both new and existing resources.

14 There were no such rejections in the second Forward Capacity Commitment Period.

- Increase large and volatile uplift charges that can be difficult for participants to hedge and which may discourage participation in the ISO-NE market.

To ensure that these issues are minimized, it is beneficial for the ISO to regularly review its assumptions and processes for determining that additional commitments are necessary to satisfy its reliability requirements. In this regard, the ISO should consider modifying the assumptions it makes regarding real-time imports and exports once it implements the CTS process to improve the physical interchange with the NYISO.

The correlation between real-time prices and the amount of surplus capacity also reinforces the importance of:

- Fully reflecting reliability needs in the market requirements for operating reserves. Procuring less operating reserves in the real-time market than needed for reliability increases the apparent surplus capacity amounts and depresses real-time prices. Ultimately, this reduces the incentive for generators to be available in real time; and

- Allowing individual generators to sell only quantities of operating reserves than they are capable of providing. Additional sales artificially raises the apparent real-time supply of operating reserves and tends to depress real-time prices.

The ISO is moving forward on initiatives to address these issues. First, the ISO is proposing to procure "replacement reserves" in the real-time market, which will better enable the real-time prices to reflect reliability concerns that have arisen recently regarding increasing fuel supply uncertainty. Currently, the ISO is proposing a fixed quantity of replacement reserves, although we recommend that the ISO seek authority to modify this quantity daily based on its concerns regarding load and fuel supply uncertainty.

Second, the ISO is revising its procedures for auditing the 10-minute and 30-minute reserve capabilities of off-line and on-line resources to improve their accuracy. This will ensure that the real-time market procures a sufficient quantity of operating reserves and that real-time prices more accurately reflect the cost of maintaining reliability.

Additionally, the ISO is working to provide generators with additional flexibility to modify their offers closer to the real time (i.e., intraday reoffers) to reflect changes in marginal costs. This will provide incentives for generators to be more available because it will better enable them to recover their operating costs. This is particularly important when gas prices are volatile in the hours leading up to real-time. Likewise, for hydroelectric resources with daily output limitations and gas-fired generators with daily fuel consumption limitations, this would provide resources with the ability to modify their offer prices throughout the operating day if their production is substantially higher or lower than anticipated. The ISO is planning to introduce hourly day-ahead energy offers and intraday reoffers as early as the fourth quarter of 2014.[15]

→ *We recommended providing generators with the flexibility to modify their offers closer in the real time (i.e., intraday reoffers) to reflect changes in marginal costs.*

We also recommend changes in Section V that would allow the real-time prices of energy and reserves to better reflect the costs of maintaining reliability during tight operating conditions. Since expectations of real-time prices are the primary determinant of day-ahead prices, these changes should increase the day-ahead market commitment of generators that can satisfy system's reliability criteria.

G. Forward Capacity Market

The FCM was introduced to provide efficient economic signals that augment those provided by the energy and ancillary services markets in order to govern long-term investment and retirement decisions. The FCM consists of annual Forward Capacity Auctions (FCA) held three years in advance of the commitment period when the capacity must be delivered. The first Forward Capacity Auction (FCA 1) was held in February 2008, facilitating the procurement of installed capacity from June 2010 to May 2011. By the end of 2012, six auctions have been held, which had competitive results and satisfied ISO-NE's planning requirements through May 2016.

1. FCM Results

In June 2010, the first Capacity Commitment Period began, allowing for the termination of the individual reliability agreements that had been used extensively to maintain the resource requirements in Connecticut, Boston, and Western Massachusetts. This has significantly improved the efficiency of the long-term incentives for suppliers compared with relying on

15 See 2013 Wholesale Markets Project Plan, page 8.

reliability agreements to retain existing capacity. Unlike markets, reliability agreements do not provide transparent prices indicating the marginal value of capacity in each area.

Each of the six FCAs has procured a significant amount of excess capacity. For example, FCA 6 procured over 36 GW of resources, exceeding the Net Installed Capacity Requirement (NICR) by 2.8 GW. The excess procurements are largely due to the effects of the price floor that prevents capacity prices from falling sufficiently to clear only the minimum requirement. When 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.

2. Facilitating New Investment and Retirement Decisions

The primary goal of deregulated wholesale markets is to facilitate market-based investment in new resources where the investment risks (and potential rewards) are borne by private firms rather than regulated investment, where the risks are borne by captive consumers. Therefore, we evaluate the effectiveness of the ISO-NE markets in facilitating new investment.

In each of the first six FCAs, an average of nearly 1,700 MW of new capacity was procured from generation, demand response resources, and imports. Imports and demand response resources accounted for 83 percent of the procured new capacity collectively, which may decrease in the future when the floor price is no longer used. Generation resources accounted for the remaining 17 percent. However, most of the new investment in generation under FCM has been motivated by out-of-market payments related to RFPs of the Connecticut Department of Public Utility Control (DPUC). A very small amount of new generation has been directly facilitated by the FCM (i.e., generation that was not already committed to enter or that received an award under Connecticut Request For Proposals (RFPs)). This fact alone does not raise any concerns regarding the FCM because there is a substantial surplus of capacity in New England and the prevailing prices in the FCM are well below most estimates of the entry costs for new generation.

It is unlikely that significant generation investment will occur until capacity clearing prices increase significantly. Hence, it will be difficult to determine whether the FCM facilitates efficient market-based investment in new generation until the current surplus of capacity diminishes.

Another goal of these markets is to facilitate the orderly departure of existing resources that are no longer economic to remain in service. However, a large share of the capacity that has attempted to go out-of-service by de-listing has been unable to do so for reliability reasons. The failure of the FCM to allow the departure of these resources (and to facilitate their replacement by new resources) has been due to:

- Inconsistencies between the Local Sourcing Requirements (LSRs) of local capacity zones and the ISO's Transmission Security criteria, which has caused the LSRs to be too low. This was substantially resolved for FCA 4.

- Local capacity zones are not modeled all of the time so the local requirements are not fully reflected in the market's selection of resources and prices. For example, 79 MW of de-list bids were rejected in NEMA in FCA 6 because the LSR could not be satisfied otherwise and the NEMA zone was not modeled. This was addressed in FCA 7 where four zones (including NEMA) were modeled. Beginning in FCA 8, eight zones will always be modeled.

Even with these changes to better reflect locational capacity requirements and the implementation of enhanced market power mitigation measures, we find that the current FCM design is not likely facilitate the efficient entry and exit of resources in New England. Hence, we believe it is critical for the ISO to introduce market reforms to address these issues before the current surplus of capacity declines. To this end, we recommend that the ISO:

→ *Replace the current vertical demand curve with a sloped demand curve that recognizes that excess capacity above the minimum planning reserve requirement provides additional benefits in the forms of increased reliability and lower energy and ancillary services prices.*

→ *Evaluate the interaction of the rules for new suppliers that are related to the Rationing Election and the Capacity Commitment Period Election to determine whether they will promote efficient investment and FCM outcomes over the long-term. These rules designed to encourage new investment can significantly affect the expected FCM outcomes over time and, therefore, its effectiveness in facilitating new investment.* [16]

16 The Rationing Election Allows A New Generating resource to elect to make its offer "rationable", meaning that it need not be wholly accepted. The owner can elect to make the offer rationable down to a specified MW level. The Capacity Commitment Period Election allows a new resource to lock-in the capacity clearing price of the FCA in which it initially sells for a period of up to five years.

3. Performance Incentive Proposal

The ISO has also introduced a Performance Incentive proposal improve suppliers' incentives to be available during 30-minute operating reserve shortages. This proposal will likely achieve the ISO's objective of increasing the incentive for suppliers to be available in real time. We have identified a number of aspects of the proposal that should be further studied, which are discussed in this report. In addition, we encourage the ISO to develop a PPR that does not substantially exceed the expected value of lost load during the shortages in which it will apply.

H. Competitive Assessment

The report evaluates the market concentration and competitive performance of the markets operated by ISO-NE in 2012. Based on our evaluation of the markets in New England we find that the markets performed competitively in 2012.

This competitive assessment has two main components. First, we utilize structural analyses to identify potential market power issues. Second, we evaluate the conduct of market participants in several areas. Although the structural analyses indicate that some suppliers may possess market power under certain conditions, our analyses do not indicate that suppliers withheld resources to raise prices in the ISO-NE markets.

1. Structural Market Power

The structural component of our assessment evaluates each geographic market primarily using a pivotal supplier analysis to determine the demand conditions under which a supplier may have market power. This analysis identifies conditions under which the energy and operating reserve requirements cannot be satisfied without the resources of a given supplier (i.e., the "pivotal supplier"). This is most likely to occur in constrained areas that can become separate geographic markets with a limited number of suppliers when congestion arises. Based on our pivotal supplier analysis, we found that one or more suppliers were pivotal in a large number of hours in 2012 in Connecticut (34 percent of hours), Boston (59 percent of hours), and All of New England (34 percent of hours).

2. Market Participant Conduct

The behavioral component of this assessment examines market participant behavior to identify potential exercises of market power. We analyzed potential economic withholding (i.e., raising offer prices to reduce output and raise prices) and physical withholding (i.e., reducing the claimed capability of a resource or falsely taking a resource out of service). Based on our evaluation in the Competitive Assessment section of this report as well as the monitoring we performed over the course of the year, we find very little evidence of attempts to exercise market power.

While there is no substantial evidence that suppliers withheld capacity from the market to raise clearing prices, suppliers can also exercise market power by raising their offer prices to inflate the NCPC payments they receive when committed for local reliability. Due to the substantial decline in commitments for local reliability in recent years compared to historical levels, this was not a significant concern in 2012.

3. Market Power Mitigation

High levels of structural market power are commonplace in wholesale electricity markets and are usually addressed through effective market power mitigation measures. Such measures address anticompetitive behavior by requiring generators that have the ability to affect LMPs to offer at competitive levels and by deterring generators from physically withholding with the potential for financial sanctions. Hence, it is not surprising that although there is significant structural market power in the New England wholesale market, there is no indication of attempts to exercise market power. Indeed, the market power mitigation measures are an important factor in producing competitive outcomes in the New England wholesale market.

In April 2012, the ISO automated the market power mitigation process in the real-time market. Under the new process (known as the Automated Mitigation Procedure, or “AMP”), the real-time market software performs the test of whether a generator’s offer has a significant effect on the LMP in parallel with the real-time dispatch software. Hence, AMP enables the ISO to identify and prevent the abuse of market power in a more timely and accurate fashion than the previous manual process.

We evaluated the performance of the AMP software after implementation in April 2012. As expected, mitigations rose considerably under the AMP, which occurred more than 250 times in 2012. On average, 120 MW of capacity was mitigated each day. The increased frequency was expected because some offer behaviors that would not trigger mitigation before are now subject to mitigation under the AMP. For example, the manual process would only offers that were marginal (set the real-time energy price). However, the new automated mitigation procedure imposes mitigation all offers that have failed the conduct test and would raise real-time energy prices by more than the applicable mitigation threshold.

We have reviewed these mitigations and find that most were appropriate. However, some of mitigation can be attributed to inaccurate reference levels. Prior to automated mitigation, market participants were under much less pressure to update their reference levels information on a timely basis and consult with the ISO when a reference level becomes inaccurate. The increase in mitigation after the process was automated has prompted participants to consult with the IMM and submit more timely updates to information. Likewise, the IMM has been responsive in working with the participants to improve the reference levels. It has also been working on improving the reference level processes so that they can better handle natural gas price volatility that can cause inappropriate mitigation.

Apart from the automated mitigation, we continue to monitor market outcomes closely for potential economic and physical withholding together with the IMM, and have found little additional conduct that would raise competitive concerns.

I. Table of Recommendations

We make the following recommendations based on our assessment of the ISO-NE's market performance in 2012. A number of these recommendations have been made previously and are now reflected in the ISO's *Wholesale Market Plan*.

Recommendation	Wholesale Mkt Plan	High Benefit[17]	Feasible in ST[18]
Energy Markets			
1. Develop pricing changes to allow the costs of fast-start units and operator actions to maintain reliability (e.g., export curtailments) to be reflected in real-time prices.	✓	✓	
2. Develop pricing changes to allow the costs of deployed demand response resources to be reflected in prices when they are needed to avoid a shortage.			
3. Develop provisions to coordinate the physical interchange between New York and New England in real-time.	✓	✓	
4. Modify allocation of “Economic” NCPC charges to make it more consistent with a “cost causation” principle.	✓	✓	✓
5. Modify inputs to the ex post pricing process to improve consistency with ex ante prices.			✓
6. Provide suppliers with the flexibility to modify their offers closer to real time to reflect changes in marginal costs.	✓	✓	
Reserve Markets			
7. Allow ISO to vary the quantity of replacement reserves in the operating day to improve consistency between the market outcomes and the ISO’s reliability needs.		✓	✓
8. Consider introducing day-ahead operating reserve markets that are co-optimized with the day-ahead energy market.		✓	
Capacity Markets			
9. Replace the current capacity requirement (i.e., vertical demand curve) with sloped demand curve that recognizes the value of additional capacity.		✓	
10. Evaluate the interaction of the Rationing Election and the Capacity Commitment Period Election to determine whether they will promote efficient investment and FCM outcomes over the long-term.			✓

17 Recommendation will likely produce considerable efficiency benefits.

18 Complexity and required software modifications are likely limited.

I. Prices and Market Outcomes

In this section, we review wholesale market outcomes in New England during 2012. This section includes an analysis of overall price trends and a review of prices in transmission congested areas. We also provide an evaluation of the performance of the day-ahead market, which includes analyses of the convergence of day-ahead and real-time markets and of virtual trading patterns.

A. Summary of Prices and Market Outcomes

Both average day-ahead and real-time energy prices fell roughly 22 percent from 2011 to 2012. At the New England Hub, average real-time energy prices fell from approximately $49 per MWh in 2011 to $38 per MWh in 2012, while average day-ahead prices were about 1 percent lower than average real-time prices in both years. The reductions in energy prices from 2011 to 2012 were mostly attributable to the decline in the average price of natural gas, which fell 21 percent from 2011.[19] Several other factors also contributed to the reduction to a lesser extent:

- Average load fell 1 percent and the annual peak load fell 7 percent in 2012, although there were more hours in 2012 when load was at moderately high levels (e.g., 20 GW);
- Net imports from Hydro Quebec and upstate New York rose in 2012 by an average of 310 MW;
- Generation from nuclear units increased in 2012 by an average of 260 MW because they experienced fewer outages in 2012; and
- The entry of a new 620 MW gas-fired combined cycle unit in Connecticut in mid-2011.

Consistent with recent years, New England experienced little congestion in 2012 into historically-constrained areas such as Boston, Connecticut, and Lower Southeast Massachusetts as a result of transmission upgrades that have been made between 2007 and 2009. Most of the price separation between net-exporting regions and net-importing regions was due to transmission losses, rather than transmission congestion. As discussed more fully in Section II, ISO-NE collected day-ahead congestion revenues of only $30 million in 2012, significantly less

19 Natural gas fuels the marginal generation that sets energy prices in most hours. Natural gas prices are based on the indices reported by Platts for the Algonquin pipeline at City Gates.

than the congestion revenues collected in other markets (e.g., more than $300 million in the NYISO and more than $700 million in the MISO).

Differences between day-ahead and real-time prices were moderate in 2012. Average real-time prices were higher than average day-ahead prices by less than 1 percent in 2012, which was a slight improvement from 2011. Good convergence is important because it leads to efficient day-ahead resource commitment, external transaction scheduling and natural gas scheduling. The real-time price premiums that have persisted over the last few years raise efficiency concerns because real-time prices tend to be understated for reasons discussed in this report. In general, it is efficient for average day-ahead prices to exceed average real-time prices by a small margin and for net schedules in the day-ahead ahead market to be close to the actual real-time load. The market response to the real-time premiums that would move toward more efficient day-ahead outcomes is inhibited by the allocation of significant NCPC charges to transactions that improve the day-ahead outcomes (e.g., virtual load).

B. Energy Price Trends

This subsection begins with an examination of the day-ahead prices at the New England Hub.[20] Figure 1 shows the load-weighted average price at the New England Hub in the day-ahead market for each month in 2011 and 2012. The figure also shows the monthly average natural gas price, which should be a key driver of electricity prices when the market is operating competitively.[21]

The figure shows that natural gas price fluctuations were a significant driver of variations in monthly average electricity prices in 2011 and 2012 as expected. In 2012, nearly 45 percent of the installed generating capacity in New England used natural gas as its primary fuel.[22] Low-cost nuclear resources and other baseload resources typically produce at full output, while natural

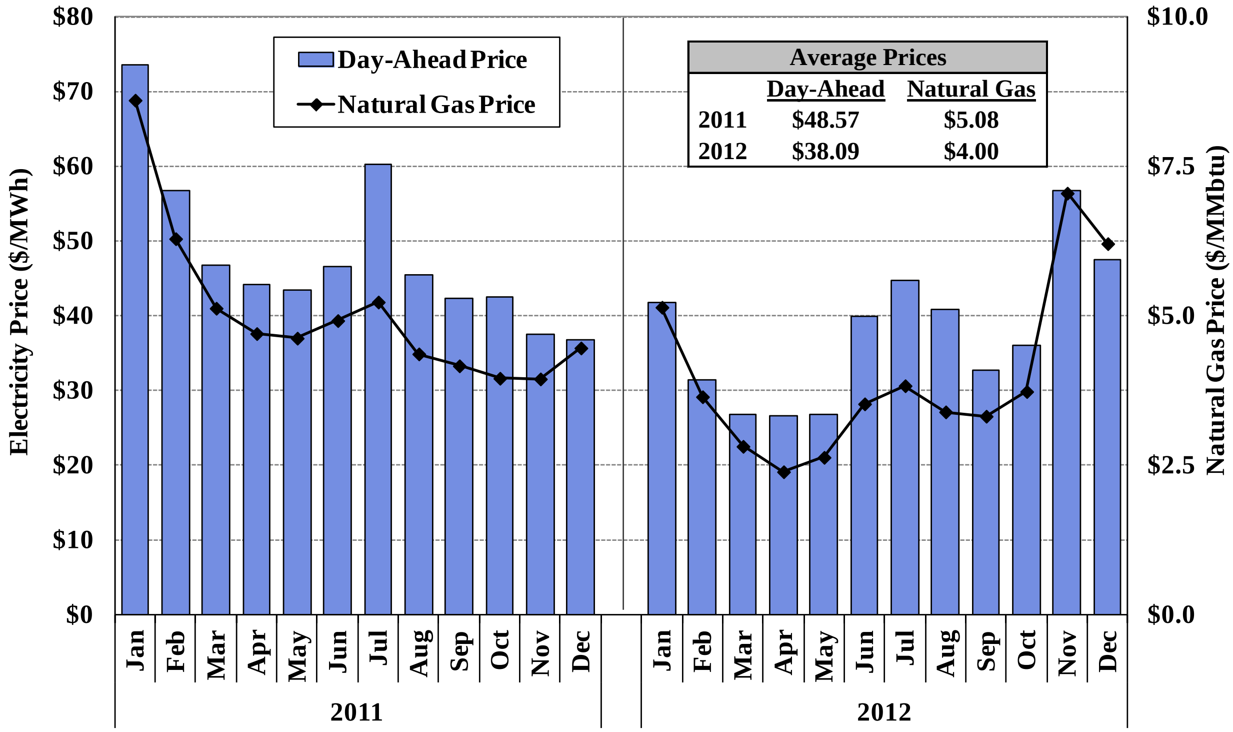
20 The New England Hub is in the geographic center of New England. The Hub price is an average of prices at 32 individual pricing nodes, which has been published by the ISO to disseminate price information that facilitates bilateral contracting. Futures contracts are currently listed on the New York Mercantile Exchange and Intercontinental Exchange that settle against day-ahead and real-time LMPs at the Hub.

21 The figure shows the gas price indices reported by Platts for the Algonquin pipeline at City Gates.

22 ISO-NE, “2012-2021 Forecast Report of Capacity, Energy, Loads, and Transmission (CELT) Report,” May 2012.

gas-fired resources, which accounted for 52 percent of all electricity production in 2012, are on the margin and set the market clearing price in most hours.[23] Therefore, electricity prices should be strongly correlated with natural gas prices in a well-functioning competitive market. Natural gas prices are typically higher during the winter months when heating demands for natural gas increases due to colder weather. Accordingly, natural gas prices decreased from January to March and rose from October to December in both 2011 and 2012, leading to concomitant changes in electricity prices over the same period. However, weather was unseasonably mild from December 2011 to April 2012, contributing to unusually low natural gas and electricity prices in these months.

Figure 1: Monthly Average Day-Ahead Energy Prices and Natural Gas Prices
New England Hub, 2011 – 2012



Energy prices usually increase during high load periods in the summer and winter when the demand for cooling and heating are highest. The effects of seasonal changes in demand were

23 According to preliminary data from EIA Form 923 for 2012, 52 percent of net generation was produced from natural gas, while 30 percent was produced from nuclear fuel, 6 percent from hydroelectric, 9 percent from other renewables sources (including refuse burning), 3 percent from coal, and 0.3 percent from fuel oil.

significant in both years during the summer months. For example, average natural gas prices increased 34 percent in June 2012 from the prior month, while the average electricity prices rose 49 percent as demand increased sharply.

Overall, the average New England Hub price in the day-ahead market decreased 22 percent from 2011 to 2012. The most significant driver of the lower prices was the 21 percent decrease in average natural gas prices from 2011 to 2012. Lower gas prices led to lower energy prices in most hours because natural gas-fired units were frequently on the margin (e.g., roughly 80 percent of all pricing intervals in 2012). Reduced load levels, increased imports, new generating capacity in Connecticut, and fewer outages of nuclear units also contributed to the reduction in prices. These factors are discussed in more detail below.

To better identify changes in energy prices that are not related to the fluctuations in natural gas prices, Figure 2 shows the marginal heat rate that would be implied if natural gas resources were always on the margin. The implied marginal heat rate is equal to the energy price divided by the natural gas price measured in MMbtu. Thus, if the electricity price is $72 per MWh and the natural gas price is $9 per MMbtu, this would imply that an 8.0 MMbtu per MWh generator is on the margin. Figure 2 shows the load-weighted average implied marginal heat rate for the New England Hub in each month during 2011 and 2012.

The implied marginal heat rate shows more clearly the seasonal variation in electricity prices due to other factors. The figure shows that implied marginal heat rates were highest in the peak summer months. This was due primarily to the higher loads and tighter market conditions that prevail on the hottest days during the summer. In 2012, the months with the highest average implied marginal heat rate were July and August, which were also the months with the hottest temperatures and highest average loads. The variations in load levels are discussed in more detail in the next sub-section.

Figure 2: Monthly Average Implied Marginal Heat Rate
Based on Day-Ahead Prices at New England Hub, 2011 – 2012

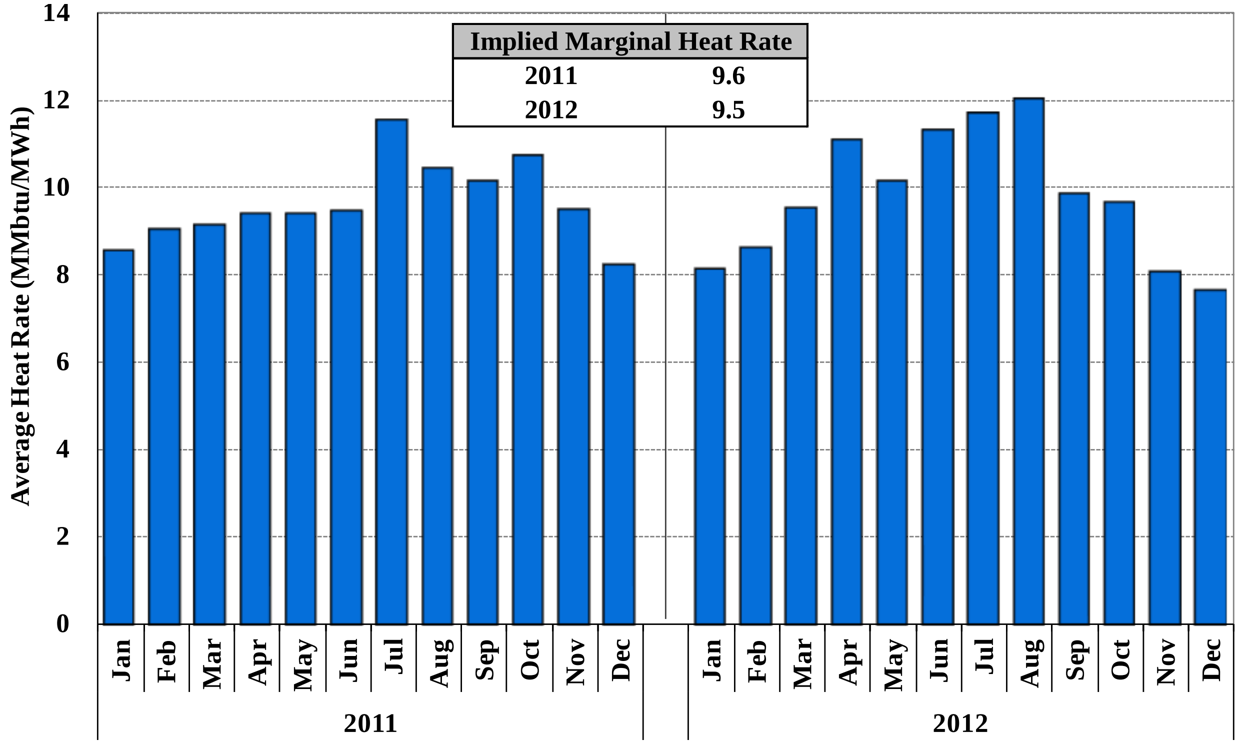


Figure 2 also shows that the average implied marginal heat rate fell approximately 1 percent from 2011 to 2012. Given that a small share of generation costs (e.g., variable operating and maintenance expenses) is not related to fuel prices, reductions in fuel prices generally cause the non-fuel costs to increase as a share of total generation costs. Consequently, implied heat rates rise when natural gas prices fall substantially. Therefore, the reduction in marginal heat rates in 2012 indicates that factors other than lower natural gas prices have also contributed to the decrease in energy prices, including:

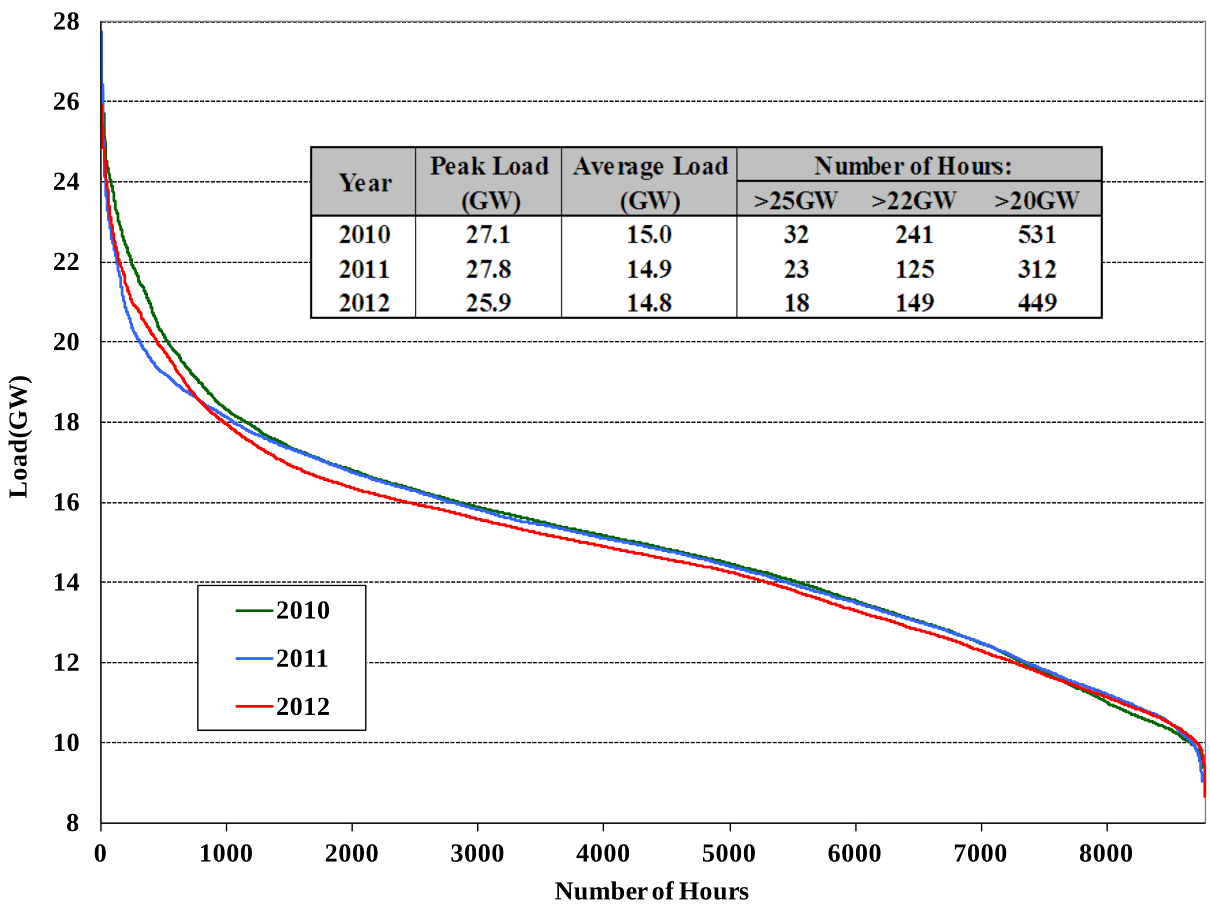
- An average load decrease of 1 percent from 2011 to 2012.
- Higher net imports in 2012 from neighboring areas, particularly Hydro Quebec and Upstate New York. Total net imports averaged approximately 1,440 MW over all hours in 2012, up nearly 300 MW from 2011.
- Increased generation from nuclear units in 2012. Fewer outages occurred on nuclear units in 2012, leading the average output from these units to rise by 260 MW from 2011.

C. Energy Demand

In addition to fuel price changes, changes in electric supply and demand also contribute to price movements in New England. The amount of available supply changes slowly from year to year, so fluctuations in electricity demand explain much of the short-term variations in electricity prices. The hours with the highest loads are important because a disproportionately large share of the market costs to consumers and revenues to generators occur in these hours.

Figure 3 illustrates the variation in demand during the year by showing load duration curves for each of the last three years. Load duration curves show the number of hours on the horizontal axis in which the system-wide load was greater than or equal to the level shown on the vertical axis. For each of the last three years, the table in the figure shows the average load level, the peak load level, and the number of hours when the system was in high load conditions.

Figure 3: Load Duration Curves
2010 – 2012



Year	Peak Load (GW)	Average Load (GW)	Number of Hours: >25GW	>22GW	>20GW
2010	27.1	15.0	32	241	531
2011	27.8	14.9	23	125	312
2012	25.9	14.8	18	149	449

In general, electricity demand grows slowly over time, tracking population growth and economic activity. However, the figure shows that average load declined modestly from 2010 to 2011 and again from 2011 to 2012. These declines were largely due to the changes in weather patterns in these years. In particular, weather has been milder in the past two years, especially in the winter months.

The frequency of extreme high load conditions declined in 2012. Load peaked at 25.9 GW on July 17, which was 4 percent lower than the annual peak in 2010 and 7 percent lower than the annual peak in 2011. Likewise, load exceeded 25 GW in 18 hours during 2012, down from 23 hours in 2011 and 32 hours in 2012. As a result, there were less frequent shortage conditions in 2012, resulting in less severe real-time price spikes. These variations were generally consistent with the weather patterns in the summer months over the past three years.

D. Prices in Transmission Constrained Areas

ISO-NE manages flows over the network to avoid overloading transmission constraints by altering the dispatch of its resources and establishing locational marginal prices (LMPs) to establish efficient, location-specific prices that are consistent with the marginal costs of serving load at that location. Transmission congestion arises because the lowest-cost resources cannot be fully dispatched due to limited transmission capability. The LMPs can vary substantially across the system, reflecting the fact that higher-cost units must be dispatched in place of lower-cost units to serve incremental load while not overloading any transmission facilities. This causes LMPs to be higher in “constrained locations”. In addition, transmission constraints may also require additional operating reserves in certain locations to maintain reliability. When generation is redispatched in real time to provide additional reserves to a local area, the marginal system cost of the redispatch is reflected in the LMPs. The reserve markets are discussed in Section III.

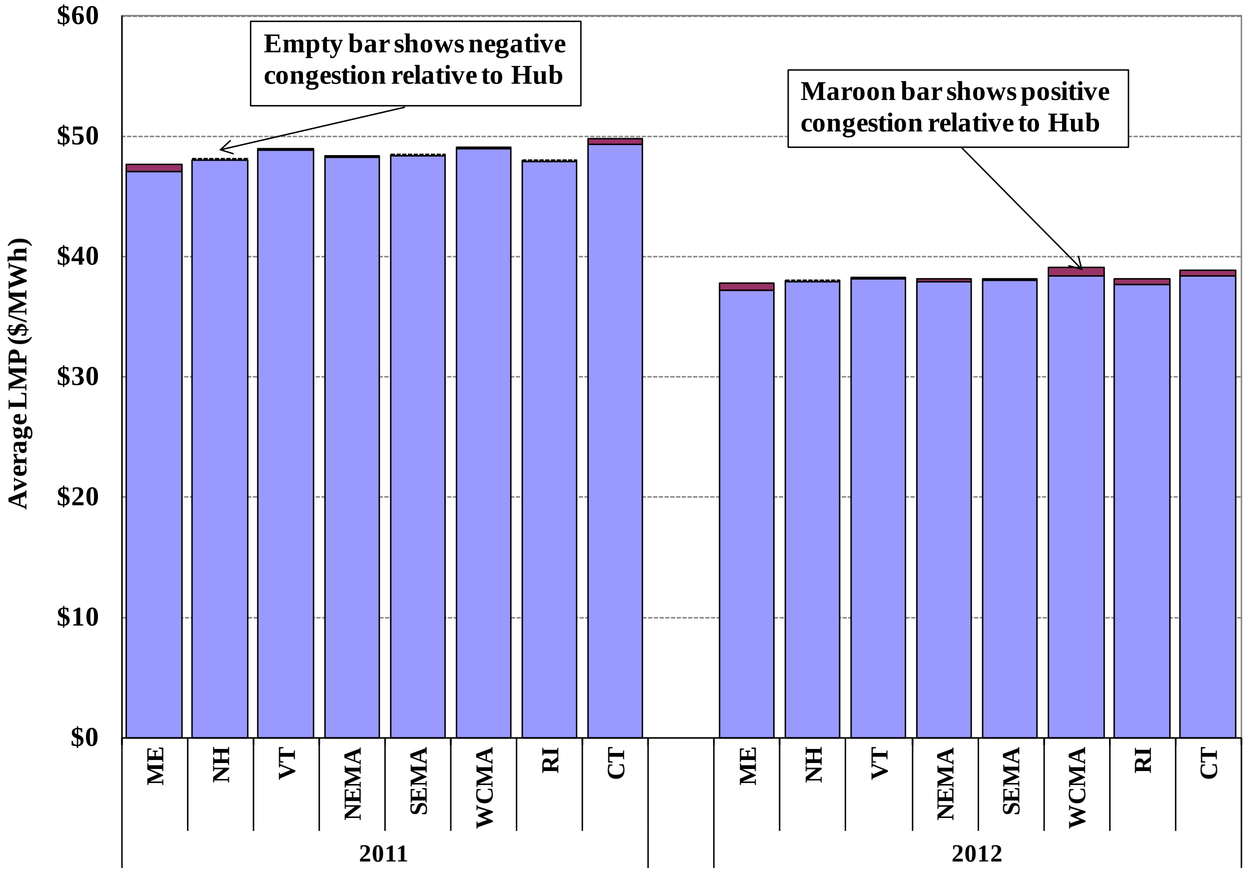
LMPs also reflect the marginal value of transmission losses. Transmission losses occur whenever power flows across the transmission network. Generally, transmission losses increase as power is transferred over longer distances or as the power flows increase, and are higher on lower voltage facilities (for a given amount of power transfer).

Historically, there have been significant transmission limitations between net-exporting and net-importing regions in New England. In particular, exports from Maine to the rest of New England have been limited by transmission constraints at times, while Connecticut and Boston were often unable to import enough power to satisfy demand without dispatching expensive local generation in the past. However, congestion has been very limited in recent years because of the transmission upgrades made in Boston, Connecticut, and Southeast Massachusetts from 2007 to 2009. These upgrades greatly increased the transfer capability into these areas and eliminated most of the congestion into these historically constrained regions. Consequently, the current levels of LMPs do not provide significant incentives for locating new resources in net-importing regions, such as Connecticut and Boston.

We examined the differences in energy prices across the system during the study period. Figure 4 shows load-weighted average day-ahead LMPs in 2011 and 2012 for the eight load zones in New England.[24] For each location, the load-weighted average LMP is indicated by the height of the solid bars. The maroon portion of the bars indicates positive congestion to the location from the New England Hub, while negative congestion is indicated by the empty bars. The blue bars show the portion of the LMP that is based on value of energy and the marginal effect from transmission losses at each zone. Thus, zones that are import-constrained (e.g., the Connecticut load zone) exhibit positive congestion from the Hub.

24 New England is divided into the following eight load zones: Maine (ME), New Hampshire (NH), Vermont (VT), Rhode Island (RI), Connecticut (CT), Western/Central Massachusetts (WCMA), Northeast Massachusetts and Boston (NEMA), and Southeast Massachusetts (SEMA).

Figure 4: Average Day-Ahead Prices by Load Zone
2011 – 2012



The figure shows that congestion levels remained very low in 2012, although it rose slightly from in 2012, particularly in Western/Central Massachusetts. The increase was due primarily to more frequent peaking conditions in 2012 and effects of some transmission outages. These outages are discussed in more detail in Section II.

Transmission losses accounted for most of the locational differences in LMPs in both 2011 and 2012. The largest average congestion-price differential between load zones was about 1.2 percent of the average LMP in 2011 and 1.8 percent in 2012.[25] On the other hand, the average loss-related price differentials between load zones were generally more significant. The largest

25 The congestion-price differentials are reported between $0.01 and $0.61 per MWh in 2011 and between $0.04 and $0.72 per MWh in 2012.

such price differential between load zones was 4.5 percent of the average LMP in 2011 and 3.0 percent in 2012.[26]

E. Day-Ahead Market Performance: Convergence with Real-Time Prices

The day-ahead market allows participants to make forward purchases and sales of power for delivery in real time. This provides a valuable financial mechanism that allows participants to hedge their portfolios and manage risks associated with the real-time market. Loads can hedge price volatility in the real-time market by purchasing in the day-ahead market. Suppliers can avoid the risk of unprofitably starting their generators, because the day-ahead market will accept their offers only when they will profit from being committed. However, suppliers that sell day-`ahead are exposed to some risk because they are committed to deliver energy in the real time. An outage or failure to secure fuel can force them to purchase replacement high-priced energy from the spot market. In addition to the value it provides to individual market participants, perhaps the greatest value of the day-ahead market is that it coordinates the overall commitment of resources that are used to satisfy the next day's needs at the lowest cost.

In well-functioning day-ahead and real-time markets, we expect that day-ahead and real-time prices will not systematically diverge by a substantial amount. If day-ahead prices were predictably higher or lower than real-time prices, participants should adjust their purchases and sales in the day-ahead market to bring the prices into convergence. However, day-ahead prices tend to be slightly higher than real-time prices in a well-functioning energy market because many buyers are willing to pay a small premium for day-ahead purchases to avoid the more volatile real-time prices.

Good convergence between day-ahead and real-time prices is important. The day-ahead market facilitates most of the generator commitments in New England. Hence, good price convergence with the real-time market helps ensure that resources are committed efficiently to satisfy the anticipated real-time operating needs of the system. Additionally, most settlements occur through the day-ahead market and its results are the basis for payments to FTR holders.

---

26 The transmission loss-price differentials are reported between $0.15 and $2.17 per MWh in 2011 and between $0.08 and $1.17 per MWh in 2012.

Persistent differences between day-ahead and real-time prices can undermine incentives for suppliers to offer their resources at marginal cost in the day-ahead market. We expect random variations resulting from unanticipated changes in hourly supply and demand between the two markets, but persistent price differences would raise potential concerns.

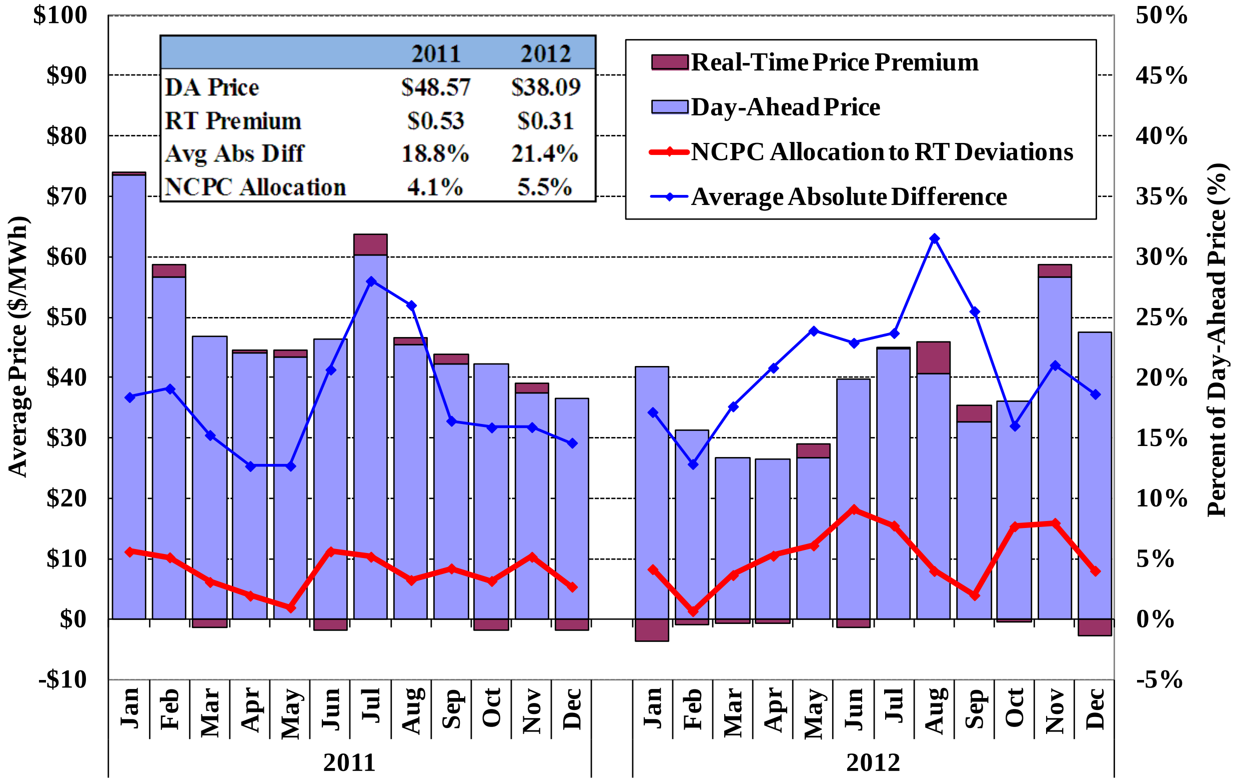
Since there was little congestion in the system, price convergence between the day-ahead and real-time markets at the New England Hub provides an indication of the overall price convergence. In this section, two measures are used to assess price convergence. The first measure reports the simple difference between the average day-ahead price and the average real-time price. The second measure reports the average absolute difference between day-ahead and real-time prices on an hourly basis. The first measure is an indicator of the systematic differences between day-ahead and real-time prices. This is the most important measure because it indicates whether the day-ahead prices reflect an accurate expectation of real-time prices. The second measure captures the overall variability between day-ahead and real-time prices.

Figure 5 summarizes day-ahead prices and the convergence between day-ahead and real-time prices at the New England Hub in each month of 2011 and 2012.[27] The first measure of convergence reported in the figure, the average real-time premium, is equal to the average real-time price minus the average day-ahead price. The sum of the average day-ahead price (blue bar) and the average real-time price premium (maroon bar) is equal to the average real-time price. The second measure of convergence, the average absolute difference between day-ahead and real-time prices, is shown by the blue line and is reported as a percentage of the average day-ahead price in the month. The figure also shows the monthly average rate of Net Commitment Period Compensation (NCPC) that is charged to real-time deviations, which is shown by the red line and is also reported as a percentage of the average day-ahead price in each month. The inset table compares these quantities on an annual basis for 2011 and 2012.

Figure 5 shows that the market generally exhibited a real-time premium in 2011 and 2012, although there were individual months exhibiting a day-ahead premium. Both years exhibited a real-time price premium overall of approximately 1 percent of the average day-ahead price.

27 Day-ahead and real-time prices are averaged on a load-weighted basis.

Figure 5: Convergence of Day-Ahead and Real-Time Prices at New England Hub
2011 – 2012



We do not believe that persistent real-time price premiums are efficient because small day-ahead premiums generally lead to a more efficient commitment of the system's resources. Section V shows that real-time energy prices frequently do not reflect the full costs of the marginal source of supply. For example, when high-cost peaking resources are committed to satisfy the real-time demand, real-time prices generally do not reflect the full costs of such resources. Because the real-time prices are understated in these cases, day-ahead prices would have to be slightly higher than the actual real-time prices in order to efficiently facilitate a day-ahead commitment of resources to fully satisfy the real-time system needs.

One reason for the real-time premiums in the past few years is the significant increase in average allocation of NCPC charges after May 2010, which is discussed in detail in the next subsection. The increased allocation of NCPC charges (per MWh) to virtual load in particular has likely inhibited the natural market response to the sustained real-time price premiums. Hence, we recommend discontinuing the allocation of real-time NCPC charges to virtual load and other deviations that generally do not cause real-time NCPC charges.

The second measure of price convergence evaluated in the figure is the average absolute difference between day-ahead and real-time prices, which is calculated by averaging the absolute value of the hourly differences between day-ahead and real-time prices on a load-weighted basis. As a percentage of the average day-ahead price in each year, the average absolute difference has increased from 18.8 percent in 2011 to 21.4 percent in 2012. These levels are higher than in the period from 2005 to 2009, when the annual average absolute difference ranged from 14.9 percent to 17.9 percent. Higher price volatility has contributed to this increase. This higher volatility occurred because the ISO must make fewer supplemental commitments for reliability, allowing it to operate with lower surplus capacity in real-time and generate far less NCPC uplift. There has also been a substantial reduction in participation by virtual traders since the beginning of 2010, which has likely contributed to the increased differentials between day-ahead and real-time prices.

F. Virtual Trading Activity and Profits

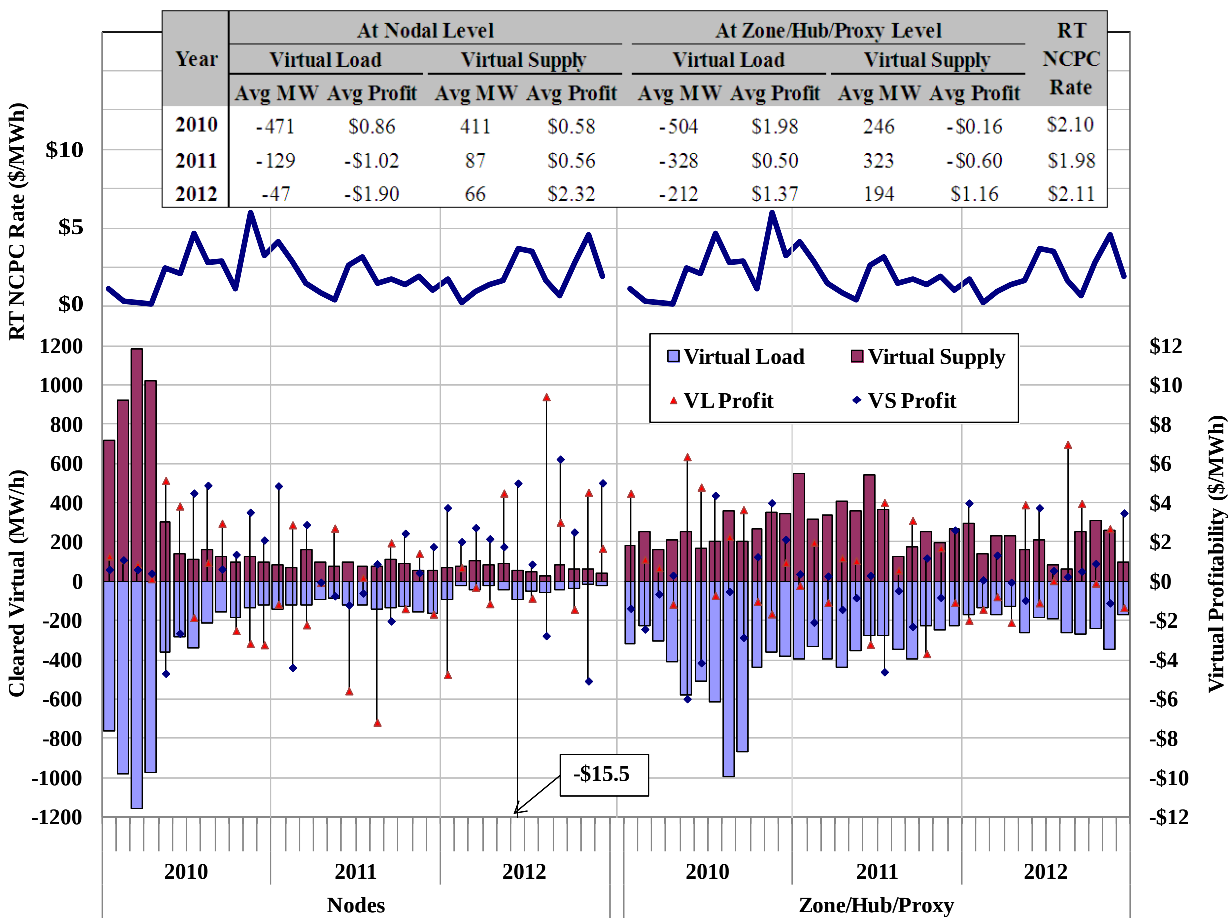
Virtual trading plays a key role in the day-ahead market by improving price convergence between day-ahead and real-time markets, thereby promoting efficient commitment and scheduling of resources in the day-ahead market. Virtual trading in the day-ahead market consists of purchases or sales of energy at individual nodes, zones or the NE Hub that are not associated with physical load or physical resources. Virtual bids and offers provide liquidity to the day-ahead market, constituting a large share of the price-sensitive supply and demand and facilitate efficient day-ahead prices.

Virtual transactions that are scheduled in the day-ahead market are settled against real-time energy prices. Virtual demand bids are profitable when the real-time energy price is higher than the day-ahead price, while virtual supply offers are profitable when the day-ahead energy price is higher than the real-time price. Accordingly, if prices are lower in the day-ahead market than in the real-time market, a virtual trader may purchase energy in the day-ahead market and sell it back in the real-time market. This will tend to increase day-ahead prices and improve price convergence with the real-time market. Hence, profitable virtual transactions improve the performance of the day-ahead market.

Figure 6 shows the average volume of virtual supply and demand that cleared the market in each month of 2010 to 2012 by location, as well as the monthly average gross profitability of virtual purchases and sales. Gross profitability is the difference between the price at which virtual traders bought and sold energy between the day-ahead and real-time market. These quantities are shown separately for transactions at individual nodes and transactions at aggregated locations (i.e., the New England Hub, load zones, and external proxy buses). The gross profitability shown here does not account for NCPC cost allocations. The upper portion of the figure shows the average real-time NCPC rate for each month.[28]

Figure 6: Virtual Transaction Volumes and Profitability

By Month and By Location, 2010 – 2012



Year	At Nodal Level				At Zone/Hub/Proxy Level				RT NCPC Rate
	Virtual Load		Virtual Supply		Virtual Load		Virtual Supply		
	Avg MW	Avg Profit	Avg MW	Avg Profit	Avg MW	Avg Profit	Avg MW	Avg Profit	
2010	-471	$0.86	411	$0.58	-504	$1.98	246	-$0.16	$2.10
2011	-129	-$1.02	87	$0.56	-328	$0.50	323	-$0.60	$1.98
2012	-47	-$1.90	66	$2.32	-212	$1.37	194	$1.16	$2.11

28 The monthly real-time NCPC rate is defined as the total NCPC charges allocated system wide divided by the total real-time deviations for each month.

