2011 ASSESSMENT OF THE ISO NEW ENGLAND ELECTRICITY MARKETS

Prepared by:

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

POTOMAC ECONOMICS

EXTERNAL MARKET MONITOR
FOR ISO-NE

JUNE 2012

Table of Contents

Executive Summary i

I. Prices and Market Outcomes 1
- A. Summary of Prices and Market Outcomes 1
- B. Energy Price Trends 2
- C. Prices in Transmission Constrained Areas 6
- D. Day-Ahead Market Performance: Energy Price Convergence 8
- E. Uplift Allocation and Virtual Trading 11
- F. Conclusions 15

II. Transmission Congestion and Financial Transmission Rights 17
- A. Congestion Costs and the Role of FTRs 17
- B. Congestion Revenue and Payments to FTR Holders 18
- C. Congestion Patterns and FTR Prices 21

III. Reserve and Regulation Markets 24
- A. Real-Time Reserve Market Results 25
- B. Forward Reserve Market 30
- C. Regulation Market 39
- D. Conclusions 45

IV. External Interface Scheduling 47
- A. Summary of Imports and Exports 48
- B. Interchange with New York 50
- C. Coordination of Interchange by the ISOs 56
- D. Conclusions and Recommendations 57

V. Real-Time Pricing and Market Performance 59
- A. Real-Time Commitment and Pricing of Fast-Start Resources 59
- B. Real-Time Operation and Pricing During Operating Reserve Shortages 67
- C. Real-Time Pricing During the Activation of Demand Response 69
- D. Ex Ante and Ex Post Pricing 75
- E. Real-Time Price Corrections 81
- F. Conclusions and Recommendations 83

VI. System Operations 85
- A. Accuracy of ISO Load Forecasts 86
- B. Commitment for Local and System Reliability 88
- C. Out-of-Merit Generation 91
- D. Surplus Capacity and Real-Time Prices 94
- E. Supplemental Commitments and Surplus Capacity 96
- F. Uplift Costs 100
- G. Conclusions and Recommendations 103

VII. Forward Capacity Market 105
- A. Proposed Changes to ISO-NE's Capacity Market 105
- B. Background on the Forward Capacity Market 106
- C. Analysis of Forward Capacity Auction Results 108

D. Capacity Market Design – Sloped Demand Curve 117
E. Capacity Market Design – Investment in a Forward Capacity Market.................. 121
F. Forward Capacity – Conclusions 126

VIII. Competitive Assessment 128
A. Market Power and Withholding 128
B. Structural Market Power Indicators 130
C. Economic Withholding.. 141
D. Physical Withholding 147
E. Conclusions 152

List of Figures

Figure 1: Monthly Average Day-Ahead Energy Prices and Natural Gas Prices 3
Figure 2: Monthly Average Implied Marginal Heat Rate 4
Figure 3: Average Day-Ahead Prices by Location 7
Figure 4: Convergence of Day-Ahead and Real-Time Prices at New England Hub 9
Figure 5: Virtual Transaction Volumes and Profitability 12
Figure 6: Congestion Revenue and Target Payments to FTR Holders 20
Figure 7: FTR Auction Prices vs. Day-Ahead and Real-Time Congestion 22
Figure 8: Quarterly Average Reserve Clearing Prices by Product and Location 29
Figure 9: Summary of Forward Reserve Auction for Local Areas 33
Figure 10: Summary of Forward Reserve Auction for Outside Local Areas 35
Figure 11: Forward Reserve Assignments by Resource Type 38
Figure 12: Regulation Market Expenses 40
Figure 13: Monthly Average Supply of Regulation 43
Figure 14: Average Net Imports from Canadian Interfaces 48
Figure 15: Average Net Imports from New York Interfaces 49
Figure 16: Real-Time Price Difference Between New England and Upstate New York 52
Figure 17: Efficiency of Scheduling in the Day-Ahead and Real-Time 53
Figure 18: Correlation Between Price Differences and Lead Time 55
Figure 20: Difference Between Real-Time LMPs and Offers of Fast-Start Generators 66
Figure 21: Real-Time Demand Response Program Enrollments 71
Figure 22: Average Difference Between Five-Minute Ex Post and Ex Ante Prices 77
Figure 23: Difference in Constraint Shadow Costs Between Ex Post and Ex Ante 80
Figure 24: Rate of Real-Time Price Corrections 82
Figure 25: Average Daily Peak Forecast Load and Actual Load 87
Figure 26: Commitment for Reliability by Zone 90
Figure 27: Average Hourly Out-of-Merit Generation 93
Figure 28: Surplus Capacity and Implied Marginal Heat Rates 95
Figure 29: Daily Supplemental Commitments and Surplus Capacity 97
Figure 30: Daily Supplemental Commitments and Surplus Capacity 97
Figure 31: FCM Auction Clearing Summary by Location 109
Figure 32: Summary of Accepted De-list Bids by Type 112