The figure shows that scheduled virtual transactions have decreased substantially over the past three years, particularly at the nodal level. Overall, scheduled virtual load fell 53 percent from 2010 to 2011 and fell 43 percent from 2011 to 2012. Likewise, scheduled virtual supply fell 38 from 2010 to 2011 and fell 36 percent from 2011 to 2012. The decline in virtual trading volumes was most significant at the nodal level, where scheduled virtual load and virtual supply fell 90 percent and 84 percent from 2010 to 2012, respectively.

The substantial drop in virtual transactions at the nodal level began in May 2010 when the ISO deployed a software solution to address an inconsistency in loss modeling at certain locations. This modeling inconsistency had motivated a significant quantity of virtual trading at the affected locations where such trades produced low levels of consistent virtual profits (due to predictable differences between day-ahead and real-time LMPs). Hence, when this inconsistency was remedied, the associated virtual trading at those nodes ceased. More recently, FERC enforcement actions against virtual traders in a number of markets have likely increased the perceived regulatory risks associated with virtual trading and contributed to the reduction in activity. Finally, the rising NCPC rates have likely also contributed to this reduction, which is described below.

Virtual Trading Profits

Figure 6 shows that virtual trading was generally profitable in 2012 (before including NCPC charges) with an overall net profit of $5 million, indicating that virtual trading improved convergence between day-ahead and real-time prices.[29] This is because virtual trades that are profitable (before including NCPC charges) generally contribute to better convergence between day-ahead and real-time prices.

However, when NCPC charges are considered, overall virtual trading was unprofitable with an overall net loss of $4 million.[30] The effects of NCPC allocations on virtual trading profits, and

29 Not including NCPC charges, profits can be tabulated for each category of virtual transactions in the figure by multiplying "Avg MW" and "Avg Profit" by the number of hours (8760 or 8784 for 2012) in each year.

30 Virtual transactions would net a loss on average after paying NCPC charges in each year of 2010 to 2012 (e.g., gross profitability of all cleared virtual transactions was $1.11 per MWh in 2012 compared to an average real-time NCPC charge rate of $2.11 per MWh in 2012).

ultimately virtual trading activity, have increased in recent years. This issue is discussed in the following sub-section.

NCPC Allocation and Virtual Trading

Since real-time NCPC charges are allocated across virtual transactions and other Real-Time Deviations, the reduced volume of nodal virtual trading has resulted in higher NCPC charges to the remaining real-time deviations since May 2010. The real-time NCPC rate increased substantially from $0.46 per MWh in the first four months of 2010 to $3.60 per MWh in the last eight months of 2010. The rate has remained high, averaging roughly $2 per MWh in 2011 and 2012. Additionally, supplemental commitment for system-wide reliability has increased in recent years, contributing to elevated NCPC rates as well. High NCPC rates provide a significant disincentive for firms to schedule virtual transactions because virtual profits tend to be relatively low. Hence, it is likely that the increased NCPC rates have reduced contributed to the reduction in virtual trading activity and, ultimately the consistency between day-ahead and real-time prices.

ISO-NE currently allocates nearly all real-time "Economic" NCPC charges to deviations between the day-ahead and real-time schedules.[31] In reality, some deviations are "harming" and tend to increase NCPC, while others are "helping" and reduce NCPC. For example, underscheduling physical load in the day-ahead market can cause the ISO to commit additional units in real-time, which are likely to increase NCPC—this is a "harming" deviation. Conversely, "helping" deviations, such as over-scheduling load (including virtual load), generally result in higher levels of resource commitments in the day-ahead market and, therefore, usually decrease the ISO's need to make additional commitments, thereby avoiding NCPC. The current allocation does not distinguish between helping and harming deviations and is, therefore, not consistent with cost causation. Hence, this allocation assigns NCPC charges to transactions that actually tend to *reduce* the need for supplemental commitments, including virtual load.

31 Real-Time Deviations include Real-Time Load Obligation Deviations, which are positive or negative differences between day-ahead scheduled load and actual real-time load, uninstructed generation deviations from day-ahead schedules, virtual load schedules, and virtual supply schedules.

NCPC charges are caused by many factors other than real-time deviations, including: peaking resources not setting real-time prices, operator actions to satisfy system reliability needs, and unforeseen events such as outages. Hence, we find that the current allocation scheme over-allocates costs to deviations relative to the portion of the NCPC they likely cause. This is particularly true of virtual load transactions, which tend to increase day-ahead commitments and, therefore, decrease the need for supplemental commitments. Given that real-time price premiums prevailed for much of 2010 through 2012, allocating substantial NCPC costs to virtual load that does not cause these costs has likely degraded the performance of the day-ahead market.

Hence, we recommend that the ISO modify allocation of Economic NCPC charges to be more consistent with a "cost causation" principle, which would generally involve not allocating NCPC costs to virtual load and other real-time deviations that cannot reasonably be argued to cause real-time economic NCPC. We continue to work with the ISO and its IMM to develop changes to the NCPC allocation that would address this issue and improve the incentives for efficient day-ahead scheduling by market participants.

G. Assessment of Virtual Trading Efficiency

The reduction in virtual trading activity that was discussed in the previous sub-section raises potential concerns regarding the efficiency of the day-ahead market because active virtual trading in the day-ahead market promotes price convergence with the real-time market. Good price convergence, in turn, facilitates an efficient commitment of generating resources, lowering the costs of satisfying the system's needs in real time. Active virtual traders also protect the day-ahead market against market manipulation and market power abuses, since they make it more difficult for a single firm to affect day-ahead clearing prices by submitting uneconomic bids and/or offers in the day-ahead market.

Just as profitable virtual trades contribute to better convergence between day-ahead and real-time prices, unprofitable virtual trades tend to degrade price convergence. Hence, uneconomic virtual transactions may indicate an attempt to manipulate prices in the day-ahead market, so we routinely evaluate the virtual trading activity of firms that exhibit a pattern of unprofitable virtual trading.

In our review of unprofitable virtual transactions in 2012, we found that the firms with the most significant virtual losses generally scheduled virtual transactions that hedged their exposure to fluctuations in real-time prices. Hence, we did not find that the unprofitable virtual transactions were anti-competitive or manipulative. Nonetheless, we and the Internal Market Monitor continue to screen market outcomes for potentially manipulative conduct. In the long-run, the best safeguard against manipulative conduct is a liquid market with participation by a large number of firms. A liquid market is relatively resistant to attempts by a single firm to push day-ahead above or below competitive levels. This highlights the importance of changing the allocation of NCPC charges improve the incentives for efficient participation by virtual traders and other firms in the day-ahead market.

H. Conclusion

Energy prices decreased 22 percent in 2012, driven primarily by substantially lower natural gas prices. Other factors, including reduced average load levels, increased net imports from neighboring areas, and increased production from nuclear units, also contributed to the decline in energy prices. Relatively little transmission congestion occurred as the transmission investments made between 2007 and 2009 continued to provide excess transfer capability into historically constrained areas in the vast majority of hours.

Differences between day-ahead and real-time prices were relatively small in 2012, but the sustained real-time price premiums have raised a potential concern that the market is unable to quickly adjust to the higher real-time prices. These market outcomes are consistent with the inefficient allocation of real-time NCPC costs to virtual load and other real-time deviations. Therefore, we recommend that the ISO revise the allocation methodology for Economic NCPC, making it more consistent with cost causation principles.

II. Transmission Congestion and Financial Transmission Rights

Congestion arises when the transmission network does not have sufficient capacity to dispatch the least expensive generators to satisfy the demands of the system. When congestion occurs, the market software establishes clearing prices that vary by location to reflect the cost of meeting load at each location. These Locational Marginal Prices ("LMPs") reflect the economic value of binding transmission constraints in causing the dispatch of higher-cost generation to manage the flow over the constrained transmission facility. These prices also establish long-term economic signals that govern investment in generation, transmission, and demand response resources. Hence, a primary focus of this report is to evaluate locational marginal prices and associated congestion costs.

Congestion costs are incurred in the day-ahead market based on the modeled transmission flows resulting from the day-ahead energy schedules. These costs result from the difference in prices between the points where power is consumed and generated on the network. A price difference due to congestion indicates the gains in trade between the two locations if additional transmission capability were available. Hence, the difference in prices between the locations represents the marginal value of transmission. The differences in locational prices caused by congestion are revealed in the congestion component of the LMP at each location.[32]

Market participants can hedge congestion charges in the day-ahead market by owning Financial Transmission Rights (FTRs).[33] An FTR entitles a holder to payments corresponding to the congestion-related difference in prices between two locations in a defined direction. For example, a participant that holds 150 MW of FTRs from point A to point B is entitled to a payment equal to 150 times the locational energy price at point B less the price at point A (a negative value means the participant must pay) assuming no losses. Hence, a participant can hedge the congestion costs associated with a bilateral contract if it owns an FTR between the same receipt and delivery points as in the bilateral contract.

32 The congestion component of the LMP represents the difference between the marginal cost of meeting load at that location versus the marginal cost of meeting load at a reference location, not including transmission losses.

33 FTRs can also be used as speculative investments for purchasers who forecast higher congestion revenues between two locations than the cost of the associated FTR.

Through the auctions it administers, ISO-NE sells FTRs with one-year terms (annual FTRs) and one-month terms (monthly FTRs). The annual FTRs allow market participants greater certainty by allowing them to lock-in congestion hedges further in advance. ISO-NE auctions 50 percent of the forecasted capacity of the transmission system in the annual auction, and all of the remaining capacity in the monthly auctions.[34] FTRs are auctioned separately for peak and off-peak hours.[35]

In this section, we summarize congestion costs and assess two aspects of the performance of the FTR markets. First, we evaluate the net payments to FTR holders, which increased approximately 55 percent from 2011 to 2012, consistent with the overall increase in congestion in the day-ahead market. Payments to FTR holders are funded by the congestion revenue collected by ISO-NE. In 2012, the congestion revenue collected by ISO-NE was sufficient to satisfy 100 percent of the obligations to FTR holders (referred to as the “target payment amount”).

Second, we compare FTR prices with congestion prices in the day-ahead and real-time markets. Since FTR auctions are forward financial markets, FTR prices should reflect the expectations of market participants regarding congestion in the day-ahead market. In 2012, FTR prices in the monthly auctions were more consistent with congestion values in the day-ahead and real-time markets than FTR prices in the annual auction. The improvement in consistency of FTR prices and congestion values from the annual auction to the monthly auctions is expected because market participants gain more accurate information about market conditions as the lead time for the auction decreases.

A. Congestion Revenue and Payments to FTR Holders

As discussed above, the holder of an FTR from point A to point B is entitled to a payment equal to the value of the congestion between the two points. The payments to FTR holders are funded

34 In the annual auction the ISO awards FTRs equivalent to 50 percent of the predicted power transfer capability of the system, and in the monthly auctions the ISO awards FTRs equivalent to 100 percent of the remaining predicted power transfer capability after accounting for planned transmission outages. See generally, the ISO-NE Manual for Financial Transmission Rights, Manual M-06.

35 Peak hours include hours ending 8 to 23, Monday through Friday, not including NERC holidays. Off-peak includes all other hours.

from the congestion revenue fund, which is primarily generated from congestion revenue collected in the day-ahead market.

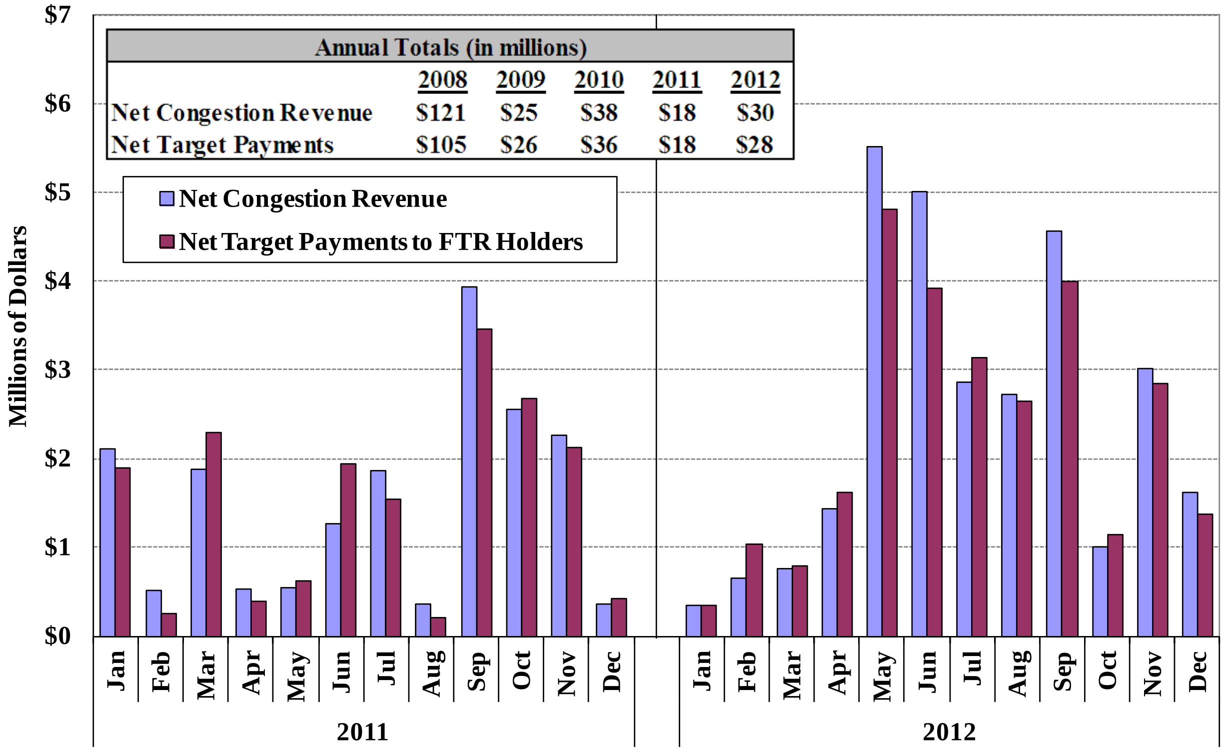
Day-ahead congestion revenue is equal to the megawatts scheduled to flow across a constrained transmission path times the day-ahead shadow price (i.e., the marginal economic value) of the transmission path. Real-time congestion revenue is equal to the change in scheduled flows (relative to the day-ahead market) across a constrained transmission path times the real-time shadow price of the transmission path. When a real-time constraint binds at a limit that is less than the scheduled flows in the day-ahead market, it results in negative congestion revenue.[36] These costs are generally recovered as a form of uplift.

When the total congestion revenue collected by the ISO-NE is not sufficient to satisfy the targeted payments to FTR holders, it implies that the quantities sold in the FTR auctions exceeded the actual capability of the transmission system. In months when this occurs, the unpaid FTR amounts are accrued until the end of the year when any excess congestion revenues remaining from months with a surplus are used to pay amounts accrued from months with a shortage, plus interest. If the end-of-year surplus is less than the total accrued shortfall amounts, the end-of-year payments on shortfall amounts are discounted pro rata. If the surplus is greater than the total accrued shortfall amounts, the excess congestion revenues are returned to transmission customers per the Tariff.

Figure 7 compares the net congestion revenue collected by the ISO-NE with the net target payments to FTR holders in each month of 2011 and 2012. The inset table compares the two quantities in the past five years. Net congestion revenue includes the sum of all positive and negative congestion revenue collected from the day-ahead and real-time markets. Net target payments to FTR holders include the sum of all positive target payments to FTR holders and all negative target payments (i.e., payments from FTR holders).

36 For example, suppose 100 MW is scheduled to flow across an interface in the day-ahead market in a given hour, and the interface is constrained when 90 MW is scheduled to flow across it in the real-time market (due to a reduction in transmission capability after the day-ahead market). If the real-time shadow price of the constraint is $50 per MWh, the 10 MW flow reduction from the day-ahead to the real-time market will result in negative $500 of congestion revenue for the hour.

Figure 7: Congestion Revenue and Target Payments to FTR Holders
2011 – 2012



The net congestion revenue rose by 62 percent from approximately $18 million in 2011 to $30 million in 2012. Likewise, the net target payments to FTR holders increased from $18 million in 2011 to $28 million in 2012. The increase in congestion in 2012 was due to several factors:

- The frequency of peaking conditions increased from 2011 to 2012, leading to more frequent congestion into import-constrained areas, particularly in the summer months. There were 449 hours in 2012 when the load exceeded 20 GW, compared to 312 such hours in 2011.
- Congestion became more frequent in some areas when planned transmission outages substantially affected the network capability into these areas. For example, LMPs were significantly elevated in Western Central Massachusetts on many days in May when the planned outages of two 115 kV transmission lines caused substantial congestion. Similarly, congestion into Maine increased significantly on several days in September that was driven largely by the planned outage of the 345 kV Buxton-Deerfield line.
- Natural gas prices rose substantially in November and December 2012, increasing redispatch costs and associated congestion-related price differences. As a result, congestion costs were higher than in the same months of 2011.

The inset table shows that net congestion revenues fell substantially from 2008 to 2009 and have remained relatively low in each subsequent year. The decline in 2009 was due primarily to transmission upgrades in Boston, Connecticut, and Southeast Massachusetts that were completed from 2007 to 2009. The patterns of congestion are evaluated in greater detail in the next subsection.

The figure also shows that net congestion revenues exceeded net target payments to FTR holders in at least half of the months in 2011 (7 months) and in 2012 (6 months). As a result, the total net congestion revenues for the 12 months in both 2011 and 2012 were sufficient to fund 100 percent of the net target payments. The correspondence of FTR obligations and day-ahead congestions indicates that ISO-NE has modeled the transmission system consistently in the FTR market and day-ahead markets. This consistency can be difficult to achieve because transmission outages and network flows caused by those on other areas (i.e., "loop flows") can be difficult to predict when the FTR auctions are being run.

B. Congestion Patterns and FTR Prices

In this section, we evaluate the performance of the FTR markets by comparing the FTR prices to the congestion prices in the day-ahead and real-time markets. FTR auctions take place in the prior month (for monthly auctions) or at the end of the preceding year (for annual auctions). Prices in the FTR auctions reflect the expectations of market participants regarding congestion in the day-ahead market. When the market is performing well, the FTR prices should converge over time with the actual congestion on the network.

Figure 8 shows day-ahead and real-time congestion prices and FTR prices for each of the eight ISO-NE load zones in 2011 and 2012. The congestion prices shown are calculated for peak hours relative to the New England Hub. Hence, if the congestion price in the figure indicates $1 per MWh, this is interpreted to mean the cost of congestion to transfer power from the New England Hub to the location averaged $1 per MWh during peak hours. The congestion price difference between any two points shown in the figure is the congestion price at the sink location less the congestion price at the source location. For example, a negative $0.50 per MWh FTR price for Maine and $2 per MWh FTR price for Connecticut would indicate a total price for an FTR from Maine to Connecticut of $2.50 per MWh. For each location, the figure shows the auction prices in chronological order leading up to real time, from left to right. The annual FTR auction occurs first, then the monthly FTR auction, and then the day-ahead market.

Figure 8: FTR Auction Prices vs. Day-Ahead and Real-Time Congestion
Average Difference from New England Hub in Peak Hours, 2011-2012

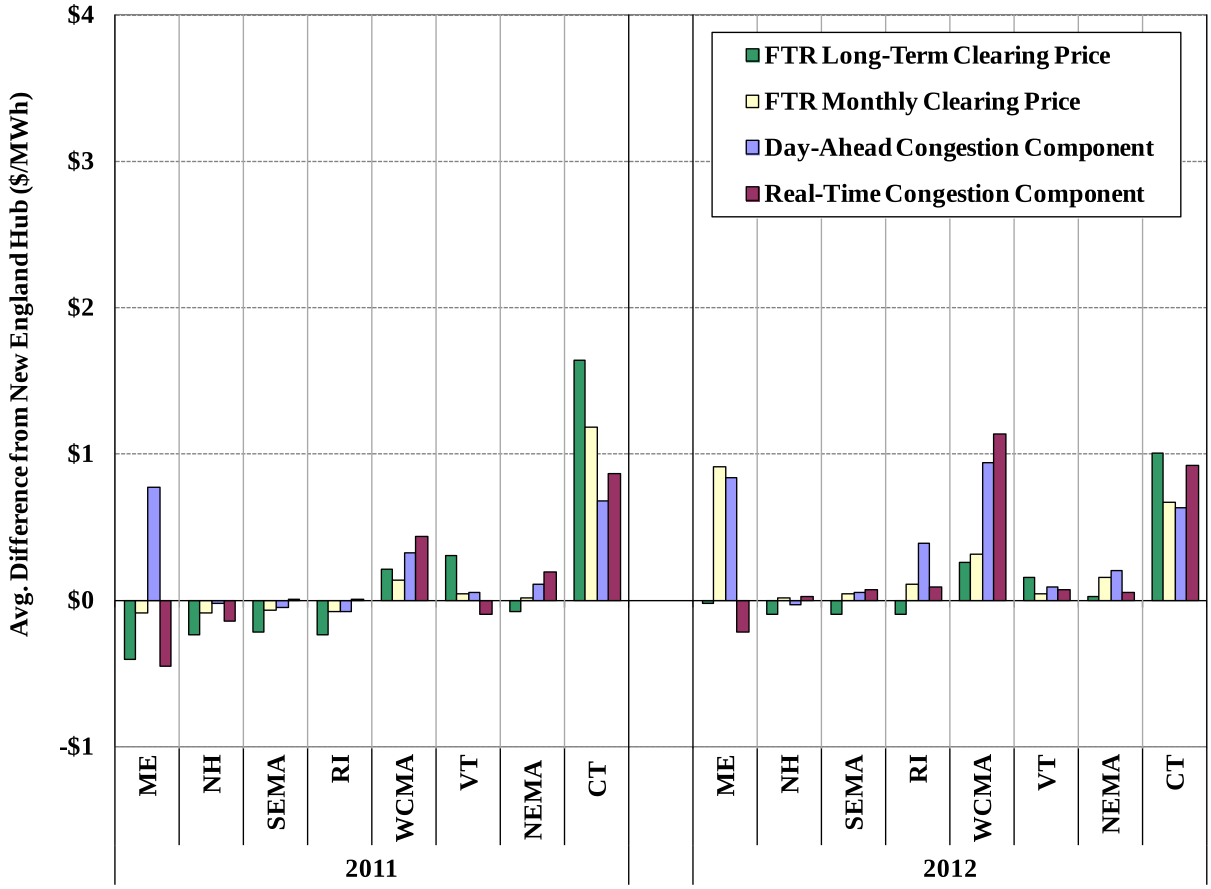


Figure 8 shows that in most areas during 2011 and 2012, monthly FTR prices were more consistent with congestion prices in the day-ahead market than were annual FTR prices. For example:

- The annual FTR prices from the New England Hub to Connecticut were $0.37 per MWh higher than the corresponding day-ahead congestion values in 2012, while the monthly FTR prices were only $0.03 per MWh higher.
- Similarly, the annual FTR prices from the New England Hub to Maine were $0.86 per MWh lower than the corresponding day-ahead congestion values in 2012, while the monthly FTR prices were only $0.07 per MWh higher.

This pattern is expected because market participants face greater uncertainty and have less information in the annual auction regarding likely congestion levels than they do at the time of monthly auctions.

The figure also shows that monthly FTR auction prices were still lower by a significant margin than the day-ahead congestion prices from the New England Hub to Western Central Massachusetts in 2012. This suggests that participants did not fully anticipate the effects of transmission outages and, therefore, substantially under-estimated day-ahead congestion into the area in the monthly auctions. Participants likely based their expectations more on the congestion that occurred in prior periods.

Given that variations in congestion patterns can be difficult to anticipate in advance, we find that FTRs were reasonably valued in the FTR auctions. The FTR market responded to changes in patterns of day-ahead congestion, which was particularly evident by the changes in the pricing of FTRs in the monthly auctions.

III. Reserve and Regulation Markets

This section evaluates the operation of the markets for operating reserves and regulation. The real-time reserve market has system-level and locational reserve requirements that are integrated with the real-time energy market. The real-time market software co-optimizes the scheduling of reserves and energy, which enables the real-time market to reflect the redispatch costs that are incurred to maintain reserves in the clearing prices of both energy and reserves. Energy-only markets (i.e., markets that do not co-optimize energy and reserves) do not recognize the economic trade-offs between scheduling a resource for energy rather than reserves. It is particularly important to consider such trade-offs during tight operating conditions because efficient scheduling reduces the likelihood of a reserve shortage. When available reserves are not sufficient to meet the requirement, the real-time model will be short of reserves and set the reserve clearing price at the level of the Reserve Constraint Penalty Factor (RCPF).

The forward reserve market enables suppliers to sell reserves into a forward auction on a seasonal basis. Similar to the real-time reserve market, the forward reserve market has system-level and locational reserve requirements. Suppliers that sell in the forward auction satisfy their forward reserve obligations by providing reserves in real-time from online resources with unused capacity or offline resources capable of starting quickly (i.e., fast-start generators that can start within 10 or 30 minutes). The forward reserve market is intended to attract investment in capacity that is able to provide reserves at relatively low cost, particularly fast-start generation.

ISO-NE runs a market for regulation service, which is the capability of specially-equipped generators to increase or decrease their output every few seconds in response to signals from ISO-NE. Regulation is used to balance actual generation with load on a moment-to-moment basis in New England. The regulation market provides a market-based system for meeting ISO-NE's regulation needs. Unlike many other ISO-run markets, the ISO-NE markets currently do not co-optimize the scheduling of regulation with reserves and energy.

This section of the report evaluates market outcomes in the real-time reserve, the forward reserve market, and the regulation market.

A. Real-Time Reserve Market Results

1. Real-Time Reserve Requirements

The real-time market is designed to satisfy the system's reserve requirements, including locational requirements to maintain minimum reserve levels in certain areas. There are four geographic areas with real-time reserve requirements: Boston, Southwest Connecticut, Connecticut, and the entire system (i.e., "All of New England"). In addition to the different locations, the reserve markets recognize three categories of reserve capacity: 10-Minute Spinning Reserves (TMSR), 10-Minute Non-Spinning Reserves (TMNSR), and 30-Minute Operating Reserves (TMOR).

Sufficient reserves must be held in the ISO-NE reserve zones to protect the system in case contingencies (e.g., generator outages) occur. The ISO used to hold an amount of 10-minute reserves (i.e., TMSR plus TMNSR) at the system level equal to the size of the largest generation contingency on the system. However, the ISO increased this requirement by 25 percent effective July 23, 2012, (i.e., now 125 percent of the largest generation contingency of the system) due to generator performance issues during past reserve activation events.[37] Based on system conditions, the operator determines how much of the 10-minute reserve requirement to hold as spinning reserves.

The ISO holds an amount of 30-minute reserves that includes the system's 10-minute reserve (TMSR and TMNSR) requirements plus an incremental quantity of 30-minute reserves (TMOR). This total used to equal to the size of the largest generation contingency on the system plus half of the size of the second-largest contingency on the system. However, the increase in the 10-minute reserve requirement on July 23, 2012 resulted in a corresponding increase in the 30-minute reserve requirement by 25 percent of the size of the largest contingency. Since higher quality reserves may always be used to satisfy requirements for lower quality products, the entire 30-minute reserve requirement can be satisfied with TMSR or TMNSR.

In 2012, the system-wide reserve requirements averaged:

---

37 The 10-minute reserve requirement was increased 20 percent to account for poor observed performance in the deployment of reserve units relative to their claimed capability. See Agenda Item 16: "http://www.iso-ne.com/committees/comm_wkgrps/mrkts_comm/mrkts/mtrls/2012/jul1112132012/index.html"

- 10-Minute Reserves: 1,370 MW prior to July 23 and 1,735 MW afterwards; and
- 30-Minute Reserves: 2,060 MW prior to July 23 and 2,430 MW afterwards.
- An average of 46 percent of the 10-minute reserve requirement was held in the form of spinning reserves during intervals with binding TMSR constraints in 2012.[38]

In each of the three local reserve zones, ISO-NE is required to schedule sufficient resources to maintain service in case the two largest local contingencies occur within a 30-minute period, resulting in two basic operating requirements. First, ISO-NE must dispatch sufficient energy in the local area to prevent cascading outages if the largest transmission line contingency occurs. Second, ISO-NE must schedule sufficient 30-minute reserves in the local area to maintain service if a second contingency occurs after the largest transmission line contingency. Alternatively, the local 30-minute reserve requirement can be met with 10-minute reserves or by importing reserves. Additional energy can be produced within the local area in order to unload transmission into the area, thus permitting the import of reserves if needed. Although ISO-NE is not the first RTO to co-optimize energy and reserves in the real-time market, it remains the only RTO that optimizes the level of imported reserves to constrained load pockets. As a result, ISO-NE is able to satisfy the local reserve requirements at a lower cost.

2. Real-Time Reserve Market Design

The real-time market software jointly optimizes reserves and energy schedules. By co-optimizing the scheduling of energy and reserves, the market is able to reflect the redispatch costs incurred to maintain reserves in the clearing prices of both energy and reserves. For example, if a $40 per MWh combined cycle unit is backed down to provide reserves when the LMP is $50 per MWh, the marginal redispatch cost is $10 per MWh and the reserve clearing price will not be lower than $10 per MWh. The marginal system cost that is reflected in the reserve clearing prices is equal to the marginal redispatch cost of the resources. When excess reserves are available without incurring any costs, reserve clearing prices will be $0 per MWh.

38 The TMSR requirement is binding when a non-zero cost is incurred by the market to satisfy the requirement. This occurred in 3.4 percent of the intervals in 2012.

Higher quality reserve products may always be used to satisfy lower quality reserve requirements, ensuring that the clearing prices of higher quality products are never lower than the clearing prices of lower quality products. For instance, if TMOR is available to be scheduled at a marginal system cost of $5 per MWh and an excess of TMNSR is available at no cost, the real-time market will fully schedule the TMNSR to meet the 30-minute reserve requirement. If the zero-cost TMNSR is exhausted before the requirement is met, the real-time market will then schedule additional TMOR and set the clearing prices of both TMNSR and TMOR at $5 per MWh.

When multiple reserve constraints are binding, the clearing price of the highest quality product will be the sum of the underlying marginal system costs for each product. For example, suppose the marginal system costs were $3 per MWh to meet the 10-minute spinning reserve constraint, $5 per MWh to meet the 10-minute reserve constraint, and $7 per MWh to meet the 30-minute reserve constraint. In this case, the TMSR clearing price would be $15 per MWh (i.e., $3 plus $5 plus $7) because a megawatt of TMSR would help satisfy all three constraints. Likewise, the TMNSR clearing price would be $12 per MWh (i.e., $5 plus $7) because a megawatt of TMNSR would help satisfy two of the constraints.

ISO-NE is the only RTO that counts imported reserves towards satisfying the local reserve requirements in the co-optimization of energy and reserves. Since local reserve requirements can be met with reserves on internal resources or import capability that is not used to import energy, allowing the real-time model to import the efficient quantity of reserves is a substantial improvement over other market designs. This enhancement is particularly important in New England where the market fulfills a significant share of its local area reserve requirements with imported reserves. For example, imported reserves satisfied 20 percent of the Connecticut requirement during constrained intervals in 2012.

The marginal system costs that the market incurs to satisfy reserve requirements are limited by RCPFs. There is an RCPF for each real-time reserve constraint. The RCPFs are:

- $500 per MWh for the system-level 30-minute reserve constraint since June 1, 2012. Previously, the RCPF was $100 per MWh;[39]
- $850 per MWh for the system-level 10-minute reserve constraint;
- $50 per MWh for the system-level 10-minute spinning reserve constraint; and
- $250 per MWh for the local 30-minute reserve constraints.[40]

When available reserves are not sufficient to meet a requirement or when the marginal system cost of maintaining a particular reserve requirement exceeds the applicable RCPF, the real-time model will be short of reserves and set clearing prices based on the RCPF. For example, if the marginal system cost of meeting the system-level 30-minute reserve requirement were $550 per MWh, the real-time market would not schedule sufficient reserves to meet the requirement and the reserve clearing price would be set to $500 per MWh. Hence, RCPFs should be set at levels that reflect the values of the reserves and the reliability implications of a shortage of each class of reserves.

Additionally, these values are additive when there are shortages of more than one class of reserves. Since energy and operating reserves are co-optimized, the shortage of operating reserves is also reflected in energy clearing prices.[41] For example, shortages of both 30-minute and 10-minute reserves would produce a clearing price of $1,350 per MWh for the system-level 10-minute reserves ($500 plus $850 per MWh) and energy prices likely exceeding $1,400 ($1,350 plus the marginal production cost of energy).

Hence, the system-level 10-minute reserve RCPF of $850 per MWh, together with the other RCPFs, would likely result in energy and operating reserve prices exceeding the ISO-NE market's energy offer cap of $1,000 per MWh during sustained periods of significant operating

39 We discuss this change in more detail in Section V.B.

40 The RCPF for local 30-minute reserve constraints was $50 per MWh before January 1, 2010.

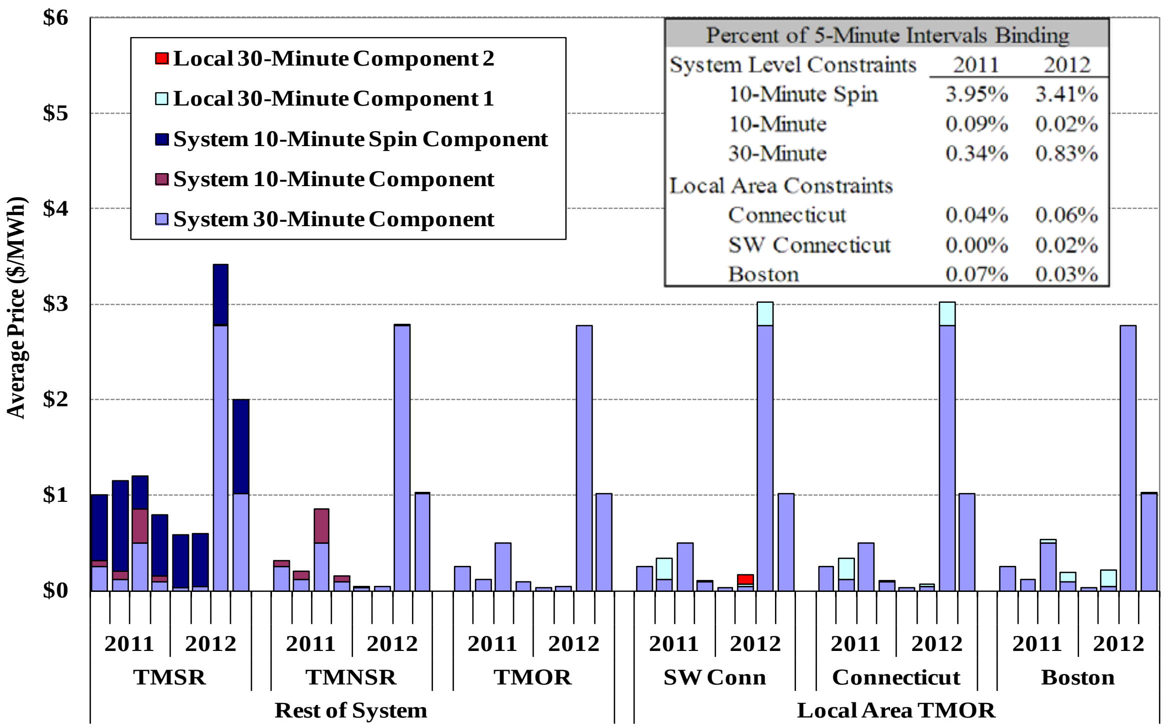
41 This assumes the operating reserve shortage results from a general deficiency of generating capacity.

reserve shortages. The use of RCPFs to set efficient prices during operating reserve shortages has been endorsed by FERC.[42]

3. Market Outcomes

Figure 9 summarizes average reserve clearing prices in each quarter of 2011 and 2012. The left side of the figure shows prices outside the local reserve zones for three service types. The right side shows prices in the three local reserve zones for TMOR only. Each price is broken into components for each underlying requirement. For example, the Southwest Connecticut price is based on the costs of meeting three requirements: the Southwest Connecticut 30-minute reserve requirement; the Connecticut 30-minute reserve requirement; and the system-level 30-minute reserve requirement. Likewise, the system-level 10-minute spinning reserve price is based on the costs of meeting three requirements: the 10-minute spinning reserve requirement; the 10-minute non-spinning reserve requirement; and 30-minute reserve requirement.

Figure 9: Quarterly Average Reserve Clearing Prices by Product and Location
2011 – 2012



42 *Wholesale Competition in Regions with Organized Electric Markets*, Order No. 719, 73 Fed. Reg. 64100 (October 28, 2008), FERC Stats. & Regs. ¶ 31,281 (2008) (Order No. 719).

The figure shows that reserve constraints bound infrequently in New England in 2011 and 2012. The most frequent binding constraint was the system-level 10-minute spinning reserve requirement, which was binding in roughly 3.4 percent of market intervals in 2012, down slightly from 3.95 percent in 2011. The other reserve requirements were binding in less than 1 percent of all intervals.

The clearing prices for operating reserves increased from 2011 to 2012, primarily in the second half of 2012. Outside the local constrained areas, the average prices for each reserve product changed as follows:

- TMSR prices increased from $1.04 per MWh in 2011 to $1.65 per MWh in 2012;
- TMNSR prices increased from $0.39 per MWh in 2011 to $0.98 per MWh in 2012; and
- TMOR price increased from $0.25 per MWh in 2011 to $0.97 per MWh in 2012.

The higher reserve clearing prices in the second half of 2012 were primarily due to two significant market rule changes. First, the RCPF for system-level 30 minute reserves increased from $100 to $500 per MWh in June 2012. This led the real-time market to set much higher clearing prices and incur higher re-dispatch costs during tight operating conditions. After June 1, the 30-minute reserve clearing price exceeded $100 per MWh in 0.8 percent of the intervals, averaging roughly $174 per MWh in these intervals. Previously, the real-time model would have simply been short of reserves in these intervals and set the clearing price at $100 per MWh, or the operators would have had to maintain adequate reserve levels through out-of-market actions.

Second, the system-level 30-minute reserve requirement rose in July 2012 consistent with the 25 percent increase in the system-level 10-minute requirement. This contributed to more frequent binding constraints and higher clearing prices.

In the local areas, TMOR clearing prices were almost identical to those in other areas because the local TMOR requirements were rarely binding in the real-time market.[43] Local reserve

43 TMNSR and TMSR clearing prices are not shown in the local areas because they can also be derived from the underlying requirements. For instance, the average clearing price of TMSR in Boston was $1.70 per MWh in 2012 ($1.65 per MWh for market-wide TMSR plus $0.05 per MWh for TMOR in Boston).

constraints have bound very infrequently (e.g., less than 0.1 percent in all three local regions during 2011 and 2012) since significant transmission upgrades were made in Boston and Connecticut between 2007 and 2009.

Despite the increases, average reserve clearing prices were still relatively low in 2012 because reserve clearing prices were $0 in the vast majority of real-time intervals. This reflects that there was surplus capacity in most hours sufficient to meet system-level and local reserve requirements with no need to redispatch generation. For example, the system-level 10-minute reserve requirement was binding in less than 0.1 percent of intervals in 2012, indicating that the requirement can be met at no cost with surplus capacity in over 99.9 percent of intervals.

4. Future Expansions in Operating Reserve Requirements

The ISO plans to increase the 30-minute operating reserve requirement to procurement to cover the entire size of the second largest contingency in addition to the 10-minute reserve requirement. Presently, the ISO procures only half of the second largest contingency. It has referred to this additional quantity as “replacement reserves” and plans to apply an RCPF to this quantity of $100 to $300 per MWh.

This is a significant improvement for a number of reasons. First, this will allow the ISO’s true reliability needs to be more fully specified and priced. Over the course of the past two years, the ISO has become increasingly concerned regarding the availability of fuel and the performance of generation. This has caused the ISO on many days to commit additional resources to ensure that it has sufficient supply available to maintain reliability. Because these manual supplemental commitments often cause the ISO’s total reserves to exceed its operating reserve requirements, they can lead both energy and operating reserve prices to be understated. This occurs because the real-time market does not recognize the increased demand for reserves that the operators are manually securing.

Second, because the prices are understated, suppliers’ incentive to be available and perform in real time is similarly understated. For example, assume that the ISO is making supplemental commitments to achieve operating reserve levels that are 500 MW in excess of the current 10 and 30-minute reserve requirements. If a supplier fails to start or cannot obtain fuel and the ISO is only able to achieve an additional reserve level of 300 MW, the ISO is effectively 200 MW short of the operating reserves it has deemed necessary for reliability. In this case, the real-time market will perceive a 300 MW surplus and will likely price its reserves at $0 per MWh and its energy in the range of $40 to $70 per MWh. However, if the shortage in replacement reserves were perceived and priced at $300 per MWh, all classes of reserves would clear at $300 per MWh or above, and energy would clear in the range of $340 to $370 per MWh. In this case, the generator that failed to start would lose substantial profit (~$300) that it would have earned by running (or if it had a day-ahead schedule, it would have to buy-back its energy at the prevailing real-time price).

Hence, the incentive to be available will be substantially improved by more fully specifying and pricing the ISO's true reliability needs. This is particularly important now as concerns regarding the availability of generation and natural gas supplies have grown substantially over the past year. These concerns have prompted the ISO propose Performance Incentives that would sharply increase incentives for suppliers to be online or providing reserves during reserve shortage conditions. These reliability concerns and recommendations to address them are discussed more fully in Section VI.

B. Forward Reserve Market

Each year, ISO-NE holds two auctions for Forward Reserves, one for the summer procurement period (the four months from June through September) and one for the winter procurement period (the eight months from October through May). Suppliers that sell in the Forward Reserve auction satisfy their obligations by providing reserves in real time from online resources or offline fast-start resources (i.e., peaking resources). This section evaluates the forward reserve auction results and examines how suppliers satisfied their obligations in real time.

1. Background on Forward Reserve Market

ISO-NE purchases two products in the Forward Reserve Market auction: 10-minute non-spinning reserves (TMNSR) and 30-minute reserves (TMOR). The forward reserve market also currently has four geographic zones: Boston, Southwest Connecticut, Connecticut, and the entire system (i.e., all of New England). Hence, the products procured in the forward reserve market are consistent with reserves in the real-time market except that the forward reserve market has no TMSR requirement.

Forward reserves are cleared through a cost-minimizing uniform-price auction, which sets clearing prices for each category of reserves in each reserve zone. Suppliers sell forward reserves at the portfolio level, which allows them the flexibility to shift where they hold the reserves on an hourly basis. Suppliers also have the flexibility to trade their obligations prior to the real-time market. The flexibility provided by portfolio-level obligations rather than unit-level and bilateral trading enables suppliers to satisfy their obligations more efficiently.

Forward reserve obligations may be satisfied in real time with reserves of equivalent or higher quality. When obligations are met with reserves of equivalent quality, the reserve provider receives the forward reserve payment instead of real-time market revenue based on the reserve clearing price. When obligations are met with reserves of higher quality, the reserve provider receives the forward reserve payment in addition to real-time market revenue based on the difference in clearing prices between the higher and lower quality products.[44]

2. Forward Reserve Auction Results

Forward Reserve auctions are held approximately one-and-a-half months prior to the first month of the corresponding procurement period. For example, the auction for the Winter 2012/13 Procurement Period (October 2012 to May 2013) was held in August 2012. Prior to each auction, ISO-NE sets minimum purchase requirements as follows:

- For the system-level, the TMNSR requirement is based on 50 percent of the forecasted largest contingency, and the TMOR requirement is based on 50 percent of the forecasted second largest contingency.[45]
- For each local reserve zone, the TMOR requirement is based on the 95th percentile of the local area reserve requirement in the daily peak hour during the preceding two like Forward Reserve Procurement Periods. The TMOR requirement is also adjusted for major changes in the topology of the system or the status of supply resources.

In the Forward Reserve Market auction, an offer of a high quality reserve product is capable of satisfying multiple requirements in the auction. In such cases, the higher quality product is

44 For example, if Boston TMOR obligations are satisfied in the real-time market with Boston TMSR, the reserve provider will receive the forward reserve payment for Boston TMOR plus the revenue from the real-time price difference between Boston TMSR and Boston TMOR.

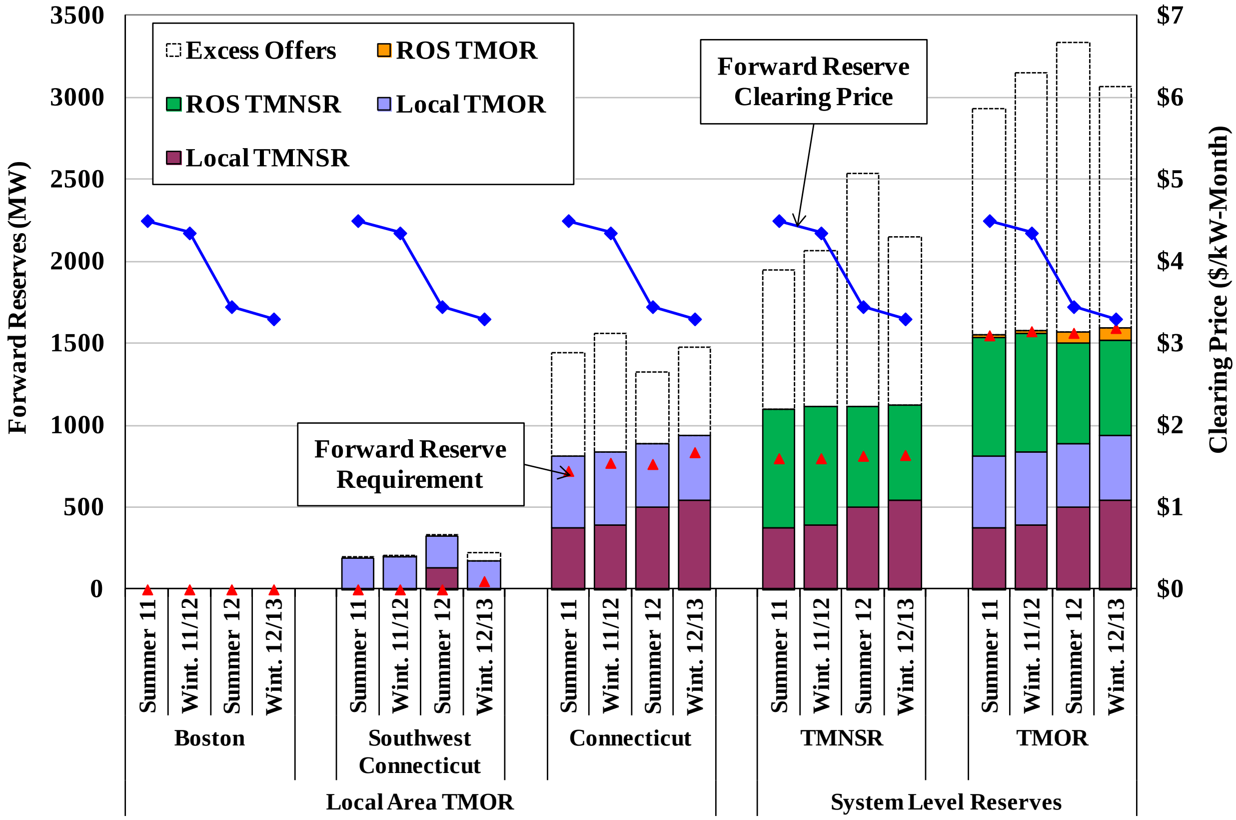
45 Usually, the forecasted largest contingency is the HQ Phase II Interconnection and the forecasted second largest contingency is the combination of the Mystic 8 and Mystic 9 generating units.

priced according to the sum of the values of the underlying products, although this is limited by the $14 per kW-month price cap.[46]

The following figure summarizes the market outcomes in the last four forward reserve auctions each of the requirements in each area. For each procurement period, Figure 10 shows:

- Forward reserve clearing price for each requirement,
- Reserves procured inside local reserve zones or outside the zones (i.e., Rest of System),
- Forward reserve requirement for each product, and
- The quantity of excess offers that was not cleared in the auctions.

Figure 10: Summary of Forward Reserve Auctions
Procurement for June 2011 to May 2013



46 For instance, 1 MW of TMNSR sold in Boston contributes to meeting three requirements: system-level TMNSR, system-level TMOR, and Boston TMOR. The Boston TMNSR clearing price equals the system-level TMNSR clearing price (which incorporates the price of system-level TMOR) plus the difference between the Boston TMOR clearing price and the system-level TMOR clearing price.

The figure shows that the TMOR prices in the three local areas cleared at the same levels as the system TMOR prices in all of the four auctions because none of the local requirements were binding. This trend is caused by the fact that transmission upgrades have substantially increased the transfer capability into the local zones and caused the local requirements to fall sharply. For example, the reserve requirements for Boston and for Southwest Connecticut were close to 0 MW in the four auctions.[47,48] Likewise, the Connecticut reserve requirement has fallen from historical levels of more than 1,300 MW to an average of less than 800 MW in recent auctions.. Nonetheless, forward reserves were procured in the local areas that were used to satisfy the system-level requirements for TMNSR and TMOR and were paid the system-wide price for these products.

Outside of the local reserve areas, the system TMOR requirement was binding in each of the four auctions.[49] However, the system TMNSR requirement was not binding because it was met by TMNSR offers that were accepted to satisfy the system TMOR requirement. Accordingly, the TMNSR prices cleared at the same levels as the system TMOR prices in all four auctions. Therefore, in each of the four auctions, the clearing price has been the same for all forward reserve products in all locations.

The clearing price fell moderately from an average of $4.40 per kW-month in the 2011/12 Capability Period to $3.35 per kW-month in the 2012/13 Capability Period. This is consistent with the reduction in Forward Capacity prices over the same period, which fell from $3.60 per kW-month in the 2011/12 Capability Period (i.e., FCA 1) to $2.95 per kW-month in the 2012/13 Capability Period (i.e., FCA 2). This is expected since forward reserve suppliers are paid according to the differential between the forward reserve clearing price and the forward capacity

47 A substantial amount of transmission capability was added into the Boston area in 2007, leading ISO-NE to assume 820 to 1,400 MW of External Reserve Support in the recent four auctions. External Reserve Support is the amount of the local reserve zone need that is assumed to be satisfied by the transmission capability into the zone, which reduces the amount that must be satisfied by internal resources. As a result, the amount of local reserves required from internal Boston resources was reduced to 0 MW (i.e., no need for local resources to provide reserves due to enough External Reserve Support) in the last four auctions.

48 Transmission upgrades into Southwest Connecticut were brought into service in early 2009.

49 Beginning in the auction for the Summer 2011 Procurement Period, the Rest of System TMOR requirement was no longer binding because it had been eliminated before the auction. Consequently, the system TMOR requirement was binding for the first time (since the Locational Forward Reserve Market was implemented in 2006).

price.[50] Hence, the forward reserve prices should move as forward capacity prices move. After deducting the forward capacity prices, the effective forward reserve clearing prices were $0.80 per kW-month in the 2011/12 Capability Period and $0.40 per kW-month in the 2012/13 Capability Period.