Figure 33: New Capacity Procurement by Location in FCAs 1 through 5 115
Figure 34: Capacity Market Outcomes 119
Figure 35: Sloped Demand Curve 120
Figure 36: Cycle of Investment with a Vertical Demand Curve 123
Figure 37: Supply Resources versus Summer Peak Load in Each Region 133
Figure 38: Installed Capacity Market Shares for Three Largest Suppliers 135
Figure 39: Frequency of One or More Pivotal Suppliers by Type of Withheld Capacity 138
Figure 40: Frequency of One or More Pivotal Suppliers by Load Level 140
Figure 41: Average Output Gap by Load Level and Type of Supplier 145
Figure 42: Average Output Gap by Load Level and Type of Supplier 146
Figure 43: Average Output Gap by Load Level and Type of Supplier 147
Figure 44: Forced Outages and Deratings by Load Level and Supplier 149
Figure 45: Forced Outages and Deratings by Load Level and Supplier 150
Figure 46: Forced Outages and Deratings by Load Level and Supplier 151

List of Tables

Table 1: Resource Availability During Activation of RT Demand Response 73
Table 2: Allocation of Uplift for Out-of-Market Energy and Reserves Costs 102
Table 3: Withholding by Type of Resource 132

Guide to Acronyms

APR	Alternative Price Rule
ASM	Ancillary Services Market
CONE	Cost of New Entry
CT DPUC	Connecticut Department of Public Utility Control
EMMU	External Market Monitoring Unit
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
IMMU	Internal Market Monitoring Unit
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 Monitoring Unit 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 2011 and our recommendations for future improvements. This report complements the Annual Markets Report, which provides the Internal Market Monitoring Unit's evaluation of the market outcomes in 2011.

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

1 The duties of the External Market Monitoring Unit are listed in Appendix A.2.2 of "Market Rule 1."

Executive Summary

This report assesses the efficiency and competitiveness of New England's wholesale electricity markets in 2011. 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 an outage or other contingency occurs;
- 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 2011, 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 2011. 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.

Energy prices fell 7 percent from 2010 to 2011, due primarily to reductions in natural gas prices and load levels as well as increases in surplus capacity in the real-time market.[2] The average price of natural gas (which is the dominant fuel in New England) fell 5 percent in 2011 from the prior year.[3] In a competitive market, suppliers have strong incentives to offer their supply at prices close to their short-run marginal costs of production.[4] Because fuel costs constitute the vast majority of the marginal costs of most generation, lower fuel costs translate to 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.

Energy prices decreased more than natural gas prices did in 2011 due to lower load levels and higher levels of surplus capacity. Average load decreased 1 percent from 2010 to 2011 and 2.5 percent from the summer of 2010 to the summer of 2011, primarily due to milder weather. Accordingly, electricity prices fell as lower demand required generating resources with high marginal costs to operate less frequently, particularly during the summer months.

Finally, the system generally operated with higher levels of surplus capacity in 2011 than in 2010. The amount of surplus capacity generally increased because a large flexible pump storage

2 Throughout this report, the term "surplus capacity" refers to the amount of online and quickstart capacity available in the real-time market in excess of the demand for energy and operating reserves.

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

4 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.

generating resource returned to service after being out of service for most of 2010. Increased imports from neighboring areas and new generating capacity additions in Connecticut also contributed to higher surplus capacity in 2011. The higher levels of surplus capacity contributed to the reduction in energy prices because it reduced the costs of satisfying fluctuations in the system's needs and the frequency of tight supply and shortage conditions.

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 2011 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.

Upgrades to the transmission system in Connecticut and Southeast Massachusetts were completed by mid-2009, leading to significant changes in market operations. The upgrades sharply reduced the need for the ISO to commit generation for local reliability. Hence, the total uplift charges from NCPC payments fell from $387 million in 2008 to $71 million in 2011. 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. Accordingly, the average level of surplus capacity decreased significantly in 2010 and 2011 from prior years.[5]

5 However, the system generally operated with higher levels of surplus capacity in 2011 than in 2010.

In 2011, market operations were significantly affected by the return of a large flexible pump storage resource from an extended outage in 2010. Since the unit provides large quantities of low-cost operating reserves, its return contributed to a 46 percent reduction in supplemental commitment for reliability and a 40 percent reduction in uplift from NCPC payments from 2010 to 2011. Furthermore, the unit contributed to increased real-time surplus capacity and less real-time price volatility.

3. Recommendations

Overall, we conclude that the markets performed competitively in 2011 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. Six of the ten were also recommended in our *2010 Annual Assessment*. This overlap 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. 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 7 percent from approximately $53 per MWh in 2010 to $49 in 2011.[6] This was due primarily to:

- Lower fuel prices – Natural gas prices fell 5 percent from the prior year, which is important because natural gas-fired resources are most frequently on the margin in New England.
- Lower load levels – Average load decreased 1 percent overall and 2.5 percent in the summer months from 2010 to 2011 because of mild summer temperatures. The number of hours during which load exceeded 20 GW decreased from 531 hours in 2010 to 312 hours in 2011, although annual peak load was nearly 3 percent higher in 2011 than in 2010.
- Increases in net imports – Average net imports in peak hours from neighboring areas, particularly Hydro Quebec and Upstate New York, increased 500 MW in 2011 to approximately 1,300 MW.

6 The average electricity price is weighted by the New England load level in each hour.

- Higher surplus capacity levels – On average, the daily minimum surplus capacity increased 24 percent (or 180 MW) from 2010 to 2011. Key factors included the return of a large, fast-starting, pump storage generator after an extended outage in 2010 and the entry of a new combined cycle generator in Connecticut.

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 2011, 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 only $18 million in 2011, down from $38 million in 2010. The decrease in congestion revenue was primarily due to two factors. First, load levels decreased and reduced the frequency of congestion into import-constrained areas, particularly during the summer. Second, lower natural gas prices reduced costs of offering generators' dispatch levels to manage the flow over a constraint and the associated congestion-related price differences.

The recent levels of congestion revenue are far below the 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 2011 (e.g., NYISO had $460 million, MISO had $503 million, and PJM had $998 million).[7] 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.

7 These markets are larger than ISO-NE. NYISO is roughly 1.5 times larger and the other markets are 4-5 times larger. However, these markets exhibited 25 to 55 times more congestion.

The ISO uses most of the congestion revenues to fund the economic property rights to the transmission system in the form of FTRs.[8] 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 2011, 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 2010 and 2011 from prior years, and we conclude that the FTR markets performed reasonably well in 2011.

Congestion revenue is used to fund the FTRs sold by ISO-NE. The congestion revenue of $18 million collected by the ISO in 2011 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 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. We found good overall convergence with the mean absolute difference between hourly day-ahead and real-time prices averaging 19 percent in 2011

8 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).

at the New England Hub. This is low relative to other markets and is primarily attributable to lower real-time price volatility in New England.

However, average real-time prices have been persistently higher than average day-ahead prices in the past two years, which is unusual since electricity markets typically exhibit slightly higher day-ahead prices. We do not believe this result is efficient because day-ahead premiums 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 (such as when high-cost peaking resources are committed to satisfy the real-time demand). 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 satisfy the real-time system needs.

One reason for the pattern of real-time premiums in the past two years is that the average allocation of NCPC charges to virtual load has increased (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 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 load or physical 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-NE 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. 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 nodal 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 and $1.98 in 2011. 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), which is allocated to "real-time deviations" between day-ahead and real-time schedules.[9] 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 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. Allocating costs in a manner that is not consistent with the causes of the costs is inefficient because it does not facilitate efficient

9 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.

conduct from market participants. Therefore, 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. We are working with the ISO and its IMMU 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 fell from 2010 to 2011. Outside the local constrained areas, the average TMSR clearing price fell from $1.76 per MWh in 2010 to $1.04 in 2011; the average TMNSR price fell from $1.16 to $0.39; and the average TMOR price fell from $0.42 to $0.25. The decreases were related to the substantial increase in the average surplus capacity that was primarily due to a large flexible pump storage resource returning to service in December 2010 following a seven-month outage. The resource's return to service significantly increased the overall ramping capability of the system and availability of operating reserves. The ISO 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 between 2007 and 2009.

2. 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. [10] This report evaluates the results of recent forward reserve auctions and examines how suppliers satisfied their obligations in the real-time market.