3. Forward Reserve Obligations in the Real-Time Market

Forward reserve providers satisfy their obligations in the real-time market by assigning individual resources to provide specific quantities of forward reserves in each hour from 7:00 AM to 11:00 PM, Monday through Friday. Resources assigned to provide forward reserves must be fast-start units or units that are online. These resources must be capable of ramping quickly enough to provide the specified quantity of reserves in 10 minutes for TMNSR and 30 minutes for TMOR. The assigned resources must offer the assigned quantity of incremental energy at a minimum price level.[51] Resources assigned to provide forward reserves forfeit any NCPC payments that they would otherwise receive. Forward reserve providers can arrange bilaterally for other suppliers to meet their obligations, although bilateral trading of obligations between non-affiliated firms was very limited in 2012. Suppliers that do not meet their forward reserve obligations incur a Failure to Reserve Penalty.[52]

There are several types of costs that suppliers consider when assigning units to provide forward reserves. First, suppliers with forward reserve obligations face the risk of financial penalties if their resources fail to deploy during a reserve pick-up.[53] Suppliers can reduce this risk by

---

50 See Market Rule 1 III.9.8

51 This level, known as the "Forward Reserve Threshold Price," is equal to the monthly fuel index price posted prior to each month multiplied by a Forward Reserve Heat Rate in MMbtu per MWh, which is based on the 2.5 percentile value of an historical analysis of "implied heat rates". For example, the monthly natural gas index price was $3.47 per MMbtu and the forward reserve heat rate was 15.747 MMbtu per MWh for October 2012. Hence, it resulted in a Forward Reserve Threshold Price of approximately $55 per MWh for this month. The monthly fuel index price is based on the lower of the natural gas or diesel fuel index prices in dollars per MMbtu. The implied heat rate analysis is based on the real-time hub LMP and the lower of the distillate or natural gas fuel price indices for New England.

52 The Failure to Reserve penalty is equal to the number of megawatts not reserved times 1.5 times the Forward Reserve Payment Rate, which is the forward reserve clearing price (adjusted for capacity payments) divided by the number of obligation hours in the month.

53 The Failure to Activate penalty is equal to the MW quantity that does not respond times the sum of the Forward Reserve Payment Rate and the Failure to Activate Penalty Rate, which is 2.25 times the higher of the LMP at the generator's location or the Forward Reserve Payment Rate.

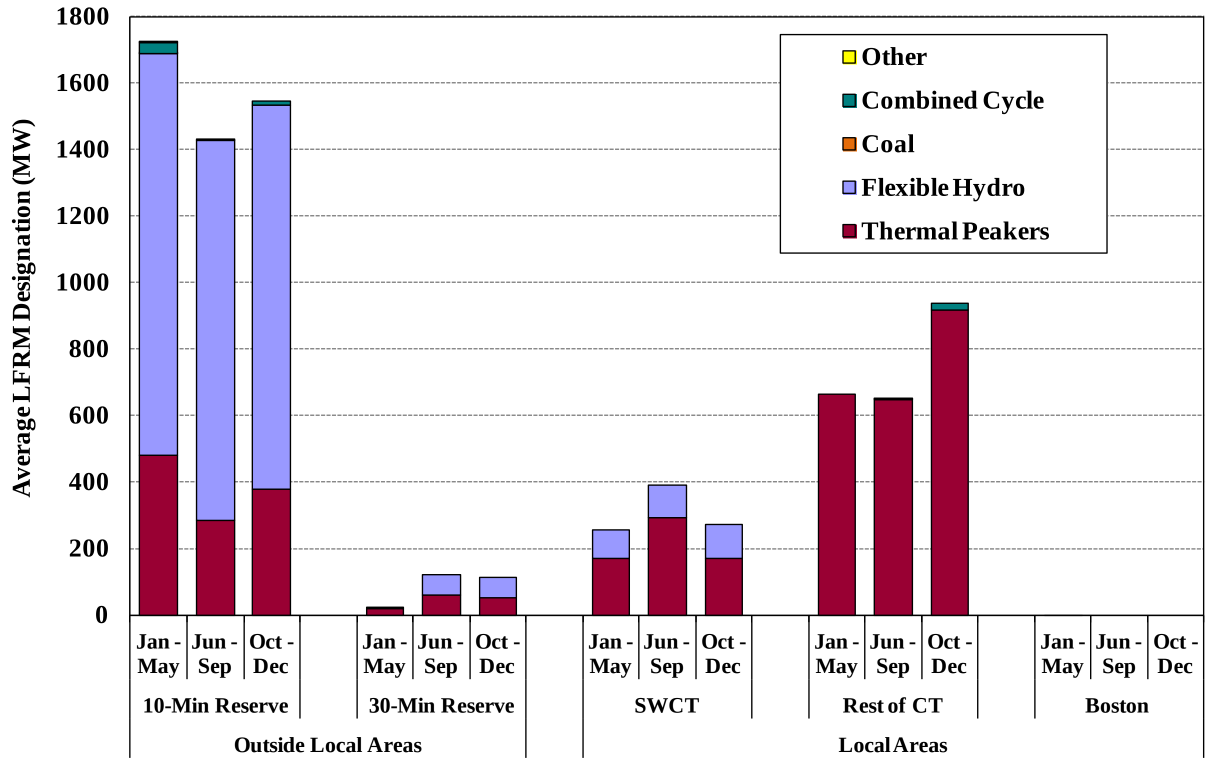
meeting their obligations with resources that are more reliable. Second, suppliers with forward reserve obligations forego the value of those reserves in the real-time market. For instance, suppose that real-time clearing prices are $10 per MWh for TMOR and $15 per MWh for TMNSR. A supplier that has TMOR obligations would not be paid if scheduled for TMOR or would be paid $5 per MWh (i.e., the price difference between TMNSR and TMOR) if scheduled for TMNSR. Hence, the foregone reserve revenues are the same regardless of whether the supplier is ultimately scheduled for TMOR, TMNSR, TMSR, or energy in the real-time market.

Third, suppliers may forego profitable energy sales as a result of offering incremental energy at the Forward Reserve Threshold Price. For instance, suppose the Forward Reserve Threshold Price is $100 per MWh and a supplier assigns a generator that has incremental costs of $60 per MWh to provide forward reserves. Because the supplier is required to offer at $100 per MWh, the supplier will not be scheduled to sell energy when the LMP is between $60 per MWh and $100 per MWh. The magnitude of this opportunity cost decreases for units that have high incremental costs (this opportunity cost is zero for units that have incremental costs greater than the Forward Reserve Threshold Price).

The previous three kinds of costs may be incurred by all units that provide forward reserves, but there are additional costs that are faced only by units that must be online to provide reserves. In order to provide reserves from a unit that is not a fast-start unit, a supplier may have to commit a unit that would otherwise be unprofitable to commit. This type of cost is zero when energy prices are high and the unit is profitable to operate based on the energy revenues. However, when energy prices are low, the commitment costs incurred by some units may far exceed the net revenue that they earn from the energy market. Because fast-start resources do not face this cost, they are generally most economic to meet forward reserve obligations.

The following analysis evaluates how market participants satisfied their forward reserve obligations in 2012 by procurement period. The figure shows the average amount of reserves assigned in each region by type of resource.

Figure 11: Forward Reserve Assignments by Resource Type
2012



Approximately 99 percent of the capacity assigned to provide forward reserves was hydro and thermal peaking capacity capable of providing offline reserves. In some cases, these units were online and providing energy (which is acceptable as long as they offer in accordance with the forward reserve rules). The frequent assignment of fast-start resources to provide forward reserves confirms that it is generally more costly to provide forward reserves from slower-starting resources.

Combined cycle units were assigned to provide a small portion (0.8 percent) of the forward reserves in 2012. Most of these units were ones that are capable of providing offline reserves within 30 minutes.

In summary, the vast majority of forward reserves were provided by fast-start units. This suggests that many slower-starting resources did not sell forward reserves because the expected costs of providing forward reserves exceeded the clearing prices in the forward reserve auctions. However, slower-starting units that could provide forward reserves at a cost below the forward reserve clearing price may be discouraged from participating because units that are frequently committed for local reliability and receive substantial NCPC payments have disincentives to provide forward reserves (they would be required to forgo the NCPC payments). Some had expected that the Forward Reserve Market would lower NCPC costs because high-cost units committed for local reliability would sell Forward Reserves. However, this has not occurred.

C. Regulation Market

Regulation service is the capability of specially-equipped generators to increase or decrease their output on a moment-to-moment basis in response to signals from the ISO. The system operator deploys regulation to maintain the balance between actual generation and load in the control area. The regulation market provides a market-based system for meeting the system's regulation needs.

The ISO determines the quantity of regulation capability required to maintain the balance between generation and load based on historical performance and ISO-NE, NERC and NPCC control standards. The ISO schedules an amount of regulation capability that ranges from 30 MW to 150 MW depending upon the season, time of day, and forecasted operating conditions. Historically, the ISO has scheduled 15 to 20 MW more regulation capability in the summer and winter than it has acquired in the spring and fall. During emergency conditions, the ISO may adjust the regulation requirement to maintain system reliability. The ISO periodically reviews regulation performance against the applicable control standards. The high level of performance in recent years has permitted a steady decline in the average quantity of regulation scheduled over the last seven years: from 143 MW in 2005 to 61 MW in 2012.

In this report, we evaluate two aspects of the market for regulation. First, we review the overall expenses from procuring regulation. Second, we explain how regulation providers are selected and examine the pattern of supply offers from regulation providers. The end of this subsection summarizes our conclusions related to the regulation market.

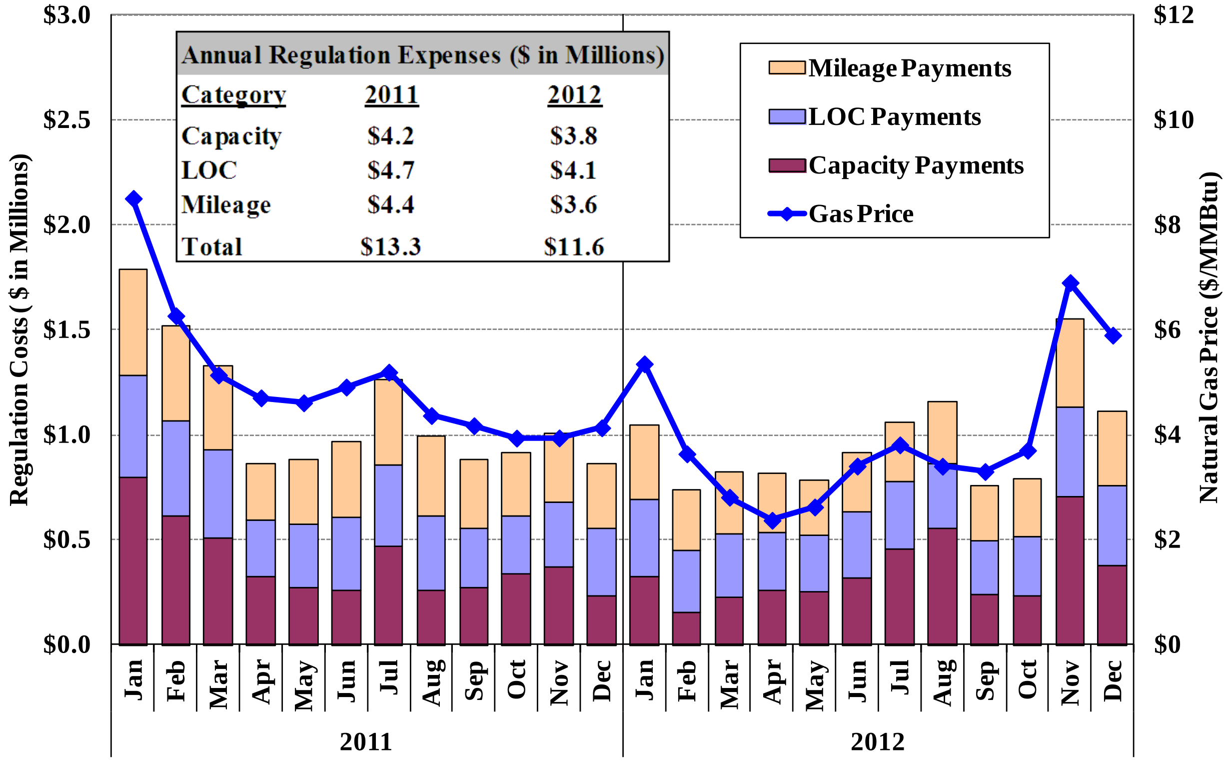
1. Regulation Market Expenses

Resources providing regulation service receive the following payments:[54]

- Capacity Payment – This equals the Regulation Clearing Price (RCP) times the amount of regulation capability provided by the resource. The RCP is based on the highest accepted offer price.
- Mileage Payment – This is equal to 10 percent of the "mileage" (i.e., the up and down distance measured in MW) times the RCP. Based on historic patterns of regulation deployment, this formula was expected to generate mileage payments and capacity payments of similar magnitude in the long term.
- Lost Opportunity Cost (LOC) Payment – This is the opportunity cost of not providing the optimal amount of energy when the resource provides regulation service.

A summary of the market expenses for each of the three categories is shown in Figure 12 by month for 2011 and 2012. The figure also shows the monthly average natural gas price.

Figure 12: Regulation Market Expenses
2011 – 2012



54 In ISO-NE Manual M-11, Capacity Payment is the "Time-on-Regulation Credit," Mileage Payment is the "Regulation Service Credit," and the Lost Opportunity Cost Payment is the "Regulation Opportunity Cost."

This figure shows that each category of expenses accounts for approximately one-third of total regulation expenses. Total regulation expenses declined 13 percent from $13.3 million in 2011 to $11.6 million in 2012, consistent with the reduction in natural gas prices over the same period. In particular, the figure shows that variations in monthly regulation market expenses were generally correlated with changes in the monthly average natural gas price.

Input fuel prices can affect regulation market expenses in several ways. First, generators may consume more fuel to produce a given amount of electricity when they provide regulation, leading the costs of providing regulation to be correlated with the price of fuel. Market participants reflect these costs in their regulation offer prices, which directly affect Capacity Payments and Mileage Payments. Second, natural gas-fired combined cycle generators are usually committed more frequently during periods of low gas prices. This increases the availability of low-priced regulation offers and leads to lower regulation expenses. Third, lower fuel prices normally reduce the opportunity costs for units to provide regulation service, which is consistent with the general decrease in regulation opportunity cost expenses in the summer months compared to the winter months.

Changes in natural gas prices and commitment patterns led to changes in offer patterns that explain some of the fluctuations in regulation market expenses in 2011 and 2012. Offer patterns are examined in more detail in the following section.

2. Regulation Offer Patterns

Competition should be robust in ISO-NE's regulation market in most hours because the amount of capability available in New England generally far exceeds the amount required by the ISO. The regulation market selects suppliers for the upcoming hour with the objective of minimizing consumer payments. Each resource offering to provide regulation is ranked according to the estimated payment it would receive if it were to provide regulation. The model selects the resources with the lowest rank price to provide regulation.

The rank price is the sum of the following four quantities:

- <u>Estimated Capacity Payment</u> – In the first iteration of the model, this is the offer price of each resource. But since the RCP is set by the highest accepted offer, the subsequent iterations set this equal to the higher of the offer price and the previous iteration's highest priced accepted offer.

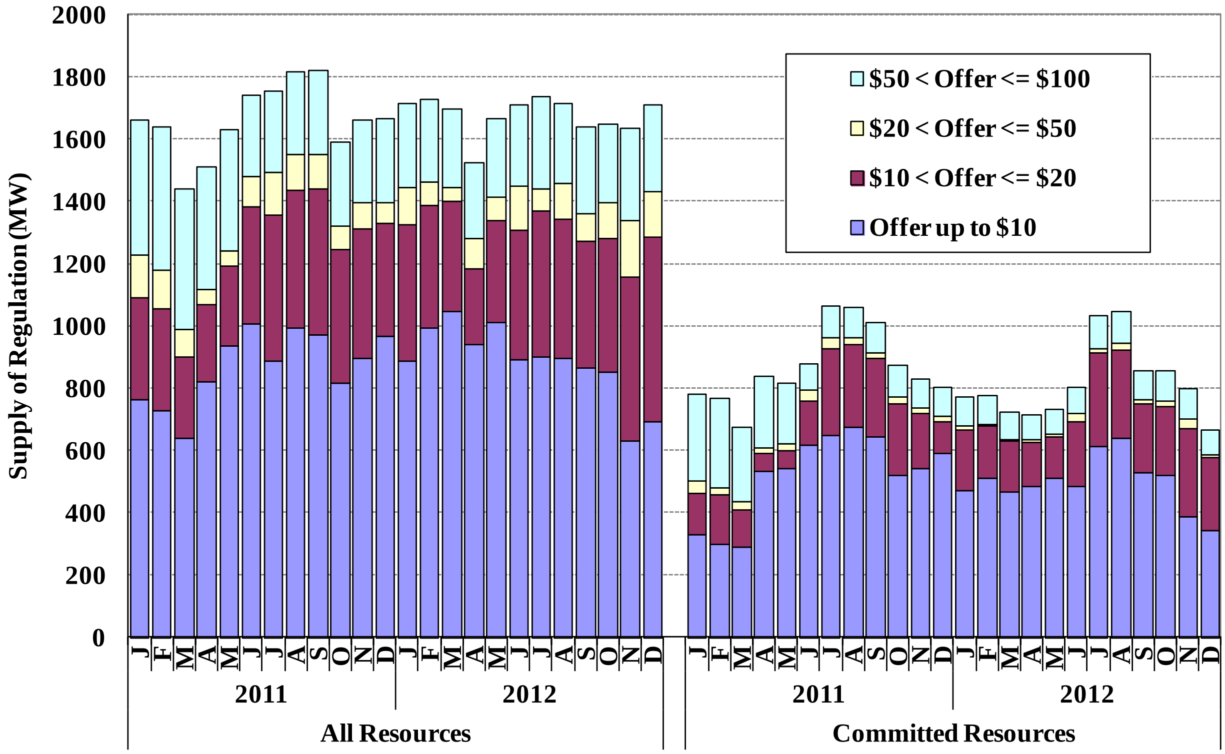
- Estimated Mileage Payment – This is equal to the estimated capacity payment.
- Estimated Lost Opportunity Cost Payment – This is the estimated opportunity cost from being dispatched at a level that allows a resource to provide regulation rather than at the most economic dispatch level given the resource's offer prices and the prevailing LMP.
- The Look-Ahead Penalty – This is equal to 17 percent of the maximum possible change in the energy offer price within the regulating range. This is included in order to avoid selecting resources that would earn large opportunity cost payments if they were to regulate into a range of their energy offer priced at extreme levels.

The ranking process iterates until the set of resources selected to provide regulation does not change for two consecutive iterations.[55]

This part of the section evaluates the offer patterns of regulation suppliers in 2012. Offline units cannot provide regulation service so selection of units is limited to units that are online at the time the service is needed. To highlight the importance of this limitation, Figure 13 examines regulation offers from all resources and from online resources by showing monthly averages of the quantity of regulation offered into the market in 2011 and 2012. The left panel in the figure shows offers from all online and offline resources, while the right panel is limited to resources that are actually available to provide regulation. The different color-coded bars in the chart show the average quantities offered within different offer price ranges.

55 However, if the RCP rises from one iteration to the next, the model will use the previous iteration to rank resources. For additional details, see Section 3.2.5 of ISO-NE Manual M-11 on Market Operations.

Figure 13: Monthly Average Supply of Regulation
2011 – 2012



The left panel of Figure 13 shows that the regulation offer prices and quantities over the past two years were relatively consistent during most of the period. The quantities of total regulation offers varied typically between 1,400 MW and 1,800 MW in most months of 2011 and 2012. The portion of regulation offers in each price range was also relatively consistent over the past two years. In 2012, 53 percent of the total regulation offers were priced below $10 per MWh, while 16 percent were priced above $50 per MWh.

The right panel shows the changes in offer quantities and prices that more directly determine market outcomes, since only offers from committed resources can be selected. In 2012, on average, approximately 49 percent of the regulation offered in the day-ahead market was available to the hourly real-time selection process. Regulation-capable capacity can be unavailable in a given hour because the capacity is on a resource that was not committed for the hour, or because the capacity is held on a portion of a resource that was self-scheduled for energy. More regulation capacity tends to be available during the high-load portion of the day because more units are online. Similarly, more regulation capacity tends to be available during the summer when loads are higher and more generation is committed.

During both 2011 and 2012, significantly more regulation capability was offered into the market than was actually procured (i.e., slightly more than 60 MW on average for both years) by the ISO. This excess supply generally limits competitive concerns in the regulation market because demand can easily be supplied without the largest regulation supplier. However, supply is sometimes tight in the regulation market when energy demand is high and the regulation market must compete with the energy market for resources. High energy prices during peak-demand periods can lead resources to incur large opportunity costs when providing regulation service, thereby increasing prices for regulation. Likewise, regulation supplies may be tight in low-demand periods when many regulation-capable resources are offline. These conditions can lead to transitory periods of high regulation prices.

On October 20, 2011, FERC issued Order 755 on Frequency Regulation Compensation, which requires ISO-NE and other ISOs to make certain modifications to their market designs.[56] Specifically, Order 755 requires ISOs to operate regulation markets that compensate generators for "actual service provided, including a capacity payment that includes the marginal unit's opportunity costs and a payment for performance that reflects the quantity of frequency regulation service provided."

ISO-NE and NEPOOL filed modifications on April 30, 2012 to comply with Order No. 755. However, these proposed changes were not accepted because the Commission found that they did not provide for uniform prices in the manner directed by Order No. 755. The Commission also indicated that the proposed bundled payment mechanism did not satisfy the requirement to have separate payments for capacity and performance.[57] On February 6, 2013, ISO-NE and

56 See Frequency Regulation Compensation in the Organized Wholesale Power Markets, Order No. 755, 137 FERC ¶ 61,064 (2011).

57 See FERC November 8, 2012 Order on compliance filing at http://www.iso-ne.com/regulatory/ferc/orders/2012/nov/er12-1643-000_11-8-12_order_rejecting_order_755_filing.pdf.

NEPOOL submitted another filing with proposed regulation market changes to address these issues and requested they become effective on or after January 1, 2015.[58]

D. Conclusions and Recommendations

In the real-time market, the scheduling of operating reserves and energy is co-optimized, enabling the real-time model to consider how the cost of energy is affected by the need to maintain operating reserves, and vice versa. Outside the local reserve areas, the average 30-minute reserves clearing price increased for two primary reasons:

- The RCPF for system-level 30-minute reserves was raised from $100 to $500 per MWh in June 2012.
- The 30-minute reserve requirement was increased by an average of 270 MW in July because the 10-minute portion of the requirement increased after poor reserve deployment performance by units in New England.

In the local reserve areas, reserve clearing prices were comparable to prices outside the local areas, reflecting that local reserve constraints have been binding very infrequently since the completion of transmission upgrades in Connecticut and Boston between 2007 and 2009.

The ISO plans to increase the 30-minute operating reserve requirement to procure additional "replacement reserves". We believe this is a valuable change because it will:

- Allow the ISO's true reliability needs to be more fully specified and priced. This has become increasingly important over the past two years as concerns regarding the availability of fuel and the performance of generation.
- Improve suppliers' incentive to be available and perform in real time.

However, this change would be even more effective if the ISO had the ability to vary the replacement reserve quantity as reliability dictates. Under cold weather conditions when the ISO's concern regarding fuel availability is heightened, it may be reasonable for it to procure a larger quantity of replacement reserves. Under mild conditions when fuel uncertainty and load

58 For a description of the proposed modifications, see http://www.iso-ne.com/regulatory/ferc/filings/2013/feb/er12-1643-001_order_755_2-6-2013.pdf.

forecast uncertainty are both minimal, it may be reasonable for the ISO to procure no replacement reserves. Allowing the ISO to determine the quantity of replacement reserves it needs will help maintain consistency between the market outcomes and the ISO's reliability requirements.

In the forward reserve market, all of the prices in both the 2011/2012 and the 2012/13 Capability Periods cleared at the same level because the system TMOR constraint was the only binding constraint. Prices fell roughly 24 percent from the 2011/12 Capability Period to the 2012/13 Capability Period. Most of the decrease was related to the reduction in the Forward Capacity Market clearing prices. As observed in previous years, we also found that 99 percent of the resources assigned to satisfy forward reserve obligations in 2012 were fast-start resources capable of providing offline reserves.

The ISO may wish to consider the long-term viability of the forward reserve market for several reasons. First, it has not achieved one of its primary objectives, which was to lower NCPC by purchasing forward reserves from high-cost units frequently committed for reliability. Second, the Locational Forward Reserve Market is largely redundant with the locational requirement in the Forward Capacity Market. Third, the forward reserve requirements are determined seasonally and the obligations of forward reserve suppliers are not consistent with the day-to-day operational needs of the system. In fact, the forward procurements do not ensure that sufficient reserves will be available during the operating day.

In the longer-term, we recommend the ISO consider introducing day-ahead reserve markets. Such markets would allow the ISO to procure the reserves it needs for the following day and to set clearing prices that reflect the costs of satisfying the operating reserve obligations. Such markets would also likely help address the ISO's concerns regarding unit availability. The day-ahead reserve schedules would be established in a timeframe in which suppliers can make arrangements for fuel and staffing to allow them to respond to reserve deployments.

Overall, the regulation market performed competitively in 2012. On average, approximately 815 MW of available supply competes to provide 60 MW of regulation service. The significant excess supply generally limits competitive concerns in the regulation market. In October 2011, FERC issued Order 755 on Frequency Regulation Compensation, which requires ISO-NE and other ISOs to operate regulation markets that compensate generators for "actual service provided, including a capacity payment that includes the marginal unit's opportunity costs and a payment for performance that reflects the quantity of frequency regulation service provided." If accepted, ISO-NE's latest proposal for complying with the Order will become effective on or after January 1, 2015. It is likely that the regulation market will continue to perform competitively given the large amount of supply in the market relative to the average demand.

IV. External Interface Scheduling

This section examines the scheduling of imports and exports between New England and adjacent regions. ISO-NE receives imports from Quebec and New Brunswick in most hours, which reduces wholesale power costs for electricity consumers in New England. Between New England and New York, power can flow in either direction depending on market conditions, although ISO-NE exported more power to NYISO than it imported in 2012. The transfer capability between New England and adjacent control areas is large relative to the typical load in New England, making it particularly important to schedule interfaces efficiently.

Consumers benefit from the efficient use of external transmission interfaces. The external interfaces allow low-cost external resources to compete to serve demand in New England. The ability to draw on neighboring systems for emergency power, reserves, and capacity also lowers the costs of meeting reliability needs in the interconnected system. Wholesale markets facilitate the efficient use of both internal resources and transmission interfaces between control areas.

ISO-NE is interconnected with three neighboring control areas: the NYISO, TransEnergie (Quebec), and the New Brunswick System Operator. ISO-NE and NYISO are interconnected by three interfaces:

- The Roseton Interface, which is the primary interface and includes several AC tie lines connecting upstate New York to Connecticut, Massachusetts, and Vermont;
- The 1385 Line, a controllable AC interface between Norwalk and Long Island; and
- The Cross-Sound Cable, a DC interface between Connecticut and Long Island.

New England and Quebec are interconnected by two interfaces: Phase I/II (a large DC interconnection), and the Highgate Interface (a smaller AC interconnection between Vermont and Quebec). New England and New Brunswick are connected by a single interface.

This section evaluates the following aspects of transaction scheduling between ISO-NE and adjacent control areas. Section A summarizes scheduling between New England and adjacent areas in 2012. Section B evaluates the efficiency of scheduling by market participants between New York and New England. Section C discusses ISO England’s recent efforts to improve the utilization of its interfaces with New York. Section D provides a summary of our conclusions and recommendations.

A. Summary of Imports and Exports

The following two figures provide an overview of imports and exports by month for 2011 and 2012. Figure 14 shows the average net imports across the three interfaces with Quebec and New Brunswick by month, for peak and off-peak periods.[59] The net imports across the two interfaces linking Quebec to New England are combined.

Figure 14: Average Net Imports from Canadian Interfaces
2011 – 2012

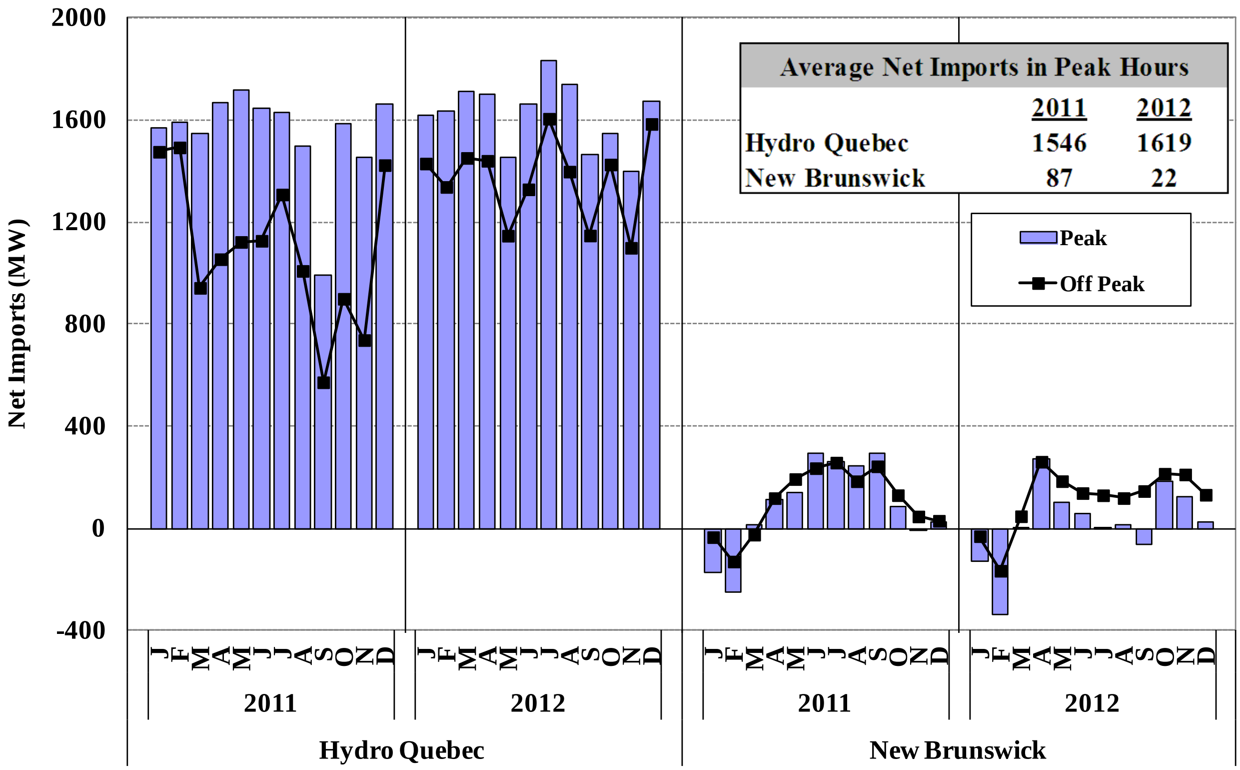


Figure 14 shows power is generally imported from Quebec. Average net imports from Quebec were higher during peak hours than during off-peak hours by roughly 450 MW in 2011 and by 250 MW in 2012. This reflects the tendency for hydro resources in Quebec to store water during low demand periods in order to make more power available during high demand periods. In the same way that the imports vary from peak to off-peak hours, imports also vary seasonally with imports rising during periods when energy prices are the highest. This was evident in both 2011

59 Peak hours include hours ending 8 to 23, Monday through Friday (not including NERC holidays), and the remaining hours are included in Off-Peak.

and 2012, when average net imports generally rose in the summer months and in periods with high natural gas prices (i.e., typically the winter months). This pattern is beneficial to New England because it tends to smooth the residual demand on New England internal resources.

Power was imported from New Brunswick in most hours, although the prevailing direction of flows changed in the winter such that New England exported to New Brunswick (e.g., 233 MW on average in peak hours in January and February 2012). Hence, scheduling with New Brunswick led to tighter supply conditions in New England in the winter and greater excess in mild periods.

Figure 15 shows average net imports across the three interfaces with New York by month in 2011 and 2012 for peak and off-peak periods. The net imports across the Cross-Sound Cable and the 1385 Line are combined.

Figure 15: Average Net Imports from New York Interfaces
2011 – 2012

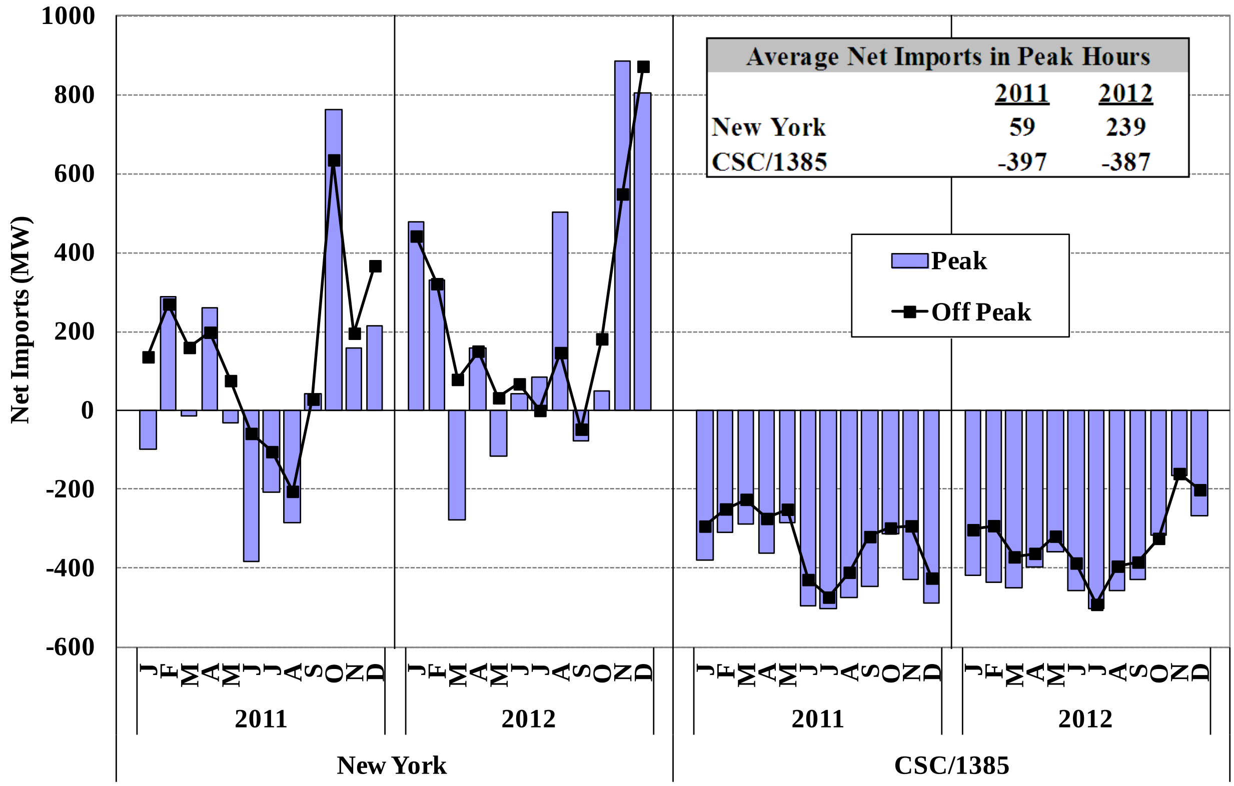


Figure 15 shows that the direction and the level of flows varied considerably across the primary interface with New York (i.e., the Roseton interface) during the past two years, reflecting the variations in relative prices in the two markets. New England tends to import more power from (or export less power to) New York in the winter months for several reasons. First, New England is more reliant on natural gas generation, which is typically most expensive in the winter months.

Second, the spread in natural gas prices between New England and New York tends to increase in the winter months when demand for heating rises. For example, ISO-NE imported an average of 625 MW from New York during peak hours in four months (January, February, November, and December) of 2012, but only an average of 45 MW in other eight months. During these four winter months, the spread in natural gas prices between New England and New York averaged $0.70 per MMBtu. In the other eight months of 2012, the spread averaged $0.10 per MMBtu. [60]

New England was a net importer from New York across the primary interface in both 2011 and 2012. Net imports averaged approximately 235 MW over all hours in 2012, up 130 MW from 2011. The increase was consistent with the increased spread in natural gas prices between New England and New York, which rose from an average of less than $0.10 per MMBtu in 2011 to $0.30 per MMBtu in 2012.

The figure also shows that flows were relatively consistent from New England to Long Island across the Cross-Sound Cable and the 1385 Line. The Cross-Sound Cable and the 1385 Line have transfer capabilities of 330 MW and 200 MW, respectively. Both lines are usually fully utilized during peak hours to export power to Long Island when they are in service. There were two notable month-to-month variations over the two-year period shown in the figure. First, the average level of exports across the 1385 Line increased in June 2011 after the completion of upgrades that increased the normal transfer capability of the interface from 100 MW to 200 MW. Second, the average level of exports across both interfaces fell substantially (by nearly 50 percent) in November and December 2012 when the natural gas price spread between New England and Long Island rose to an average of $1.45 and $0.75 per MMBtu, respectively.

60 The spread is based on the difference between the Algonquin City Gates index, which is representative of natural gas prices in most of New England, and the Iroquois Zone 2 index, which is representative of natural gas prices in eastern upstate areas of New York and in Long Island.

B. Interchange with New York

The performance of ISO-NE's wholesale electricity markets depends not only on the efficient use of internal resources, but also the efficient use of transmission interfaces with adjacent areas. This section evaluates the efficiency of scheduling between New England and New York. Since both regions have real-time spot markets, market participants can schedule market-to-market transactions based on transparent price signals in each region. In this sub-section, we evaluate the extent to which the interface is scheduled efficiently.

When an interface is used efficiently, prices in adjacent areas should be consistent unless the interface is constrained. For example, when prices are higher in New England than in New York, imports from New York should continue until prices have converged or until the interface is fully scheduled. A lack of price convergence indicates that resources are being used inefficiently. In other words, higher-cost resources are operating in the high-priced region that could have been supplanted by increased output from lower-cost resources in the low-priced region. It is especially important to schedule flows efficiently between control areas during peak demand conditions or shortages when flows between regions have the largest economic and reliability consequences.

However, one cannot expect that trading by market participants alone will optimize the use of the interfaces. Several factors prevent real-time price differences between New England and New York from being fully arbitraged.

- Market participants do not operate with perfect foresight of future market conditions (e.g., may not be able to predict which side of the interface will have a higher real-time price) at the time when transaction bids and offers must be submitted.
- Differences in the procedures and timing of scheduling in each market serve as barriers to full arbitrage.
- There are transaction costs associated with scheduling imports and exports that diminish the returns from arbitrage. Participants will not schedule additional power between regions unless they expect a price difference greater than these costs.
- The risks associated with curtailment and congestion reduce participants' incentives to schedule external transactions when the expected price difference is small.

Given these considerations, one cannot reasonably expect that trading by market participants will fully optimize the use of the interface. Nevertheless, we expect trading to improve the efficiency of power flows between regions.

1. Price Convergence Between New England and New York

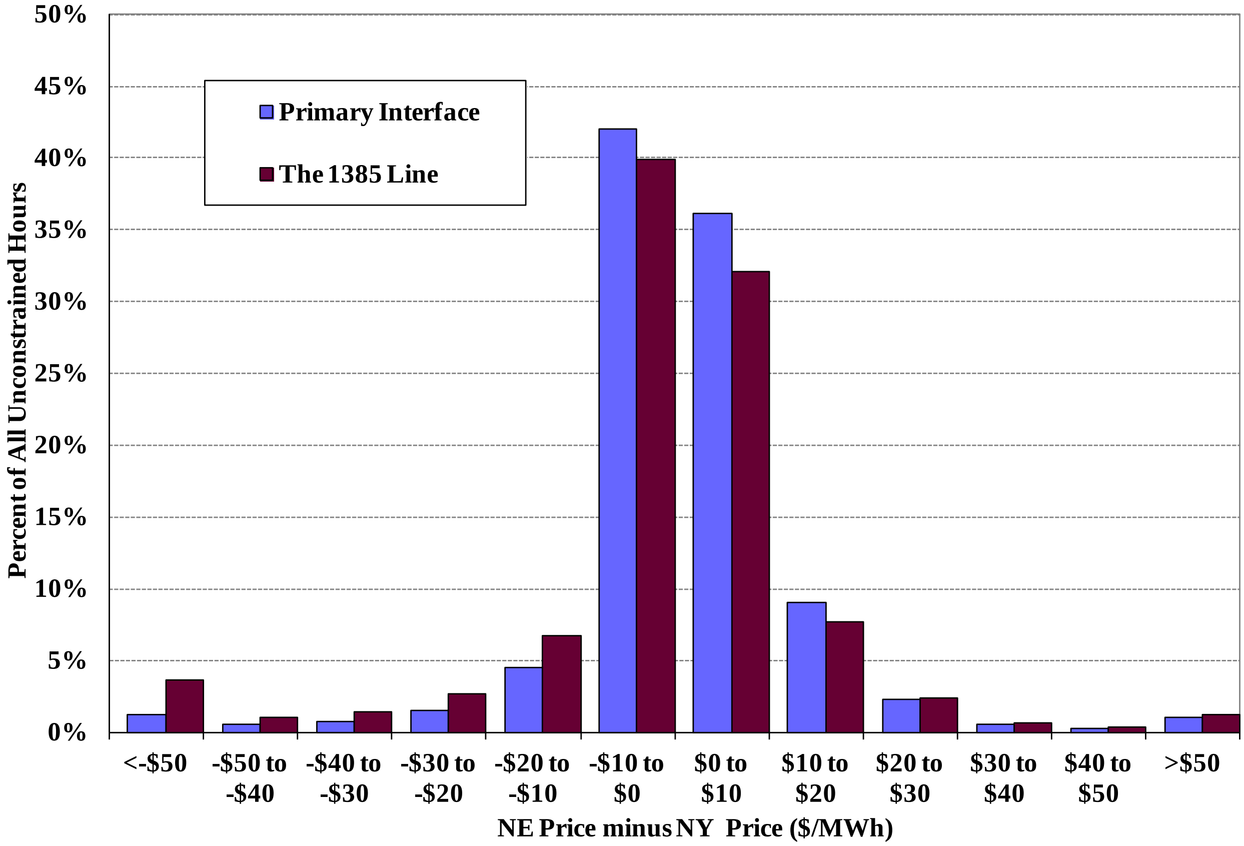
The following figure evaluates scheduling between New England and New York across the primary interface and the Northport Norwalk Scheduled Line (i.e., the 1385 Line). The Cross-Sound Cable is omitted because it is scheduled under separate rules.[61] Figure 16 shows the distribution of real-time price differences across the primary interface between New England and New York and the 1385 Line in hours when the interfaces were not constrained.[62]

While the factors described above prevent complete arbitrage of price differences between regions, trading should help keep prices in the neighboring regions from diverging excessively. Nonetheless, Figure 16 shows that although the price differences were evenly distributed around $0 per MWh, a substantial number of hours had price differences more than $10 per MWh for each interface. In 2012, the price difference between New England and New York exceeded $10 per MWh in 22 percent and 28 percent of the unconstrained hours for the primary interface and the 1385 Line, respectively. Additionally, the price difference was greater than $30 per MWh in 4 percent of the unconstrained hours for the primary interface and in 8 percent of the unconstrained hours for the 1385 Line.

61 Service over the Cross-Sound Cable is provided under the Merchant Transmission Facilities provisions in Schedule 18 of ISO-NE's Tariff, which is separate from the transmission service provisions governing use of the Pool Transmission Facilities. Access to the MTF requires Advance Reservations on the CSC, recommended to be acquired in advance of submitting transactions to the day-ahead market, and energy transactions accepted in ISO-NE and NYISO market systems. Scheduling limits restrict the ability to use the CSC interface for short-run arbitrage transactions between Connecticut and Long Island.

62 The prices used in this analysis are the prices at the New England proxy bus in the New York market (i.e., New York price) and the prices at the New York proxy bus in the New England market (i.e., New England price).

Figure 16: Real-Time Price Difference Between New England and New York
Unconstrained Hours, 2012



These results indicate that the current process does not fully utilize the interface. Given the pattern of price differences shown, there are many hours when increasing flows from the lower priced region to the higher priced region would have significantly improved the efficiency of clearing prices and production in both regions. This failure to fully arbitrage the interfaces leads to market inefficiencies that could be remedied if the ISOs were to coordinate interchange.

2. Efficiency of Scheduling Between New England and New York

Although market participants have not fully arbitraged the interface between New York and New England, the following analyses evaluate whether the direction of participants' transaction schedules have been consistent with the relative prices in the two regions and have, therefore, improved price convergence and efficiency.

Table 1 evaluates the relationship between real-time schedules and clearing prices for New England and New York across the primary interface and two scheduled lines (i.e., the 1385 Line and the Cross Sound Cable). The table shows: (a) the average hourly real-time flows between New England and New York (a positive number indicates a net import from New York); (b) the average real-time price differences between markets for each interface (a positive number indicates that the average price was higher on the New England side of the interface); and (c) the share of the hours when power was scheduled in the efficient direction (i.e., from the lower-price market to the higher-priced market).

Table 1: Efficiency of Real Time Schedules Between New York and New England
Over Primary Interfaces and Scheduled Lines, 2012

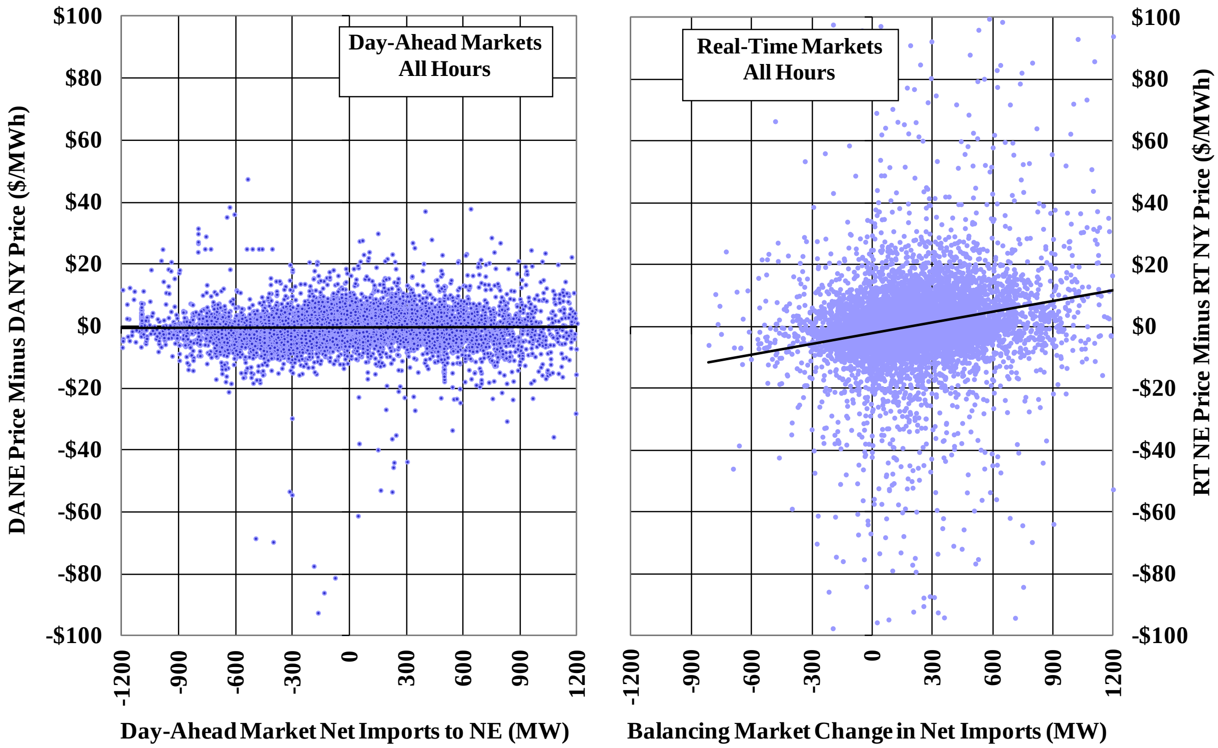
	Average Net Imports (MW/h)	Avg Internal Minus External Price ($/MWh)	Percent in Efficient Direction
Free-flowing Ties			
Northen New England	237	$0.59	52%
Controllable Ties			
1385 Line	-108	-$4.18	57%
Cross Sound Cable	-251	-$7.74	58%

The table shows that transactions scheduled by market participants flowed in the efficient direction in slightly over half of the hours on the three interfaces between New England and New York during 2012. The share of hours with efficient scheduling ranged from 52 percent over the primary interface to 58 percent over the Cross Sound Cable Scheduled Line. Large share of hours remained when power flowed in the inefficient direction on all three interfaces. Furthermore, there were many hours when power flowed in the efficient direction, but additional flows would have been necessary to fully arbitrage between markets.

The next analysis focuses on whether the incremental changes in participants' schedules (i.e., real-time adjustments from day-ahead schedules) have been consistent with the relative prices in the two regions. Figure 17 shows a scatter plot of net scheduled flows across the primary interface versus the difference in prices between New England and upstate New York for each hour in 2012. The left side of the figure shows price differences in the day-ahead market on the vertical axis versus net imports scheduled in the day-ahead market on the horizontal axis. The right side of the figure shows hourly price differences in the real-time market on the vertical axis versus the *change* in the net scheduled imports after the day-ahead market on the horizontal axis.

For example, if day-ahead net scheduled imports for an hour are 300 MW and real-time net scheduled imports are 500 MW, the change in net scheduled imports after the day-ahead market would be 200 MW (= 500 – 300).

Figure 17: Efficiency of Scheduling in the Day-Ahead and Real-Time
Primary Interface Between New England and New York, 2012



The trend lines in the left and right panels show statistically significant positive correlations between the price difference and the direction of scheduled flows in the day-ahead and real-time markets. However, the correlation in the day-ahead market is extremely weak, which indicates the difficulty participants have in scheduling transactions efficiently. The correlation is much stronger in the real-time market, reflecting larger price variations in each real-time market. These positive relationships indicate that the scheduling of market participants generally respond to price differences by increasing net flows scheduled into the higher-priced region. However, this response is incomplete and the interface remains substantially under-utilized as a result.

The difficulty of predicting changes in market conditions in real-time is reflected in the wide dispersion of points on the right side of Figure 17. More than 45 percent of the points in the real-time market panel are in inefficient quadrants – upper left and lower right – indicating hours when the net real-time adjustment by market participants shifted scheduled flows in the unprofitable direction (from the high-cost market to the low-cost market). Although market participant scheduling has helped converge prices between adjacent markets, Figure 17 highlights that considerable room for improvement remains.

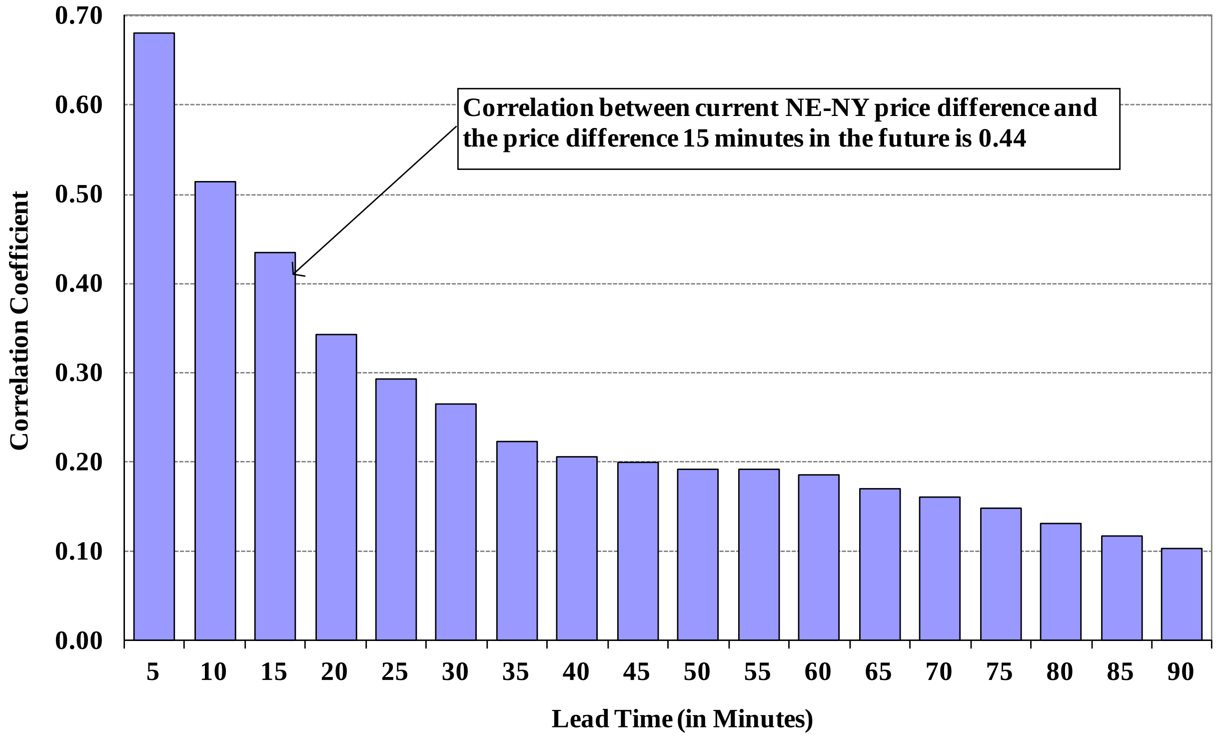
Although the arbitrage is not complete, the positive correlation between the price differences and the schedule changes indicate that participants generally respond rationally to the price differences in the real-time market. Additionally, total net revenues from cross-border scheduling in 2012 were $0.4 million in the day-ahead market and $8.5 million in the real-time market (not accounting for transaction costs).[63] The fact that significant profits were earned from the external transactions provides additional support for the conclusion that market participants generally help improve market efficiency overall by facilitating the convergence of prices between regions.