In Connecticut, the TMOR clearing prices fell 49 percent from an average of $8.65 per kW-month in the 2010/11 Capability Period (June 2010 to May 2011) to $4.40 per kW-month in the 2011/12 Capability Period (June 2011 to May 2012). Adjusting for the fact that forward capacity payments are deducted from payments to forward reserve providers, the effective price of forward reserves in Connecticut fell 71 percent from an average of $4.40 per kW-month in the 2010/11 Capability Period to $1.28 per kW-month in the 2011/12 Capability Period. Connecticut prices fell because the Connecticut reserve requirement fell from an average of 1,025 MW in the 2010/11 Capability Period to 756 MW in the 2011/12 Capability Period. The local reserve requirement decreased in 2011 because the most recent transmission upgrades in Connecticut were completed in 2009 (the forward reserve market incorporates the impact of transmission upgrades on the local requirement gradually over a two year period). Offer quantities increased with the introduction of new peaking capacity and increased participation by existing peaking capacity.

Outside Connecticut, the TMOR clearing prices fell 22 percent from an average of $5.65 per kW-month in the 2010/11 Capability Period to $4.40 per kW-month in the 2011/12 Capability Period in the forward reserve market. Prices fell partly because of the elimination of the Rest of System TMOR requirement, which reduced the amount of reserves that must be purchased from outside Connecticut and Boston.

Finally, 98 percent of the resources assigned to satisfy forward reserve obligations in 2011 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.

10 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 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, so the obligations of forward reserve suppliers are not consistent with the day-to-day operational needs of the system. Consequently, excess reserves are available on some days in local areas and system-wide, while insufficient reserves are available on other days, necessitating out-of-market commitments for reliability.

3. Regulation Market

The regulation market performed competitively in 2011, with an average of approximately 870 MW of available supply competing to serve an average of 60 MW of regulation demand.[11] 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.[12]

Regulation market expenses fell modestly from $14 million in 2010 to $13 million in 2011. This reduction was due in part to the reduction in regulation requirement in 2011 (down roughly 9 percent from 2010) and the increase in low-cost regulation offers following new entry in Connecticut and the return from outage of a large flexible pump storage.

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

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

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

regulation service provided."[13] To comply with the Order, ISO-NE plans to incorporate a separate movement offer (i.e., the cost of moving a unit up and down) in the criteria for minimizing the total expected regulation cost, but not set a separate clearing price for this component of the regulation offer.

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 1,630 MW during peak hours and 1,200 MW during off-peak hours in 2011. 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 western 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 smaller interfaces (averaging roughly 400 MW during peak hours in 2011), while participants schedule power flows that can alternate directions on the larger interface depending on the relative prices. In 2011, ISO-NE imported an average of 60 MW from NYISO across the larger interface during

---

13 *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).

peak hours and 140 MW during off-peak hours. In contrast, ISO-NE was a net exporter of power to NYISO in 2010, which contributed to the lower energy prices observed in 2011.

On an hourly basis, market participants should arbitrage the prices in 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. In fact, we found that power was scheduled in the wrong direction (*from* the high-priced market *to* the low priced-market) 41 percent of the time. This results in substantial inefficiencies and higher costs in both areas.

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 being implemented to allow intra-hour changes in the interchange between control areas. Under CTS, the ISOs will schedule intra-hour interchange transactions based on short-term forecasts of prices. We continue to recommend that ISO-NE and NYISO place a high priority on implementing 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 the second initiative are substantially lower than the benefits of the CTS initiative.

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 2011 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 2011, 75 percent of the fast-start capacity that was started in the real-time market did not recoup its offer. These resources received $12 million in NCPC payments as a result. 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.

- *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 marginal cost of meeting system-level 30-minute reserve requirements can exceed the ISO's Reserve Constraint Penalty Factor (RCPF) of $100 per MWh, requiring the ISO to curtail exports and take other manual actions outside the market. This can lead to inefficiently low real-time prices that do not properly reflect the cost of maintaining reliability. For these reasons, the ISO will replace the $100 RCPF for system-level 30-minute reserves with the $500 RCPF on June 1, 2012. This will reduce the need for manual actions to maintain reserves and provide more efficient price signals during reserve shortages.

3. Real-Time Pricing During Demand Response Activations

Demand response resources were activated to maintain operating reserves on 2 days in 2011 during OP-4 conditions. These contributed to system reliability and reduced the overall cost of satisfying demand. However, the inflexibility of demand response resources led real-time clearing prices not to adequately reflect the marginal costs of the demand response resources during most of the periods in which they were activated.

- *We recommend that the ISO develop rules for allowing the activation of 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 2011, 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 2011.

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 2011. 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 in the past two years because significant transmission upgrades in historically import-constrained areas were completed in 2009. These upgrades have allowed additional imports to these areas, reducing the amount of online and quick start capacity that must be available internally. In 2011, the amount of capacity committed for most local reliability issues averaged 90 MW, down 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 2011. 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 recent transmission upgrades have substantially reduced the need for the ISO to commit generation to satisfy local reliability requirements since July 2009, the ISO still needs to make supplemental commitments periodically 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 more than 400 MW in 2010 and 2011.