3. Correlation of Price Differences and Lead Time

The next analysis examines the correlation between the lead times for scheduling transactions and the predictability of price differences between adjacent markets. Figure 18 reports the correlation coefficient of the real-time price difference between New England and upstate New York between the current period and each subsequent five-minute period over 90 minutes. For example, the correlation of the price difference at the current time and the price difference 15 minutes in the future was 0.44 in 2012.

63 This likely underestimates the actual profits from scheduling because it assumes that day-ahead exports from one market are matched with day-ahead imports in the other market. However, market participants have other options such as matching a day-ahead export in one market with a real-time import in the other market. This flexibility actually allows participants to earn greater profits from more efficient trading strategies than those represented in the figure.

Figure 18: Correlation Between Price Differences and Lead Time
Interface between Upstate NY and New England, 2012



Not surprisingly, Figure 18 shows that actual price differences are more strongly correlated to price differences in periods near in time than to price differences in periods more distant in time. Hence, the further in advance a participant schedules a transaction, the less likely it is that the transaction will be efficient. Currently, to schedule transactions between New York and New England, market participants must submit their offers 75 minutes before the start of an hour, which is 75 to 135 minutes before the power actually flows since transactions are scheduled in one-hour blocks beginning at the top of the hour.

This analysis shows that reducing the lead times for scheduling would improve participants' ability to forecast the price differences and determine their schedules. However, the correlation remains relatively low at lead times of 15 minutes or more. The correlation was 0.44 at 15 minutes ahead of real time, which is the shortest scheduling lead time currently used by any RTO, indicating the difficulty of practically predicting changes in real-time market prices. Further, even if market participants were able to schedule much closer to real-time, they would still be at risk when scheduling in the efficient direction because the total quantity scheduled could lead to a reversal of the price spread between markets. Hence, the likely benefits of reducing scheduling lead-times are modest relative to the benefits from more direct coordination of the interchange. The next section describes how these issues can be more completely addressed through explicit coordination.

C. Coordination of Interchange by the ISOs

Incomplete price convergence between New England and New York suggests that more efficient scheduling of flows between markets would lead to production cost savings and substantial benefits to consumers. Although past efforts to reduce barriers to market participant scheduling between regions have improved the efficiency of flows and additional such efforts would lead to further improvements, uncertainty and risk are inherent in the market participant scheduling process. Hence, even with improvements, one cannot reasonably expect the current process to fully utilize the interface. As is the case for efficient scheduling of the transmission capability within ISO regions, optimal use of transmission capability between ISO regions requires explicit coordination of the interchanges by the ISOs.

In July 2010, ISO-NE and NYISO commenced a joint effort known as the Inter-Regional Interchange Scheduling project to address the issue of inefficient scheduling between the two markets. The RTOs proposed two solution options:

- Tie Optimization- The ISOs exchange information 15 minutes in advance and optimize the interchange based on a prediction of market conditions. The interchange would be adjusted every 15 minutes.
- Coordinated Transaction Scheduling (CTS)- Identical to Tie Optimization, except the interchange schedule is only adjusted to the extent that market participants have submitted intra-hour Interface Bids priced below the predicted price difference between the markets.

We employed simulations to estimate the benefits of these two initiatives. The benefits of efficient scheduling include reduced production costs and lower prices for consumers. The production cost net savings represent the increased efficiency of generator operations in both regions as additional production from lower-cost generators one ISO displaces production from higher-cost generators in the other ISO. The net consumer savings arise because improved coordination between the ISOs tends to lower prices on average in both regions.

The simulation results indicated significant potential benefits from fully optimizing the interchange, including roughly $17 million per year in production cost savings and $200 million per year in consumer savings. Both proposals would capture a large share of these potential benefits (60 to 70 percent). The Tie Optimization proposal performed slightly better in our simulations than the Coordinated Transaction Scheduling proposal. However, the benefits are very similar if participants submit relatively low-cost interface bids. [64]

Through their respective stakeholder processes, ISO-NE and NYISO decided to move forward with the CTS proposal to improve coordination between markets.[65] Accordingly, a market design project for CTS is currently under way and is scheduled to be effective in 2015.[66] Given the potential benefits from more efficient coordination with other control areas, we recommend that the ISO-NE continue to place a high priority on this initiative.

D. Conclusions and Recommendations

Efficient use of transmission interfaces between regions allows customers to be served by lower-cost external resources. New England imports large amounts of power from Quebec, which reduces wholesale power costs for electricity consumers in New England. Power flows in either direction between New England and New York, depending on market conditions in each region.

We find that the external transaction scheduling process was functioning properly and that scheduling by market participants tended to improve convergence, but significant opportunities remain to improve scheduled interchange between regions. Improving the efficiency of flows between regions is particularly important during shortages or very high-priced periods because modest changes in the physical interchange can substantially affect the market outcomes in both New England and New York.

ISO-NE and the NYISO are planning two initiatives that are intended to improve the efficiency of scheduling between the two control areas. First, the Coordinated Transaction Scheduling

64 For a detailed description of simulation models and results, see our *2010 Assessment of Electricity Markets in New England*, Section IV.C.

65 ISO-NE and NEPOOL filed the proposed tariff changes on February 24, 2012 in Docket ER12-1155-000. These were accepted by FERC on April 19, 2012.

66 See the *2013 ISO-NE Wholesale Markets Project Plan*, page 27.

(CTS) process is under development to coordinate the interchange between control areas. Under CTS, the ISOs will schedule interchange based on short-term forecasts of market conditions and new bidding procedures that will allow market participants to submit bids that are jointly evaluated by the ISOs. Second, market-to-market congestion management coordination will institute procedures for enabling one ISO to redispatch its internal resources to relieve congestion in the other control area when it is efficient to do so. The estimated benefits of the second initiative are substantially lower than the benefits of the coordinated interchange initiative given the current low levels of congestion in New England. We continue to recommend that ISO-NE and the NYISO place a high priority on implementing CTS.

V. Real-Time Pricing and Market Performance

The goal of the real-time market is to efficiently procure the resources required to meet the reliability needs of the system. To the extent that reliability needs are not fully satisfied by the market, the ISO must procure needed resources outside of the market process. This tends to distort the real-time prices and may indicate that there are reliability needs that are not fully priced. Both of these issues are significant because they undermine the efficiency of the real-time price signals. Efficient real-time price signals are essential because they encourage competitive conduct by suppliers, efficient participation by demand response, and investment in new resources or transmission where it is needed most. Hence, it is beneficial to regularly evaluate whether the market produces efficient real-time price signals.

In this section, we evaluate several aspects of the market operations related to pricing and dispatch in the real-time market in 2012. This section examines the following areas:

- Prices during the deployment of fast-start generators;
- Prices during shortages of operating reserves;
- Prices during the activation of real-time demand response;
- Efficiency of real-time ex post prices; and
- Frequency of price corrections.

At the end of this section, we provide a list of our conclusions and recommendations regarding the efficiency of real-time prices.

A. Real-Time Commitment and Pricing of Fast-Start Resources

Fast-start generators are capable of starting from an offline status and ramping to their maximum output within 30 minutes of notification, which enables them to provide valuable offline reserves. Areas without significant quantities of fast-start generation must maintain more reserves on online units, which can be very expensive. Another benefit of fast-start units is that they can ramp more quickly than most baseload units, better enabling the system operator to respond rapidly to unexpected changes in operating conditions. During such conditions, it is particularly important to operate the system efficiently and to set prices that accurately reflect the cost of satisfying demand and reliability requirements.

This section of the report discusses the challenges related to efficient real-time pricing when fast-start generators are the marginal supplier of energy in the market. It also evaluates the efficiency of real-time prices when fast-start generators were deployed by the real-time market in 2012. This can be an issue because fast-start peaking units are relatively inflexible once they are started. This causes them to frequently not set the real-time price, even when they are the marginal source of supply).

1. Treatment of Fast-Start Generators by the Real-Time Dispatch Software

This subsection describes how fast-start peaking units are committed by the real-time market dispatch software. The ISO's real-time dispatch software, called Unit Dispatch System (UDS), is responsible for scheduling generation to balance load and satisfy operating reserve requirements, while not exceeding the capability of the transmission system. UDS provides advance notice of dispatch instructions to each generator for the next dispatch interval based on a short-term forecast of load and other operating conditions.[67] Most commitment decisions are made in the day-ahead timeframe prior to the operation of UDS. UDS' primary function is to adjust the output levels of online resources. The only resources that UDS can commit (i.e., start from an offline state) are fast-start generators.[68] It is more efficient to allow UDS to start fast-start generators than to rely exclusively on operators to manually commit such units because UDS performs an economic optimization.[69]

When determining dispatch instructions for most online generators, UDS considers only incremental offers. However, for fast-start generators, UDS also considers commitment costs (since they must be committed from an offline state) and uses various assumptions regarding the dispatchable range of the generator. The treatment of commitment costs and the dispatchable

67 Generators are usually given instructions 15 minutes in advance, but this can be set higher or lower by the operator.

68 Fast-start units are units that are capable of providing 10-minute or 30-minute non-synchronous reserves and have a minimum run time and a minimum down time of one hour or less.

69 Based on its real-time optimization, UDS recommends that individual fast-start units be started. However, the final decision to start a unit remains with the real-time operator.

range have important implications for price setting by the real-time software (i.e., how real-time LMPs are determined).

UDS schedules fast-start generators using the following criteria:

- Offline fast-start generators – UDS considers commitment costs by adding the start-up offer (amortized over 1 hour) and "no-load" offers to the incremental offer. UDS treats the generator as having a dispatchable range from 0 MW to its maximum output level.
- Online fast-start generators during the minimum-run time – UDS considers only the incremental offer. UDS treats the generator as having a dispatchable range from its minimum output level to its maximum output level.
- Online fast-start generators after the minimum-run time has elapsed – UDS considers only the incremental offer. UDS treats the generator as having a dispatchable range from 0 MW to its maximum output level.

In the first phase of commitment listed above (when the unit is offline), real-time LMPs usually reflect the full cost of deploying the fast-start generator, partly because UDS considers the no-load offer and the start-up offer of the generator. Furthermore, UDS allows the fast-start generator to "set price" when the generator is economic to be online by treating the generator as dispatchable between 0 MW and the maximum output level.

However, in the second and third phases of commitment (i.e., once the unit is online), real-time LMPs frequently do not reflect the full cost of deploying the fast-start generator, even if the generator is still economic to be online. Since UDS does not consider the start-up and no-load offers, the real-time price-setting logic incorporates only the incremental offer. Furthermore, since the minimum output level of most fast-start generators is within 90 percent of their maximum output level, fast-start generators are frequently dispatched at their minimum output levels where they do not set price during the second phase of commitment. In such cases, the resulting LMP may be lower than the incremental offer of the fast-start generator.

The following example illustrates the pricing challenges when fast-start generators are deployed economically by the real-time market. Suppose UDS needs to schedule an additional 15 MW in an import-constrained area and the lowest cost supply is an offline fast-start generator with an incremental offer price of $75 per MWh, a no-load offer of $300 per hour, a start-up offer of $500 per start, a minimum output level of 18 MW, and a maximum output level of 20 MW. In this case, the average total offer of the offline unit is $115 per MWh ($75 per MWh + $300/hour

÷ 20 MW + $500/hour ÷ 20 MW) when it runs at full output for one hour. This total offer is used in the price-setting logic during the first phase of commitment.

In the start-up interval, UDS treats the fast-start generator as flexible and schedules 15 MW from the fast-start generator. This generator is the marginal generator and, therefore, sets the LMP at $115 per MWh. Since 15 MW is lower than the minimum output level of the generator, the generator is instructed to produce at its minimum output level. Once the generator is running (but before its minimum run period has expired) it is no longer possible to schedule 15 MW from the fast-start generator since the minimum output level (18 MW) is enforced. As a result, the fast-start generator is dispatched at 18 MW rather than 15 MW, and the output level of the next most expensive generator is reduced by 3 MW to compensate for the additional output from the fast-start generator. In this case, the fast-start generator is no longer eligible to set the LMP since it is at its minimum output level, so the next most expensive generator sets the LMP at a price lower than the incremental offer of the fast-start generator ($75 per MWh).

After the minimum run time elapses, UDS can schedule 15 MW from the fast-start generator if that is most economic, because the minimum output level is not enforced in this phase. In this case, the fast-start generator sets the LMP at its incremental offer of $75 per MWh. However, when the UDS solution reduces the output of the unit below its economic minimum, the operator must decide whether to decommit the resource.

In this example, the fast-start generator is dispatched in merit order, although the full cost of the decision is not reflected in real-time LMPs. The fast-start generator costs $115 per MWh to operate in the first hour and $90 per MWh thereafter; however, the LMP is set to $115 per MWh in the first UDS interval (usually approximately 10 minutes), less than $75 per MWh for the remainder of the first hour, and $75 per MWh thereafter. This issue is worse when an operator commits a fast-start generator for reliability. In this case, the unit will generally operate at its economic minimum and not set prices because they tend to be higher cost than the units committed economically. In both cases, the owner of the fast-start unit would receive NCPC payments to make up the difference between the total offer and the real-time market revenue.

2. Evaluation of Fast-Start Deployments by UDS in 2012

The following two analyses assess the efficiency of real-time pricing during periods when fast-start units were deployed in merit order. The first analysis summarizes how consistent the real-time prices are with the offer costs of the fast-start resources. The second analysis evaluates how LMPs would be affected if the average total offers were fully reflected in real-time prices.

Figure 19 summarizes the consistency of the real-time LMP with the average total offer for fast-start units committed economically by UDS. The average total offer includes no-load and start-up costs amortized over one hour and the comparison is made over the units' over the initial commitment period, which is usually one hour. When the average real-time LMP is greater than the average total offer, the figure shows the associated capacity in the category labeled "Offer (including Startup) < LMP". However, when the average real-time LMP is less than the average total offer, LMPs do not fully reflect the cost to the system of deploying the fast-start generator. Figure 19 shows hydroelectric and thermal units separately, and categorizes such occurrences in five categories based on the relative economics of the units.

Figure 19: Comparison of Real-Time LMPs to Offers of Fast-Start Generators
First Hour Following Start-Up by UDS, 2012

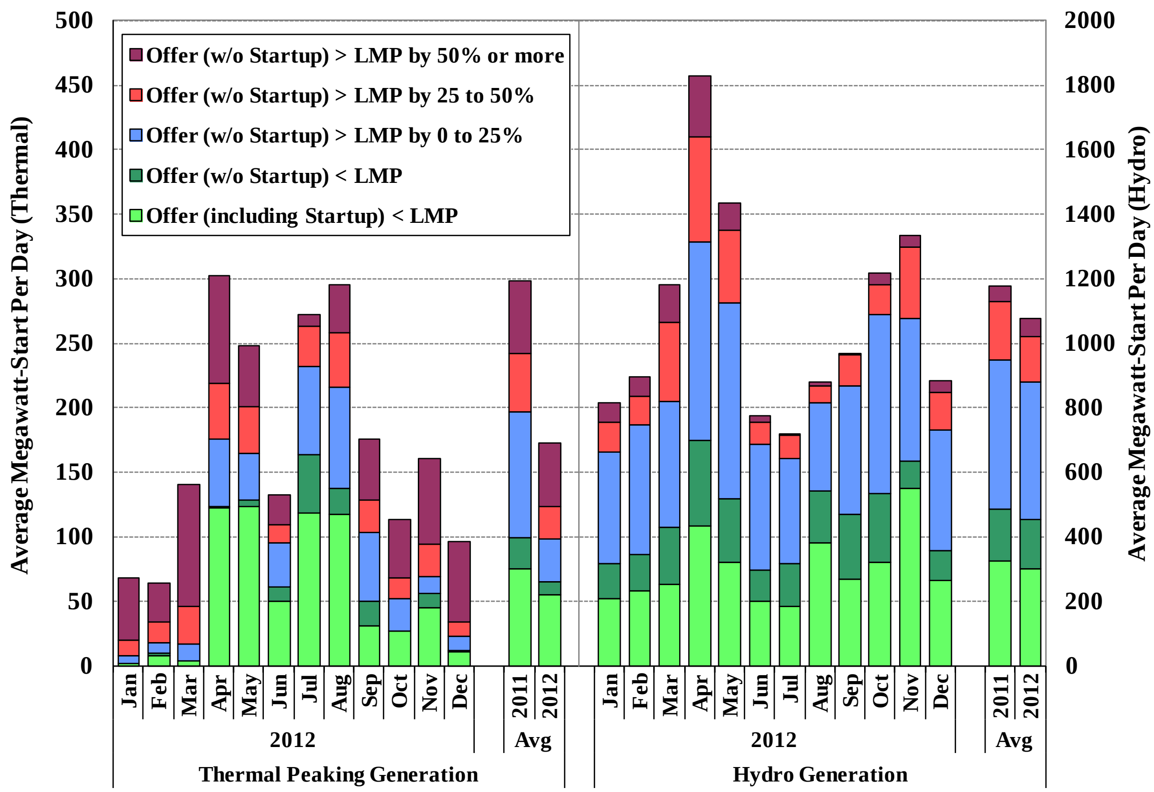


Figure 19 shows that flexible hydro generation accounted for over 80 percent of fast-start generation that was started in merit order by UDS in 2011 and 2012. This indicates that hydro generators are generally less costly and have sufficient water to operate on a daily basis. Many of the thermal peaking units in New England have low capacity factors because of their high production costs.

The amount of thermal peaking generation that was started in merit order by UDS fell 42 percent from 2011 to 2012. Part of this decrease may be due to the increase in supplemental commitment in 2012. This increase can reduce the need to commit thermal peaking units for ramping needs.

The overall pricing efficiency during hours when fast-start resources are committed by UDS did not change significantly from 2011 to 2012. The average total offer (including start-up costs) was higher than the real-time LMP in 72 percent of starts in 2012, compared to 73 percent in 2011. This ratio was similar for thermal peaking resources and hydro generation resources. Hence, real-time prices do not usually reflect the full cost of satisfying load when fast-start resources are deployed.

Although thermal peaking generators are deployed in a relatively limited number of hours, they are frequently the marginal source of supply in the hours that they run. This makes it particularly important to reflect the full cost of their deployment in real-time LMPs when they are deployed efficiently in merit order. Even when start-up costs are excluded, there were still 62 percent of the thermal peaking generation that exhibited offers greater than the real-time LMP. Hence, even though these units are economic, they often relied on NCPC payments to recoup their full as-bid operating costs. More importantly, these results indicate that real-time prices do not accurately reflect the marginal cost of serving real-time demand, which affects the economic signals provided by the day-ahead and forward markets in New England. The following analysis examines how real-time energy prices would be affected if the average total offers of such units were reflected in real-time LMPs.[70]

70 If a gas turbine from the earlier example was started with a total offer of $115/MWh when the LMP was $75/MWh, this analysis would assume the unit would increase the LMP by $40 per MWh. Other lower-

Figure 20 summarizes the portion of the fast-start units' costs that were not fully reflected in real-time LMPs in 2012. The lower portion of Figure 20 shows how frequently thermal and hydro fast-start units were started economically by UDS when their average total offers were greater than the LMP during the minimum run time in 2012.[71] The figure excludes fast-start units that were started in import-constrained areas since these LMPs would be representative of only a limited area of New England.[72] The upper portion of the figure shows the difference between the average total offer and the real-time LMP from such periods averaged over the year by time of day. The figure also shows potential market impacts separately for hydro and thermal fast-start units by dividing all examined market intervals into two groups: (a) "Hydro Only" intervals if only flexible hydro resources are started and running in that interval; and (b) "Thermal & Hydro" intervals if at least one thermal fast-start unit is started and running in that interval.[73] The table in the chart summarizes our estimates of the reduction in NCPC if the average total offers of such units were fully reflected in real-time LMPs.

Figure 20 shows that fast-start units were deployed economically by UDS when their average total offer was greater than the real-time LMP in a substantial portion of hours. Such hours were most frequent from hours-ending 7 to 22, particularly around the morning peak (hours-ending 8 to 12) when load picks up rapidly and the evening peak (hours-ending 18 to 21). Ramping needs are highest on the system during these periods, so fast-start generation is sometimes needed to satisfy load during these periods.

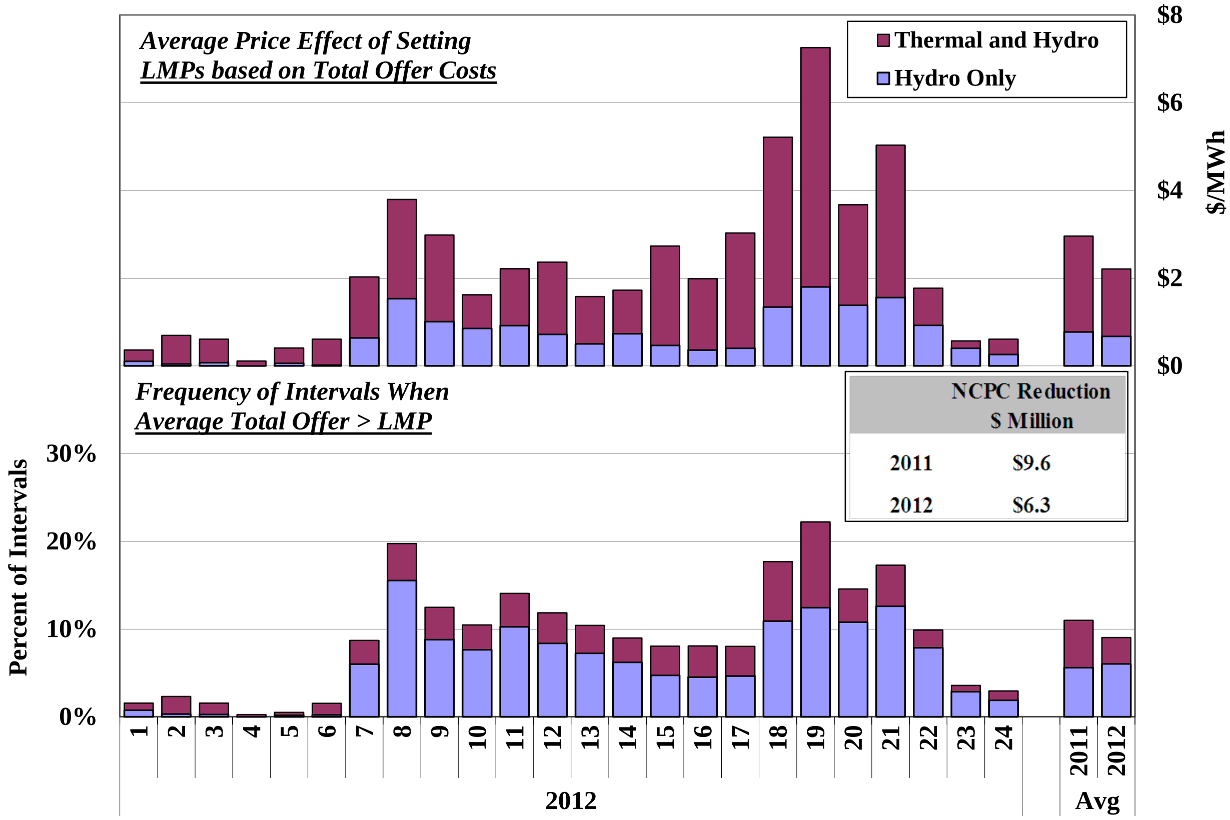
cost gas turbines or flexible hydro resources started in the same hour would not affect prices because they are inframarginal.

71 If multiple fast-start units are started at one time, the analysis uses the one with the largest difference between the average total offer and the real-time LMP, which is usually the highest-cost unit.

72 The area is treated as import-constrained if the congestion component of the LMP at the fast-start unit's node is greater than the congestion component at New England Hub by $1 per MWh or more.

73 Since thermal fast-start units typically have higher offer costs, lower-cost hydro fast-start units started in the same hour would generally not affect prices because they are inframarginal.

Figure 20: Difference Between Real-Time LMPs and Offers of Fast-Start Generators
First Hour Following Start-Up by UDS, 2012



Overall, fast-start units were started economically by UDS when their average total offer exceeded the real-time LMP during the minimum run time in 9 percent of all hours in 2012. This is slightly lower than in 2011 due to the reduced operation of thermal peaking units. In 2012, hydro-only fast-start units were started during two-thirds of these hours (i.e., 6 percent of all hours in 2012), while at least one thermal fast-start unit was started in the remaining one-third of these hours.

If the average total offers of these units were fully reflected in the energy price in these hours, the average real-time LMP would increase approximately $2.21 per MWh in 2012. Almost three-quarters of this increase is attributable to allowing thermal peaking generators to set prices, even though these units are started much less frequently than the hydroelectric units. Overall, the estimated price increase would be largest in hour-ending 19 when the average LMP would rise by $7.25 per MWh. If the estimated price increases were reflected in the calculation of

NCPC uplift charges, we estimate that they would be reduced by $9.6 million in 2011 and $6.3 million in 2012.

These differences likely overstate the impact from more efficient real-time pricing during fast-start resource deployments because they do not consider the likely market responses to the higher real-time prices:

- Incentives to purchase more in the day-ahead market would increase, which would increase the amount of lower-cost generation committed in the day-ahead market.
- Net imports would increase from neighboring control areas, particularly New York.

Hence, the actual effect on real-time LMPs from more efficient pricing during fast-start deployments and resulting reductions in NCPC uplift charges would be smaller than the results shown in Figure 20. However, these responses would substantially improve efficiency because higher-cost peaking generation would be displaced by lower-cost intermediate generation and net imports. Allowing peaking resources to set prices when marginal would also improve the incentives governing longer-term investment and retirement decisions by participants.

Therefore, we continue to recommend that the ISO evaluate potential changes in the pricing methodology that would allow the deployment costs of fast-start units to be more fully reflected in the real-time market prices. The NYISO has a methodology for allowing fast-start resources to set the real-time LMP, and the Midwest ISO is preparing to implement a similar methodology. We recommend ISO-NE consider using a similar mechanism to improve the efficiency of real-time pricing when fast-start resources are the marginal sources of supply.[74]

B. Real-Time Operation and Pricing During Operating Reserve Shortages

In the real-time market, the Reserve Constraint Penalty Factors ("RCPFs") limit the costs that the model may incur to meet the reserve requirements (i.e., marginal dispatch actions that would exceed the relevant RCPF are foregone). Consequently, if the cost of maintaining the required

74 The MISO is currently planning to implement this methodology, which is known as "ELMP" in 2014.

level of a particular reserve exceeds the applicable RCPF, the real-time market model will allow a reserve shortage and set the reserve clearing price based on the level of the RCPF.[75,76]

The RCPF levels are important because they determine how the real-time market responds under tight operating conditions. When it is not possible to meet the reserve requirements, the RCPFs prevent the model from incurring extraordinary costs for little or no reliability benefit. However, if RCPFs are not sufficiently high, the model may not schedule all available resources to meet the reliability requirements and real-time clearing prices may not adequately reflect the market conditions when this occurs. In such cases, the operator will likely intervene to maintain reserves and significantly affect market clearing prices in the process. Hence, it is important to evaluate the RCPF levels periodically to determine whether modifications are warranted.

Accordingly, ISO-NE recently assessed the appropriateness of the $100 per MWh RCPF for the system-level 30-minute reserves requirement and determined to raise it to a higher level of $500 per MWh, effective on June 1, 2012. Figure 21 shows the number of 5-minute market intervals during which the shadow price of the system-level 30-minute reserve requirement was at least $100 per MWh in 2011 and 2012.

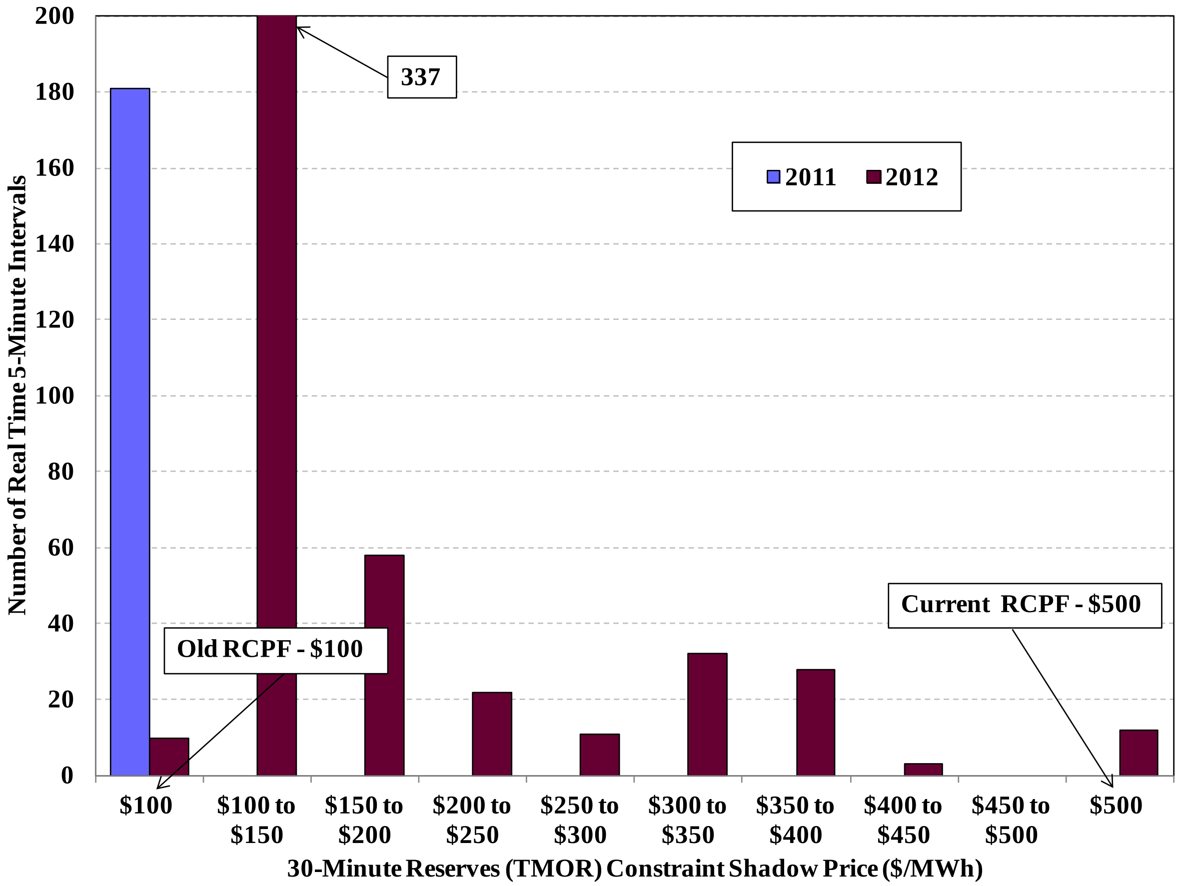
The figure shows that operating reserve shortages occurred very rarely after the increase in the RCPF. Of the 503 five-minute intervals shown in 2012 when the shadow price exceeded $100 per MWh, reserve shortages occurred in just 12 intervals. This is because the real-time model dispatches all of the available resources to satisfy the 30-minute reserve requirement when it is possible to do so at a cost of less than $500 per MWh.

---

75 For example, suppose an online generator with a $60 per MWh incremental offer could be backed down to provide reserves when the LMP is $160 per MWh. In this case, the marginal cost to the system of providing reserves from this unit is the opportunity cost of the unit not providing energy at the LMP. This opportunity cost is equal to the difference between the LMP and the incremental offer of the unit or $100 per MWh in this example ($160 per MWh LMP minus $60 per MWh incremental cost). If the RCPF is $50 per MWh, the market will not back the unit down to provide reserves and the system would be short of reserves since the marginal system cost of doing so ($100 per MWh) exceeds the RCPF ($50 per MWh).

76 If only one reserve constraint is binding, the reserve clearing price will be set equal to the RCPF of the reserve that is in shortage. However, if multiple reserve constraints are binding, the reserve clearing price will be set equal to the sum of shadow prices of the binding constraints.

Figure 21: Distribution of System-level TMOR Constraint Shadow Cost
Intervals with Shadow Prices of $100 or More, 2011-2012



Before the RCPF was increased, 30-minute reserve shortages occurred much more frequently (in 181 five-minute intervals during the period shown). To maintain adequate reserves before the RCPF was increased, the ISO would have to take more frequent out-of-market actions, including curtailing exports to neighboring areas, manually dispatching online generators with available capacity that was not providing 30-minute reserves, and manually committing slow-start generators to bring additional capacity online. Out-of-merit actions undermine the efficiency of the market because: (a) they artificially lower energy and operating reserve prices below levels that reflect the costs of maintaining reliability, and (b) they are more costly than dispatching the available resources in the real-time market.

In addition, the new RCPF provides more efficient price signals during reserve shortages, which will provide better incentives for resources to be available and perform reliably under high load conditions in at least two ways. First, higher prices will provide better incentives for imports from New York and other areas with available capacity. Second, more efficient prices will improve the incentives for slow-starting generators to be committed in the day-ahead market, thereby increasing the availability of resources in real-time.

C. Real-Time Pricing During the Activation of Demand Response

Price-responsive demand has the potential to enhance wholesale market efficiency in theory. Modest reductions in consumption by end-users in high-price periods can significantly reduce the costs of committing and dispatching generation. Furthermore, price-responsive demand reduces the need for new investment in generating capacity. Indeed, the majority of new capacity procured in the first six Forward Capacity Auctions was composed of demand response capability rather than generating capability. As interest increases in demand response programs and time-of-day pricing for end-users, demand will play a progressively larger role in wholesale market outcomes. This part of the section discusses the effects of demand response programs on the efficiency of real-time prices in the wholesale market.

1. Real-Time Demand Response Programs and Participation

Prior to the beginning of the first Forward Capacity Commitment Period on June 1, 2010, the ISO was operating the following four active real-time demand response programs:

- <u>Real-Time 30-Minute Demand Response Program</u>. These resources could be deployed for anticipated capacity deficiencies with 30 minutes of notice and received the higher of the LMP or $500 per MWh for a minimum duration of 2 hours.

- <u>Real-Time 2-Hour Demand Response Program</u>. These resources could be deployed for anticipated capacity deficiencies with 2 hours of notice and received the higher of the LMP or $350 per MWh for a minimum duration of 2 hours.

- <u>Real-Time Profiled Response Program</u>. These resources could be interrupted for anticipated capacity deficiencies within a specified time period and received the higher of the LMP or $100 per MWh for a minimum duration of 2 hours.

- <u>Real-Time Price Response Program.</u> These resources could reduce load (but are not required to do so) when they received notice on the previous day. If they reduced their load, they received the higher of the LMP or $100 per MWh for the eligibility period.

The first three programs were reliability-based programs that activated emergency demand response resources according to the OP-4 protocol during a capacity deficiency, and the resources received capacity payments for being available to do so.[77] The fourth program was a price-based program that provided a mechanism for loads to respond when the wholesale price was expected to be greater than or equal to $100 per MWh, and it was the only one of the four that was originally extended beyond the start of the first Capacity Commitment Period under FCM.[78]

Many resources transitioned from one of the above programs to one of the following programs under the FCM:

- Real-Time Demand Response. Demand resources comprising installed measures (e.g., products, equipment, system, services, practices, and/or strategies) at end-use customer facilities. These resources may be deployed by the ISO with 30 minutes of notice.

- Real-Time Emergency Generation. Distributed generation whose federal, state and/or local air quality permit(s) limit their operation to hours when the ISO dispatches Real-Time Emergency Generation Resources. These resources may be dispatched by the ISO with 30 minutes of notice.

- On-Peak Demand Resource. These typically consist of non-dispatchable measures that are not weather sensitive and reduce load across the per-defined hours. On-Peak Demand Resources measure their load reduction during (i) summer on-peak hours (1:00pm – 5:00pm on non-holiday weekdays from June to August), and (ii) winter on-peak hours (5:00pm – 7:00pm on non-holiday weekdays in December and January).

- Seasonal Peak Demand Resource. This is designed for non-dispatchable, weather sensitive measures (e.g., energy efficient HVAC measures). These resources must reduce load during non-holiday weekdays when the real-time system hourly load is equal to or greater than 90 percent of the most recent "50/50" system peak load forecast for the applicable Summer or Winter season.

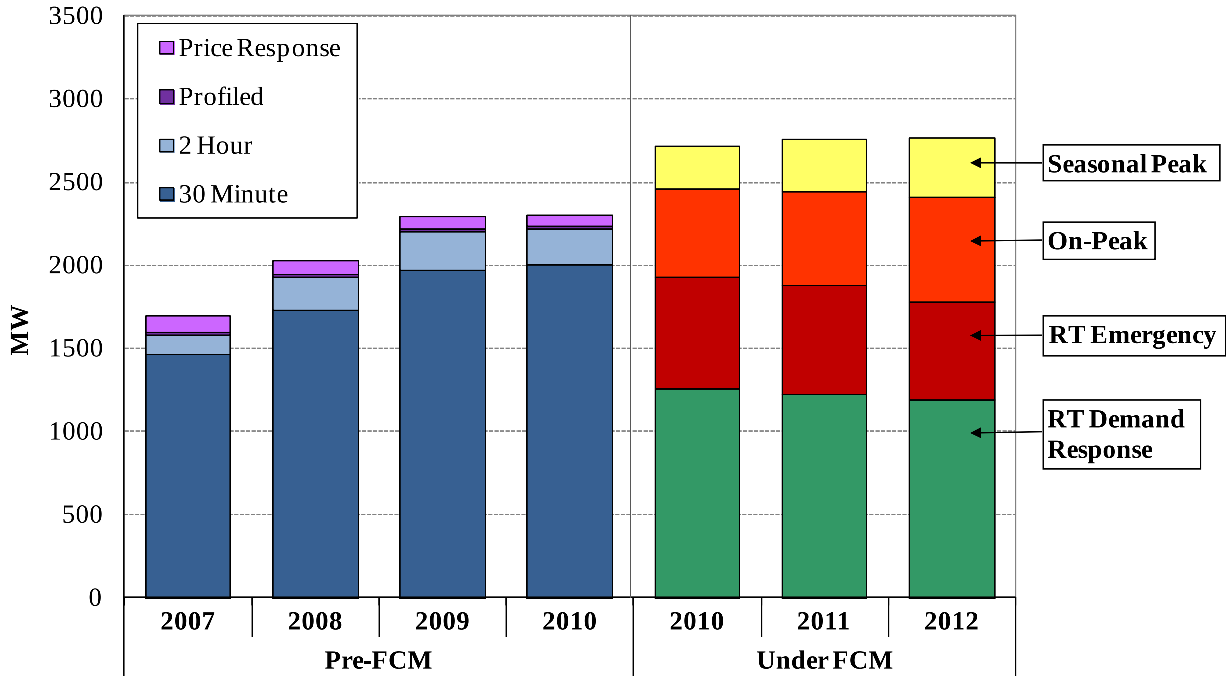
The first two are *active* (i.e., dispatchable) demand resources that operate based on real-time system conditions via dispatch by the ISO. They are defined at the Dispatch Zone level and

77 Real-Time 30-Minute Demand Response Program resources are activated under OP-4 Actions 9 and 12. Real-Time 2-Hour Demand Response Program resources and Real-Time Profiled Response Program resources are activated under OP-4 Action 3.

78 Resources in the Real-Time Price Response Program do not receive capacity payment. This program expired on May 31, 2012. Beginning June 1, 2012, the Transition Period Price Responsive Demand Program paid demand response resources that curtailed the real-time LMP rather than the higher of the LMP or $100 per MWh.

reduce energy demand during OP-4 conditions.[79] The last two are *passive* (i.e., non-dispatchable) demand resources that are defined at the Load Zone level and reduce energy demand during peak hours.[80] Demand response participation has surged in New England in recent years. Figure 22 shows the quantity of resources enrolled in each of the real-time demand response programs from 2007 to 2012. The quantities reported in this figure represent enrollments at the end of each year, except the quantities reported for pre-FCM periods during 2010 represent enrollments on May 31, 2010.

Figure 22: Real-Time Demand Response Program Enrollments
2007 – 2012



During the periods before the first FCM Capacity Commitment Period commenced, the quantity of enrolled resources increased from 1,694 MW in 2007 to 2,298 MW in 2010. Most demand response capacity was enrolled in the Real-Time 30-Minute Demand Response Program (87

79 There are 19 dispatch zones defined in New England: Northwest Vermont, Vermont, New Hampshire, Seacoast, Maine, Bangor Hydro, Portland ME, Western MA, Springfield MA, Central MA, North Shore, Boston, SEMA, Lower SEMA, Norwark-Stamford, Western CT, Northern CT, Eastern CT, and Rhode Island. Real-time demand response resources can be called under OP-4 Action 2, and real-time emergency generation resources can be called under OP-4 Action 6.

80 There are eight load zones defined in New England: Vermont, New Hampshire, Maine, Southeast Massachusetts, West Central Massachusetts, North East Massachusetts, Connecticut, and Rhode Island.

percent at the end of May 2010). The enrollment in the Real-Time Price Response Program decreased over the period, from 98 MW in 2007 to 65 MW in 2010.

The FCM has attracted more passive demand response resources and less active demand response than the previously existing programs. Nonetheless, a total of 2,769 MW of demand resources were enrolled by the end of 2012, with 64 percent (or 1,781 MW) being active resources. Hence, capacity payments before and under FCM have encouraged the development of demand response resources, which is discussed in detail in Section VII.

2. Real-Time Pricing During Activation of Real-Time Demand Response

The rise in demand response participation is beneficial in many ways, but it also presents significant challenges for efficient real-time pricing. Active demand resources procured in the forward capacity market (i.e., Real-Time Demand Response and Real-Time Emergency Generation) are currently not dispatchable within the real-time dispatch software and cannot, therefore, set real-time energy prices. Instead, they are dispatched as part of the OP-4 procedures under Actions 2 and 6. [81,82]

The activation of demand response in real time can inefficiently depress real-time prices substantially below the marginal cost of the foregone consumption by the demand response resources, particularly during shortages or near-shortage conditions. Although there is little information available on the marginal cost of foregone consumption for demand response resources, the marginal costs of most demand response resources are likely to be much higher than the marginal costs of most generators. Hence, real-time prices should be very high when demand response resources are activated.

In 2012, there were no capacity deficiencies that required activating emergency demand response resources, so the market outcomes were not affected by these pricing issues. Additionally, participation in the Real-Time Price Response Program and the Transitional Price Response

81 "Dispatchable" refers to resources that are able to modify their consumption or generation in response to remote dispatch instructions from the ISO generated by the real-time market.

82 Loads that are dispatchable in the real-time market are able to participate in the Asset Related Demand (ARD) programs. ARDs are paid according to day-ahead and real-time LMPs. ARDs are not paid for capacity, however, they are also not charged for capacity obligations.

Program is still relatively limited. Nevertheless, the activation of demand response is likely to occur more frequently in the future, making it important to address these pricing issues in the development of new demand response programs.

To this end, ISO-NE filed with the Commission in August 2011 to allow active demand response resources to submit multi-part offers into the day-ahead and real-time markets and for the ISO to schedule demand response resources in merit order as it would for a generating resource.[83] These new demand response programs are scheduled for implementation on June 1, 2016. Since the new demand response programs will allow resources to offer based on their marginal willingness to consume and be scheduled in economic merit order (rather than be activated based on an operating procedure), it should have a better basis for allowing demand response resources to set prices. However, most demand response resources will still likely be relatively inflexible on a five-minute basis, so market developments that allow fast start resources to set prices should be applicable to demand response resources as well. To the extent that a portion of the demand response resources continue to be available only during emergencies (i.e., not economically through the ISO markets), the ISO should consider additional provisions to allow these resources to set prices.

D. Ex Ante and Ex Post Pricing

Ex ante prices are produced by the real-time dispatch model (UDS) when it determines dispatch instructions, although the ISO uses ex post prices to settle with market participants in the real-time market. In this section, we examine inconsistencies between the ex ante and ex post prices, and we identify several factors that can undermine the efficiency of the ex post prices.

Ex ante prices are produced by the real-time dispatch model (UDS) and are consistent with the cost-minimizing set of dispatch instructions produced by UDS. They are consistent in the sense that the offer prices of dispatched resources are less than or equal to the LMP and the offer prices of un-dispatched resources are greater than or equal to the LMP. Hence, ex ante prices are set to levels that give generators an incentive to follow their dispatch instructions (assuming they are

83 See the Commission order accepting ISO-NE's compliance filing to Order 745: ISO-NE Inc., 138 FERC ¶ 61,042 (January 19, 2012).

offered at marginal cost). Because these prices are consistent with the optimized dispatch, they are an efficient reflection of the prevailing market conditions.

Ex post prices are produced by the LMP Calculator. At the end of each interval, the LMP Calculator re-calculates dispatch quantities and prices using inputs that are different in several respects from the inputs used by UDS. For each flexible resource, a "real-time offer price" is used in place of its offer curve.[84] For a resource following dispatch instructions, its real-time offer price equals the ex ante price at its location or, if it is operating at its maximum output level, the offer price corresponding to its actual production level. For a resource that is under-producing, the real-time offer price equals the offer price corresponding to the resource's actual production level. Each flexible resource is treated as having a small dispatchable range around its actual production level, where the upward range is much smaller than the downward range (e.g., approximately 0.1 MW up and 2 MW down). The purpose of the ex post pricing method is to generate a set of prices that is consistent with the actual production levels of generators in the market, rather than their dispatch instructions. This is intended to improve the incentives of generators to follow dispatch instructions.

The evaluation in this section identifies three inconsistencies between ex ante and ex post prices in 2012:

- The current implementation of ex post pricing results in a small (0.4 percent) but persistent upward bias in real-time prices;
- Inconsistencies between ex ante and ex post prices do not improve the incentives of generators to follow dispatch instructions; and
- Occasional distortions in the ex post prices lead to inefficient pricing in congested areas.

The end of this section provides a summary of the conclusions and recommendations from the evaluation of ex post pricing.

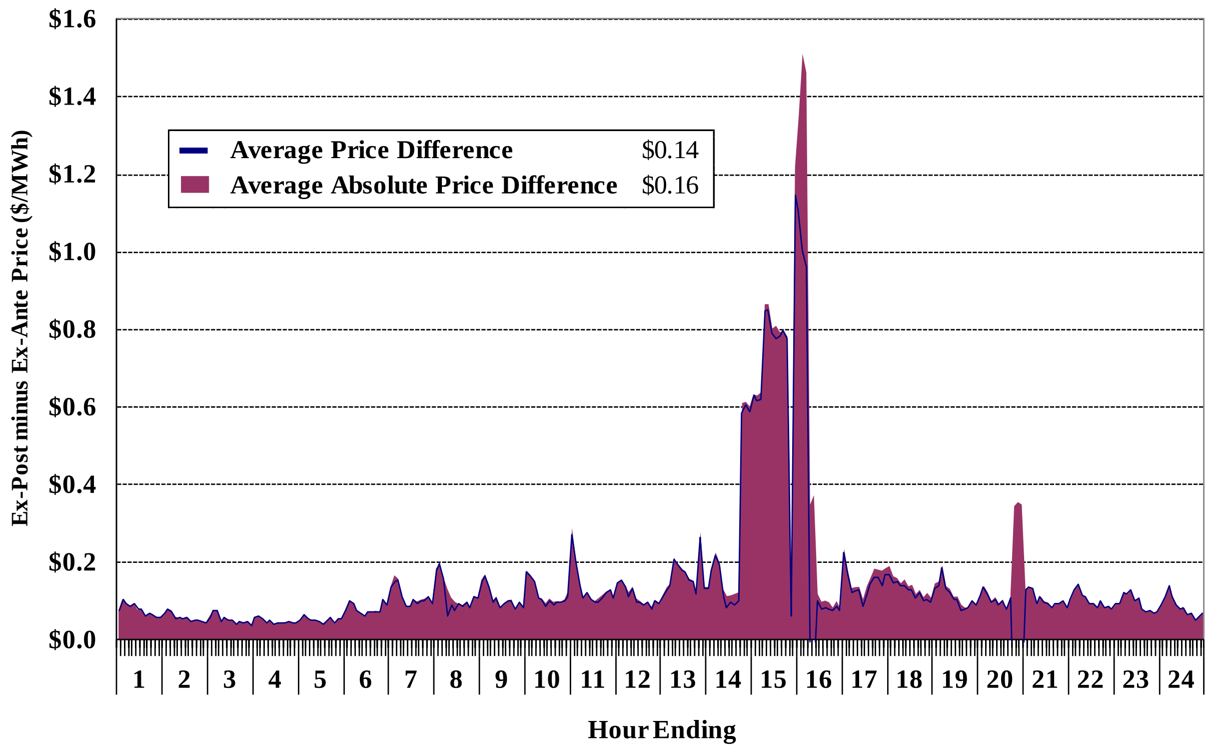
84 For most resources, they are treated as flexible if they are producing more than 0 MW and they meet one of the following conditions: (i) being committed for transmission, (ii) being dispatchable and producing less than 110 percent of their dispatch instruction, and (iii) being dispatchable and having a real-time offer price at their actual production level that is less than or equal to the ex ante price.

1. Persistent Differences Between Ex Ante and Ex Post Prices

The first analysis highlights an issue with the current implementation of ex post pricing that leads to a small but persistent upward bias in real-time prices. Figure 23 summarizes differences between ex ante and ex post prices in 2012 at a location close to the New England Hub.[85] This location is relatively uncongested, making it broadly representative of prices throughout New England. The blue line shows average ex post price minus average ex ante price by the time of day. The purple area shows the average absolute price difference by the time of day.

The average differences between the ex post and ex ante prices were relatively small in 2012. However, the line shows a persistent bias that causes the ex post prices to be slightly higher than ex ante prices in the vast majority of intervals. As a result, average ex post prices were $0.14 per MWh higher than ex ante prices at this location in 2012.

Figure 23: Average Difference Between Five-Minute Ex Post and Ex Ante Prices
2012



85 The MillBury station was selected because it is near the New England Hub. The New England Hub was not chosen because UDS does not calculate ex ante prices for load zones or the New England Hub.

Figure 23 shows that the average ex post price is greater than the average ex ante price in 98 percent of intervals. This persistent bias is the result of the interaction between the following two factors. First, loss factors change slightly due to the time lag between the calculation of the ex ante and ex post prices. Even though many units' real-time offer prices are equal to the ex ante price (which should make them economically equivalent), these changes in loss factors affect the offer costs of some resources relative to others, which causes the ex post pricing model to move resources. Second, the dispatchable range of each resource is generally 20 to 40 times larger in the downward direction than the upward direction.

In a typical interval, there may be 100 or more flexible resources. At locations where the loss factors increase the most from the ex ante model to the ex post model, resources will appear most costly and be ramped downward in the ex post model. Since the downward dispatchable range is much larger than the upward dispatchable range, many resources will be ramped up to their maximum to replace the unit that is ramped down. In a typical interval without congestion, four or five units are ramped down and 100 or so units are ramped up. As units that are ramped up in the ex post model reach their maximums, increasingly expensive units set ex post prices. Hence, the resource that is marginal in the ex post calculation usually has a loss factor that is higher than in the ex ante calculation, thereby leading to an upward bias in prices.

2. Theoretical Problems with Ex Post Pricing

Proponents have justified ex post pricing partly as a means to provide resources with incentives to follow dispatch instructions. However, ex post pricing does not efficiently provide such an incentive for several reasons. First, suppliers that are primarily scheduled day-ahead will not be substantially harmed by small adjustments in the real-time price because very little of their output is settled at real-time prices. Second, with the exception of the episodic price effects in congested areas, which are discussed in Part 3 of this subsection, the pricing methodology will not usually result in significant changes in prices when a unit does not follow dispatch instructions. In general, this is the case because many other units will have real-time offer prices in the ex post model that are very close to the offer price of the unit failing to following dispatch. Further, any slight change in the ex post price will not affect the unit failing to follow dispatch in a manner that has any relationship to the cost to the system of its actions. Hence, it is very unlikely that the ex post pricing enhances incentives to follow dispatch instructions. In fact, because ex post pricing can, on occasion, substantially affect prices in congested areas, it can diminish suppliers' incentives to follow ex ante dispatch instructions when prices in the congested area are volatile. A much more efficient means to send targeted incentives to respond to dispatch instructions is the use of "uninstructed deviation" penalties.[86]

A final theoretical concern is that ex post prices are theoretically less efficient than ex ante prices. The ex ante dispatch and prices represent the least cost dispatch of the system, given bids, offers, and binding constraints. If a unit is unable to respond to the dispatch instruction, then it implies that less supply is available to the market, and thus, the price should have been set by a more expensive offer. In other words, a higher-cost offer would have been taken if the market had known the unit could not respond. In such a case, however, the ex post pricing method would reduce the energy prices from the ex ante level because the marginal unit loses its eligibility to set prices. Due to the specific implementation in New England, this theoretical concern is rarely manifested.

3. Ex Post Pricing in Congested Areas

On occasion, there are large differences between ex ante prices and ex post prices in congested areas. Such occasions arise when the marginal unit for the binding constraint becomes inflexible or flexible but with a reduced offer price in the ex post pricing.[87]

For example, suppose a combustion turbine with an incremental offer of $150 per MWh and an amortized start-up and no-load cost of $100 per MWh is started in order to resolve a load pocket constraint. Suppose that there is also a $50 per MWh unit in the load pocket that is dispatched at its maximum level. The ex ante LMPs in the load pocket will be $250 per MWh. Two pricing inefficiencies can occur in the ex post calculation. First, if the combustion turbine has not started because its start-up time has not elapsed or because it comes on late, the turbine will be deemed inflexible in the ex post calculation. This causes the $50 per MWh unit to set prices because it is

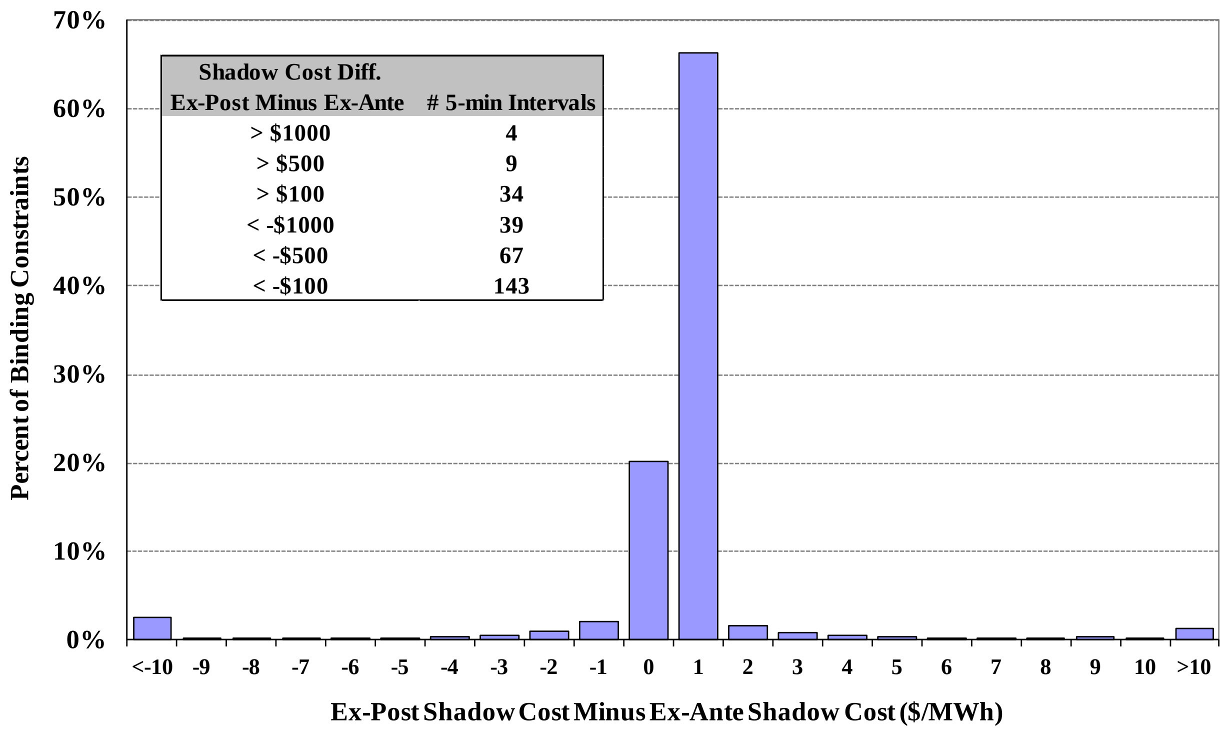
86 Uninstructed deviation penalties are penalties applied to suppliers that are not within a specified range of the dispatch instruction sent by ISO-NE.

87 When a fast-start unit is committed by UDS, its combined offer that adds its start-up and no-load offers on top of its incremental energy offer is used. In the ex post pricing, however, when the unit's offer is used, the start-up and no-load offers are not included.

the only flexible resource in the load pocket. Second, if the combustion turbine does start-up and is deemed flexible, the amortized start-up and no-load offers are not reflected in the current ex post pricing. As a result, the turbine would set a $150 per MWh ex post price in the load pocket. In either case, the ex post congestion value is substantially reduced, causing significant discrepancy between ex ante and ex post prices in the load pocket. In both cases, the marginal source of supply costs $250 per MWh and the ex ante price is therefore the efficient price.

The significance of this issue depends on the frequency of such instances. Figure 24 summarizes differences in constraint shadow prices between ex post and ex ante calculations in 2012. A positive value indicates a higher shadow cost in the ex post calculation. For example, the value "2" on the x-axis means the ex post shadow cost is $1-$2 per MWh higher than the ex ante cost.

Figure 24: Difference in Constraint Shadow Costs Between Ex Post and Ex Ante
All Binding Constraints, 2012



The average difference was not significant in 2012. About 96 percent of all differences were within $10 per MWh. However, there were a small number of intervals with substantial differences in congestion costs between the ex ante and ex post calculations. There were 34 intervals during which ex post shadow prices were at least $100 per MWh higher than ex ante prices, and 143 intervals during which ex post shadow prices were at least $100 per MWh lower than ex ante prices. These results can be attributed partly to the very low levels of congestion that currently prevail in the ISO-NE markets. However, as load grows and transmission congestion increases, we expect that these instances will also increase.

4. Conclusions regarding Ex-Post Pricing

Our evaluation of the ex post pricing results indicates that the real-time ex post prices:

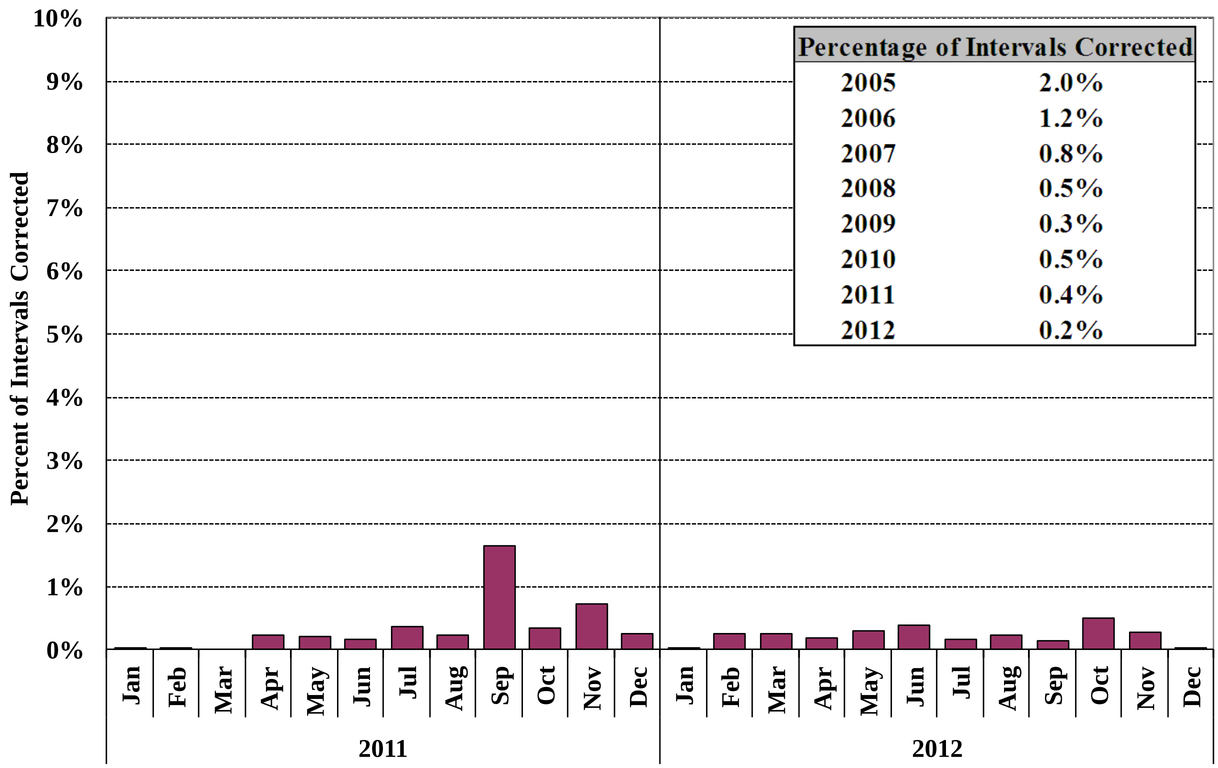
- Are slightly biased in the upward direction in uncongested areas;
- Introduce small potential inefficiencies when they are not consistent with dispatch instructions; and
- Sometimes distort the value of congestion into constrained areas.

The primary benefit of ex post pricing is that it allows the ISO to correct the real-time prices when the ex ante prices are affected by corrupt data or communication failures. Given that ex post prices are sometimes set at inefficient levels, we recommend that the ISO consider modifying the inputs from UDS to the ex post pricing model to improve the consistency of the ex post and ex ante prices.

E. Real-Time Price Corrections

This subsection evaluates the rate of real-time price corrections during 2012. Price corrections are necessary to address a variety of issues, including software flaws, operations or data entry errors, system failures, and communications interruptions. Although they cannot be completely eliminated, a market operator should aim to minimize price corrections. Substantial and frequent corrections raise ISO and market participant costs and can harm the integrity of the market. Figure 25 shows the rate of real-time price corrections in New England in each month of 2011 and 2012. The inset table shows the annual rate of price corrections in the past eight years.

Figure 25: Rate of Real-Time Price Corrections
2011 – 2012



The figure shows that real-time price corrections were infrequent in both 2011 and 2012. The rate was less than one-half percent in all but three months during 2011 and 2012. October exhibited the highest rate of price correction of any month in 2012 at 0.52 percent. This was primarily caused by price corrections for 12 hours due to software errors on one day and for 12 hours due to a planned software outage on another day. The annual rate of price corrections has declined since 2004 and has been at or below 0.5 percent in recent years. It is also notable that about 65 percent of the intervals that experienced price corrections in 2011 and 2012 were due to issues with the real-time software's Dead Bus Logic, which affects the LMPs at very few pricing nodes.[88] Hence, during many of the real-time intervals with price corrections, the effect of the price correction on the market was very limited.

88 Due to equipment outages, the main transmission system may consist of several islands, of which only one is a viable sub-system and the others are considered dead. The market clearing problem is solved only for the viable island and the LMPs are determined in the LMP Calculator. LMPs at dead buses are not directly available from the LMP Calculator. However, there is need for market settlement purposes to determine

Overall, the frequency of price corrections has been very low over the past four years, supporting the conclusion that the real-time market software for the ISO-NE wholesale market has functioned well.

F. Conclusions and Recommendations

Efficient price formation is an important function of real-time market operations. Efficient real-time price signals provide incentives for generators to be available, for demand response to participate in the wholesale market, and for investors to build capacity in areas where it is most valuable. Hence, efficient prices provide market participants with incentives that are compatible with the ISO's mandate to maintain the reliability of the system.

This section evaluates several aspects of real-time pricing in the ISO-NE market during 2012. Our evaluation leads to the following conclusions and recommendations:

- Fast-start generators are routinely deployed economically, but the resulting costs are often not fully reflected in real-time prices. This leads to inefficiently low real-time prices, particularly under tight operating conditions when thermal peaking generators are needed to satisfy real-time demand. During such conditions, efficient price signals will provide incentives for suppliers to make sufficient capacity available to meet the needs of the system.
    - ✓ We recommend that the ISO evaluate potential changes in the pricing methodology that would allow the deployment costs of fast-start generators to be more fully reflected in the real-time market prices.
- The marginal cost of meeting system-level 30-minute reserve requirements can exceed the $100 per MWh RCPF. Before the RCPF was increased to $500 per MWh, the ISO often had to curtail exports and take other manual actions outside the market in order to maintain adequate reserves. This led to inefficiently low real-time prices that did not properly reflect the cost of maintaining reliability. The new RCPF level provides market participants better incentives to schedule in the day-ahead market and schedule net imports from external areas that will lower the costs of maintain reliability.
- Demand response programs help reduce the cost of operating the system reliably, particularly during peak periods. However, the inflexibility of demand response resources creates challenges for setting efficient prices that reflect scarcity during periods when emergency demand response resources are activated.

the LMPs at dead buses. The algorithm, referred to as LMPc Dead Bus Logic, has been used to facilitate this need.

- ✓ Hence, we recommend that the ISO allow the costs of non-dispatchable demand response resources to be reflected in clearing prices when there is a capacity deficiency or when a deficiency is avoided by the activation of the demand response resources.

- Given that ex post prices are sometimes set at inefficient levels, we recommend that the ISO consider modifying the inputs from UDS to the ex post pricing model to improve the consistency of the ex post and ex ante prices.

- Price corrections were very infrequent in 2012, which reduces uncertainty for market participants in the ISO-NE wholesale market. Further, a large share of the price corrections that did occur affected a very small number of pricing nodes.

VI. System Operations

To maintain the reliability of the system, sufficient resources must be available in the operating day to satisfy forecasted load and reserve requirements without exceeding the capability of the transmission system. The wholesale market is designed to satisfy these requirements at the lowest cost. In particular, the day-ahead market and the forward reserve market are intended to provide incentives for market participants to make resources available to meet these requirements. The day-ahead market clears physical and virtual load bids and supply offers, and produces a coordinated commitment of resources. The forward reserve market provides suppliers with incentives to make reserve capacity available, particularly from offline fast-start resources.

When the wholesale market does not satisfy all forecasted reliability requirements for the operating day, the ISO performs the Reserve Adequacy Assessment (RAA) to ensure sufficient resources will be available. The primary way in which the ISO makes sufficient resources available is by committing additional generation. Such commitments generate expenses that are uplifted to the market and increase the amount of supply available in real time, which depresses real-time market prices and leads to additional uplift. Hence, out-of-market commitment tends to undermine market incentives for meeting reliability requirements. Out-of-market commitments can also indicate that there are important reliability requirements that are not fully reflected in the wholesale market requirements, so the cost of satisfying these requirements is not fully reflected in market clearing prices. Therefore, we evaluate supplemental commitments and other operating actions in this section of the report.

The rising demand for natural gas in recent years has reduced the availability of gas to electricity generators during severe winter weather conditions, creating new challenges for the design of wholesale electric markets. The primary challenge is for the market to coordinate the scheduling of electric resources in a manner that satisfies the system’s reliability needs and leads to efficient procurement and scheduling of natural gas and other fuels, both for electric generation and other uses. During severe winter weather, the amount of installed capacity is more than adequate, but the limited supply of natural gas reduces the availability of installed capacity. Moreover, the ISO has limited information about the fuel supplies of individual generators, increasing uncertainty about whether the available capacity will be adequate under tight conditions. When the market does not have mechanisms to fully reflect the limited availability of generating capacity, market prices will be depressed and, therefore, will not provide efficient incentives for generation to be available. This leads the ISO to make additional resources available through supplemental commitment and other out-of-market actions, which further depress prices further and undermine the market incentives for reliable generator performance.

In this section, we evaluate several aspects of market operations that are related to the ISO's process to ensure that sufficient resources are available to meet the forecasted reliability requirements. In particular, we evaluate the following:

- Accuracy of Load Forecasting – The ISO's load forecasts are used by market participants to inform scheduling in the day-ahead market and by the ISO to determine the forecasted reliability requirements;
- Reliability Commitment and Out-of-Merit Generation – Reliability commitments make additional resources available to operate in real time, and they increase the amount of generation that runs out-of-merit in real time;
- Surplus Generation – The amount of capacity from online or available offline fast-start resources in excess of the system's energy and operating reserve requirements; and
- Uplift Expenses – This examines the financial charges that result from out-of-market commitment and reliability agreements.

A. Accuracy of ISO Load Forecasts

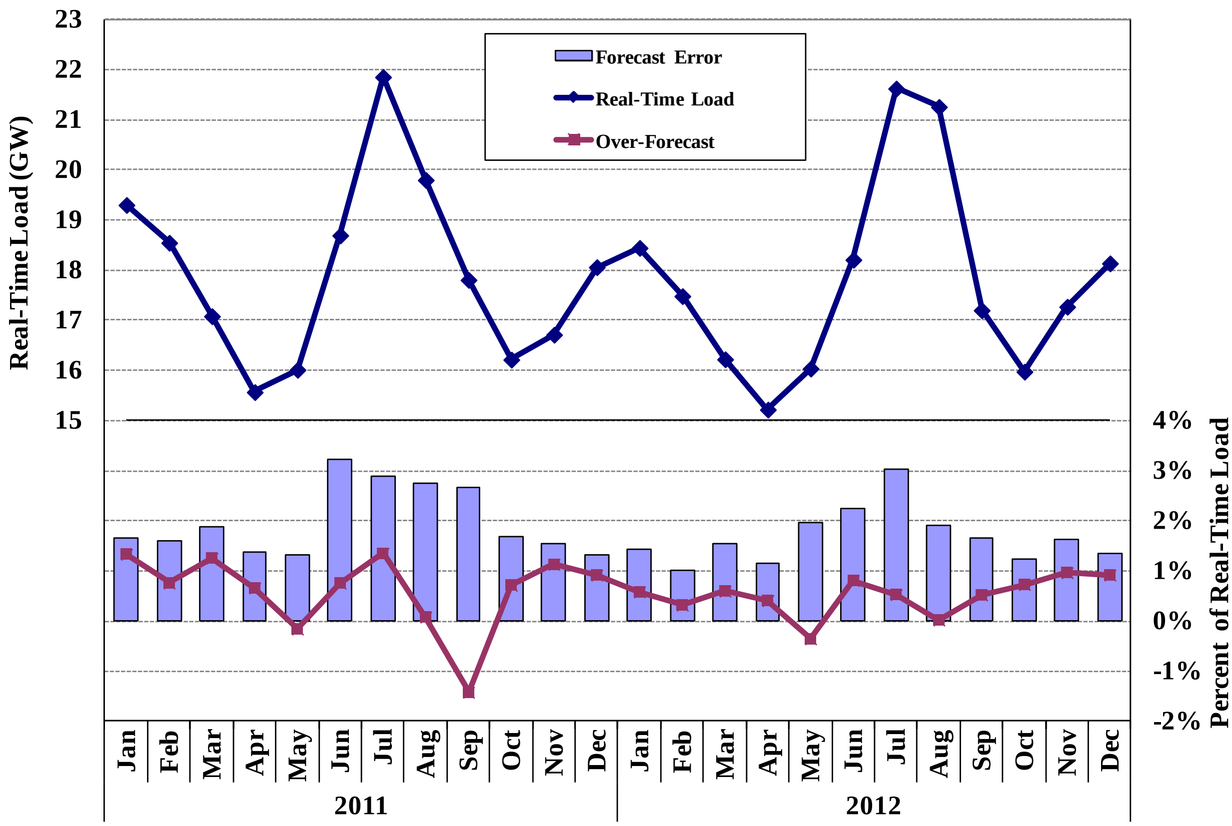
The ISO produces a load forecast seven days into the future and publishes the forecast on its website. This forecast is significant because market participants may use it and other available information to inform their decisions regarding fuel procurement, management of energy limitations, formulation of day-ahead bids and offers, or short-term outage scheduling.

In addition, the ISO uses the forecast to estimate the amount of resources that will be needed to satisfy load and reserve requirements without exceeding the capability of the transmission system. The day-ahead forecast is most important because most scheduling and unit commitment takes place on the day prior to the operating day (either in the day-ahead market or in the RAA).

Accurate load forecasts promote efficient scheduling and unit commitment. Inaccurate load forecasts can cause the day-ahead market and/or the ISO to commit too much or too little capacity, which can affect prices and uplift. Therefore, it is desirable for the day-ahead forecast to accurately predict actual load.

Figure 26 summarizes daily peak loads and two measures of forecast error on a monthly basis during 2011 and 2012. The *Over-Forecast* is the percentage by which the average day-ahead forecasted daily peak load exceeded the average real-time daily peak load in each month. [89] Positive values indicate over-forecasting on average and negative values indicate under-forecasting on average. The *Forecast Error* is the average of the absolute difference between the day-ahead forecasted daily peak load and the actual daily peak load, expressed as a percentage of the average actual daily peak load.

Figure 26: Average Daily Peak Forecast Load and Actual Load
Weekdays, 2011 – 2012



89 The real-time daily peak load is based on the average load in the peak load hour of each day. Thus, the instantaneous peak load of each day is slightly higher than the values used in Figure 25.

The figure shows a seasonal pattern of high loads during the winter and summer and mild loads during the spring and fall. Overall, load decreased modestly from 2011 to 2012. The annual peak load of 25.9 GW occurred on July 17, 2012, down approximately 7 percent from the peak load of 27.7 GW in 2011, which was the second-highest all-time peak load level.[90] The average load declined nearly 1 percent, from 14.9 GW in 2011 to 14.8 GW in 2012. However, the frequency of actual load conditions exceeding 20 GW increased from 312 hours in 2011 to 449 hours in 2012. The decline in load levels was particularly notable in the first quarter when average load fell 5 percent from the previous year, which was primarily due to milder winter weather in early 2012.

The ISO's day-ahead load forecasts are very consistent with actual load, although the ISO tends to slightly over-forecast load on average. The average over-forecast was comparable in the two years: 0.6 percent in 2011 and 0.5 percent in 2012. The ISO regularly evaluates the performance of its load forecasting models to ensure there are no factors that bias the forecast unjustifiably.[91]

The figure also shows the average forecast error, which is the average of the absolute value of the difference between the daily forecasted peak demand and the daily actual peak demand. For example, a one percent over-forecast on one day and a one percent under-forecast on the next day would result in an average forecast error of one percent, even though the average forecast load would be the same as the average actual load. The average forecast error was roughly 1.7 percent in 2012, consistent with 2011. The forecast error tends to increase during the summer months. In 2012, the forecast error averaged 2.4 percent in the summer months (June to August) and just 1.4 percent in other months. Nonetheless, these levels of forecast error are still relatively small, and the load forecasting performance of the ISO remains good overall.

90 New England's all-time peak is 28,130 MW, recorded on August 2, 2006.

91 A small bias toward over-forecasting may be justifiable because the costs of under-forecasting (i.e., under-commitment and potential for shortages) are likely larger than the costs of over-forecasting. Furthermore, it may be appropriate when the instantaneous peak load is expected to be substantially higher than the hourly average peak load.

B. Commitment for Local and System Reliability

In ISO-NE, sufficient resources must be available to satisfy local and system reliability requirements. To ensure reliability at the system level, sufficient online and offline quick-start resources are needed to satisfy forecasted load, to recover from the largest single contingency, and to recover from 50 percent of the second-largest single contingency. To ensure that local areas can be served reliably, a minimum amount of capacity must be committed in each load pocket (i.e., import-constrained area). Specifically, sufficient online capacity is required to: (i) meet forecasted load in the load pockets without violating any first contingency transmission limits (i.e., ensure the ISO can manage congestion on all of its transmission interfaces); (ii) ensure that reserves are sufficient in local constrained areas to respond to the two largest contingencies; (iii) support voltage in specific locations of the transmission system; and (iv) manage constraints on the distribution system that are not modeled in the market software (known as Special Constraint Resources (SCRs)).

In the day-ahead market, generators are scheduled based on the bids and offers submitted by buyers and sellers. A generator is committed when demand bids from load serving entities and virtual traders are high enough for the unit to be economic given its start-up, no-load, and incremental offer components. The willingness of load serving entities and virtual traders to buy or sell power in the day-ahead market is partly based on their expectations of LMPs in the real-time market on the following day. Thus, the resulting day-ahead market commitment is strongly affected by expectations of real-time prices.

After the day-ahead market, the ISO may need to commit additional generators with high commitment costs to meet local and system-level reliability requirements. Once the commitment costs have been incurred, these generators may be inexpensive providers of energy and reserves. Because these commitment costs are not reflected in the market prices, the real-time LMPs frequently do not reflect the full value of online and fast-start capacity when generators are committed for reliability. Like any other forward financial market, the day-ahead market LMPs tend to converge with the real-time LMPs. Hence, day-ahead LMPs also do not reflect the full value of online and fast-start capacity, which reinforces the tendency of the day-ahead market-based commitment to not satisfy reliability requirements.

Given the effects of supplemental commitment on market signals, it is important to minimize these commitments while still maintaining reliability. Periodically, the ISO evaluates refinements to the procedures and tools used in the RAA to make the process more efficient. The ISO has also made market enhancements that better reflect reliability requirements in the real-time market, reducing the need for supplemental commitment. Nonetheless, supplemental commitments are still needed to meet reliability requirements, so it is important to continue evaluating potential market improvements. This section summarizes the pattern of supplemental commitment for reliability in the past two years.

Figure 27 shows the average amount of capacity committed to satisfy local and system-level requirements in the daily peak load hour in each region in 2011 and 2012.[92,93] The category "*RAA/RT – First Contingency & System Reserves*" shows capacity committed for local first contingency protection and for system-level reserve requirements together since the ISO does not maintain data that distinguishes between these two reasons for commitment. The figure shows the entire capacity of these units, although their impact on prices depends on the amounts of energy and reserves they provide to the real-time market.

Supplemental commitment increased 40 percent from an average of 350 MW in 2011 to 490 MW in 2012. In particular, commitment for local reliability rose from an average of 90 MW in 2011 to 155 MW in 2012. The increase occurred primarily in Maine and Western Central Massachusetts where planned transmission outages led to commitments for second contingency protection (Maine) and local voltage support (Western Central Massachusetts).

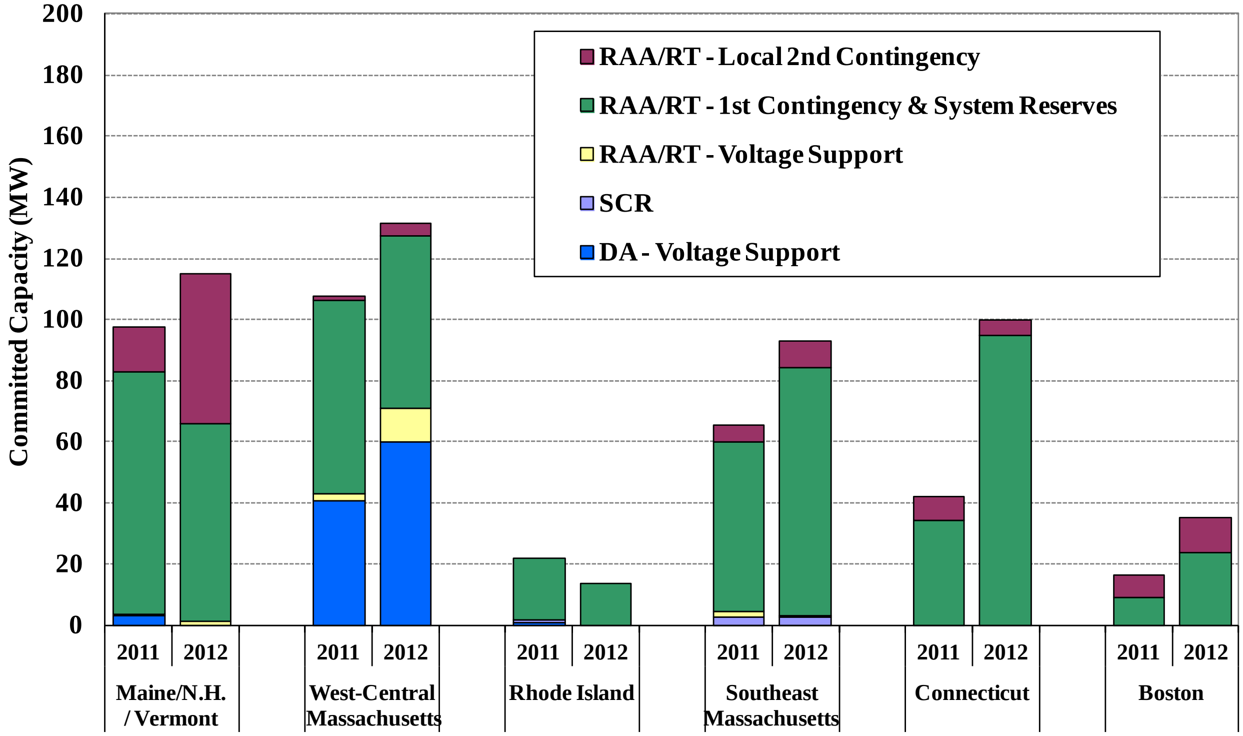
Additionally, supplemental commitment for system-wide reserves (and local first contingencies) increased 29 percent, from an average of 260 MW in 2011 to 335 MW in 2012. The ISO does not systematically distinguish supplemental commitments for local first contingencies from those for system-wide reserves, but the vast majority of commitments in this category were for system-

92 In accordance with its Tariff, ISO-NE classifies certain day-ahead commitments as Local Second Contingency commitments even though they occur as the result of market-based scheduling activity. Since these are not out-of-market commitments, we exclude them from our analyses of supplemental commitment in this section.

93 Capacity committed day-ahead for voltage support that would have been economically committed in the day-ahead market is excluded from the figure.

wide reserves. This category accounted for more than two-thirds of total reliability commitments in 2012. Later in this section, our analysis shows that the sharp rise in supplemental commitment in the weeks following the arrival of Superstorm Sandy on October 30, 2012 accounts for a substantial share of the increase in this category of supplemental commitment.

Figure 27: Commitment for Reliability by Zone
Daily Peak Hour, 2011 – 2012



Despite the increase in supplemental commitments, the amount of supplemental commitment was still significantly lower than in the years prior to 2010. [94] We evaluate the need for these commitments and their effects on real-time energy prices later in this section. Variations in the pattern of supplemental commitments have substantially affected operations in several ways that are discussed later in this section. Subsection C illustrates how the quantities of out-of-merit dispatch (i.e., capacity producing output at a cost greater than the LMP) have changed.

94 The average amount of supplemental commitment was well above 1,000 MW in the years prior to 2010 (e.g., 1,670 MW in 2007). The reduction in recent years resulted primarily from transmission upgrades in Boston, Connecticut, and Southeast Massachusetts between 2007 and 2009 that reduced the capacity needed to satisfy local first and second contingency requirements. See our 2007 and 2008 Assessments for details.

Subsections D and E show that the amount of surplus online capacity has decreased, and they analyze the effect on real-time prices. Subsection F reports the uplift charges resulting from reliability-committed units.

C. Out-of-Merit Generation

Out-of-merit generation occurs in real time when energy is produced from an output range on a unit whose energy offer is greater than the LMP at its location. Out-of-merit generation tends to reduce energy prices by causing lower-cost resources to set the energy price. In a very simple example, assume the two resources closest to the margin are a $60 per MWh resource and a $65 per MWh resource, with the market clearing price set at $65 per MWh in the absence of congestion and losses. When a $100 per MWh resource is dispatched out-of-merit, it will be treated by the software as a must-run resource with a $0 per MWh offer. Assuming the energy produced by the $100 per MWh resource displaces all of the energy from the $65 per MWh resource, the energy price will decrease to $60 per MWh.

Out-of-merit generation occurs for several reasons. First, a unit may run at its EcoMin to satisfy its minimum run time after having run in-merit in previous hours or in anticipation of running in an upcoming hour. This is efficient because the software is minimizing cost over the total run-time of the unit. Second, a unit committed for reliability reasons during or after the day-ahead market may be out-of-merit at its EcoMin. Units are committed for reliability when they are not economic in the day-ahead market, so their incremental energy offer tends to be higher than the LMP. Third, a unit may be dispatched out-of-merit in real time to satisfy reliability requirements, although this accounts for a very small share of the total out-of-merit generation.[95]

Figure 28 summarizes the average out-of-merit generation by location during peak hours (i.e., weekdays 6 AM to 10 PM, excluding holidays) in 2011 and 2012. The figure shows five categories of out-of-merit generation on units that are committed (and occasionally dispatched)

95 Similar to the supplemental commitments, operators may request certain units to run at higher levels than would result from their energy offers. This can be necessary for a number of reasons, including: (a) providing voltage support on transmission or distribution facilities; (b) managing congestion on local facilities that are not represented in the dispatch model; or (c) providing local reserves to protect against second contingencies.

for reliability reasons.[96] The figure also shows an "other dispatch" category that includes generation from units that were economically committed but are running at their EcoMin.

Figure 28: Average Hourly Out-of-Merit Generation
Weekdays 6 AM to 10 PM, 2011 – 2012

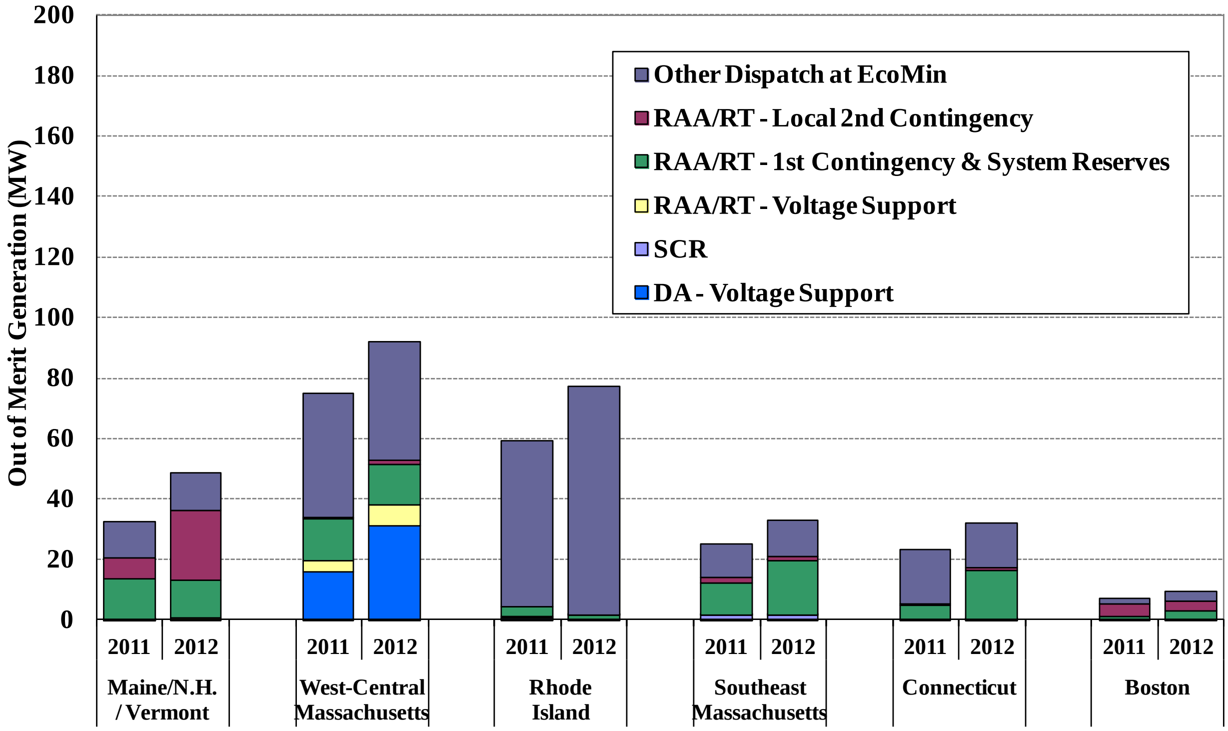


Figure 28 shows that in most regions, most of the out-of-merit dispatch was from units committed through the RAA process for reliability in 2012. However, this was not the case in Rhode Island where the majority of the out-of-merit generation was attributable to non-reliability units being dispatched at EcoMin.

The average quantity of out-of-merit generation from units committed for reliability rose from 83 MW in 2011 to 135 MW in 2012. The increase in out-of-market generation from units committed for reliability tracked the rise in supplemental commitments and was caused by the

96 Day-ahead commitments that are flagged for Local Second Contingency are excluded from this category if they occur as the result of market-based scheduling activity. Likewise, day-ahead commitments that are flagged for Voltage Support are excluded from this category if they would have been economically committed.

same underlying factors. The increased commitment for reliability in most areas led to proportionate increases in out-of-merit energy in those zones.

The amount of out-of-merit energy from units that were committed economically (i.e., Other Dispatch at EcoMin) has also increased modestly from an average of 139 MW in 2011 to 157 MW in 2012, which was due primarily to the increase in Rhode Island.

D. Surplus Capacity and Real-Time Prices

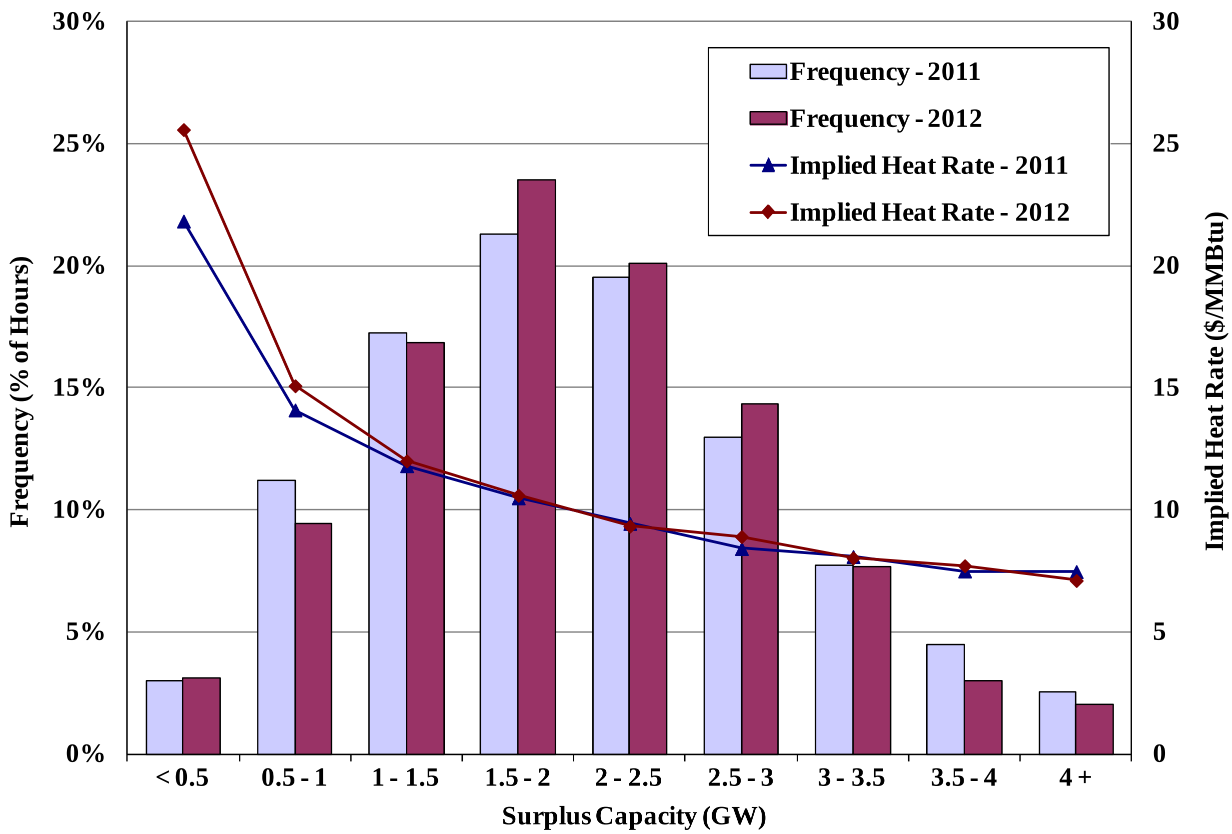
Under normal operating conditions, the available online and fast-start capacity is more than sufficient to satisfy load and reserve requirements, which suggests that some surplus capacity will exist in almost every hour. This is a normal outcome in a properly functioning market. Surplus capacity does not raise concerns unless inflated by inefficient commitments by the ISO or market participants.

Surplus capacity is also important because it constitutes the resources that are available to respond to unexpected changes in real-time operating conditions. Accordingly, the quantity of surplus capacity exhibits a strong negative correlation with real-time energy prices. This section evaluates the pattern of surplus capacity and real-time energy prices. In this report, we define "Surplus Capacity" as the amount of capacity that is online or capable of starting within 30 minutes in excess the amount required to meet load and reserve requirements. Hence, surplus capacity is equal to:

Online Reserves + Offline Reserves Deployable in 30 minutes – TMOR Requirement

Figure 29 summarizes the relationship of surplus capacity to real-time energy prices at ISO-NE Hub in each peak hour of 2011 and 2012. Each bar shows the frequency of peak hours when Surplus Capacity was in the range of values shown on the horizontal axis. For example, there was 0.5 GW to 1.0 GW of surplus capacity in approximately 11 percent of the peak hours in 2011 and 9 percent in 2012. The lines show the average real-time implied marginal heat rate at New England Hub in the hours that correspond to each range of surplus capacity. For example, in hours when there was 0.5 GW to 1.0 GW of surplus capacity, the average real-time implied marginal heat rate was 14.1 MMbtu per MWh in 2011 and 15.1 MMbtu per MWh in 2012. The implied marginal heat rate is shown in order to normalize real-time energy prices for changes in natural gas prices during 2011 and 2012.[97]

Figure 29: Surplus Capacity and Implied Marginal Heat Rates
Based on Real-Time LMPs at the Hub in Peak Hours, 2011-2012 [98]



The figure shows a strong correlation between the quantity of surplus capacity and the implied marginal heat rate in real time. In 2012, the average implied marginal heat rate ranged from approximately 25.6 MMbtu per MWh in hours with less than 0.5 GW of surplus capacity to 7.1 MMbtu per MWh in hours with more than 4 GW of surplus capacity.

Overall, the average implied heat rate during peak hours rose modestly from 10.6 MMbtu per MWh in 2011 to 10.9 MMbtu per MWh in 2012, This was mostly attributable to increased implied heat rates during hours with less than 0.5 GW of surplus capacity. Although the

---

97 In this section, the implied marginal heat rate in a particular hour is equal to the real-time LMP divided by the natural gas index price.

98 In this figure, "peak hours" includes hours-ending 7 through 22 on weekdays.

frequency of such hours was consistent from 2011 to 2012 (roughly 3 percent in both years), the average implied heat rate during such hours rose from 21.8 MMbtu per MWh in 2011 to 25.6 MMbtu per MWh in 2012. The increase was primarily due to higher LMPs in hours with tight operating conditions following the changes in the RCPF for 30-minute reserves and in the required amount of 30-minute reserves.[99] Furthermore, the reduction in natural gas prices from 2011 to 2012 has increased the size of non-gas related production costs when measured in proportion to the price of natural gas (as is the case for the implied heat rate).[100]

The average implied heat rates were more consistent in hours with more than 1 GW of surplus capacity from 2011 to 2012. This was because the effects of lower natural gas prices were offset by the combined effects from increased imports, increased nuclear generation, and overall lower load levels.

The figure also shows that although distribution of surplus capacity changed from 2011 to 2012, the average amount of surplus capacity did not change significantly. In particular, the share of hours with less than 1.0 GW or more than 3.5 GW of surplus capacity decreased from 21 percent in 2011 to 17 percent in 2012.

Two important conclusions can be drawn from the analysis shown above. First, it reinforces the importance of minimizing out-of-market commitment for reliability, since the commitment of even one additional generator has substantial effects on the amount of surplus capacity, market clearing prices, and the resulting NCPC charges. This is evaluated further in the following subsection.

Second, to the extent that reliability criteria require additional resources to be online beyond the quantities reflected in the market rules, it leads to an upward bias in the amount of surplus capacity in the real-time market. This increased surplus depresses real-time prices below

---

99 These two rule changes are discussed further in Section III.A.

100 For example, if the natural gas price is $5 per MMbtu and $10 per MWh of the LMP is related to the non-gas related production costs of the marginal generator, the non-gas related costs will contribute 2 MMbtu per MWh of the implied heat rate (i.e., $10 per MWh divided by $5 per MMbtu). If the natural gas price falls to $2.50 per MMbtu, the non-gas related costs will contribute 4 MMbtu per MWh of the implied heat rate (i.e., $10 per MWh divided by $2.50 per MMbtu).

efficient levels. Hence, we are supportive of several ISO initiatives that will improve the recognition by the market of the amount of capacity that is required to maintain reliability:

- Replacement Reserves Procurement – Due to concerns regarding certain large contingencies and the reliability of individual generators, the ISO sometimes requires higher quantities of operating reserves than are explicitly reflected in the 30-minute minute reserve requirement. The ISO proposes to reflect these needs in an additional Replacement Reserve requirement before the 2013/14 Winter period, which will result in the procurement of additional 30-minute reserves using an RCPF that reflects the relative importance of the additional reserve needs (e.g., $250 per MWh).[101]

- Off-line Reserve Auditing – The ISO is improving its methods for determining the off-line reserve capability of fast-start resources to help ensure consistentency between the offered performance and the actual performance. Fast-start resources provide considerable benefits by allowing the system to respond quickly to unexpected system conditions,[102]

- On-line Reserve Auditing – The ISO is also improving its methods for determining the capability of on-line resources to provide reserves. This project will also result in more accurate estimates of the amount of reserves and reduce the tendency for the market to over-estimate the available reserves. This will lead to more efficient real-time pricing of reserves and energy, and it is scheduled for implementation June 1, 2013.[103]

These initiatives are expected to lead to more accurate calculation of the amount of available resources relative to the amount of resource required for reliability. This should lead to higher real-time clearing prices for energy and reserves during tight operating conditions when reliable generator performance is most important for system reliability. This should, in turn, reduce NCPC uplift charges, improve the incentives for generator commitment in the day-ahead market, and provide signals for investments to improve performance by both new and existing resources.

E. Supplemental Commitments and Surplus Capacity

Given the effect of surplus capacity on prices, it is important to evaluate the supplemental commitments made by the ISO and self-commitments made by market participants after the day-

---

101 The proposal is described in: "http://www.iso-ne.com/committees/comm_wkgrps/mrkts_comm/mrkts/2013/apr9102013/a08_iso_presentation_04_09_13.ppt"

102 See ISO New England Inc. and New England Power Pool, *Market Rule 1 Revisions Relating to Auditing of Generation Resources*, Docket No. ER13-323-000 (November 6, 2012), which may be found at: "http://www.iso-ne.com/regulatory/ferc/filings/2012/nov/er13-323-000_11-6-2012_audit_claim.pdf"

103 *Id.*

ahead market. Both types of commitments can depress real-time prices inefficiently, while supplemental commitments by the ISO also lead to increased uplift costs.

As discussed earlier, transmission upgrades have substantially reduced the need for the ISO to commit generation to satisfy local reliability requirements. Since July 2009, the ISO's need to make supplemental commitments for local reliability has largely been eliminated. However, the ISO must still periodically make commitments to satisfy ISO-NE's system-wide reliability requirements. To evaluate the effectiveness of this process, the following two figures show the supplemental commitments and self-scheduled commitments by day in the bottom panel, and the surplus capacity in the peak load hour and the minimum surplus capacity in any hour of each day in the upper panel. Figure 30 shows the first six months of 2012, and Figure 31 shows the last six months of 2012.