Despite the overall increase in the recent two years, the amount of supplemental commitment for system-wide reliability requirements fell from 590 MW in 2010 to 260 MW in 2011, accounting for 91 percent and 75 percent of total reliability commitments in the two years.[14] The need to commit resources for system-wide reliability was greater in 2010 than in 2011 due to the extended outage of a large flexible pump storage resource during most of 2010. The return of this unit increased the amount of fast start resources available to satisfy system capacity requirements in the Reserve Adequacy Assessment (RAA) process, thereby reducing the need to supplementally commit non-fast start resources. This pattern underscores the importance of setting price signals in the day-ahead and real-time markets that reflect the full value that such resources provide to the system.

After reviewing the supplemental commitments and the surplus capacity levels that resulted from real-time operating conditions, we found that roughly 42 percent of the capacity that was supplementally committed in 2011 was actually needed to maintain system level reserves in retrospect.[15] 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 simultaneously become unavailable due to the limitations of the natural gas network. Consequently, the ISO may commit oil-fired and/or dual-

14 These quantities include small amounts of commitments for local first contingencies.

15 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".

fueled capacity for reliability in order to protect the system in the event that some generators are unavailable due to limited gas supplies.

4. Uplift Charges

Uplift charges fell significantly from $119 million in 2010 to $71 million in 2011, primarily for two reasons. First, out-of-market capacity payments (including payments under reliability agreements and payments to rejected delist capacity under FCM) fell from $24 million in 2010 to $1.4 million in 2011. All of the reliability agreements expired on June 1, 2010, when the first Forward Capacity Commitment Period began. The remaining out-of-market capacity payments were paid to two units in Connecticut in 2010 and 2011 under FCM because their de-list requests were rejected in the first Forward Capacity Commitment Period for reliability purposes.

Second, the "Economic" category of uplift payments fell from $85 million in 2010 to $54 million in 2011. The reduction in supplemental commitment for system-wide reliability requirement has led to concomitant reductions in the NCPC payments to such resources. These reductions are largely attributable to the return to service in December 2010 of a large flexible pump storage resource following an extended outage. The reduction in NCPC payments highlights the value of flexible fast-start resources because they reduce the ISO's need to avoid making costly supplemental commitments to maintain system reliability.

5. Conclusions

Our assessment of system operations indicates that the ISO has operated the system relatively well, and we found no major concerns. Additionally, the cost of satisfying the system's local reliability and system-wide reliability requirements continued to decrease in 2011.

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 (FCA1) was held in February 2008, facilitating the procurement of installed capacity from June 2010 to May 2011. By the end of 2011, five auctions have been held, which had competitive results and satisfied ISO-NE's planning requirements through May 2015.

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 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 five FCAs has procured a significant amount of excess capacity. For example, FCA5 procured nearly 37 GW of resources, exceeding the Net Installed Capacity Requirement (NICR) by 3.7 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, the price will likely fall close to zero due to the level of existing capacity and the vertical demand curve implicit in the FCM design.

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 the first five FCAs, nearly 8 GW of new capacity was procured from generation, demand response resources and imports. 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 the 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.

However, the FCM has supported the development of a substantial amount of new demand response resources. Based on experience to date in New England, demand response resources may not provide response comparable to supply resources during shortage or emergency conditions. When demand response resources have been 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.[16] 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 for the ISO to assess 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.

In addition to facilitating efficient development of new resources, 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 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 FCA4.
- 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. Four zones will be modeled in FCA7 and eight zones should be modeled thereafter.

The Commission's April 2011 Order directed the ISO to work with stakeholders to resolve a number of issues in the FCM, including modeling the eight zones and addressing the local

16 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.

market power that may result by developing effective seller-side and buyer-side market power mitigation measures.[17]

Seller-side measures ensure that suppliers offer to sell in the capacity market when it is economic for them to do so (i.e., when the expected revenues from remaining in service exceed the unit's operating costs and going forward costs). Buyer-side measures deter large buyers from deliberately building uneconomic new supply in order to depress market prices below competitive levels. Because we believe these changes are critical, we continue to recommend the ISO place a high priority on completing and implementing these changes. We also intervened in response to the ISO's request in January 2012 for more time for the stakeholder discussions, and recommending that the Commission establish a deadline for the ISO to file these changes that would ensure that they are in place by FCA8.[18] The Commission agreed and has required the ISO to make a compliance filing by December 2012.

Even with these changes, however, we find that the current FCM design is flawed and will 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.[19]

---

17 *ISO New England Inc. and New England Power Pool Participants Committee*, 135 FERC ¶ 61,029 (2011) ("April Order").