Figure 30: Daily Supplemental Commitments and Surplus Capacity
January to June, 2012

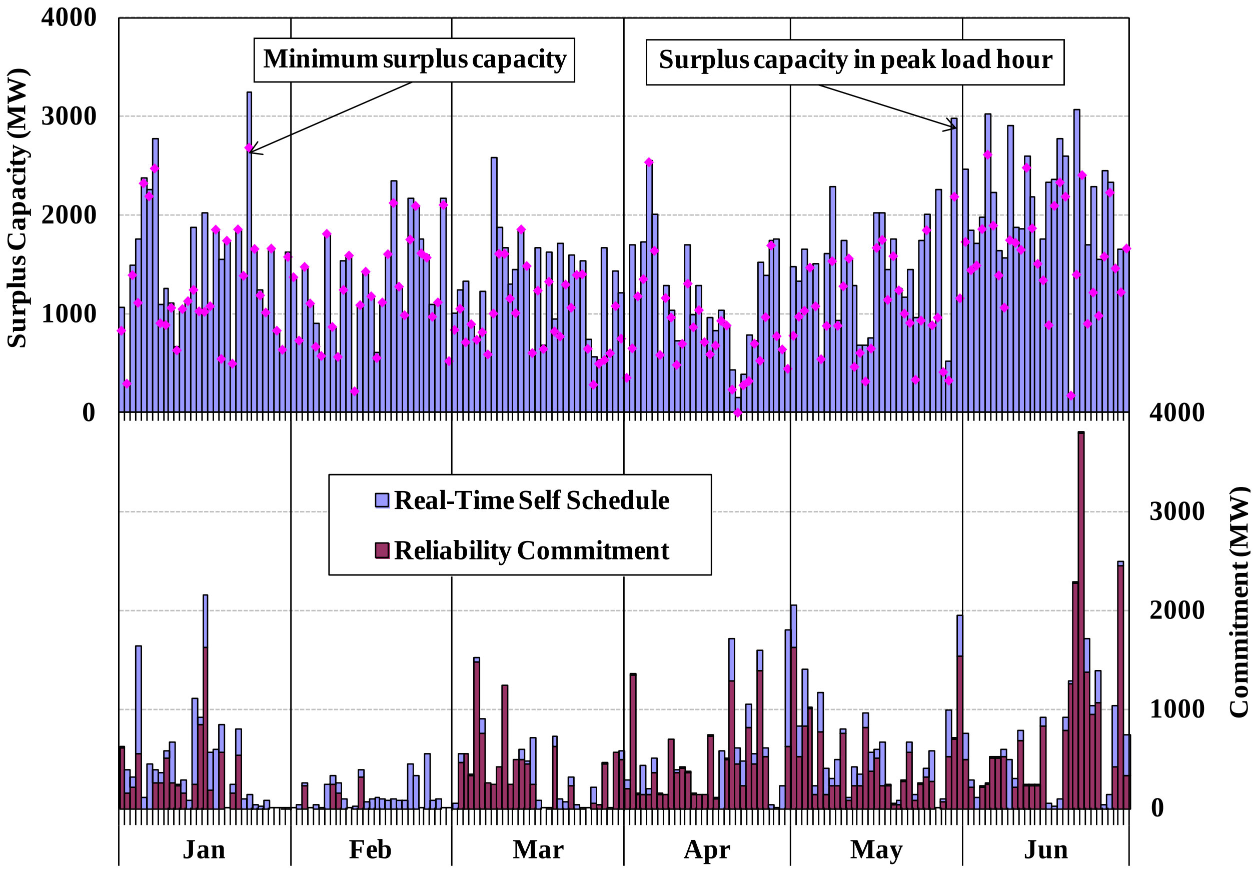
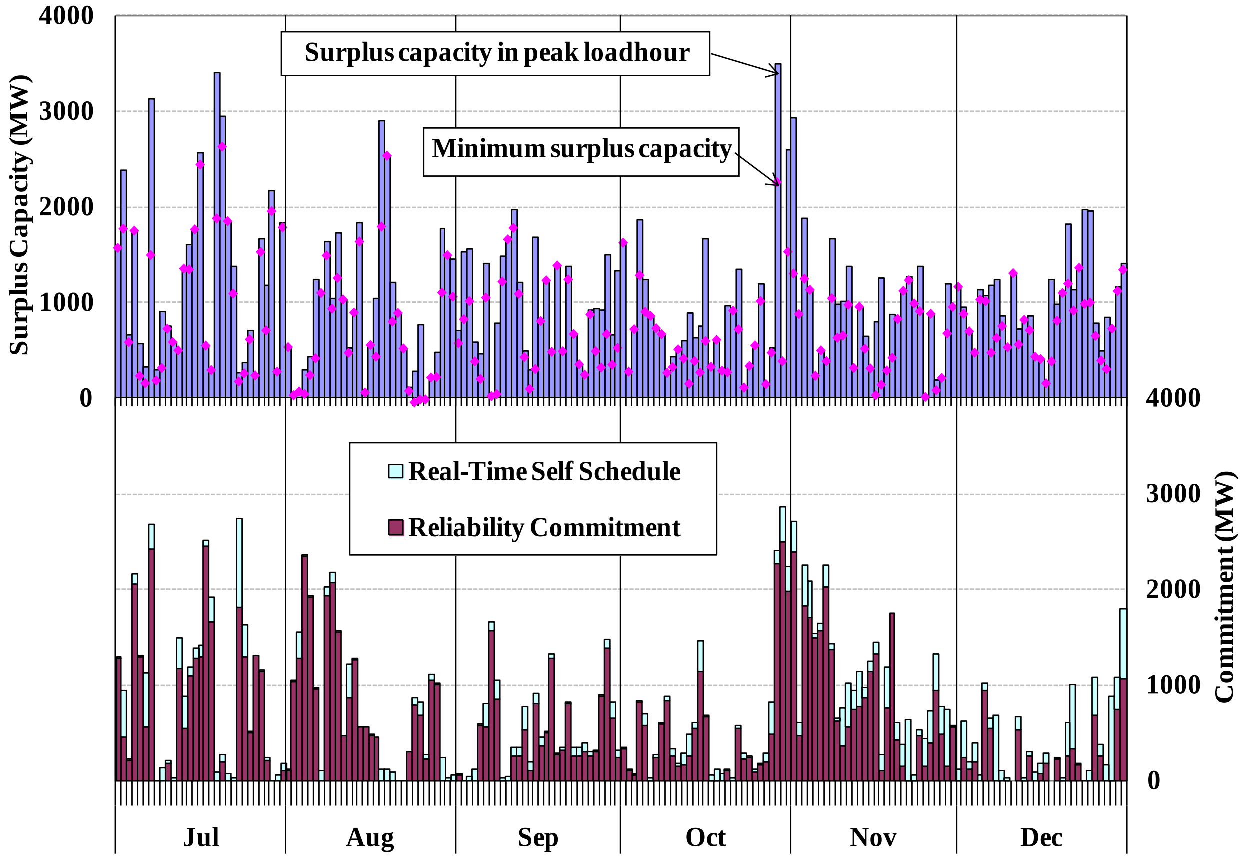


Figure 31: Daily Supplemental Commitments and Surplus Capacity
July to December, 2012



We evaluate the need for supplemental commitments because unnecessary commitments for system-wide reliability requirements that lead to large surplus capacity levels generally raise costs to ISO-NE's customers and distort real-time prices. Although the figures show that the minimum surplus capacity levels were relatively low on most days when the ISO made supplemental commitments for system-wide capacity needs, there were some days when large quantities of supplemental commitments resulted in large quantities of surplus capacity. After reviewing the supplemental commitments and the surplus capacity levels that resulted from real-time operating conditions, we found that roughly 44 percent of the supplemental resource commitments in 2012 were needed to maintain system level reserves in retrospect.[104] The fact

104 This is a simple evaluation that treats any surplus capacity (online and available offline capacity less the need to meet system load and reserve requirements) as "not needed" for the system. This simple evaluation tends to understate the necessity of supplemental commitments because: 1) the evaluation is based on hourly integrated peak rather than the higher instantaneous peak, and 2) the ISO cannot commit just a

that some of the reliability-committed capacity was not needed in retrospect is typically due to the following factors.

First, ISO-NE has a limited quantity of fast-start generating resources, which help ensure that sufficient capacity will be available if unexpected conditions arise. This leads the ISO in some cases to rely on slower-starting units that must be notified well in advance of the operating hour when uncertainty regarding load, imports, and generator availability is high. Most of the commitments of slow-starting units are made overnight, more than 12 hours before the forecasted peak.

Second, ISO-NE is heavily reliant upon gas-fired generating capacity, which can become unavailable due to the limitations of the natural gas system. Consequently, the ISO may commit oil-fired and/or dual-fueled capacity in order to protect the system in the event that the supply of natural gas is interrupted to some units.

Third, there are two assumptions in the reliability commitment process that can make large contributions to the over-commitment on some days:

- The "desired capacity surplus" that operators have the discretion to determine to account for concerns regarding generator availability, load forecast errors, or other factors; [105] and
- The assumed level of imports and exports. When evaluating the need for commitments in advance, the ISO generally assumes day-ahead scheduled transactions will flow.

In general, the desired capacity surplus should be minimized since the operating reserve requirements are set at levels that should ensure reliability. Adding a non-zero desired capacity surplus introduces an inconsistency between the market requirements and the operating requirements, although we recognize that conditions can sometimes arise that would justify an increase in the desired capacity surplus. The ISO's proposal to procure Replacement Reserves

portion of a unit. For example, if the ISO needs an additional 200 MW of capacity to satisfy system reliability needs and commits the most economic unit with a capacity of 300 MW. In this evaluation, 100 MW of capacity would be deemed as "not needed".

105 The operators have the discretion to commit surplus generation when they believe it is necessary to deal with uncertainty as stated in the System Operating Procedure, Perform Reserve Adequacy Assessment, Section 5.3.2.3, "The Forecaster may commit additional Generators as needed for reliability (anticipated storms, hurricanes or other conditions that affect Bulk Power System reliability)."

will allow it to reduce the desired capacity surplus by a comparable amount, which should improve real-time pricing by reducing the size of inconsistencies between the market requirements and the operating requirements.[106] We would go further in this regard and allow operators to vary the quantity of the replacement reserve requirement based on their concerns regarding load and fuel supply uncertainty.

With regard to the import and export assumptions, we believe that substantial improvements are possible. In general, the assumptions regarding imports and exports are that the day-ahead scheduled transactions will flow in real time. By committing generation to support day-ahead exports, they are treated as firm and we understand from the ISO that the operators generally do not curtail day-ahead exports. This treatment of the day-ahead exports in the capacity evaluation process raises potential efficiency concerns because:

- The participants are not obligated to schedule the exports in real time, which could render the units committed to support them unnecessary;
- The value of the day-ahead exports may not justify the costs of the supplemental commitments made to support them; and
- Assuming a fixed schedule substantially understates the ability of adjustments in interchange to help maintain reliability.

This is particularly true when exports are scheduled to New York when the difference in price on the New York side of the border is not significantly higher than on the New England side (which represents the value of the export).

Hence, the ISO should consider whether its assumptions regarding imports and exports in its capacity evaluation process could be improved. The ISO-NE is moving forward with the NYISO in implementing Coordinated Transaction Scheduling (CTS), which should rationalize the physical flow between the two markets in real-time. This should, in turn, allow the ISO to rely more heavily on the markets to cause power to flow in the efficient direction, making it unnecessary to commit generation to support day-ahead exports.

---

106 For example, suppose that the desired capacity surplus would have been 300 MW on a particular day. If the ISO procures 200 MW of Replacement Reserves, it will be able to reduce the desired capacity surplus to 100 MW without reducing reliability.

F. Uplift Costs

To the extent that the wholesale market does not satisfy ISO-NE's reliability requirements, the ISO takes additional steps to ensure sufficient supplies are available. The ISO primarily makes supplemental commitments of resources that were not economic in the market ensure that it can satisfy its reliability needs. [107] Such generators receive NCPC payments, which make up the difference between their accepted offer costs and the market revenue. The costs associated with these payments are recovered from market participants through uplift charges. This section describes the main sources of uplift charges and how they are allocated among market participants.

The following table summarizes several categories of uplift in 2011 and 2012. The main categories of uplift are:

- FCM Reliability Credits – The uplift from these out-of-market capacity payments are allocated to Network Load in the zone where the generator is located.[108] Generators that are prevented from delisting for reliability reasons receive Reliability Credits under FCM, which are equal to the difference between their rejected delist bid and the FCA clearing price.

- Local Second Contingency Protection Resources – In 2012, 96 percent of the uplift from these units was allocated to Real-Time Load Obligations and Emergency Sales in the zone where the generator is located.[109] The remaining uplift associated with day-ahead rather than real-time commitments was allocated to day-ahead load schedules in the local zone.

- Special Constraint Resources – The uplift paid to these resources is allocated to the Transmission Owner that requests the commitment.

- Voltage Support Resources – The uplift paid to these resources is allocated to Network Load throughout New England, export transactions, and wheel-through transactions.

107 Historically, the ISO also used reliability agreements, which give the owners of uneconomic generating facilities supplemental payments to keep them in service, to ensure reliability, particularly in local import-constrained areas, All reliability agreements expired on June 1, 2010 when the first Forward Capacity Commitment Period began. However, since the first Forward Capacity Commitment Period began, several generators have still received out-of-market payments (which are known as "FCM Reliability Credits") to remain in service.

108 Network Load includes transmission customers that are served by the Transmission Owner.

109 Real-Time Load Obligations include load customers that are served by the Load Serving Entity.

- Economic and First Contingency Protection Resources – In 2012, 84 percent of this uplift was allocated to Real-Time Deviations throughout New England.[110] The remaining uplift associated with units committed in the day-ahead market is allocated to day-ahead scheduled load throughout New England. Non-fast-start units are typically started in the RAA process to maintain adequate reserves, while the fast-start units are typically started in economic merit order by the real-time dispatch model but do not recover the full as-offered cost (i.e., start-up, no-load, and incremental offer costs).[111]

When uplift charges are incurred to address local supply inadequacies, it is generally appropriate to allocate these charges to the local customers who benefit directly from the service. For this reason, the first three of these categories are allocated to local customers, while the uplift charges for Voltage Support Resources and other supplemental commitment are allocated to customers throughout New England.

The following table summarizes the total costs of uplift associated with NCPC charges and out-of-market capacity payments under FCM in 2011 and 2012.[112] The "Economic" category is broken into NCPC charges for quick-start resources versus non-quick-start resources (which are primarily committed for reliability). The year-over-year changes in uplift are shown as well.

110 Real-Time Deviations include Real-Time Load Obligation Deviations, which are positive or negative differences between day-ahead scheduled load and actual real-time load, uninstructed generation deviations from day-ahead schedules, virtual load schedules, and virtual supply schedules.

111 Section V.A discusses further the tendency for fast start units to be committed in economic merit order but not set the LMP during most of the period for which they are committed.

112 The numbers in the table are based on information available at time of reporting, which may be different from the numbers in final settlements.

Table 2: Allocation of Uplift for Out-of-Market Energy and Reserves Costs
2011 – 2012

	Millions of Dollars		% Change
Category of Uplift	**2011**	**2012**	**2011-2012**
FCM Reliability Credits	**$1.4**	**$11.4**	**714%**
Local Second Contingencies	**$6.0**	**$8.8**	**47%**
Special Case Resources	**$3.4**	**$3.7**	**8%**
Voltage Support	**$5.9**	**$14.9**	**152%**
Economic*			
Quick-Start Resources	**$11.9**	**$6.3**	**-47%**
Other Resources	**$47.1**	**$53.5**	**14%**
Total	**$75.7**	**$98.5**	**30%**

The table shows that total uplift charges increased from $76 million in 2011 to almost $99 million in 2012. Several factors contributed to the increase:

- First, out-of-market capacity payments (i.e., FCM reliability credits) rose by $10 million in 2012. In 2011, the reliability credits were paid to two units in Connecticut in the first five months because their de-list bids were rejected in the first Forward Capacity Commitment Period (i.e., June 2010 to May 2011) for reliability purposes. In the last seven months of 2012, larger reliability credits were paid to two units in Boston whose de-list bids were rejected for reliability in the third Forward Capacity Commitment Period (i.e., June 2012 to May 2013).[113]
- Uplift payments for voltage support rose by $9 million in 2012, which was primarily due to increased reliability commitments for voltage support in Western Central Massachusetts in 2012. The increased need was attributable to several planned transmission outages, some of which were required to incorporate transmission upgrades to the area.
- The "Economic" category of uplift associated with non fast-start resources increased more than $6 million in 2012. The increase in supplemental commitment for system-wide reserves (which is quantified in more detail in Subsection B above) led to concomitant increases in the NCPC payments to these resources.

113 There were no such rejections in the second Forward Capacity Commitment Period.

- Uplift payments for local second contingency protection increased by roughly $3 million in 2012, which was largely due to increased supplement commitments in Maine.

The increases related to commitment of gas-fired resources would have been higher if natural gas prices had not decreased significantly in 2012, which led to reduced commitment costs for these units. This is one reason that the "Economic" NCPC payments associated with fast-start resources fell by almost half in 2012, along with the reduction in the economic dispatch of such resources by UDS. As described above, fast-start resources dispatched economically by the real-time dispatch model (i.e., the UDS) can require NCPC payments to cover their costs because they generally do not set the LMP at the level of their total offer cost.[114] This underscores the importance of efforts to modify the real-time pricing and dispatch software to allow fast start resources to set the clearing price when they are the marginal source of supply (i.e., when their deployment enables the real-time model to avoid scheduling more expensive resources).[115]

G. Conclusions and Recommendations

When the market does not schedule sufficient resources to maintain reliability, the ISO must take out-of-market actions to make additional reserves available, such as supplemental commitments. The majority of supplemental commitments in 2012 were made to maintain reserves at the system-level rather. It is not surprising that a relatively small portion of the supplemental commitments are made for local areas because congestion has been very limited within New England. We conclude that the ISO's operations to maintain adequate reserve levels in 2012 were reasonably accurate and consistent with the ISO's procedures.

Our analysis of surplus online and fast-start capacity shows that market clearing prices are highly dependent on the amount of surplus capacity that is available in the real-time market, especially under relatively tight operating conditions.[116] Hence, factors that lead to artificially high levels of surplus capacity tend to:

- Reduce the incentive for units to be available in real time;

---

114 See Section V.A. for a detailed analysis and discussion of this issue.

115 See Section V.A. for a discussion of this recommendation.

116 In this section, Surplus Capacity refers to the amount of available on-line reserves plus available off-line reserves on fast-start resources minus the 30-minute operating reserve requirement.

- Dampen economic signals to invest in better performance and availability for both new and existing resources.
- Increase large and volatile uplift charges that can be difficult for participants to hedge and which may discourage participation in the ISO-NE market.

To ensure that these issues are minimized, it is beneficial for the ISO to regularly review its assumptions and processes for determining that additional commitments are necessary to satisfy its reliability requirements. In this regard, the ISO should consider modifying the assumptions it makes regarding real-time imports and exports once it implements the CTS process to improve the physical interchange with the NYISO.

The correlation between real-time prices and surplus capacity quantities also reinforces the importance of:

- Fully reflecting reliability needs in the market requirements for operating reserves. Procuring less operating reserves in the real-time market than needed for reliability increases the apparent surplus capacity amounts and depresses real-time prices, which reduces the incentives for generators to be available and perform reliably; and
- Allowing individual generators to sell only quantities of operating reserves than they are capable of providing. Additional sales artifically raises the apparent real-time supply of operating reserves and tends to depress real-time prices.

The ISO is moving forward on initiatives to address these issues. First, the ISO is proposing to procure “replacement reserves” in the real-time market, which will better enable the real-time prices to reflect reliability concerns that have arisen recently regarding increasing fuel supply uncertainty. Currently, the ISO is proposing a fixed quantity of replacement reserves, although we recommend that the ISO seek authority to modify this quantity daily based on its concerns regarding load and fuel supply uncertainty.

Second, the ISO is revising its procedures for auditing the 10-minute and 30-minute reserve capabilities of off-line and on-line resources to improve their accuracy. This will ensure that the real-time market procures a sufficient quantity of operating reserves and that real-time prices more accurately reflect the cost of maintaining reliability.

In addition, we have recommended the ISO provide generators with additional flexibility to modify their offers closer in the real time (i.e., intraday reoffers) to reflect changes in marginal costs. This will provide incentives for generators to be more available, since it will better enable them to recover their operating costs, particularly when gas prices are volatile in the hours leading up to real-time. The ISO is planning to introduce hourly day-ahead energy offers and intraday reoffers as early as the fourth quarter of 2014.[117]

We also recommend changes in Section V that would allow the real-time prices of energy and reserves to better reflect the costs of maintaining reliability during tight operating conditions. Since expectations of real-time prices are the primary determinant of day-ahead prices, these changes should increase the day-ahead market commitment of generators that can satisfy system's reliability criteria.

117 See 2013 Wholesale Markets Project Plan, page 8.

VII. Forward Capacity Market

ISO-NE has had an installed capacity market since it began operations in 1998, but the original market design lacked several features now recognized as important to the success of capacity markets. In particular, the original capacity market did not reflect the locational value of capacity resources, nor did it provide stable capacity price signals that potential investors could use to accurately predict investment returns for new resources. The Forward Capacity Market (FCM), the design of which was filed with FERC and approved in 2006, established a new market mechanism to attract and maintain sufficient resources to satisfy ISO-NE's long-term resource planning requirements efficiently.

The first Forward Capacity Auction (FCA 1) was held in February 2008, facilitating the procurement of installed capacity for the period from June 2010 to May 2011. Seven auctions have been held to date, which have satisfied ISO-NE's planning requirements through May 2017.[118] In June 2010, the start of the first Capacity Commitment Period allowed for the cessation of the individual reliability agreements that had been used extensively to maintain the resource requirements in Connecticut, Boston, and Western Massachusetts.

This section provides background on the FCM rules and evaluates the outcomes of the first six auctions. This section also discusses certain market reform proposals and rule changes that are underway. A summary of our conclusions and recommendations is at the end of the section.

A. ISO-NE's Forward Capacity Market Re-design

In recent years, the ISO and stakeholders have considered a number of significant reforms to the FCM. The Commission has issued several orders addressing FCM design topics.[119] These included several significant directives for the ISO to work with stakeholders to:

- Model eight capacity zones corresponding to its eight Load Zones – Ultimately, the ISO modeled four zones in FCA 7 and will model eight zones beginning in FCA 8.[120]

118 The latest auction, FCA 7, was held on February 4, 2013 for the Capacity Commitment Period of June 1, 2016 to May 31, 2017. However, the auction results are not included in this report.

119 See Order on Paper Hearing and Order on Rehearing, Docket ER10-787-000, et al. (Issued April 13, 2011). See also Order Compliance Filing, Docket ER12-953-001. (Issued February 12, 2013).

- Strengthen the supply-side market power mitigation rules – The ISO will implement supply-side mitigation measures requiring existing suppliers to justify offers exceeding $1 per kW-month beginning in FCA 8.

- Develop buyer-side market power mitigation rules in order to address the shortcomings of the proposed Alternative Price Rule [121] – The ISO will implement buyer-side mitigation measures with technology-specific offer floors for resources that are not economic without out-of-market revenues beginning in FCA 8.

- Extend the price floor in the auction until appropriate buyer-side market power mitigation measures can be implemented. Accordingly, the ISO will eliminate the price floor beginning in FCA 8.

As the External Market Monitor ("EMM"), we submitted a filing in February 2012 that made several recommendations, including that the Commission require the ISO and its stakeholders to evaluate and justify the slope of the demand curve for capacity (the current FCM implicitly employs of vertical demand curve, which raises significant concerns discussed later in this section). Although the Commission has did not require the ISO to modify the slope of the demand curve, we continue to recommend the ISO and its stakeholders modify the slope of the demand curve because it will reflect the incremental reliability that is provided by additional capacity, provide more competitive incentives to buyers and sellers, and promote capacity price stability over the long-term. This recommendation is discussed later in this section.

B. Background on the Forward Capacity Market

Capacity markets are generally designed to provide incentives for efficient investment in new resources. A prospective investor estimates the cost of investment over the life of the project minus the expected variable profits from providing energy and ancillary services (after netting the associated variable costs). This difference between investment costs and variable profits,

120 The ISO has requested rehearing on the Commission's directive to model eight zones in FCA 8.

121 The Alternative Price Rule was a provision designed to set the clearing price at a more efficient level when Out-Of-Merit capacity sales (i.e., new capacity entry from resources selling below their costs) distort the outcome of the auction.

which is known as Net Cost of New Entry (Net CONE), is the estimated capacity revenue that would be necessary for the investment to be profitable.[122]

In an efficient market, the investments with the lowest Net CONE will be the first to occur. The capacity price should clear at a level that is higher than the Net CONE of the investments that are needed and lower than the Net CONE of investments that are not needed. In this manner, the market facilitates investment in efficient capacity resources to meet system planning requirements. The resulting clearing price provides a signal to the market of the value of capacity.

FCM was designed to efficiently satisfy ISO-NE's resource adequacy requirements by using competitive price signals to retain existing resources and attract new supply. FCM has several key elements that are intended to work together to accomplish this goal. Some of the key elements are:

- Installed Capacity Requirement – The FCM procures the Net Installed Capacity Requirement (NICR)[123] of the New England Control Area and the capacity judged necessary to achieve regional reliability standards in the Capacity Commitment Period, which begins three years after the auction.

- Local Sourcing Requirement – Before each auction, the existing installed capacity[124] in each zone, less retirement and export bids, are compared to the zone's Local Sourcing Requirement (LSR).[125] Until FCA 7, if the amount of capacity was greater than the LSR, the zone would not be modeled as a separate import-constrained zone in the auction.[126] Export-constrained zones are always modeled in the auction. When the zonal requirements are modeled, the FCM produces locational prices that reflect the value of capacity in each zone.

---

122 Although the term "Net Cost of New Entry" is used here in a generic sense, Cost of New Entry has a specific meaning in the context of FCM, which is defined in Market Rule 1, Section 13.2.4.

123 The NICR is equal to the Installed Capacity Requirement minus the HQICC. This treats a portion of the capacity from Hydro Quebec as a load reduction rather than as supply.

124 This includes capacity that was sold in previous FCAs, but that is not yet in operation.

125 The LSR is the minimum amount of capacity that is needed in the load zone. Since FCA 1, the LSR has been sufficiently high to satisfy Resource Adequacy criteria (i.e., to reduce the probability per year of firm load shedding below 10 percent). Since FCA 4, the LSR is also set sufficiently high to satisfy Transmission Security Analysis criteria (i.e., to have sufficient capacity such that the system can be restored to a normal state after the largest two contingencies).

126 This rule has been modified so that four zones were modeled in FCA 7 and all eight zones will be modeled beginning in FCA 8. The effects of this rule prior to FCA 7 are discussed later in this section.

- New Capacity Treatment – Existing capacity participates in the FCM each year and has only a one-year commitment, while new capacity resources can choose an extended commitment period from one to five years at the time of qualification.[127] Both new and existing capacities are paid the same market clearing price in the first year, provided there is sufficient competition and sufficient supply. The price paid to new capacity after the first year is indexed for inflation.

The FCM design also includes several provisions that are intended to guard against the abuse of market power. Demand resources and intermittent generation resources compete with traditional generation to provide capacity, limiting supply-side market power in the capacity and energy market and enhancing economic efficiency. Certain de-list bids (the price below which a supplier will not sell its capacity) and export bids are subject to review by the Internal Market Monitor ("IMM") prior to the FCA in order to address potential withholding by suppliers. New capacity qualification rules and the three-year advance procurement feature allow new capacity projects to compete in the FCA.

C. Analysis of Forward Capacity Auction Results

Five FCAs were held before 2012 and FCA 6 was held in April 2012 for the commitment period of 2015/2016 (i.e., June 1, 2015 to May 31, 2016).[128] In each of the six auctions, there was a substantial surplus of capacity over the NICR. Accordingly, each auction cleared at the floor price: $4.50 per kW-month in FCA 1, $3.60 in FCA 2, $2.95 in FCA 3 and FCA 4, $3.21 in FCA 5, and $3.43 in FCA 6.

No import-constrained zones were deemed necessary in any of the six auctions because the amount of existing capacity exceeded the LSR in each area. Maine was modeled as an export-constrained zone in all six auctions, but there was no price separation between Maine and the rest of New England. This section summarizes and evaluates the overall results of the first six FCAs, the de-list bids of existing suppliers, and the procurement of new capacity.

---

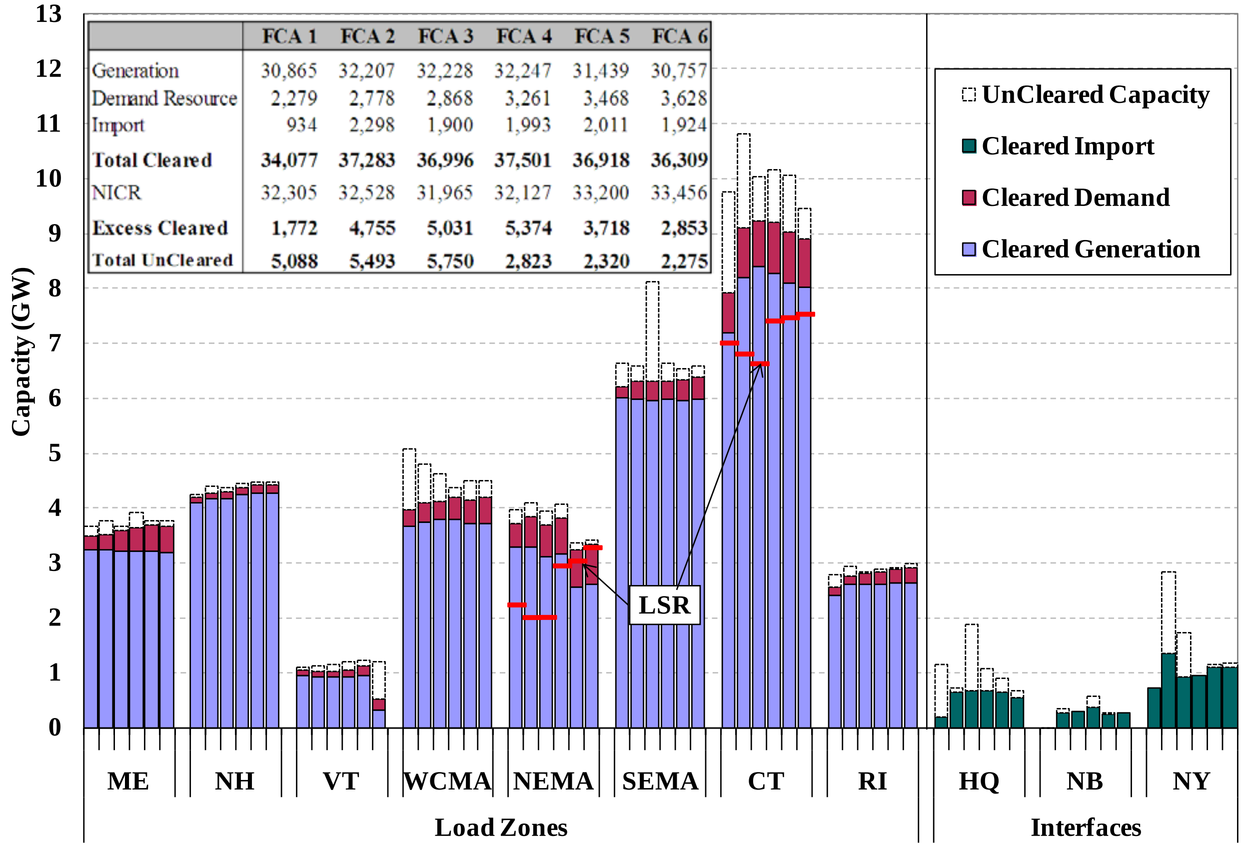
127 We have recommending modifications to the treatment of new capacity that are discussed later in this section.

128 Five FCAs were held before 2012: the first in February 2008 for the commitment period of 2010/2011 (FCA 1), the second in December 2008 for the commitment period of 2011/2012 (FCA 2), the third in October 2009 for the commitment period of 2012/2013 (FCA 3), the fourth in August 2010 for the commitment period of 2013/2014 (FCA 4), and the fifth in June 2011 for the commitment period of 2014/2015 (FCA 5).

1. Summary of Capacity Auction Results

Figure 32 summarizes the procurements in the first six FCAs, showing the distribution of cleared and un-cleared capacity by location. Cleared resources are divided into generating resources, demand response resources, and imports from external areas.[129] The amounts of cleared resources are shown relative to the LSRs for Connecticut and NEMA (in the figure) and relative to the NICR for all of New England (in the table).

Figure 32: FCM Auction Clearing Summary by Location
FCA 1 – FCA 6



	FCA 1	FCA 2	FCA 3	FCA 4	FCA 5	FCA 6
Generation	30,865	32,207	32,228	32,247	31,439	30,757
Demand Resource	2,279	2,778	2,868	3,261	3,468	3,628
Import	934	2,298	1,900	1,993	2,011	1,924
Total Cleared	**34,077**	**37,283**	**36,996**	**37,501**	**36,918**	**36,309**
NICR	32,305	32,528	31,965	32,127	33,200	33,456
Excess Cleared	**1,772**	**4,755**	**5,031**	**5,374**	**3,718**	**2,853**
Total UnCleared	**5,088**	**5,493**	**5,750**	**2,823**	**2,320**	**2,275**

Prior to each of the FCAs, it was determined that the existing capacity was sufficient to satisfy the local requirements, so no import-constrained zones were modeled. Accordingly, the amount of procured capacity in NEMA and Connecticut exceeded their LSRs by a significant margin in most of the six FCAs.

129 The amount of cleared demand response resources shown in the figure has been adjusted to exclude Real-Time Emergency Generation resources in excess of 600 MW.

The amount of excess capacity rose significantly in Connecticut after FCA 1 due to several significant new capacity additions. These new capacity additions apparently resulted from two Requests for Proposals (RFPs) that were conducted by the Connecticut Department of Public Utility Control (DPUC). These RFPs account for most of the new generation that has been added since the FCM has been in place.

The amount of excess capacity fell in Connecticut and in NEMA from FCA 3 to FCA 4 as a result of changes in the method calculating the LSRs. Until FCA 4, LSRs were based on Resource Adequacy criteria (i.e., one day in ten years criteria) and did not reflect Transmission Security criteria (i.e., the system can withstand two contingencies without load shedding).[130] Since Transmission Security criteria required a higher level of resources in each of these areas, the LSRs were not sufficiently high to meet the local planning criteria for Connecticut and for Boston until FCA 4. From FCA 4 to FCA 5, the amount of excess capacity fell substantially in NEMA due to the planned retirement of the last two units at the Salem Harbor plant totaling nearly 600 MW.

Although the amount of excess capacity in NEMA was very low based on the existing resources that were qualified to participate in FCA 6, the NEMA capacity zone was not modeled in the auction because the existing resources were adequate to satisfy the LSR. Consequently, the ISO rejected 79 MW of de-list bids in NEMA for reliability rather than accepting offers from a new capacity resource, since this would have required a higher clearing price for NEMA. This problem delayed economic entry and resulted in out-of-market payments to maintain adequate capacity levels. This was resolved after FCA 6 by requiring that NEMA and other capacity zones be modeled regardless of whether existing resources are adequate to satisfy the LSR.

The amount of capacity procured in each FCA has been more than sufficient to satisfy the system level reliability requirements. The procured excess capacity has ranged from 1.8 GW in FCA 1 to 5.4 GW in FCA 4. Substantial excess capacity cleared in the first six auctions as a result of

130 The determination of the Local Sourcing Requirements, including the modeling assumptions used to determine the Local Resource Adequacy Requirement and the Transmission Security Analysis Requirement are described in Tariff Section III.12.2.

the price floor. Originally, the price floor was supposed to be eliminated after FCA 3, but it has been extended through FCA 7.[131]

Generating resources provided the vast majority of capacity in each auction. However, the portion of the NICR satisfied by demand response resources has gradually risen from 7 percent in FCA 1 to 11 percent in FCA 6. Roughly 55 to 60 percent of the cleared demand response resources were *active* demand resources, which reduce load in response to real-time system conditions or ISO instructions. The rest were *passive* resources, which also reduce load, but not in response to real-time conditions or instructions (e.g., energy efficiency). Imports from Hydro Quebec, New Brunswick, and NYISO also accounted for a significant portion of the procured capacity, averaging more than 2,000 MW in recent FCAs.[132]

In each auction, a substantial amount of qualified resources did not clear. New proposed resources accounted for more than 70 percent of the un-cleared capacity. The presence of competitively-priced offers from potential new entrants is an aspect of the FCM that should motivate suppliers to behave competitively in the FCAs.

2. Evaluation of De-list Bids

FCM provides a mechanism to retain existing resources in New England. Stable price signals encourage existing resources to stay in-service, reducing the need to satisfy reliability requirements using out-of-market payments (e.g., payments from reliability agreements). Relying on out-of-market payments is undesirable because doing so provides the most compensation to the least efficient resources in the market. Hence, the use of out-of-market payments tends to reduce the efficiency of investment in the wholesale market.

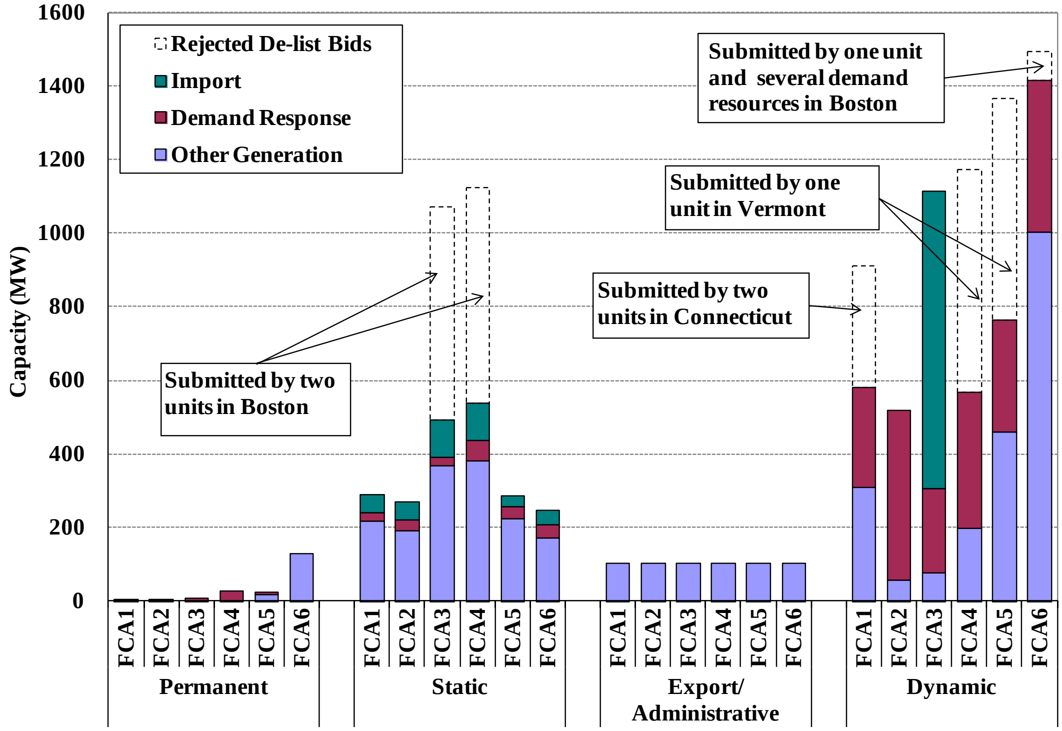
Under FCM, existing resources have the option to submit de-list bids to indicate they intend to de-list (i.e., not sell capacity) all or part of their capacity during the commitment period if the capacity price is less than their de-list bid price. The ISO reviews de-list bids and may reject them for reliability needs or in accordance with the supplier-side mitigation rules.

131 The price also cleared at the floor in FCA 7 for the areas outside NEMA.

132 A large portion of the import capability from Hydro Quebec is included in the HQICC, which is treated as a load reduction in the NICR rather than as supply.

Figure 33 evaluates several categories of accepted de-list bids in the first six FCAs. The figure shows four categories of de-list bids: permanent, static, export or administrative, and dynamic.[133] Accepted de-list bids are also separated according to the type of resource: generation, demand response resources, and imports. The figure also shows the de-list bids that were rejected by the ISO-NE for reliability reasons.

Figure 33: Summary of Accepted De-list Bids by Type
FCA 1 – FCA 6



On average, approximately 1,100 MW of generation resources attempted to de-list (excluding export de-list bids), which relieves them of their capacity obligations and allows them to go out of service for the commitment period. An average of 58 percent of this capacity was able to de-list without delay in each FCA, while the ISO prevented 42 percent of the resources attempting

133 Each category of de-list bid is defined in Tariff Section III.13.2.5.2. Permanent de-list bids are submitted by resources intending to retire; static de-list bids are known in advance of the auction and must be approved by the IMM as consistent with the resource's going forward costs if they exceed 80 percent of CONE, export de-list bids are associated with resources whose capacity will be exported if not selected in New England; and dynamic de-list bids are not known in advance of the auction, but are associated with resources that may de-list at any time once prices fall below 80 percent of CONE. The April 2011 Order directed the ISO to file changes that would require the IMM to approve any bids above $1 per kW-month as consistent with the resource's going forward costs, although this directive has not been implemented yet.

to de-list from doing so for reliability reasons. The fact that a large share of the supply attempting to de-list was unable to do so for reliability reasons raises concerns about the effectiveness of the FCM in facilitating efficient entry and exit of resources in New England.

Ideally, when capacity is needed in a particular area to satisfy local planning criteria, the capacity market should provide economic signals for capacity to enter in the area. However, the ISO's rejection of the following de-list bids in all but one of the FCAs demonstrates that this is not the case:

- 330 MW of generation de-list bids in Connecticut in FCA 1;
- 585 MW of generation de-list bids in Boston in FCA 3 and FCA 4;
- 604 MW of generation de-list bids in Vermont in FCA 4 and FCA 5; and
- 27 MW of generation de-list bids and 52 MW of demand de-list bids in Boston in FCA 6.

All of these de-list bids were rejected when the ISO determined in its Transmission Security Analysis that the units were needed for reliability. Since the rejected de-list bids were substantially smaller than the excess cleared capacity for all of New England in each of the five auctions, the price was unaffected (the auctions would have cleared at the price floor with or without the rejected bids). However, the rejection of the de-list bids has highlighted several issues with the original FCM market design.

First, the Connecticut and Boston LSRs were initially much lower than the capacity requirements that were implied by the Transmission Security Analysis, leading to the rejection of de-list bids in Connecticut in FCA 1 and in NEMA in FCA 3. In principle, markets should be designed to satisfy the full reliability needs of the system, which allows market prices to accurately reflect these needs. Accordingly, the ISO modified the LSR criteria to consider the Transmission Security Analysis that is used to determine whether a de-list bid should be rejected for zone-level reliability beginning in FCA 4.

Second, although the Boston LSR was raised in FCA 4 to be consistent with the local requirement implied by the Transmission Security Analysis, but generator de-list bids were still rejected because of local reliability requirements within the zone.

Third, de-list bids were rejected in NEMA in FCA 6 because they were needed to satisfy the LSR and the NEMA zone was not modeled. This was addressed by requiring that the NEMA zone be,modeled always beginning in FCA 7.

Fourth, the rejection of a de-list bid of a resource in Vermont shows that the need could arise for zonal price separation in areas other than Connecticut, Boston, and Maine. If all eight load zones are modeled in each FCA, then the clearing price in each zone should reflect the true capacity needs of the system. To address this concern, the Commission directed the ISO to modify its rules to always model eight capacity zones beginning in FCA 8.

3. New Capacity Procurement

A key objective of the FCM is to provide efficient market incentives for investment in new resources. The FCA provides a mechanism for prospective investors to build new resources that will be profitable based on the auction clearing price. As a result of competition between prospective investors, the investment projects that have the lowest Net CONE should clear in the auction and result in the most efficient investment over time. Figure 34 shows the amounts of new capacity that were procured in the first six FCAs by load zone or external interface. Capacity is divided by resource type: generation, demand response, and import capacity. We also distinguish the capacity based on whether it received existing treatment in FCA 1 or it cleared in FCA 1 through FCA 6.[134]

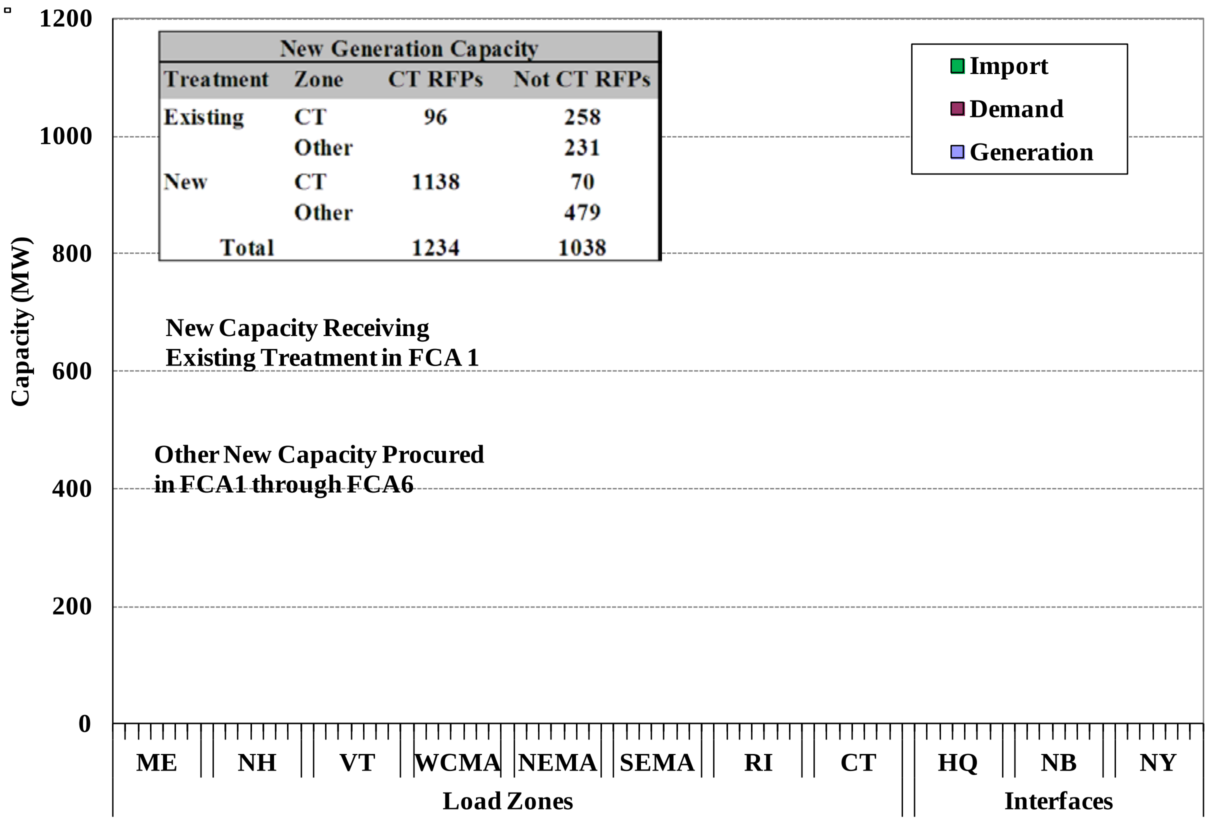
To determine whether new capacity entered due to the FCM revenue, the table in the figure identifies the quantity of capacity contracted under the Connecticut DPUC RFPs that may receive additional capacity payments beyond those from the FCM.[135] In each of the first six FCAs, an average of nearly 1,700 MW of new capacity was procured from generation, demand

134 Resources expected to be in-service prior to the first Capacity Commitment Period could elect to be treated as existing resources in FCA 1. Accordingly, they are able to submit de-list bids rather than supply offers.

135 See State of Connecticut DPUC Investigation of Measures to Reduce Federally Mandated Congestion Charges (Long-Term Measures), May 3, 2007, Docket No. 05-07-14PH02, page 2. See also State of Connecticut, DPUC Review of Peaking Generation Projects, June 25, 2008, Docket No. 08-01-01, page 64.

response resources, and imports.[136] The discussion following the figure reviews and evaluates the procurements of new capacity by resource type that are shown in Figure 34.

Figure 34: New Capacity Procurement by Location
FCA 1 – FCA 6



Import Capacity

Approximately 56 percent of new capacity was sold by importers in the first six FCAs, indicating that the suppliers expected the revenues from providing capacity to New England during these Capacity Commitment Periods to be greater than the revenues from providing capacity to another market during the same period. Many of the capacity importers to New England have the option to sell capacity into New York in future periods. Hence, the amount of capacity imports may decrease in the future if the floor price is eliminated. Similarly, the amount of capacity that de-lists in order to export may increase in the future if the floor price is removed.

136 This excludes new resources treated as existing resources because they were already committed to enter.

Demand Response Capacity

Demand response resources have sold substantial amounts of capacity under FCM, accounting for roughly 27 percent of new capacity sold in the first six FCAs. This indicates that the Net CONE of many demand response resources is lower than the capacity clearing prices. However, if demand response activation becomes more frequent in the future, the Net CONE of many demand response resources should increase. This increase would arise if the heavier reliance on demand response were to result in much more frequent emergency load curtailments that are costly for demand response providers to satisfy. If this were to happen, it would put upward pressure on capacity clearing prices or reduce the amount of capacity provided by demand response resources.

Additionally, demand response resources may not provide response comparable to supply resources during shortage or emergency conditions. When demand response resources were deployed, the performance of the resources varied widely. Often only a small portion of resources curtailed an amount of load within 10 percent of the instructed amount, which is the performance threshold used for assessing uninstructed deviation penalties to generators.[137] These results raise significant concerns about whether the demand response resources selling capacity in New England provide the same level of reliability benefits as internal generators and imports. It may be appropriate to reassess whether the performance criteria and settlements with demand response resources that do not perform as instructed should be more consistent with the criteria used for generation and imports.

Generation Capacity

A substantial amount of new generation capacity (2,272 MW) has entered the market under FCM. Entry of generation resources would generally not be expected when the price clears at the price floor as it did in the first six FCAs. The floor price is generally believed to be substantially lower than the Net CONE for new investment in most types of generation. However, the table in the figure above shows that more than 1,200 MW of the new investment in

137 For example, only 22 percent of resources curtailed an amount of load within 10 percent of instructed amount when they were activated on June 24, 2010. See 2010 Annual Markets Report, ISO-NE, June 2011, Figure 3-31 for details.

generation received additional payments under the RFPs of the Connecticut DPUC and nearly 500 MW are resources that received existing treatment, which indicates that their entry decisions were not contingent on the outcome of the FCA. We distinguish these two types of new investment because the FCA did not directly facilitate the entry, although the existence of the FCM may have motivated the processes that resulted in the entry.

Entry that occurs only because its offer is accepted in the FCA (not because the supplier was awarded a contract under a state RFP or was already building the unit) is entry that the FCM must efficiently facilitate over the long-run for the FCM to be effective. For this reason, we seek to determine how the FCM market has affected this class of capacity investment. The table in the figure shows that only 550 MW of new generating resources cleared in the six FCAs that were not under the CT RFP or treated as existing resources. Most of these resources are facilities powered by renewable fuels, designed to up-rate existing resources, or made to re-power existing power plants. Such projects may have a lower Net CONE than most of the potential investments in new generation, which explains why they would clear at the floor prices in the first six FCAs. Given the prevailing surplus in New England, it would have been surprising if a substantial amount of new generating resources had cleared in the FCM.

The amount of capacity committed to ISO-NE procured under FCM exceeded the ISO-NE capacity requirement by a significant margin in each of the first six FCAs.[138] FCM has provided strong incentives for the sale of new capacity by demand response resources and importers. However, once the price floor is discontinued, the price will likely drop significantly, and the FCA will not procure a significant amount of excess capacity.

It is still too early to determine whether the FCM will efficiently facilitate investment in new generation when it is needed. The prevailing surplus has caused the auction to clear at the floor price, which is well below most estimates of the Net CONE for new generation. Therefore, the market has not needed to facilitate investment in new generation resources.

138 The margin is 5.5 percent in FCA 1, 14.6 percent in FCA 2, 15.7 percent in FCA 3, 16.7 percent in FCA 4, 11.2 percent in FCA 5, and 8.5 percent in FCA 6.

D. Capacity Market Design – Sloped Demand Curve

Absent the price floor, the demand in the FCM market is implicitly defined by the minimum capacity requirement and the maximum price.[139] These requirements result in a vertical demand curve (i.e., demand that is insensitive to the price, buying the same amount of capacity at any price).

1. Attributes of Demand in a Capacity Market

The demand for any good is determined by the value the buyer derives from the good. For capacity, the value is derived from the reliability provided by the capacity to electricity consumers. The implication of a vertical demand curve is that the last MW of capacity needed to satisfy the minimum requirement has a value equal to the deficiency price, while the first MW of surplus has no value. In reality, each unit of surplus capacity above the minimum requirement will increase reliability and lower real-time energy and ancillary services costs for consumers (although these effects diminish as the surplus increases). This value can only be accurately reflected in the FCM market framework by a sloped demand curve. The fact that a vertical demand curve does not reflect the underlying value of capacity to consumers is the source of a number of the concerns described later in this section.

2. Attributes of Supply in a Capacity Market

In workably competitive capacity markets, the competitive offer for existing capacity (i.e., the marginal cost of selling capacity) is generally close to zero.[140] A supplier's offer represents the lowest price it would be willing to accept to sell capacity. This is determined by whether there are costs the supplier will incur to satisfy the capacity obligations for the resource, the foregone opportunity cost of exporting capacity, and any "going-forward costs" (GFC) not covered by the expected net revenues from energy and ancillary services markets. Since each of these factors tends to be very low for most existing resources, most suppliers are price-takers in the capacity market, accepting virtually any non-zero price to sell capacity. Experience in the FCM and in

139 The maximum price in FCM is the starting price for the descending clock auction.

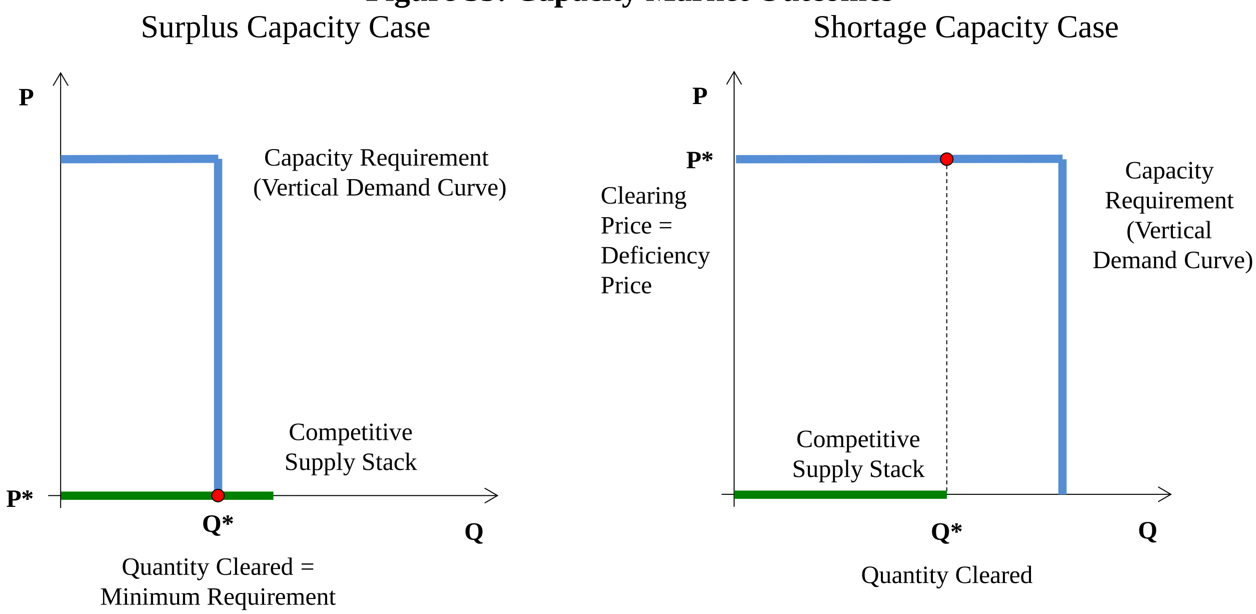
140 Assuming no opportunity costs of exporting capacity to a neighboring market.

other capacity markets with vertical demand curves confirms that most suppliers are essentially price-takers, willing to sell capacity at very low prices.

3. Implications of the Vertical Demand Curve for Performance of the Capacity Market

When the low-priced supply offers clear against a vertical demand curve, two general outcomes are possible. If the market is not in a shortage, the price will clear very low—this is illustrated the left panel in Figure 35 below—and would likely be the result for ISO-NE absent the price floor. If the market is in shortage (so the supply and demand curves do not cross), the price will clear at the deficiency price (right panel).

Figure 35: Capacity Market Outcomes



This pricing dynamic and the associated market outcomes raise significant issues regarding the long-term performance of FCM. First, this market will result in significant volatility and uncertainty for market participants. This can hinder long-term contracting and investment by making it difficult for potential investors to forecast the capacity market prices and revenues. In fact, it may be difficult for an investor to forecast that the market will be short in the future with enough certainty that its forecasted capacity revenues will be substantially greater than zero. This can undermine the effectiveness of the capacity market in maintaining adequate resources.

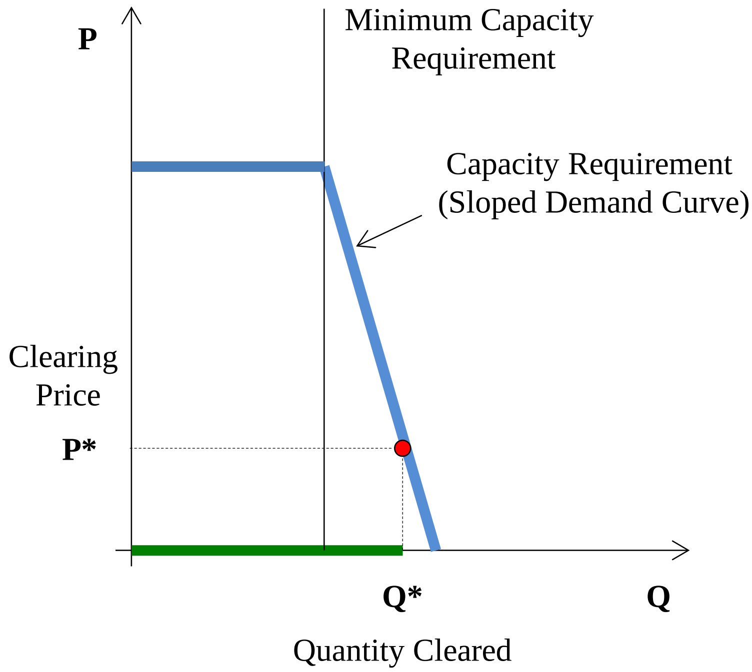
Second, since prices produced by such a construct do not accurately reflect the true marginal value of capacity, the market will not provide efficient long-term economic signals to govern investment and retirement decisions.

Third, a market that is highly sensitive to such small changes in supply around the minimum requirement level creates a strong incentive for suppliers to withhold resources to raise prices. Withholding in such a market is nearly costless since the foregone capacity sales would otherwise be priced close to zero. Therefore, market power is of greater potential concern, even in a market that is not highly concentrated. These concerns grow when local capacity zones are introduced where the ownership of supply is generally more concentrated.

4. Benefits of a Sloped Demand Curve

A sloped demand curve addresses each of the shortcomings described above. Importantly, it recognizes that the initial increments of capacity in excess of the minimum requirement are valuable from both a reliability and economic perspective. Figure 36 illustrates the sloped demand curve and the difference in how prices would be determined.

Figure 36: Sloped Demand Curve



When a surplus exists, the price would be determined by the marginal value of additional capacity as represented by the sloped demand curve. This provides a more efficient price signal from the capacity market. In addition, the figure illustrates how a sloped demand curve would serve to stabilize market outcomes and reduce the risks facing suppliers in wholesale electricity markets. Stabilizing capacity prices in a manner that reflects the prevailing marginal value of capacity would improve the incentives of suppliers that rely upon these market signals to make investment and retirement decisions.

A sloped demand curve will also significantly reduce suppliers' incentives to withhold capacity from the market by increasing the opportunity costs of withholding (foregone capacity revenues) and decrease the price effects of withholding. This incentive to withhold falls as the surplus falls because the cost of withholding increases. While it would not likely be completely effective in mitigating potential market power, it would significantly improve suppliers' incentives.

Based on both the theoretical and practical concerns with the current vertical demand curve, we recommend that the Commission require that the ISO work with its stakeholders to consider all of the relevant parameters that characterize the demand for capacity in the FCM.[141]

E. Capacity Market Design – Investment in a Forward Capacity Market

Capacity markets solve the "missing money" problem in wholesale capacity. Missing money exists when the wholesale markets do not provide sufficient revenues to support needed investment to satisfy the ISO's planning requirements. This problem arises because:

- Planning reserve requirements are higher than the installed capacity levels that an energy-only market would provide even if shortages were priced efficiently, requiring additional revenues to prompt investment;
- Increased levels of installed capacity reduce the frequency of shortages and associated shortage pricing in the energy market; and
- The out-of-market operating actions that are used to maintain reliability (mainly through the RAA process) reduce real-time prices by preventing shortages.

141 We also made this recommendation in comments filed to FERC on the ISO-NE FCM design in February 2012 in Docket ER10-787-000.

Most of the incentive to invest in wholesale electricity markets is attributable to price signals during shortages and to capacity markets. As shortage pricing rises (or becomes more frequent), the amount of revenue that must be produced by the capacity market will fall. Assuming shortage prices are not set inefficiently high (above the expected value of lost load), a capacity market will generally be necessary to ensure that the markets will maintain adequate resources.

Regardless of the relative reliance on shortage pricing and capacity market revenues, greater stability and predictability of market revenues will tend to lower the cost of capital, reducing the cost of new entry and prices for consumers. One of the virtues of a sloped demand curve (which is discussed in Subsection D) is that it tends to produce more stable price signals by instituting a transparent and predictable relationship between clearing prices and the planning reserve margin. Another virtue of a sloped demand curve is that it is likely to work better with other key provisions of FCM to produce efficient market outcomes. This section uses several examples to illustrate how a sloped demand curve is likely to produce more stable and predictable market outcomes than a vertical demand curve in light of several significant provisions related to new entry.

The ISO is also pursuing a Performance Incentive proposal that would effectively shift more of the long-term economic signal to a form of shortage pricing. This proposal is discussed in the subsection E.

1. Existing Rules to Facilitate New Entry

The capacity market has two rules that are specifically intended to facilitate new entry of generating resources, including:

- The Rationing Election – This allows a new generating resource to elect to make its offer rationable, meaning that it need not be wholly accepted. The owner can elect to make the offer rationable down to a specified MW level.[142]

- The Capacity Commitment Period Election – This allows a new resource to lock-in the capacity clearing price of the FCA in which it initially sells for a period of up to five years.[143]

142 See Tariff Section III.13.1.1.2.2.3(b).

It is important to consider how the incentives that arise from these two rules are likely to interact with the shape of the demand curve. This is important for long-term resource adequacy because new entrants will still be influenced by their expectations of how the market will perform after the fifth year their resource has been in service. Specifically, poor anticipated market performance or high expected price volatility after the first five years of the investment will provide an economic disincentive that would affect the new supplier's offer price.

The first part of this subsection discusses the anticipated cycle of investment under the current FCM rules and how this will lead to volatile capacity clearing prices, which may in turn provide disincentives for new entry and for the capital expenditures necessary to keep existing resources in service. The second part of this subsection discusses how capacity price volatility will be reduced by replacing the current vertical demand curve with a sloped demand curve, and discusses additional rule changes that might further improve market performance.

2. Investment Under a Vertical Demand Curve

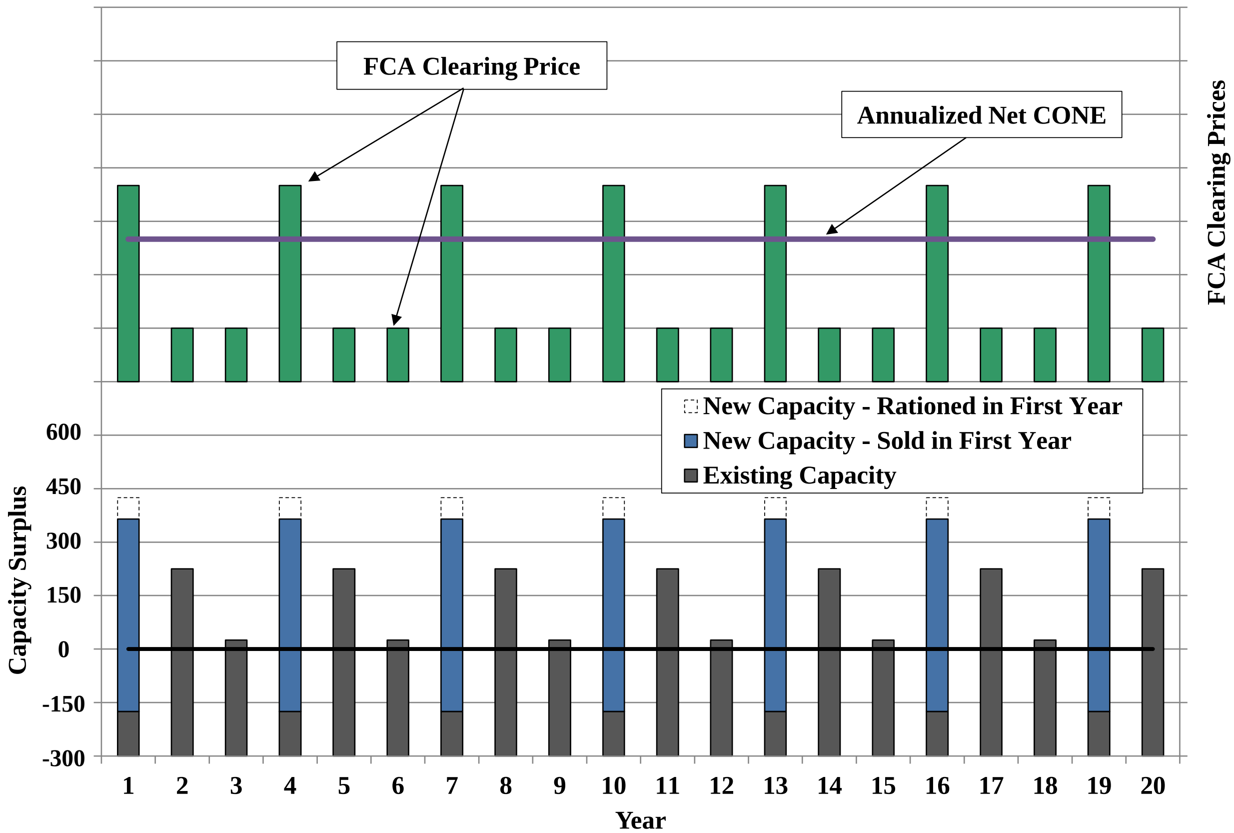
Before evaluating how prices are likely to behave in a capacity market, it is important to consider the efficient scale of investment and the amount of new resources that are likely to be needed in each year. Over the last decade, most new generation investment has been in combined cycle technology rather than peaking technology. Typically, new combined cycle installations have been 500 to 600 MW, while peaking installations have ranged from 50 to 250 MW. Consequently, it is likely that the factors that lead to the need for new resources (e.g., load growth) will not necessitate new entry in every year. Hence, in some areas, a new combined cycle generator might be expected to enter every few years.

Figure 37 illustrates how the cycle of investment would likely evolve if the price floor were eliminated and the excess capacity margin was diminished. The bottom portion of the figure shows the excess capacity margin in a particular capacity region. These bars show the growth of load and retirement of old resources out-pacing the entry of demand response by 200 MW each year. In the years when the excess capacity margin would fall below zero, a new 600 MW combined cycle is assumed to enter the market, leading to a substantial increase in the excess

143 See Tariff Section III.13.1.1.2.2.4.

capacity margin. Accordingly, this leads to new entry every three years. The top portion of the figure shows the clearing price in each year and the flat line shows the annual inflation-adjusted levelized net CONE.

Figure 37: Cycle of Investment with a Vertical Demand Curve



In the years when no new capacity investment occurs (e.g., Years 2, 3, 5, 6, 8, etc.), the auction clears at a price that reflects of marginal cost of capacity for existing resources. This would include the going forward costs of existing generation, the opportunity cost of selling in a neighboring capacity market, the costs of satisfying the ISO's capacity obligations, and/or the participation costs of demand response resources. This would likely lead to clearing prices that are substantially lower than the Annualized Net CONE as shown in Figure 37.

In the years when new entry does occur (e.g., Years 1, 4, 7, etc.), the auction clears at a price that reflects the offer price of the new entrant. In this example, the new combined cycle unit offers to sell at a price substantially higher than its annualized net CONE because it anticipates that clearing prices will be lower than the annualized net CONE over the balance of the life of the investment. Hence, offering to sell at its annualized net CONE in Year 1 would lead the supplier to invest unprofitably. The supplier would therefore raise its offer in Year 1 to compensate for the lower expected capacity prices in future years.

Figure 37 shows that if the efficient scale of new entry is sufficiently large such that new entry does not occur in each year, the current capacity market rules (and the vertical demand curve in particular) are likely to produce volatile swings in the clearing price that range well above and below the annualized net CONE for the most economic new entrant.

The figure also shows how the Capacity Commitment Period Election and the Rationing Election would likely affect market performance when a vertical demand curve is used. First, the Capacity Commitment Period Election would allow the new entrant to lock-in the Year 1 price for five years, which would insulate the new entrant from the boom-and-bust cycle for several years. The effect would be to induce the new entrant to enter at a lower offer price in Year 1, although this offer price would still likely be substantially higher than the unit's annualized net CONE. Since these considerations would apply only in years when a new unit enters the market, the average clearing price over the long-term would be substantially lower than the annualized net CONE of the new entrant. Figure 37 shows an example where the average clearing price over twenty years is roughly 25 percent lower than the annualized net CONE of the new entrant.[144]

Second, a large new entrant would use the Rationing Election to compete more effectively with other competing projects. The Rationing Election would increase the ability of large scale projects to compete with smaller scale projects in the capacity auction. However, large scale projects that elected to be rationed in the initial year would sell all of their capacity in subsequent years. This would increase the price in Year 1 relative to the price in subsequent years, increasing the tendency of the market to exhibit apparent boom and bust pricing cycles. The following section describes how the same market would likely perform if a sloped demand curve were used.

144 As illustrated in the figure, if the auction price clears at roughly $3.75 in the years when new entry occurs (e.g., Years 1, 4, 7, etc.) and at $1.0 in the years when no new capacity investment occurs (e.g., Years 2, 3, 5, 6, 8, etc.), then the average clearing price over twenty years is approximately $2 ($3.75*7/20+$1*13/20), which is roughly 25% lower than the annualized net CONE of $2.75.

3. Investment with a Sloped Demand Curve

If a sloped demand curve were adopted for use in FCM as recommended in subsection D, it would substantially reduce the price volatility illustrated in Figure 37.[145] With a sloped demand curve, the price would generally decrease much less sharply in the year following the new entry and rise as the surplus falls. However, it is still important to consider how the sloped demand curve would interact with the Rationing Election and the Capacity Commitment Period Election.

With a sloped demand curve, the Capacity Commitment Period Election should have a smaller effect because the difference between the prices paid to new and existing resources would be substantially reduced. New suppliers would not expect nearly as large a reduction in revenues after the initial 5-year period and, therefore, would likely submit lower-priced offers.

However, a large new entrant may have an incentive to: (i) elect for its offer to be rationable in the first year in order to make its offer more competitive, and (ii) lock-in the higher clearing price for five years. This pair of elections would allow the new entrant to collect a higher clearing price for a portion of its capacity for five years. The new entrant could then export the remaining capacity to a neighboring control area in the first year, and it could sell the remaining capacity internally in subsequent years. Although this incentive would be moderated by the offers from competing projects, the incentive would not be eliminated.

In utilizing a simple model to illustrate the pattern of prices and new investment under both vertical and sloped demand curves, we are able to draw two overall conclusions. First, a sloped demand curve leads to much more stable and predictable prices than a vertical demand curve, and smaller effects of the ability of entrants to lock-in their price for five years. Second, new entrants may have an incentive under the sloped demand curve to elect rationing in order to raise the clearing price (to the benefit of both its lock-in election and its existing resources).

145 For the purposes of this subsection, we assume that the sloped demand curve would be set such that average capacity clearing prices over the long-term would be equal to the annualized net CONE of an efficient new entrant. Hence, it would be set such that the average clearing price would be: (i) equal to the Annualized Net CONE when there was a small surplus; (ii) slightly higher than the Annualized Net CONE with no surplus.

Hence, in addition to our primary recommendation to implement a sloped demand curve, we recommend the ISO evaluate the interaction of the rules for new suppliers that are related to the Rationing Election and the Capacity Commitment Period Election to determine whether they will promote efficient investment and FCM outcomes over the long-term.

F. Performance Incentives for Capacity Resources

In October 2012, the ISO issued a proposal that would create new performance incentives for generators to be available and operate reliably during real-time shortage events. The proposal would use a shortage price mechanism during 30-minute reserve shortages called a Performance Payment Rate ("PPR"). The PPR will produce a settlement that is substantially similar to having higher real-time reserves and energy pricing during the shortage, but the PPR would not be included in the real-time prices and would instead be settled separately.

Generators or importers that provide more energy and reserves than their capacity obligation implies would be paid the PPR, while those that provide less would be charged the PPR. We provided an preliminary assessment of the proposal, which is briefly summarized in this subsection.[146]

1. FCM Performance Incentives versus Conventional Shortage Pricing

There are several notable differences between the FCM Performance Incentive proposal and conventional shortage pricing. First, FCM Performance Incentives will produce much more stable net revenues over time. The FCM delist bids and resulting FCA clearing prices will reflect the expected value of the shortage pricing. This is the case because suppliers will forego the shortage pricing (i.e., the incentive payment) when it sells capacity. Hence, the capacity revenues will include the *expected* value of shortages, which will be much more stable than the *actual* value of the shortages. The current Peak Energy Rent ("PER") deduction has a similar effect, but is compromised by a number of exclusions that cause the PER not to be generally applied to FCM resources during shortages.

146 See letter to NextERA Energy Resources LLC, dated February 19, 2013, *RE: Questions on ISO New England Performance Incentives Proposal* at: "http://www.iso-ne.com/committees/comm_wkgrps/mrkts_comm/mrkts/mtrls/2013/mar11122013/a14_potomac_economics_memo_02_19_13.pdf"

Second, loads will not be directly exposed to the Performance Incentives. Failing to provide these incentives to loads would be significant if the real-time load could efficiently respond to the shortage. However, since most responsive load will participate in the markets through the ISO's demand response programs, we do not believe this is a substantial problem with the proposed performance incentives.

Third, the Performance Incentive proposal may not perform as efficiently in coordinating the commitment of resources in the day-ahead market. As the probability of experiencing a real-time shortage increases, we would expect the following actions by participants:

- Under enhanced shortage pricing, day-ahead prices and net load purchases in the day-ahead market would rise, resulting in increased commitment of supply and scheduling of net imports.
- Under the performance incentives proposal, day-ahead prices and net load purchases in the day-ahead market would not rise, but generators would face increasing incentives to self-commit their resources to ensure they are online.

As a result, the day-ahead market may be less effective in coordinating supply commitments. We cannot say how significant this effect would be, but are working with the ISO to assess this issue. The significance of this issue is affected by the magnitude of the PPR,

2. Establishing a Performance Payment Rate

The ISO should develop an appropriate basis for setting a Performance Payment Rate. The proposal has advocated developing a PPR based on the entry cost for a new unit, which would effectively vesting most of the incentive to invest in the Performance Payment Rate. This raises efficiency concerns to the extent that the Performance Payment Rate substantially exceeds the fundamental value of reliability during shortages (the expected value of lost load). We recognize that there is substantial uncertainty regarding the value of lost load, but still believe that this is the appropriate basis for the PPR.

3. Other Considerations

Since the FCM Performance Incentives are based on 30-minute reserve shortage periods, the practices of the ISO to take out-of-market actions to prevent a 30-minute shortage will limit the effectiveness of the Performance Incentives (just as it does under the current market design that uses a $500 per MWh RCPF for shortage pricing). Hence, it will still be important devise ways to reflect the cost of such actions in real-time clearing prices, which we discuss in Section V.

4. Market Power Mitigation

Because the PPR is paid or charged based on a supplier's FCM obligation, the performance incentives proposal will compel sellers in the FCM to submit offers that include the expected value of the PPR (selling capacity will be similar to selling a call option on the shortage revenue associated with the PPR). Hence, when the option is called during shortages, the supplier must satisfy the option by delivering the energy or reserves, or buying them at the Performance Payment rate (via the penalty that is deducted from the supplier's FCM payment). This will make administration of the market power mitigation measures more difficult because the review of delist bids for economic withholding will have to account for a reasonable expectation of the cost of selling this option. We believe that the IMM should be able reasonably address this issue.

G. Forward Capacity – Conclusions

The Forward Capacity Market introduced by ISO-NE in 2008 has operated with no significant operational issues or procedural problems. The qualification processes and the auctions have occurred on schedule. Furthermore, the results of the auctions have been competitive, and sufficient capacity is planned to be in-service to satisfy the needs of New England through May 2016. The use of out-of-market payments by the ISO to retain existing resources has been reduced considerably. This has significantly improved incentives to capacity suppliers compared with the reliance on reliability agreements to retain existing capacity before June 2010.

The primary goal of deregulated wholesale markets is to facilitate market-based investment in new resources where the investment risks (and potential rewards) are borne by private firms rather than regulated investment where the risks are borne by captive consumers. However, most of the new investment in generation under FCM has been motivated by supplemental payments under the RFPs of the Connecticut DPUC. It is unlikely that substantial amounts of additional generation investment will occur until capacity clearing prices rise significantly. Therefore, it will be difficult to evaluate the FCM's effectiveness in facilitating efficient market-based investment until the current capacity surplus dissipates.

Another goal of these markets is to facilitate the orderly departure of existing resources that are no longer economic to remain in service. However, a large share of the capacity that has attempted to go out-of-service by de-listing has been unable to do so for reliability reasons. The failure of the FCM to allow the departure of these resources (as evidenced by the ISO rejection of de-list bids) has been due to three issues. First, the LSRs of local capacity zones did not originally reflect Transmission Security criteria so they were not set sufficiently high to satisfy the local requirements. This issue was addressed in time for FCA 4.

Second, local capacity zones are not modeled all of the time so the local requirements are not reflected in the market's selection of resources. This issue was partially resolved for FCA 7 by modeling NEMA, Connecticut, and Maine. The issue was resolved for other zones for FCA 8 by requiring that all eight zones be modeled in every auction.

Third, although ISO-NE modified the LSRs to be consistent with the Transmission Security criteria in the ISO's reliability delist bid review, local reliability issues can still justify the rejection of a de-list bid. This occurred in FCA 4.

Before the current surplus of capacity declines, it will be important to put in place market reforms that will enable the FCM to facilitate the efficient entry and exist of capacity resources. To this end we recommend the ISO:

- Replace the current vertical demand curve with a sloped demand curve that recognizes that excess capacity above the minimum planning reserve requirement provides additional benefits in the forms of increased reliability and lower energy and ancillary services prices.
- Evaluate the interaction of the rules for new suppliers that are related to the Rationing Election and the Capacity Commitment Period Election to determine whether they will promote efficient investment and FCM outcomes.

The ISO has also introduced a Performance Incentive proposal improve suppliers' incentives to be available during 30-minute reserve shortages. This proposal will likely achieve the ISO's objective of increasing the incentive for suppliers to be available in real time. We have identified a number of aspects of the proposal that should be further studied, which are discussed in this Section. Most importantly, we recommend that the ISO develop a PPR that does not substantially exceed the expected value of lost load during the shortages in which it will apply.

VIII. Competitive Assessment

This section evaluates the competitive performance of the ISO-NE wholesale markets in 2012. This type of assessment is particularly important for LMP markets. While LMP markets increase overall system efficiency, they can provide incentives for the localized exercise of market power in areas with limited generation resources or transmission capability. Most market power in wholesale electricity markets is dynamic, existing only in certain some areas and under particular conditions. The ISO has market power mitigation measures that are employed to prevent suppliers from exercising market power under these conditions. Although these measures have generally been effective, it is still important to evaluate the competitive structure and conduct in the ISO-NE markets because participants with market power may still have the incentive to exercise it at levels that would not warrant mitigation.

Based on the analysis presented in this section, we identify geographic areas and market conditions that present the greatest potential for market power abuse. We use a methodology for measuring and analyzing potential withholding that was developed in prior assessments of the competitive performance in the ISO-NE markets.[147]

We address five main areas in this section:

- Mechanisms by which sellers exercise market power in LMP markets;
- Structural market power indicators to assess competitive market conditions;
- Potential economic withholding;
- Potential physical withholding; and
- Market power mitigation.

A summary of our conclusions regarding the overall competitiveness of the wholesale market is included at the end of this section.

147 See, e.g., Section VIII, "2011 Assessment of Electricity Markets in New England", Potomac Economics.

A. Market Power and Withholding

Supplier market power can be defined as the ability to profitably raise prices above competitive levels. In electricity markets, this is generally done by economically or physically withholding generating resources. Economic withholding occurs when a resource is offered at prices above competitive levels to reduce its output or otherwise raise the market price. Physical withholding occurs when all or part of the output of a resource is not offered into the market when it is available and economic to operate. Physical withholding can be accomplished by "derating" a generating unit (i.e., reducing the unit's high operating limit).

While many suppliers can cause prices to increase by withholding, not every supplier can profit from doing so. The benefit from withholding is that the supplier will be able to sell into the market at a clearing price above the competitive level. However, the cost of this strategy is that the supplier will lose profits from the withheld output. Thus, a withholding strategy is only profitable when the price impact overwhelms the opportunity cost of lost sales for the supplier. The larger a supplier is relative to the market, the more likely it is that the supplier will have the ability and incentive to withhold resources to raise prices.

Other than the size of the market participant, there are several additional factors that affect whether a market participant has market power. First, if a supplier has already sold power in a forward market, then it will not be able to sell that power at an inflated clearing price in the spot market. Thus, forward power sales by large suppliers reduce their incentive to raise price in the spot market.[148, 149] Second, the incentive to withhold partly depends on the impact the withholding is expected to have on clearing prices. The nature of electricity markets is such that when demand is high, a given quantity of withholding has a larger price impact because the supply is substantially less elastic in the higher cost ranges. Thus, large suppliers are more likely to possess market power during high demand periods than at other times.

Third, in order to exercise market power, a large supplier must have sufficient information about the physical conditions of the power system and actions of other suppliers to know that the

148 When a supplier's forward power sales exceed the supplier's real-time production level, the supplier is a net buyer in the real-time spot market, thus, benefits from low rather than high prices.

149 However, some incentive still exists because spot prices will eventually affect prices in the forward market.

market may be vulnerable to withholding. Since no supplier has perfect information, the conditions that give rise to market power (e.g., transmission constraints and high demand) must be reasonably predictable. The next section defines market conditions where certain suppliers possess market power.

B. Structural Market Power Indicators

The first step in a market power analysis is to define the relevant market, which includes the definition of a relevant product and the relevant geographic market where the product is traded. Once the market definition is established, it is possible to assess conditions where one or more large suppliers could profitably raise price. This subsection of the report examines structural aspects of supply and demand affecting market power. We examine the behavior of market participants in later sections.

1. Defining the Relevant Market

Electricity is physically homogeneous, so each megawatt of electricity is interchangeable even though the characteristics of the generating units that produce the electricity vary substantially (*e.g.*, electricity from a coal-fired plant is substitutable with electricity from a nuclear power plant). Despite this physical homogeneity, the definition of the relevant product market is affected by the unique characteristics of electricity. For example, it is not generally economic to store electricity, so the market operator must continuously adjust suppliers' output to match demand in real time on a moment-to-moment basis. The lack of economic storage options limits inter-temporal substitution in spot electricity markets.

In defining the relevant product market, we must identify the generating capacity that can produce the relevant product. In this regard, we consider two categories of capacity: (i) online and fast-start capacity available for deployment in the real-time spot market, and (ii) offline and slower-starting capacity available for commitment in the next 24-hour timeframe. While only the former category is available to compete in the real-time spot market, both of these categories compete in the day-ahead market, making the day-ahead market less susceptible to market power. In general, forward markets are less vulnerable to market power because buyers can defer purchases if they expect prices to be lower in the spot market. The market is most vulnerable to the exercise of market power in the real-time spot market, when only online and fast-start capacity is available for deployment. The value of energy in all forward markets, including the day-ahead market, is derived from the expected value of energy in the real-time market. Hence, we define the relevant product as energy produced in real time for our analysis.

The second dimension of the market to be defined is the geographic area in which suppliers compete to sell the relevant product. In electricity markets, the relevant geographic market is generally defined by the transmission network constraints. Binding transmission constraints limit the extent to which power can flow between areas. When constraints are binding, a supplier within the constrained geographic area faces competition from fewer suppliers. There are a small number of geographic areas in New England that are recognized as being historically persistently constrained and, therefore, restricted at times from importing power from the rest of New England. When these areas are transmission-constrained, they constitute distinct geographic markets that must be analyzed separately. The following geographic markets are evaluated in this section:[150]

- All of New England;
- All of Connecticut;
- West Connecticut;
- Southwest Connecticut;
- Norwalk-Stamford, which is in Southwest Connecticut; and
- Boston.

This subsection analyzes the six geographic areas listed above using the following structural market power indicators:

- Supplier market shares;
- Herfindahl-Hirschman indices; and
- Pivotal supplier indices.

150 Lower SEMA was evaluated in prior reports, but is excluded from recent reports because the transmission constraints into the area was virtually eliminated since July 2009 when network upgrades were completed.

The findings from the structural market power analyses in this section are used to focus the analyses of potential economic and physical withholding in Subsections C and D.

2. Installed Capacity in Geographic Markets

This section provides a summary of supply resources and market shares in the geographic submarkets identified above. Each market can be served by a combination of native resources and imports. Native resources are limited by the physical characteristics of the generators in the area, while imports are limited by the capability of the transmission grid. The analysis in this subsection shows several categories of supply and import capability relative to the load in each of the six regions of interest.

We differentiate between different types of supply because some types cannot feasibly be withheld to exercise market power. For convenience, the table below shows different categories of supply and provides comments regarding the feasibility of withholding them.

Table 3: Withholding by Type of Resource

Type of Resource	Comment
Nuclear	Nuclear resources pose fewer market power concerns than other types of resources because they typically cannot be dispatched down substantially. This limits their owner's ability to withhold once a unit is online. They also generally have the lowest marginal production costs making them costly to withhold.
Hydroelectric	Hydroelectric resources that can vary their output (i.e., reservoir and pump storage units) may be able to withhold. Smaller "run-of-river" hydroelectric facilities are generally more limited in their ability to change output level.
Fossil-Fired	Fossil-fired units have relatively wide dispatch ranges and marginal production costs that are closer to the prevailing LMP. Hence, they are generally the easiest and least costly resources to withhold.

Figure 38 shows import capability and two categories of installed summer capability for each region: nuclear units and all other generators in 2011 and 2012.[151] These supplies are shown as

151 The import capability shown for each load pocket is the transfer capability during the peak load hour, reduced to account for local reserve requirements.

a percentage of 2011 and 2012 peak loads, respectively, although a substantial quantity of additional capacity is also necessary to maintain operating reserves in New England.[152] The figure shows that while imports from neighboring control areas can be used to satisfy 12 to 13 percent of the load in the New England area under peak conditions in 2011 and 2012, the five load pockets can serve larger shares of their peak load with imports. Norwalk-Stamford, which has the largest import capability relative to its size, was able to rely on imports to serve more than 100 percent of its load under peak conditions. This effectively eliminates it as an area of significant market power concern.

Figure 38: Supply Resources versus Summer Peak Load in Each Region
2011 – 2012

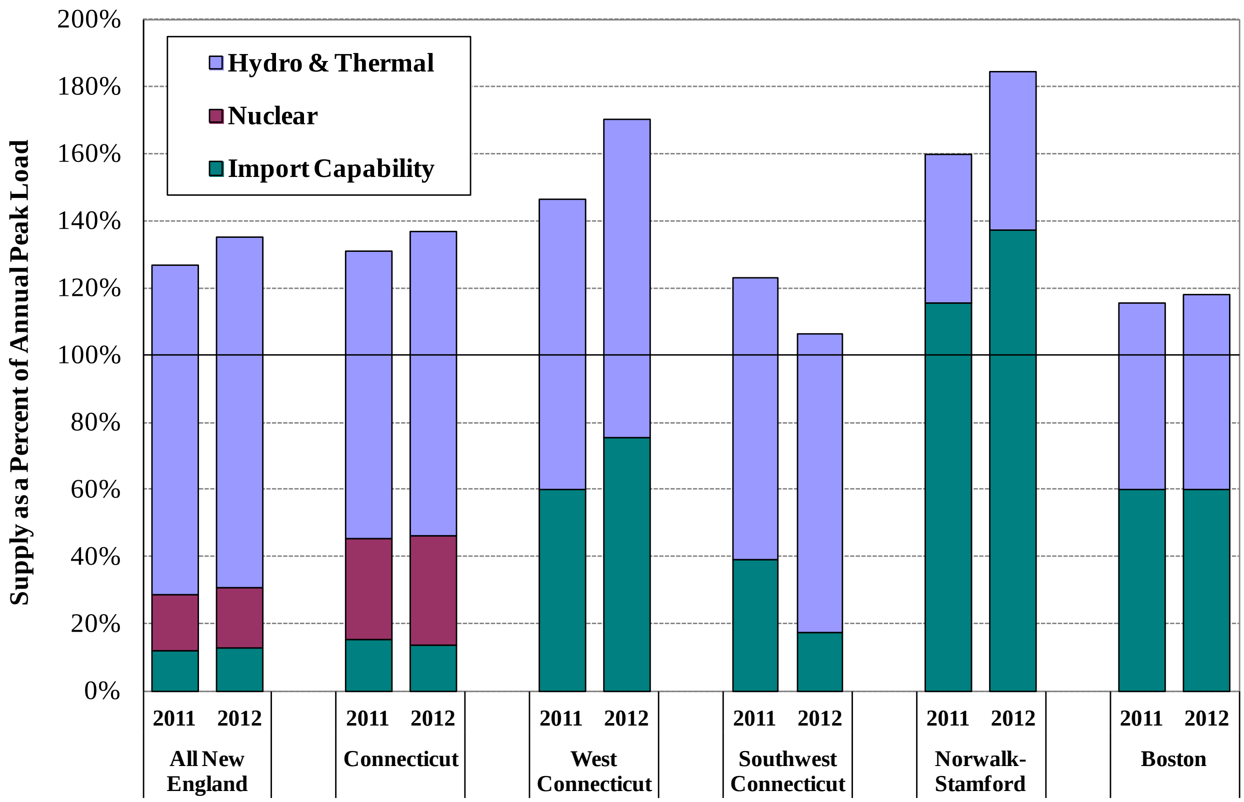


Figure 38 shows that the internal supply as a share of peak load increased modestly from 2011 to 2012 in all regions. This is because the summer peak load fell 7 percent from 2011 to 2012,

152 Roughly 2,050 to 2,100 MW of additional capacity was needed to maintain operating reserves in New England prior to July 23, 2012. After that, close to 2,500 MW of additional capacity was needed (due to the increase of 10-minute reserve requirement).

while there were very few changes to the supply of internal resources in these regions. The amount of import capability into each region did not change significantly from 2011 to 2012.[153] The variations in import capability were primarily attributable to the differences in network topology (e.g., line outages), generation patterns, and load patterns during the peak load hours in the two years.

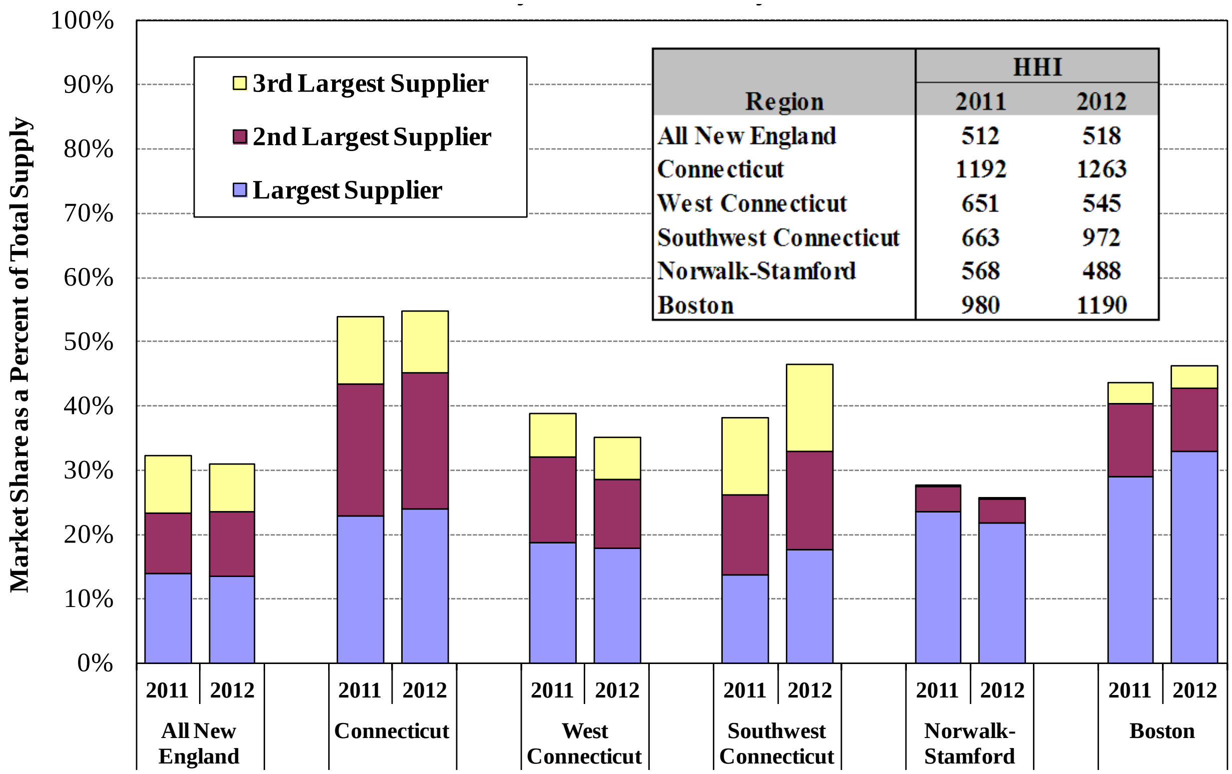
Therefore, supply conditions were generally consistent in most areas from 2011 to 2012. Figure 38 also shows the margin between peak load and the total available supply from imports and native resources. In 2012, the total supply exceeded peak load in each region, ranging from 6 percent in Southwest Connecticut to 84 percent in Norwalk Stamford. Areas with lower margins may be more susceptible to withholding than other areas.

3. Market Shares and Market Concentration

Market power is generally of greater concern in areas where capacity margins are small. However, the extent of market power also depends on the market shares of the largest suppliers. For each region, Figure 39 shows the market shares of the largest three suppliers in the annual peak load hours in 2011 (on July 22) and in 2012 (on July 17). The remainder of supply to each region comes from smaller suppliers and import capability. We also show the Herfindahl-Hirschman Index (HHI) for each region. The HHI is a standard measure of market concentration calculated by summing the square of each participant's market share. In our analysis, we assume imports are highly competitive by treating the market share of imports as zero in the HHI calculation. For example, in a market with two suppliers and import capability, all of equal size, the HHI would be close to 2200 = $[(33\%)^2 + (33\%)^2 + (0\%)^2]$. This assumption tends to understate the true level of concentration, because, in reality, the market outside of the area is not perfectly competitive, and because suppliers inside the area may be affiliated with resources in the market outside of the area.

153 The transmission system in New England has evolved significantly over the past several years, particularly from 2006 to 2009 when several major transmission upgrades were completed in the historically constrained areas such as Boston, Connecticut, and Lower SEMA. These upgrades significantly improved the transmission system infrastructure and increased the transfer capability into affected regions.

Figure 39: Installed Capacity Market Shares for Three Largest Suppliers
July 22, 2011 and July 17, 2012



Region	HHI 2011	HHI 2012
All New England	512	518
Connecticut	1192	1263
West Connecticut	651	545
Southwest Connecticut	663	972
Norwalk-Stamford	568	488
Boston	980	1190

Figure 39 indicates a substantial variation in market concentration across New England. In all New England, the largest supplier had a 13 percent market share in 2012. In the load pockets, the largest suppliers had market shares ranging from 18 percent in Southwest Connecticut and West Connecticut to 33 percent in Boston in 2012. Likewise, there is variation in the number of suppliers that have significant market shares. For instance, Norwalk-Stamford had just two native suppliers with very different market shares in 2012, while Southwest Connecticut had three native suppliers with more comparable market shares.

The figure shows that market shares of the largest three suppliers in New England changed modestly from 2011 to 2012 due to changes in the ownership of several generating units. The largest supplier lost about 750 MW of capacity over the year by retiring two units (160 MW) in early 2012 and selling two remaining units in the same plant (590 MW) to another participant. Likewise, the portfolio of the third largest supplier decreased about 600 MW when it sold a combined cycle unit to another firm. However, the second largest supplier increased its portfolio in New England by more than 200 MW by acquiring assets from another company through a merger. Despite these notable changes, the top three suppliers remained the three largest suppliers in New England in 2012.[154]

There were also some changes in the ownership of assets that affected market shares in the local areas. Most notably, one firm acquired 550 MW of combined cycle capacity from another firm when it exited the market in early 2012, making the acquiring firm one of the top three suppliers in West Connecticut. Otherwise, there were very few changes to the supply of internal resources in each region, and the import capability into each region remained similar. The modest differences in market shares on the peak load day between 2011 and 2012 in some of the local areas were generally attributable to the variations in the import capability associated with differences in network topology, generation patterns, and load patterns on the two days.

The HHI figures suggest that no areas in New England were highly concentrated in 2012.[155] The HHI for Norwalk-Stamford is 488, which is relatively low for most product markets. This is counter-intuitive since there are only two major suppliers in the area. However, because its load can be entirely served by imports, the need for local suppliers is very limited. Of the remaining areas, Connecticut and Boston had the highest HHI statistics in 2012, with 1263 and 1190, respectively.

While HHI statistics can be instructive in generally indicating the concentration of the market, they alone do not allow one to draw reliable conclusions regarding potential market power in wholesale electricity markets due to the special nature of the electricity markets. In particular, they do not consider demand conditions, load obligations, or the heterogeneous effects of generation on transmission constraints based on their location. In the next subsection, we evaluate the potential for market power using a pivotal supplier analysis, which addresses the shortcomings of concentration analyses.

---

154 The merger of NRG and Genon also made the new company one of the top three suppliers in New England. This occurred at the end of 2012, so it was not reflected in the analyses in this report for most of 2012.

155 The antitrust agencies and the FERC consider markets with HHI levels above 1800 as highly concentrated for purposes of evaluating the competitive effects of mergers.

4. Pivotal Supplier Analysis

While HHI statistics can provide reliable competitive inferences for many types of products, this is not generally the case in electricity spot markets.[156] The HHI's usefulness is limited by the fact that it reflects only the supply-side, ignoring demand-side factors that affect the competitiveness of the market. The most important demand-side factor is the level of load relative to available supply-side resources. Since electricity cannot be stored economically in large volumes, production needs to match demand in real time on a moment-to-moment basis. When demand rises, an increasing quantity of generation is utilized to satisfy the demand, leaving less supply that can respond by increasing output if a large supplier withholds resources. Hence, markets with higher resource margins tend to be more competitive, which is not recognized by the HHI statistics.

A more reliable means to evaluate the competitiveness of spot electricity markets and recognize the dynamic nature of market power in these markets is to identify when one or more suppliers are "pivotal". A supplier is pivotal when the capacity of some of its resources is needed to meet demand and reserve requirements in the market. A pivotal supplier has the ability to unilaterally raise the spot market prices to very high levels by offering its energy at a very high price level. Hence, the market may be subject to substantial market power abuse when one or more suppliers are pivotal and have the incentive to take advantage of their position to raise prices. The Federal Energy Regulatory Commission has adopted a form of pivotal supplier test as an initial screen for market power in granting market-based rates.[157] This section of the report identifies the frequency with which one or more suppliers were pivotal in various areas within New England.

Even small suppliers can be pivotal for brief periods. For example, all suppliers are pivotal during periods of shortage. This does not mean that all suppliers should be deemed to have market power. As described above, suppliers must have both the *ability* and *incentive* to raise

---

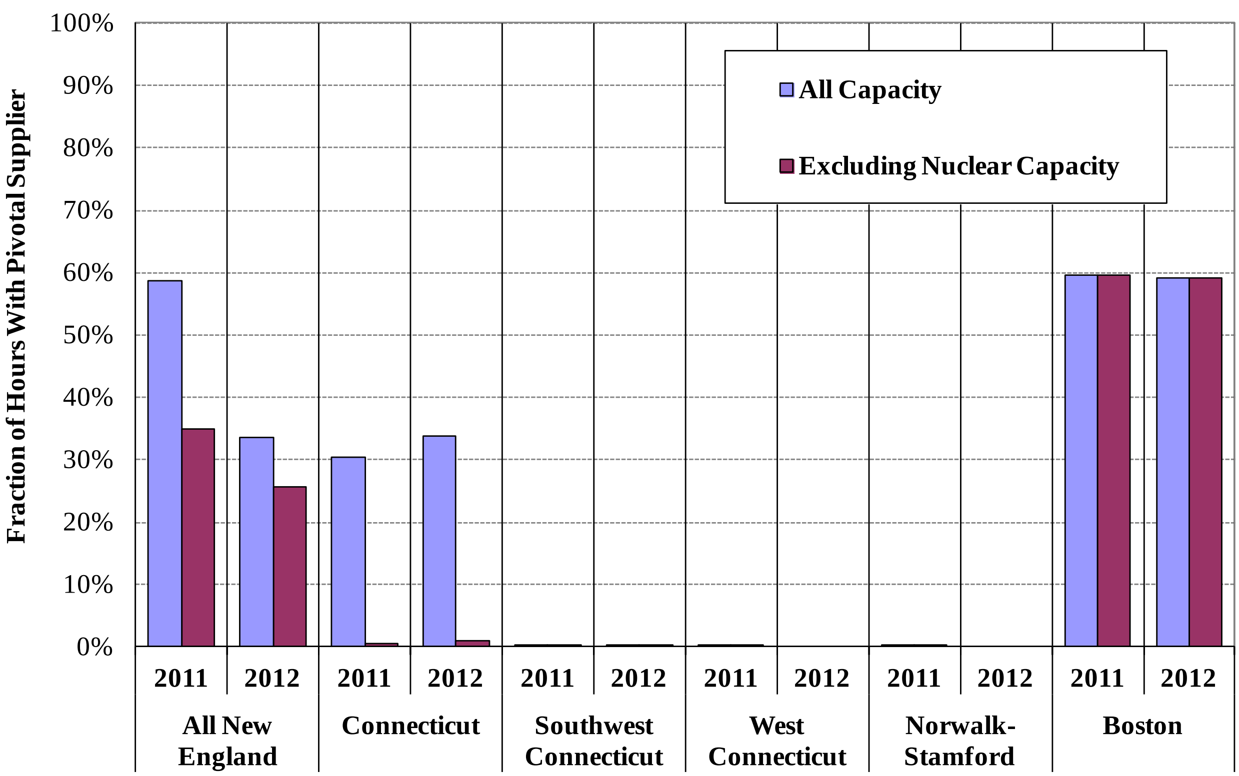
156 The DOJ and FTC evaluate the change in HHI as part of their merger analysis. However, this is only a preliminary analysis that would typically be followed by a more rigorous simulation of the likely price effects of the merger. Also, the HHI analysis employed by the antitrust agencies is not intended to determine whether a supplier has market power.

157 The FERC test is called the "Supply Margin Assessment". For a description, see: Order On Rehearing And Modifying Interim Generation Market Power Analysis And Mitigation Policy, 107 FERC ¶ 61,018, April 14, 2004.

prices to have market power. For a supplier to have the ability to substantially raise real-time energy prices, it must be able to foresee that it will likely be pivotal. In general, the more frequently a supplier is pivotal, the easier it will be for it to foresee circumstances when it can raise the clearing price.

To identify the areas where market power is a potential concern most frequently, Figure 40 shows the portion of hours where at least one supplier was pivotal in each region during 2011 and 2012. [158] The figure also shows the impact of excluding nuclear units. As discussed above, owners of nuclear units are less likely to engage in economic or physical withholding.

Figure 40: Frequency of One or More Pivotal Suppliers by Type of Withheld Capacity
2011 – 2012



158 The IMM also does a pivotal supplier assessment in its Annual Market Report. Compared with our assessment, the IMM's pivotal supplier assessment produces lower frequencies of one or more pivotal supplier due to differences in the underlying objectives of each assessment. Our analysis is a real-time test, focusing on capacity that is online (i.e., committed) or capable of starting within 30 minutes (i.e., available offline quick-start) in the real-time market, since this indicates when a single supplier could cause a shortage by withholding in the real-time market. The IMM's assessment includes capacity that is available to the day-ahead market, including longer lead time units, since this assessment is used in the execution of the market power mitigation measures for the day-ahead market,

Including all categories of capacity, the pivotal supplier analysis raises potential concerns regarding three of the six areas shown in Figure 40. The areas that do not raise potential concerns are Norwalk-Stamford, Southwest Connecticut, and West Connecticut, where imports typically serve a large share of load and the ownership of internal capacity is much less concentrated than the other load pockets.

The figure shows that potential local market power concerns were most acute in Boston, where one supplier owns nearly 70 percent of the internal capacity in 2012 and was pivotal in nearly 60 percent of hours. In addition, none of the largest supplier's capacity in Boston was nuclear capacity.

Although Connecticut had a pivotal supplier in 34 percent of hours in 2012 and 30 percent of hours in 2011, the largest supplier in Connecticut owns only nuclear capacity. In order to exercise market power, the largest supplier would need to withhold from non-nuclear resources in order to raise the clearing prices paid for its nuclear production. Therefore, it is appropriate to exclude the nuclear capacity from the pivotal supplier frequency for Connecticut. This leaves very few hours when a supplier was pivotal in Connecticut in the past two years.

For the entirety of New England, excluding nuclear capacity from the pivotal supplier analysis would substantially reduce the pivotal frequency (from 34 percent to 26 percent of hours in 2012 and from 59 percent to 35 percent of hours in 2011). However, the rationale for excluding nuclear capacity from the analysis does not apply to the largest suppliers in New England. These suppliers have large portfolios with a combination of nuclear and non-nuclear capacity, and while they are not likely to physically withhold their nuclear capacity from the market, their nuclear capacity would earn more revenue if they withheld their non-nuclear capacity. Accordingly, New England as a whole warrants further review.

In all of New England, the pivotal frequency declined from 59 percent in 2011 to 34 percent in 2012. The decrease was attributable to at least two factors:

- First, the size of some large suppliers decreased during 2012. As discussed earlier, the portfolio of the largest supplier in New England decreased about 750 MW over the year and the third largest supplier lost about 600 MW as well.

- Second, a significant portion in the largest supplier's portfolio is coal-fired capacity, which was economically committed less frequently in 2012 than in prior years due to lower natural gas prices.