18 *Motion To Intervene Out Of Time And Comments of ISO New England's External Market Monitoring Unit*, Docket No. ER12-953-000, filed February 22, 2012.

19 The Rationing Election allows a new supplier to allow its offer to be partially accepted. 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.

H. Competitive Assessment

The report evaluates the market concentration and competitive performance of the markets operated by ISO-NE in 2011. Based on our evaluation of the markets in New England we find that the markets performed competitively in 2011.

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.

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 2011 in Connecticut (30 percent of hours), Boston (60 percent of hours), and All of New England (59 percent of hours).

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, this was not a significant concern in 2011.

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 should enable the ISO to identify and prevent the abuse of market power in a more timely and accurate fashion than the current manual process. Nonetheless, we and the IMMU will continue to monitor market outcomes closely for potential economic and physical withholding.

I. Table of Recommendations

We make the following recommendations based on our assessment of the ISO-NE's market performance in 2011. 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[20]	Feasible in ST[21]
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.	✓	✓	
Capacity Market			
7. Replace the current capacity requirement (i.e., vertical demand curve) with sloped demand curve that recognizes the value of additional capacity.		✓	
8. Implement effective supplier-side and buyer-side market power mitigation measures.	✓	✓	
9. Model a full set of capacity zones and capacity transfer rights (CTRs) in order to fully reflect locational requirements in the FCA clearing prices.	✓	✓	
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.			✓

20 Recommendation will likely produce considerable efficiency benefits.

21 Complexity and required software modifications are likely limited (does not include consideration of stakeholder or regulatory issues).

I. Prices and Market Outcomes

In this section, we review wholesale market outcomes in New England during 2011. This section includes an analysis of overall price trends and a review of prices in transmission congestion 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

Average day-ahead and real-time energy prices fell 6 to 7 percent from 2010 to 2011. At the New England Hub, average real-time energy prices fell from approximately $53 per MWh in 2010 to $49 per MWh in 2011, while average day-ahead prices were 1 to 2 percent lower than average real-time prices in both years. Several factors contributed to the reductions in energy prices:

- The average price of natural gas, which fuels the marginal generation that sets energy prices in most hours, decreased 5 percent from 2010;
- Average load fell 1 percent in 2011, although the annual peak load was nearly 3 percent higher in 2011 than in 2010;
- Average net imports from Hydro Quebec and upstate New York rose substantially from 2010 to 2011; and
- A large pump storage facility returned to service following a nearly seven-month outage in 2010 (including the summer), and a new combined cycle unit (620 MW) entered in Connecticut in June 2011.

New England experienced very little congestion in 2011 into historically-constrained areas such as Boston, Connecticut, and Lower Southeast Massachusetts (Lower SEMA) as a result of transmission upgrades that have been made in recent years. 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 of only $18 million, compared to more than $500 million in New York and MISO.

Differences between day-ahead and real-time prices were moderate in 2011. Average real-time prices were higher than average day-ahead prices by 1 percent in 2011, which was an improvement from 2010 when the average real-time price premium was 2 percent higher. 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 two years raise efficiency concerns because real-time prices tend to be understated for reasons discussed in this report. In general, it is efficient for day-ahead prices to exceed real-time prices 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.[22] Figure 1 shows the load-weighted average price at the New England Hub in the day-ahead market for each month in 2010 and 2011. The figure also shows the average natural gas price, which should be a key driver of electricity prices when the market is operating competitively.[23]

The figure shows that natural gas price fluctuations were a significant driver of variations in monthly average electricity prices in 2010 and 2011 as expected. In 2011, more than 40 percent of the installed generating capacity in New England used natural gas as its primary fuel.[24] Low-cost nuclear resources and other baseload resources typically produce at full output, while natural gas-fired resources, which accounted for 52 percent of all electricity production in 2011, are on the margin and set the market clearing price in most hours.[25] Therefore, electricity prices should be 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

22 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.

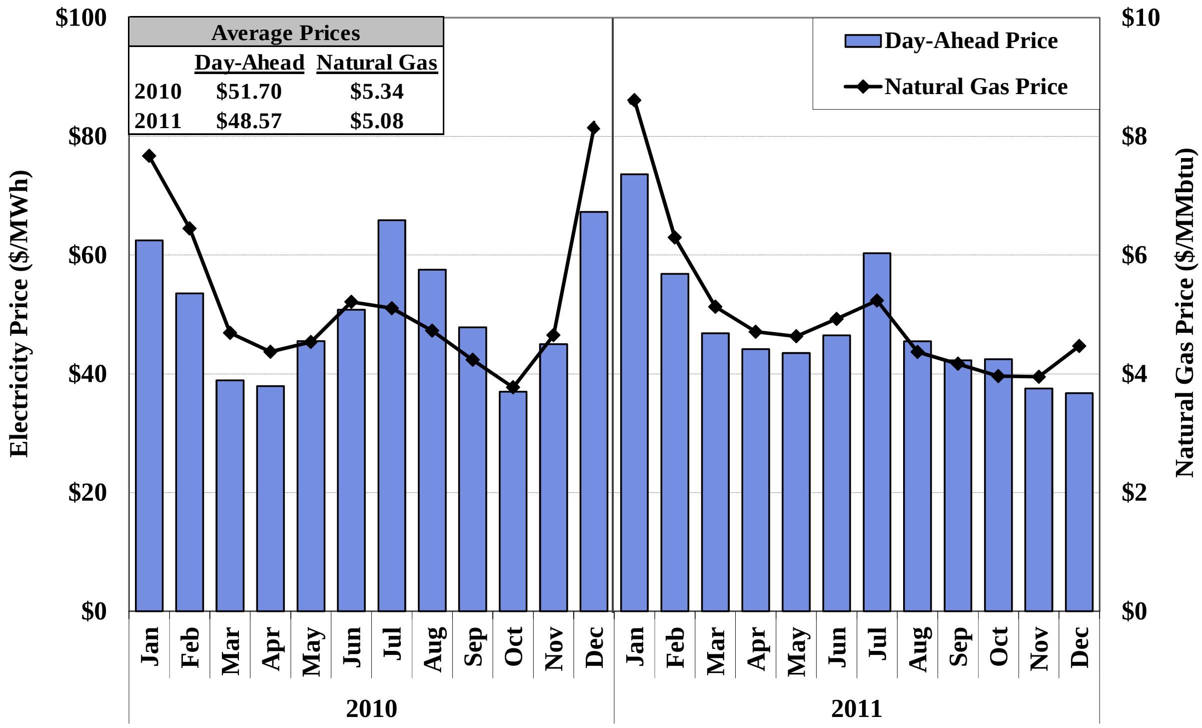
23 The figure shows the day-ahead price reported by Platts for the Algonquin pipeline at City Gates.

24 ISO-NE, “2011-2020 Forecast Report of Capacity, Energy, Loads, and Transmission (CELT) Report,” September 2011.

25 According to preliminary data from EIA Form 923 for 2011, 52 percent of net generation was produced from natural gas, while 28 percent was produced from nuclear fuel, 7 percent from hydroelectric, 6 percent from other renewables sources (including refuse burning), 6 percent from coal, and 1 percent from fuel oil.

increases due to colder weather. Accordingly, natural gas prices decreased from January to March and rose from October to December in both 2010 and 2011, leading to concomitant changes in electricity prices over the same period. However, weather was unseasonably warm in December 2011, leading to unusually low natural gas prices and electricity prices in this month.

Figure 1: Monthly Average Day-Ahead Energy Prices and Natural Gas Prices
New England Hub, 2010 – 2011