These changes led top suppliers to have lower shares of real-time dispatchable capacity (i.e., online and offline fast-start capacity) and consequently to be pivotal less frequently in 2012.

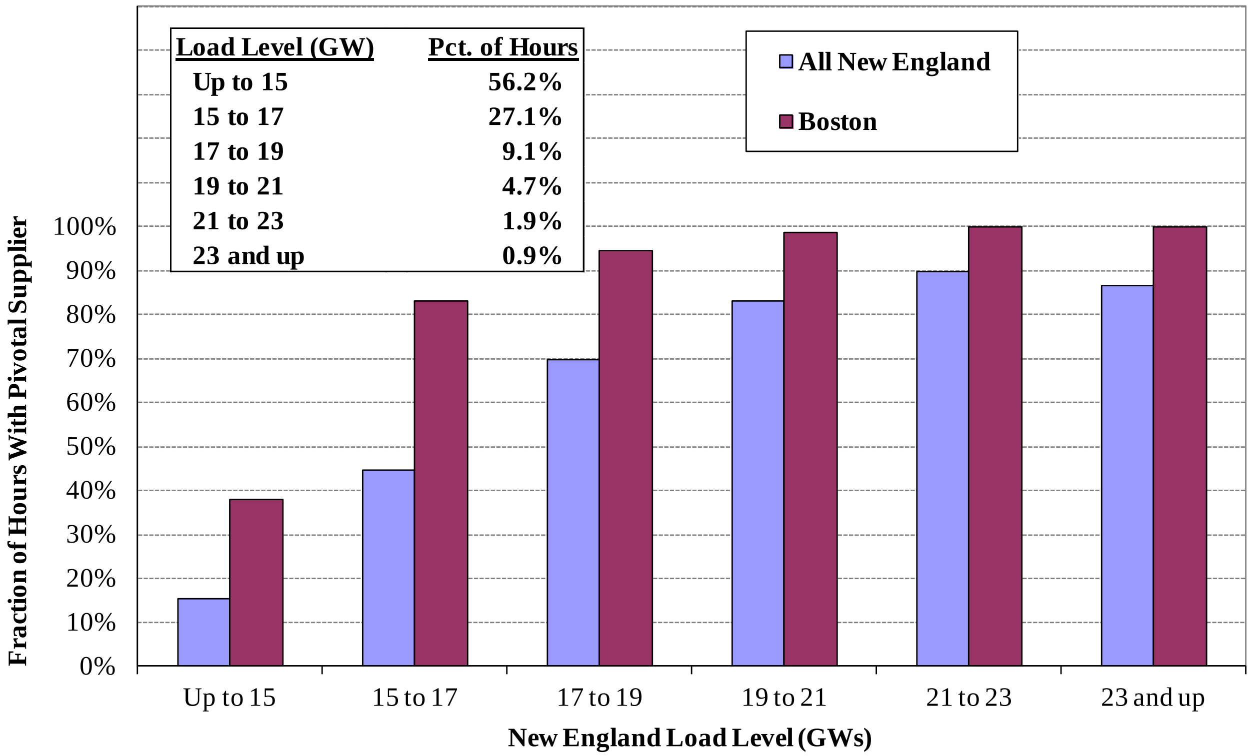
The pivotal supplier summary in Figure 40 indicates the greatest potential for market power in Boston. A close examination is also warranted for all of New England, while Connecticut raises less concern. Each area had a single supplier that was most likely to have market power. Accordingly, Sections C and D closely examine the behavior of the largest single supplier by geographic market.

As described above, market power tends to be more prevalent as the level of demand rises. In order to strategically withhold, a dominant supplier must be able to reasonably foresee its opportunities to raise prices. Since load levels are relatively predictable, a supplier with market power could focus its withholding strategy on periods of high demand. To assess when withholding is most likely to be profitable, Figure 41 shows the fraction of hours when a supplier is pivotal at various load levels.

The bars in each load range show the fraction of hours when a supplier was pivotal in All of New England and Boston. West Connecticut, Southwest Connecticut, and Norwalk-Stamford are not shown because there were very few instances of a supplier being pivotal during 2012. Connecticut is not shown because the largest pivotal supplier had exclusively nuclear capacity, which is not expected to provide that supplier with an incentive to withhold.

A supplier in Boston was pivotal in at least 83 percent of hours when the load exceeded 15 GW in New England. In all of New England, the largest supplier was pivotal in 45 percent of the hours when load exceeded 15 GW. The pivotal frequency fell to 38 percent in Boston and 16 percent in all of New England during hours when load was below 15 GW in New England.

Figure 41: Frequency of One or More Pivotal Suppliers by Load Level
2012



In 2012, based on the pivotal supplier analysis in this subsection, market power was most likely to be a concern in Boston and all of New England when load exceeded 15 GW. The pivotal supplier results are conservative for "All of New England" because the analysis assumed that imports would not change if the largest supplier were to withhold. In reality, there would likely be some increase in imports. The following sections examine the behavior of pivotal suppliers under various load conditions to assess whether the behavior has been consistent with competitive expectations.

C. Economic Withholding

Economic withholding occurs when a supplier raises its offer prices substantially above competitive levels to raise the market price. Therefore, an analysis of economic withholding requires a comparison of actual offers to competitive offers.

Suppliers lacking market power maximize profits by offering resources at marginal costs. A generator's marginal cost is the incremental cost of producing additional output, including inter-temporal opportunity costs, incremental risks associated with unit outages, fuel, additional O&M, and other incremental costs attributable to the incremental output. For most fossil-fuel resources, marginal costs are closely approximated by their variable production costs (primarily fuel inputs, labor, and variable operating and maintenance costs). However, at high output levels or after having run long periods without routine maintenance, outage risks and expected increases in O&M costs can create substantial additional incremental costs. Generating resources with energy limitations, such as hydroelectric units or fossil-fuel units with output restrictions as a result of environmental considerations, must forego revenue in a future period when they produce in the current period. These units incur an inter-temporal opportunity cost associated with producing that can cause their marginal costs to be much larger than their variable production costs.

Establishing a proxy for units' marginal costs as a competitive benchmark is a key component of this analysis. This is necessary to determine the quantity of output that is potentially economically withheld. The ISO's Internal Market Monitor ("IMM") calculates generator cost reference levels pursuant to Appendix A of Section III of the ISO's Tariff. These reference levels are used as part of the market power mitigation measures and are intended to reflect the competitive offer price for a resource. The IMM has provided us with cost reference levels, which we can use as a competitive benchmark in our analysis of economic withholding.

1. Measuring Economic Withholding

We measure economic withholding by estimating an output gap for units that fail a conduct test for their start-up, no-load, and incremental energy offer parameters indicating that they are submitting offers in excess of competitive levels. The output gap is the difference between the unit's capacity that is economic at the prevailing clearing price and the amount that is actually produced by the unit. In essence, the output gap shows the quantity of generation that is withheld from the market as a result of having submitted offers above competitive levels. Therefore, the output gap for any unit would generally equal:

$Q_i^{econ} - Q_i^{prod}$ when greater than zero, where:

Q_i^{econ} = Economic level of output for unit i; and

Q_i^{prod} = Actual production of unit i.

To estimate Q_i^{econ}, the economic level of output for a particular unit, it is necessary to evaluate all parts of the unit's three-part reference level: start-up cost reference, no-load cost reference, and incremental energy cost reference. These costs jointly determine whether a unit would have been economic at the clearing price for at least the unit's minimum run time. We employ a three-stage process to determine the economic output level for a unit in a particular hour. In the first step, we examine whether the unit would have been economic *for commitment* on that day if it had offered at its marginal costs – i.e., whether the unit would have recovered its actual start-up, no-load, and incremental costs running at the dispatch point dictated by the prevailing LMP (constrained by its EcoMin and EcoMax) for its minimum run time. If a unit was economic for commitment, we then identify the set of contiguous hours during which the unit was economic to have online. Finally, we determine the economic level of incremental output in hours when the unit was economic to run. In all three steps, the marginal costs assumed for the generator are the reference levels for the unit used in the ISO's mitigation measures plus a threshold.

In hours when the unit was not economic to run and on days when the unit was not economic for commitment, the economic level of output was considered to be zero. To reflect the timeframe in which commitment decisions are actually made, this assessment is based on real-time market outcomes for fast-start units and day-ahead market outcomes for slower-starting units.

Q_i^{prod} is the actual observed production of the unit. The difference between Q_i^{econ} and Q_i^{prod} represents how much the unit fell short of its economic production level. However, some adjustments are necessary to estimate the actual output gap because some units are dispatched at levels lower than their three-part offers would indicate. This can be due either to transmission constraints, reserve considerations, or changes in market conditions between the time when unit commitment is performed and real time. Therefore, we adjust Q_i^{prod} upward to reflect three-part offers that would have made a unit economic to run, even though the unit may not have been fully dispatched. For example, if the ISO manually reduces the dispatch of an economic unit, the reduction in output is excluded from the output gap. Hence, the output gap formula we use is:

$Q_i^{econ} - \max(Q_i^{prod}, Q_i^{offer})$ when greater than zero, where:

Q_i^{offer} = offer output level of i.

By using the greater of actual production or the output level offered at the clearing price, portions of units that are constrained by ramp limitations are excluded from the output gap. In addition, portions of resources that are offered above marginal costs due to a forward reserve market obligation are not included in the output gap.

It is important to recognize that the output gap tends to overstate the amount of potential economic withholding because some of the offers that are included in the output gap reflect legitimate responses by the unit's owner to operating conditions, risks, or uncertainties. For example, some hydroelectric units are able to produce energy for a limited number of hours before running out of water. Under competitive conditions, the owners of such units have incentives to produce energy during the highest priced periods of the day. They attempt to do this by raising their offer prices so their units will be dispatched only during the highest-priced periods of the day. However, the owners of such units submit offers prior to 6 pm on the previous day based on their expectations of market conditions. If real-time prices are lower than expected, it may lead the unit to have an output gap. Hence, output gap is not necessarily evidence of withholding, but it is a useful indicator of potential withholding. We generally seek to identify trends in the output gap that would indicate significant attempts to exercise market power.

We have observed that some units that expect to be committed for local reliability and receive NCPC payments also produce above average output gap. One explanation is that these units raise their offers in expectation of receiving higher NCPC payments and are not dispatched as a result. Such instances are flagged as output gap, even though the suppliers are not withholding in an effort to raise LMPs.

In this section we evaluate the output gap results relative to various market conditions and participant characteristics. The objective is to determine whether the output gap increases when those factors prevail that can create the ability and incentive for a pivotal supplier to exercise market power. This allows us to test whether the output gap varies in a manner consistent with attempts to exercise market power. Based on the pivotal supplier analysis from the previous subsection, the level of market demand is a key factor in determining when a dominant supplier is most likely to possess market power in some geographic market. In this section, we examine output gap results by load level separately in Boston, Connecticut, and all of New England. Our analyses apply the most stringent thresholds that are used in the current market power mitigation measures, which are: a) the lower of $25 per MWh or 50 percent over the reference level for the energy offer; and b) 25 percent over the reference level for the no-load and startup offers.

2. Output Gap in Boston

Boston is a large net-importing region, which can cause transmission interfaces into the region to bind periodically. When this occurs, competition can be limited so it is particularly important to evaluate the conduct of its suppliers. Furthermore, the pivotal supplier analysis raises concerns regarding the potential exercise of market power in Boston where one supplier owns the majority of capacity.

Figure 42 shows output gap results for Boston by load level. Output gap statistics are shown as the percentage of the portfolio size for the largest supplier compared with all other suppliers in the area. Based on the pivotal supplier analysis in the previous subsection, the largest supplier can expect that its capacity will be pivotal in most hours when load exceeds 15 GW.

Figure 42: Average Output Gap by Load Level and Type of Supplier
Boston, 2012

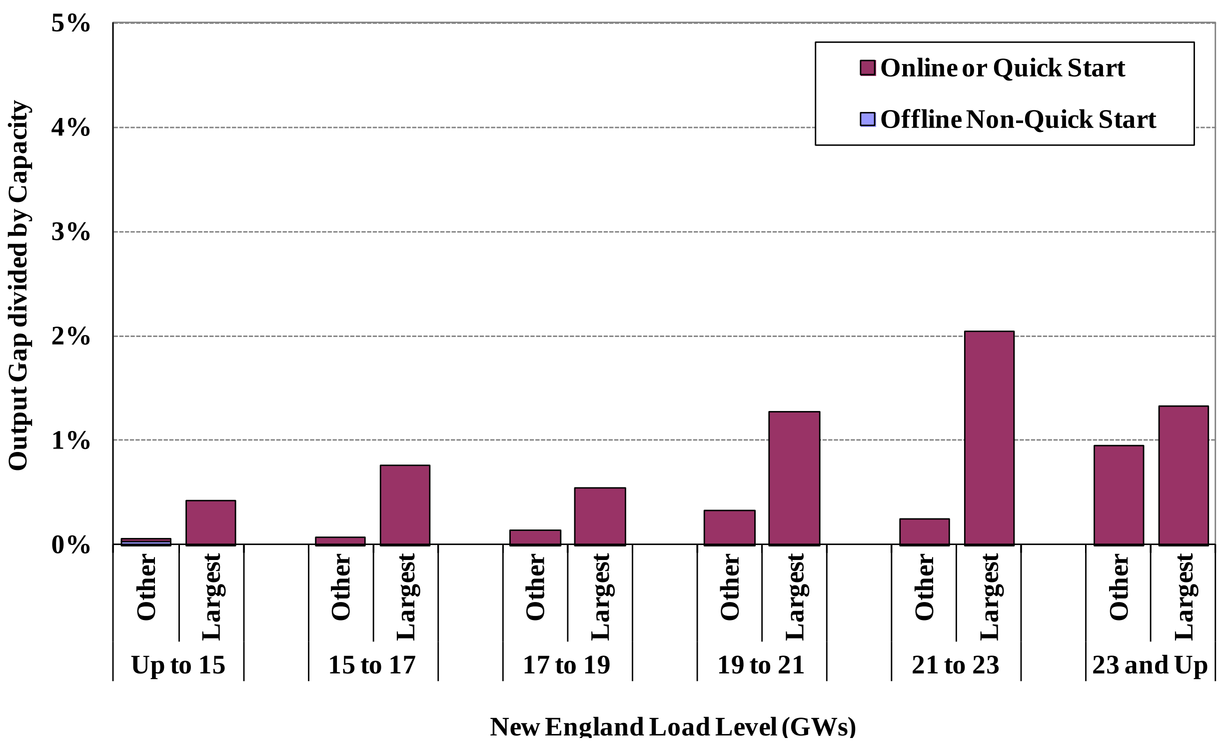
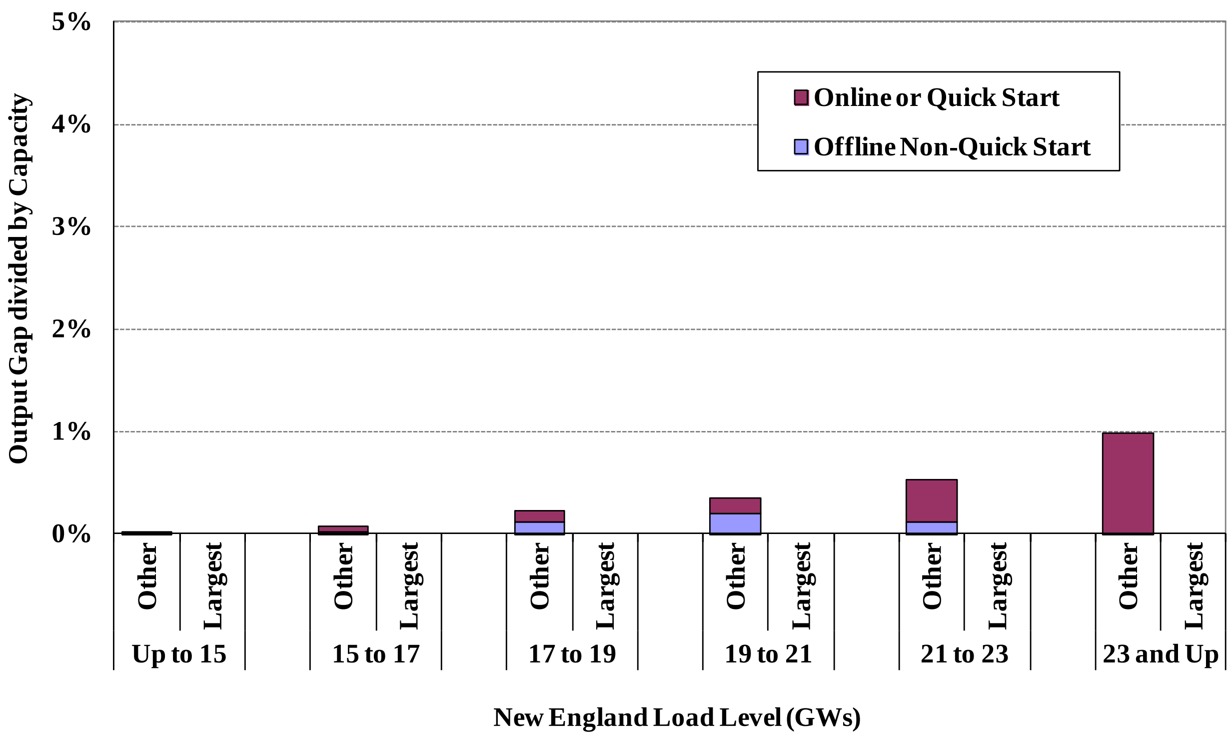


Figure 42 shows that the overall amount of output gap for the largest supplier in Boston was small as a share of its total capacity in 2012, ranging from 0.4 percent when load was below 15 GW to 2 percent when load was between 21 and 23 GW. Although the output gap for the largest supplier increased when load rose, the amount was still low. It averaged less than 40 MW and was associated with the duct-firing output range of two combined cycle units. In addition, the capacity was typically scheduled for operating reserves, which diminishes the potential effect on prices. This capacity would likely have been mitigated if significant congestion did occur during these periods. Therefore, these results do not raise significant competitive concerns.

3. Output Gap in Connecticut

In this subsection, we examine potential economic withholding in Connecticut. Historically, Connecticut has been import-constrained, although the pivotal supplier analysis does not raise significant concerns about the potential exercise of market power in 2012 in Connecticut. Figure 43 shows output gap results for Connecticut by load level. Output gap statistics are shown for the largest supplier compared with all other suppliers in the area.

Figure 43: Average Output Gap by Load Level and Type of Supplier
Connecticut, 2012

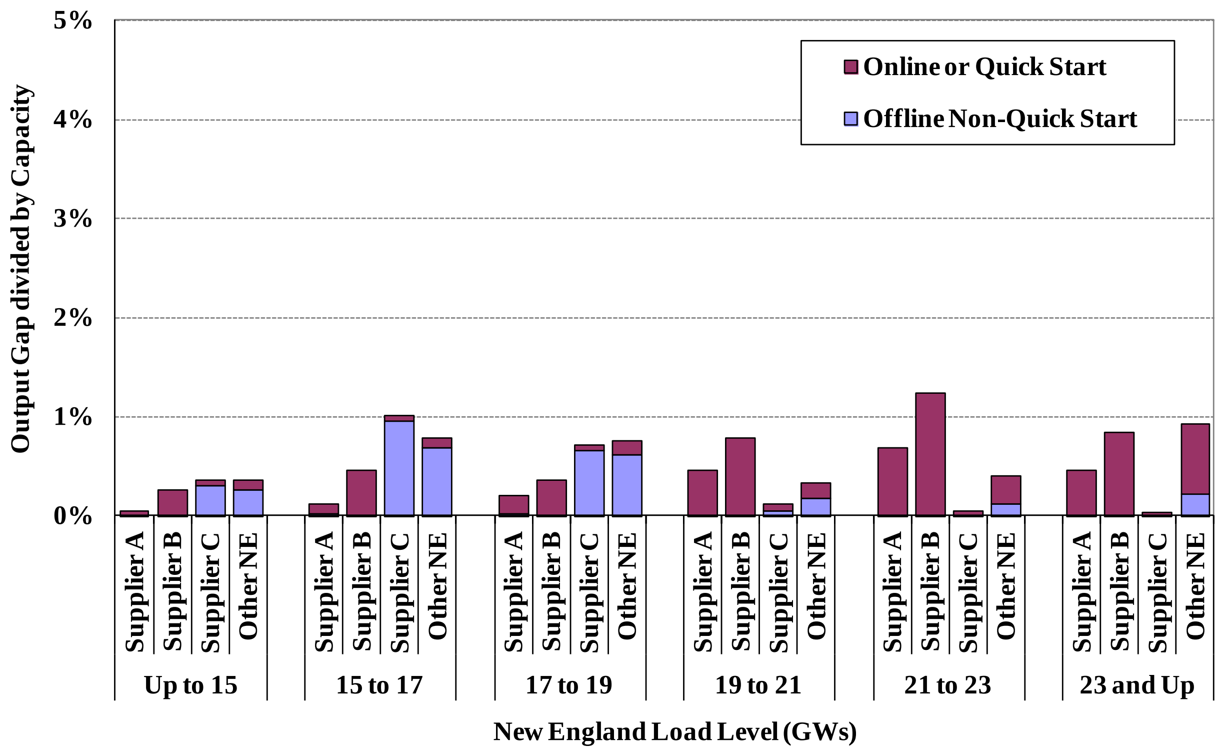


The pivotal supplier analysis indicated that the largest supplier in Connecticut was pivotal in about 34 percent of all hours when all capacity is considered, although the largest supplier owns exclusively nuclear capacity and had no output gap in 2012. Figure 43 also shows that the total output gap of all other suppliers was very low (< 1 percent) relative to the total capacity in Connecticut. Given these amounts, the results do not raise concerns regarding economic withholding in Connecticut.

4. Output Gap in All New England

Figure 44 summarizes output gap results for all of New England by load level for four categories of supply. Supplier A had the largest portfolio in New England and was pivotal in approximately 34 percent of the hours during 2012. Suppliers B and C are the second and third largest suppliers in New England and were pivotal during 22 percent and 12 percent of the hours, respectively. All other suppliers are shown as a group for reference.

Figure 44: Average Output Gap by Load Level and Type of Supplier
All New England, 2012



The figure shows that the region-wide output gap was generally low for each of the four categories of supply. Suppliers A, B, and C exhibited small output gap levels (< 1 percent of their portfolio sizes) under most load conditions. Supplier B (also the largest supplier in Boston) exhibited an output gap slightly more than one percent of its portfolio when load was between 21 and 23 GW. However, this increased quantity does not raise significant concerns for the reasons discussed in Subsection C.2. Additionally, it is a positive indication that the output gap levels for the three largest suppliers were comparable to the output gap levels of all other suppliers, which serve as a benchmark for conduct of smaller suppliers that are much less likely to have market power. Hence, these output levels are likely to reflect only measurement error in the output gap metric.

Because these output gap levels are very low and the largest suppliers' output gap amounts are comparable to the levels for other suppliers (which are not likely to have market power), especially at high load levels (when withholding is most likely to occur and be profitable), economic withholding was not a significant concern in New England in 2012.

D. Physical Withholding

This section of the report examines declarations of forced outages and other non-planned deratings to determine if there is any evidence that the suppliers are exercising market power. In this analysis, we evaluate the three geographic markets examined in the output gap analysis above: Boston, Connecticut, and all of New England.

In each market, we examine forced outages and other deratings by load level. The "Other Derate" category includes any reduction in the hourly capability of a unit from its maximum seasonal capability that is not logged as a forced outage or a planned outage. These deratings can be the result of ambient temperature changes or other factors that affect the maximum capability of a unit.

1. Potential Physical Withholding in Boston

Figure 45 shows declarations of forced outages and other non-planned deratings in Boston by load level. Based on the pivotal supplier analysis, the capacity of the largest supplier can be expected to be pivotal in most hours when New England load exceeds 15 GW. We compare these statistics for the largest supplier to all other suppliers in the area.

The figure shows the largest supplier's physical deratings as a percentage of its portfolio. The rate of other non-planned outages ('Other Derate' Category) was high at low load levels in 2012, especially when load was less than 15 GW. This was primarily driven by units that were frequently online in special operating modes (where a portion of the capacity is not available) in early morning hours. Under low load conditions, this operating practice does not raise competitive concerns and is consistent with competitive conduct.

Figure 45: Forced Outages and Deratings by Load Level and Supplier
Boston, 2012

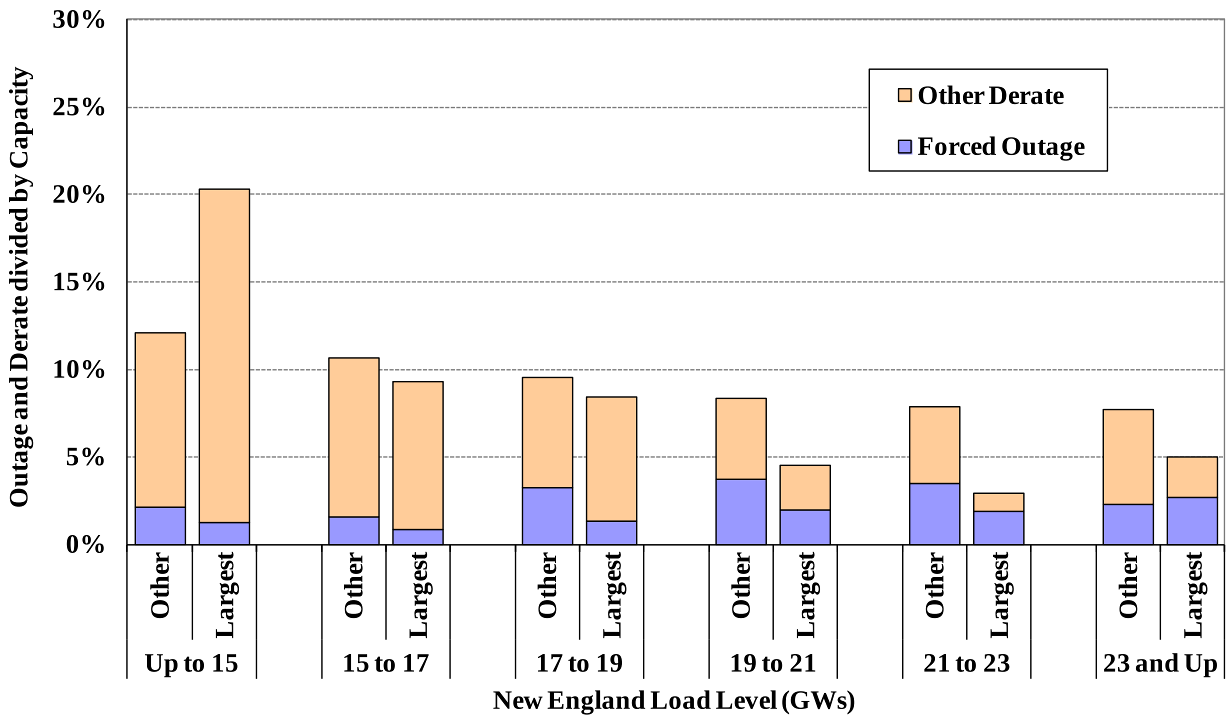


Figure 45 shows a pattern of deratings and outages consistent with expectations in a competitive market. Although levels of outages and deratings for the largest supplier were high at low load levels, they were lower than other suppliers when load exceeded 15 GW (when withholding is most likely to be profitable). Furthermore, the largest supplier showed a relatively low level of outages and deratings as load increased to the highest load levels. Even though running units more intensely under peak demand conditions increases the probability of an outage, the results shown in the figure suggest that the largest supplier increased the availability of its capacity during periods of high load when capacity was most valuable to the market. Overall, the outage and deratings results for Boston do not raise concerns of strategic withholding.

2. Potential Physical Withholding in Connecticut

Figure 46 summarizes declarations of forced outages and other deratings in Connecticut by load level in 2012. The figure shows these statistics for the largest supplier in the area and compares them with statistics for other suppliers.

Figure 46: Forced Outages and Deratings by Load Level and Supplier
Connecticut, 2012

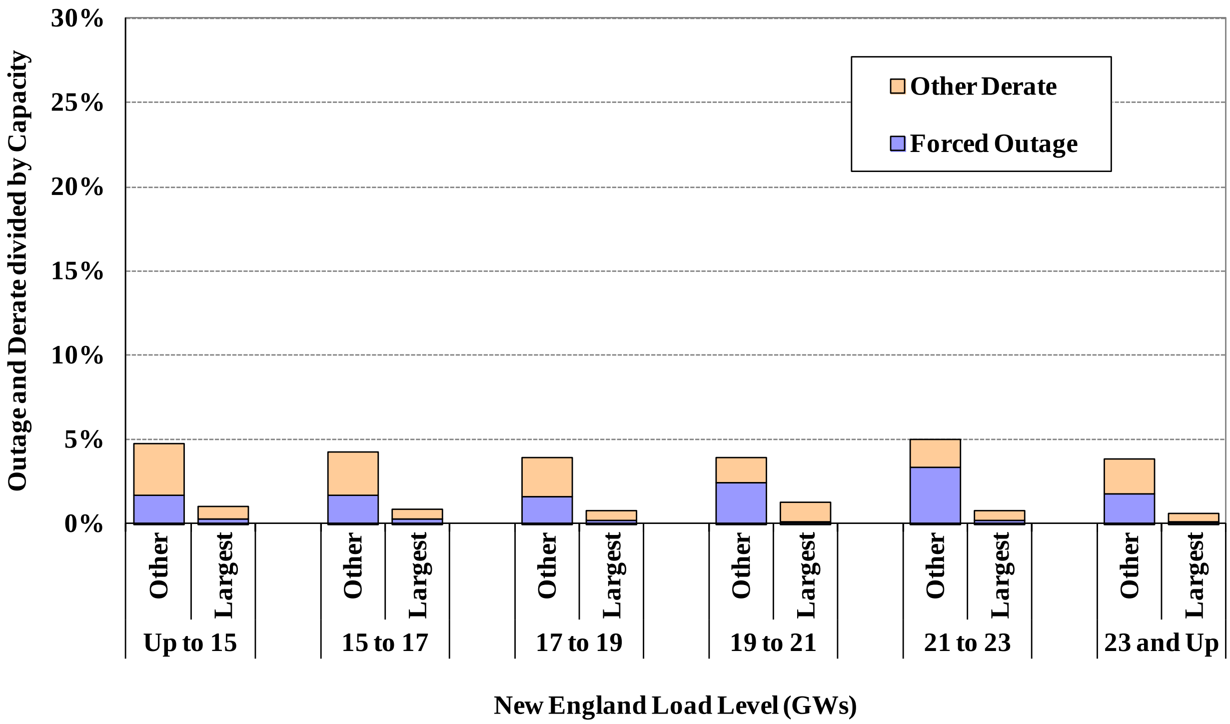


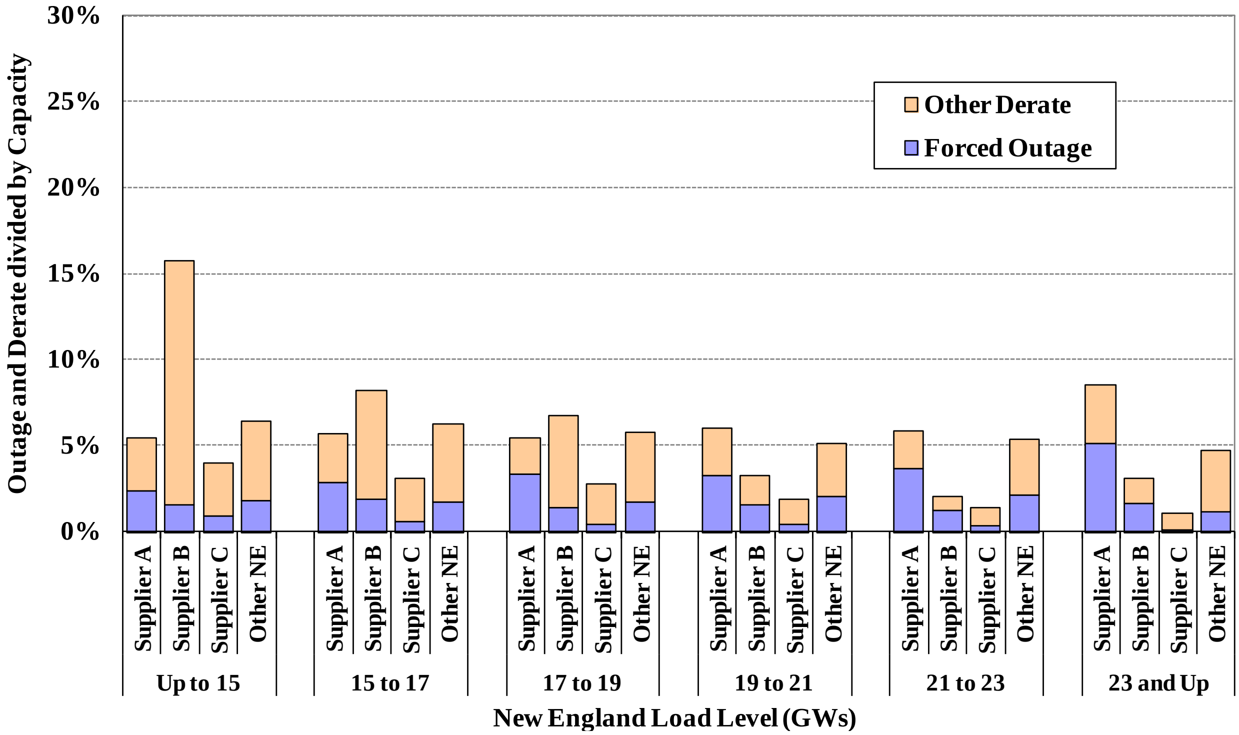
Figure 46 shows that the physical derating and forced outage quantities for the largest supplier in Connecticut were very low under all load conditions in 2012. Furthermore, the levels of deratings and outages for the largest supplier were much lower than the levels of all other suppliers, which serve as a benchmark since small suppliers are much less likely to have market power. Hence, these deratings and outages do not raise concerns about physical withholding in Connecticut in 2012.

3. Potential Physical Withholding in All New England

Having analyzed the two major constrained areas in New England, Figure 47 summarizes the physical withholding analysis for all of New England by load level in 2012. The results of this analysis are shown for four groups of supply. Supplier A had the largest portfolio in New

England and was pivotal in approximately 34 percent of the hours during 2012. Suppliers B and C are the second and third largest suppliers in New England and were pivotal during about 22 percent and 12 percent of the hours, respectively. All other suppliers are shown as a group for comparison purposes.

Figure 47: Forced Outages and Deratings by Load Level and Supplier
All New England, 2012



Supplier A and Supplier C exhibited rates of forced outages and other non-planned deratings that were moderate under all load conditions. Supplier B exhibited rates of forced outages and other non-planned deratings that were comparable to other New England suppliers when loads exceeded 17 GW, but were substantially higher at lower load levels, especially when load was less than 15 GW. Supplier B is also the largest supplier in Boston. The pattern for Supplier B was explained earlier by factors that do not raise competitive concerns.

As a group, the other New England suppliers' derating levels generally decreased as load levels increased. These patterns generally suggest that New England suppliers increased the availability of their resources under peak demand conditions. The increased availability is particularly notable when we consider the effects of high ambient temperatures on thermal generators. Naturally, ambient temperature restrictions on thermal units vary along with load and are difficult to distinguish from physical withholding through a review of market data. It is beyond the scope of this report to determine whether individual outages and other deratings were warranted. However, the overall quantity of capacity subject to the deratings was consistent with expectations for a workably competitive market, so we do not find evidence to suggest that these deratings constituted an exercise of market power.

E. Market Power Mitigation

Mitigation measures are intended to mitigate abuses of market power while minimizing interference with the market when the market is workably competitive. The ISO-NE applies a conduct-impact test that can result in mitigation of a participant's supply offers (i.e., incremental energy offers, start-up and no-load offers). The mitigation measures are only imposed when suppliers' conduct exceeds well-defined conduct thresholds and when the effect of that conduct on market outcomes exceeds well-defined market impact thresholds. This framework prevents mitigation when it is not necessary to address market power, while allowing high prices during legitimate periods of shortage. The mitigation measures are designed to allow prices to rise efficiently to reflect legitimate supply shortages while effectively mitigating inflated prices associated with artificial shortages that result from economic withholding, particularly in transmission constrained areas.

When a transmission constraint is binding, one or more suppliers may be in the position to exercise market power due to the lack of competitive alternatives in the constrained area. For this reason, more restrictive conduct and impact thresholds are used for import-constrained areas. The ISO has two structural tests (i.e., Pivotal Supplier Test and Constrained Area Test) to determine which mitigation thresholds are applied to a supply offer in the following five categories: [159]

- Market-Wide Energy Mitigation ("ME") – ME mitigation is applied to any resource that is in the portfolio of a pivotal Market Participant. The conduct test is applied to all offer blocks that are greater than $25 per MWh and uses a threshold of the lower of $100 per

159 See Market Rule 1 Appendix A Section III.A.5.2 for more details of these two structural tests.

MWh or 300 percent over the reference level. The impact test uses a slightly different threshold, which is the lower of $100 per MWh or 200 percent of the LMP.

- Market-Wide Commitment Mitigation (“MC”) – MC mitigation is applied to any resource whose Market Participant is determined to be pivotal supplier. The conduct test is applied to all individual start-up (including cold, intermediate, and hot starts) and no-load offers and has a threshold of 200 percent over the reference level.

- Constrained Area Energy Mitigation (“CAE”) – CAE mitigation is applied to a resource that is determined to be within a constrained area. The conduct test is applied to all offer blocks and has a threshold of the lower of $25 per MWh or 50 percent over the reference level. The impact test uses the same threshold relative to the LMP.

- Constrained Area Commitment Mitigation (“CAC”) – CAC mitigation is applied to a resource that is determined to be within a constrained area. The conduct test is applied to all individual start-up and no-load offers and has a threshold of 25 percent over the reference level.

- Local Reliability Commitment Mitigation (“RC”) – RC mitigation is applied to a resource that is committed or kept online for local reliability. [160] The conduct test is applied to all individual start-up and no-load offers and has a threshold of 10 percent or $80 per MWh over the low-load commitment reference level. [161]

There are no impact tests for the three types of commitment mitigation (i.e., MC, CAC, and RC), so suppliers are mitigated if they fail only the conduct test in these three categories. On the other hand, for the other two categories (i.e., ME and CAE), suppliers are mitigated if they fail both the conduct and impact tests. Once a generator is mitigated, all the financial parameters of the supply offer (i.e., including all energy offer blocks, all types of start-up, and no-load offers) are set to their reference levels from the time when the mitigation decision is made until the end of the day or the end of the minimum run time, whichever is greater.

The ISO preformed mitigation manually until April 2012 when the real-time mitigation process was automated and incorporated into the market software.[162] The automated mitigation process

160 This includes local first or second contingency protection, voltage support, or special constraint resource service.

161 The low-load commitment reference level is based on the start-up See Market Rule 1 Appendix A Section III.A.5.5.5 for more details of the calculation of the low load commitment reference and the formulation of the conduct threshold.

162 On April 17, 2012, the ISO automated mitigations in real-time, but the day-ahead mitigation is still a manual process.

("AMP") allows for a more timely and accurate assessment of the estimated price impact of offers that violate the conduct test in real-time for the following reasons:

- Before the mitigation process was automated, only resources that were "on the margin" and setting the LMP were identified as candidates for mitigation.[163]
  - → After automation, all suppliers failing the conduct test and affecting prices by more than the impact threshold are mitigated, even if they are extra-marginal.
- Before automation, a consultation between the resource and the ISO was required before making mitigation decisions. In some cases, the resource was no longer marginal when the consultation process was complete, so mitigation would not be implemented.
  - → After automation, offers are automatically mitigated so consultation must occur before the hour.

Since there is no manual review of offers and reference levels before mitigation is imposed, it is more important to ensure that the inputs to mitigation, especially the generators' cost reference levels, are complete and accurate when the process is automated.

Under the new automated-mitigation rules, the mitigation is effective without delay. The automated new process uses a parallel dispatch model to measure the price impact of supply offers that fail conduct, which is a great improvement from the usage of an overly-simplified model in the manual process.[164]

The following analyses examine the frequency and quantity of mitigation in the real-time energy market under the new automated process. Figure 48 shows the frequency of automated mitigation for each type of mitigation in 2012 on a monthly basis. Any mitigation changes made after the automated mitigation process were not included in this analysis. The lower portion of the figure shows the total number of mitigations that occurred in each month on three categories of resources: (a) hydroelectric units; (b) thermal peaking units; and (c) thermal combined cycle and steam units. The upper portion of the figure shows the total number of hours in each month

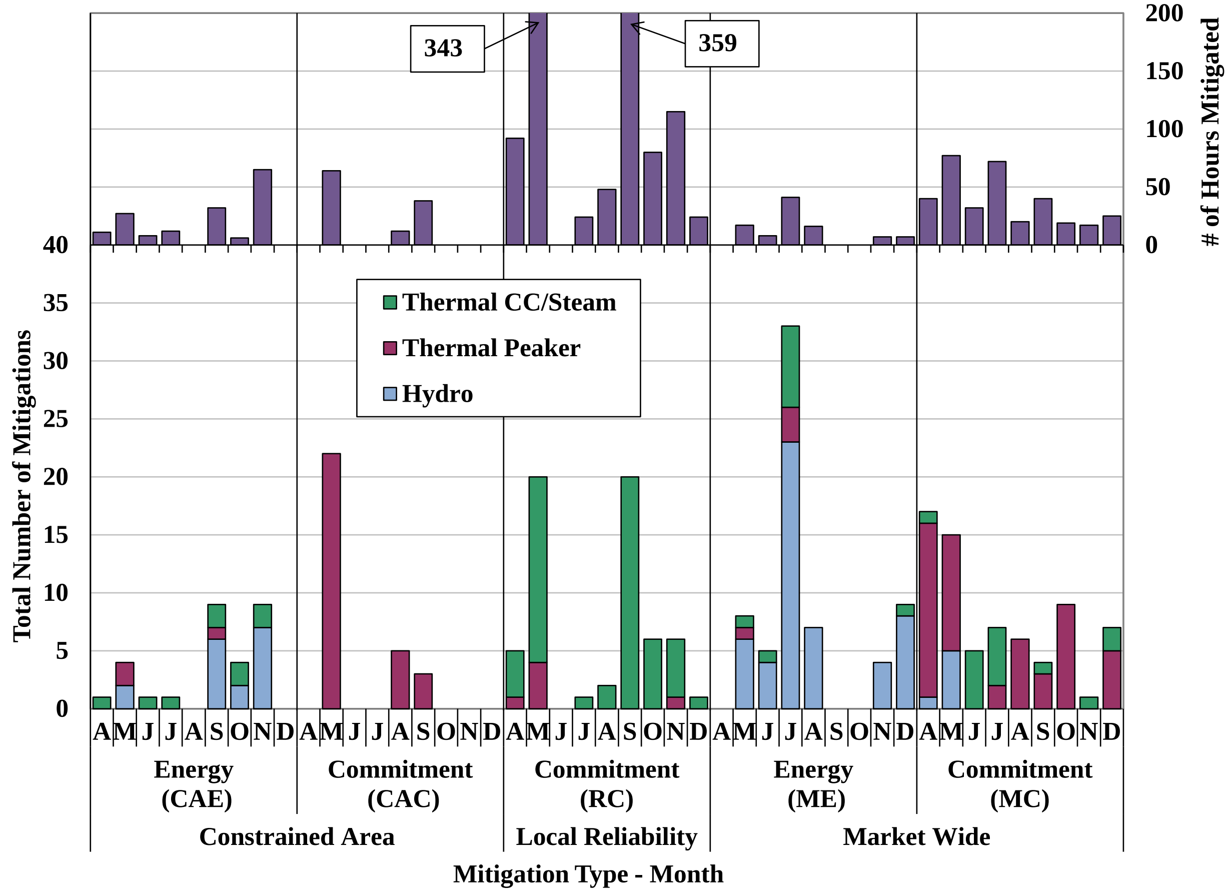
---

163 Hence, a $50 resource that caused an increase in the LMP from $75 to $150 by offering $500 would not be mitigated before April 2012.

164 The parallel dispatch model (SSPD) uses the same inputs and the same network model as in the actual dispatch model (UDS), except that UDS uses original supply offers while SSPD uses references to replace the supply offers that fail conduct. The LMPs calculated in SSPD are compared with the LMPs from UDS. If the difference exceeds the applicable impact threshold, the supply offer is mitigated.

that were affected by each type of mitigation. If multiple resources were mitigated during the same hour, only one hour was counted in the figure.

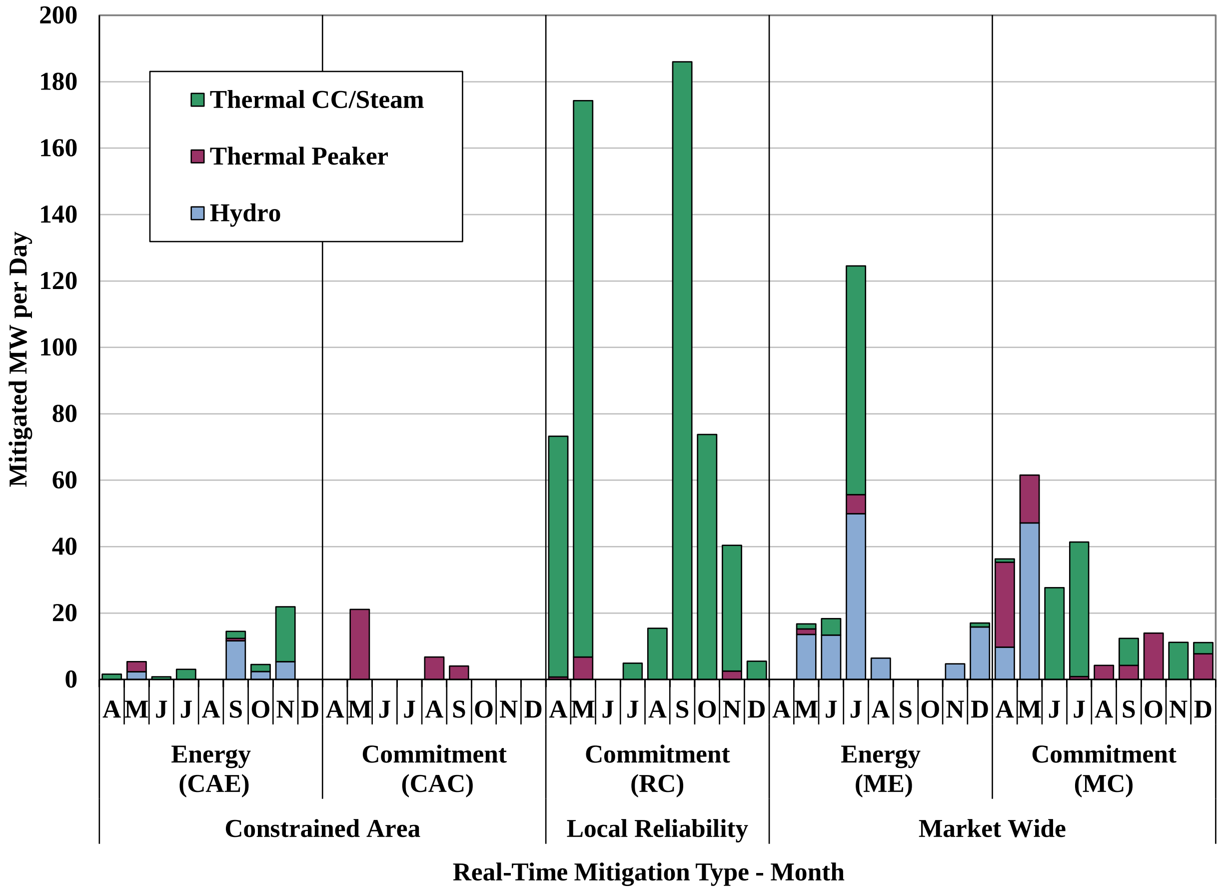
Figure 48: Frequency of Real-Time Mitigation by Mitigation Type and Unit Type
April – December, 2012



Mitigation occurred more than 250 times under the new automated process during 2012. Commitment mitigation (i.e., CAC, MC, and RC) accounted for 63 percent of all mitigation while energy mitigation (i.e., CAE and ME) accounted for the remaining 37 percent. For each type of mitigation, there was a dominant resource category. In particular, approximately 90 percent of local reliability commitment mitigation occurred on non-peaking thermal resources (i.e., fossil-fueled combined cycle units and steam units). Likewise, 80 percent of other commitment mitigation (i.e., CAC and MCM) occurred on thermal peaking resources (i.e., gas turbines) and 73 percent of energy mitigation (i.e., CAE and MEM) occurred on hydroelectric resources.

Figure 49 shows the quantity of automated mitigation for each type of mitigation in 2012 on a monthly basis. The figure shows the average amount of capacity that was mitigated each day separately for hydroelectric resources, thermal peakers, and other thermal generating resources.

Figure 49: Quantity of Real-Time Mitigation by Mitigation Type and Unit Type
April – December, 2012



Local reliability commitment mitigation (i.e., RC) accounted for the largest share (54 percent) of capacity that was mitigated in 2012. General threshold commitment and energy mitigations were the next two largest categories, accounting for 21 percent and 18 percent, respectively. Constrained area commitment and energy mitigations were much less significant, collectively accounting for roughly 7 percent in 2012. On average, roughly 120 MW of capacity was mitigated each day in 2012, of which 71 percent was for thermal non-peaking resources, 17 percent was for hydroelectric resources, and the remaining 12 percent was for thermal peaking resources.

Although it was relatively uncommon, the mitigation of hydroelectric resources raises some potential concerns, since it can be difficult to accurately reflect the marginal costs of a hydroelectric resource in the reference level calculation. This increases the potential for mitigating a competitive offer to a level below the resource's marginal cost. Most hydroelectric resources, and energy-limited resources in general, typically formulate offers that will enable them to produce output when it is valuable to the system (normally peak hours) and consequently morst profitable. Since intra-day offers are not currently allowed, when one offer curve is mitigated it remains mitigated for the rest of the day. Therefore, if a mitigation of a hydroelectric resource occurs during the early morning hours, it can potentially lead the resource to generate more than intended during off-peak hours. This issue can be challenging to address in the mitigation process.

The figures also show that mitigation rose immediately after the automated process became effective, occurring nearly 100 times from April 18 to the end of May, which affected markets in nearly 55 percent of hours during the period. An increase was expected because some offer behaviors that would not trigger mitigation before are now subject to mitigation under the new rule.[165] However, some of this increase was due to reference levels that did not accurately reflect suppliers' marginal costs.

Prior to automated mitigation, market participants were under much less pressure to update their reference levels information on a timely basis and consult with the ISO when a reference level becomes inaccurate. The increase in mitigation after the process was automated has prompted participants to consult with the IMM and submit more timely updates to information. Likewise, the IMM has been responsive in working with the participants to improve the reference levels. It has also been working on improving the reference level processes so that they can better handle natural gas price volatility that can cause inappropriate mitigation.

165 The manual process did not mitigate resources that were not setting the LMP, while the new automated mitigation procedure imposes mitigation whenever applicable conduct raises the LMP by the threshold amount.

F. Conclusions

Based on the analyses of potential economic and physical withholding in this section, we find that the markets performed competitively with little evidence of market power abuses or manipulation in 2012. The pivotal supplier analysis suggests that market power concerns exist in several areas in New England. However, the abuse of this market power is addressed by the ISO-NE's market power mitigation measures, which limits the ability of a generator to offer above competitive levels when would doing so would have a substantial impact on LMPs in an import-constrained area.

The ISO substantially strengthened the market power mitigation measures in April 2012 when it implemented Automated Mitigation Procedures. Under AMP, the market software is used to measure LMP impact in parallel with the real-time dispatch, so mitigation can be performed in a more timely and accurate fashion. While some of the mitigation under this automated process may be attributable to reference levels that did not accurately reflect some of the suppliers' marginal costs. However, the IMM has worked with suppliers to address these issues and improve the automated mitigation process. Apart from the automated mitigation, we continue to monitor market outcomes closely for potential economic and physical withholding together with the IMM, and have found little additional conduct that would raise competitive concerns.