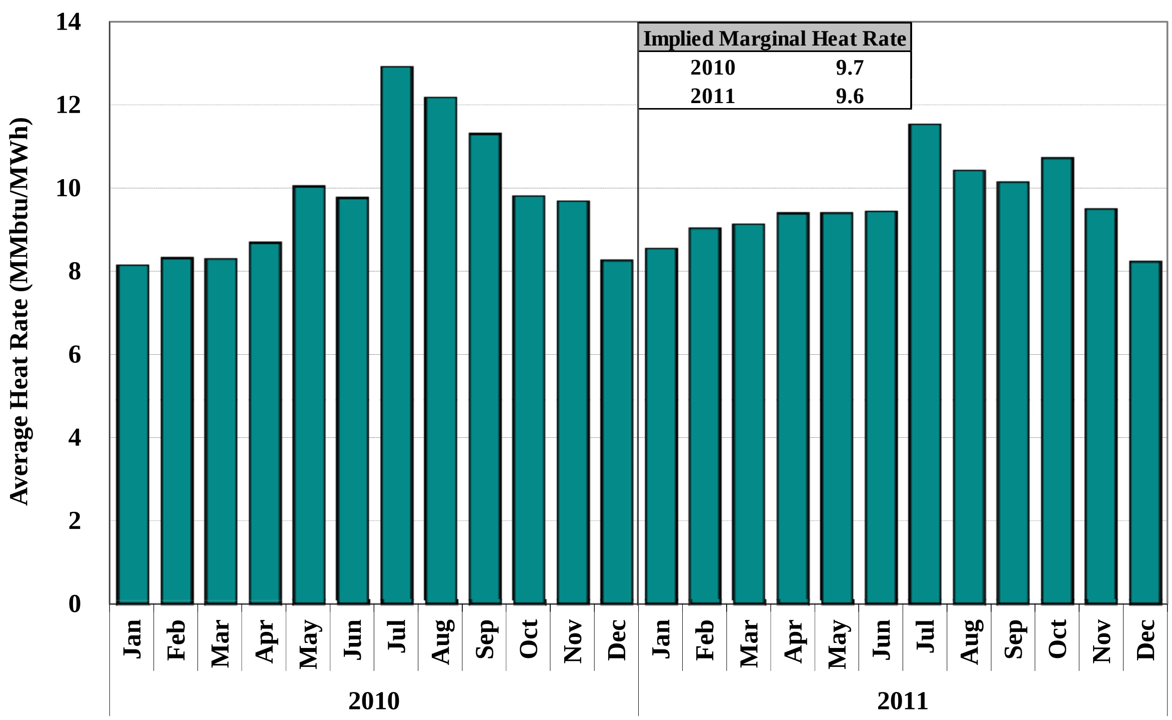


Energy prices usually increase during high summer and winter load periods when the demand for cooling and heating are highest. The effects of seasonal changes in demand were significant in both years during the summer months. For example, average natural gas prices increased only 6 percent in July 2011 from the prior month, while the average electricity prices rose 30 percent as demand increased sharply.

Overall, the average New England Hub price in the day-ahead market decreased 6 percent from 2010 to 2011. The most significant driver of the lower prices was the 5 percent decrease in average natural gas prices from 2010 to 2011. Lower gas prices resulted in lower energy prices in most hours because natural gas-fired units are frequently on the margin. However, the reduction was partly offset by increases in other fossil fuel prices. Average No. 2 fuel oil prices rose 39 percent from 2010 to 2011, average No. 6 fuel oil prices rose 48 percent, and average coal prices rose 5 percent. These increases led to higher energy prices in a smaller portion of hours. Reduced load levels, increased imports, new generating capacity in Connecticut, and the return of a large resource after a lengthy outage 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. A lower marginal heat rate indicates that factors other than lower natural gas prices have contributed to the decrease in energy prices. Figure 2 shows the load-weighted average implied marginal heat rate for the New England Hub in each month during 2010 and 2011.

Figure 2: Monthly Average Implied Marginal Heat Rate
Based on Day-Ahead Prices at New England Hub, 2010 – 2011



By adjusting for the variation in natural gas prices, 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. However, this seasonality was less pronounced in 2011 than in 2010 because the average load was lower in the summer months. In 2011, the month with the highest average implied marginal heat rate was July, which was also the month with the hottest temperatures and highest average loads.

Figure 2 shows that the average implied marginal heat rate fell approximately 1 percent from 2010 to 2011. The decrease was most notable during the summer months (June to August) when the implied marginal heat rate fell 9 percent from an average of 11.7 MMbtu per MWh in 2010 to 10.6 in 2011. The decrease in the implied heat rate was primarily due to:

- Lower load in 2011, particularly during the summer. Average load decreased 1 percent in 2011 and 2.5 percent in the summer months of 2011. The number of hours when load exceeded 20 GW decreased from 531 hours in 2010 to 312 hours in 2011.

- Higher net imports in 2011 from neighboring areas, particularly Hydro Quebec and Upstate New York. Total net imports averaged approximately 1,300 MW during peak hours in 2011, up 500 MW from 2010.

Less frequent real-time price spikes in 2011. The implied heat rate exceeded 20 MMbtu per MWh in 141 hours in 2011 compared to 272 hours in 2010. This decrease was partly due to an increase in the average quantity of surplus capacity (available online or quick-start capacity in excess of energy and reserve needs). This is discussed Section VI.D. The increase in surplus capacity was partly due to the return of a large pump storage generator after an extended outage in 2010.

The implied heat rate reduction was partly offset by other factors. First, a small share of generation costs (e.g., variable operating and maintenance expenses) is not related to fuel prices, so reductions in fuel prices lead the non-fuel costs to increase as a share of total generation costs. Consequently, implied heat rates rise when natural gas prices fall substantially. Second, natural gas prices fell significantly relative to most other fuels in 2011, which increases the implied heat rate in periods when generation fired by other fuels was on the margin setting energy prices.

C. 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 the serving load at that location. Transmission congestion arises because the lowest-cost resources cannot be fully dispatched because transmission capability is limited. 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.

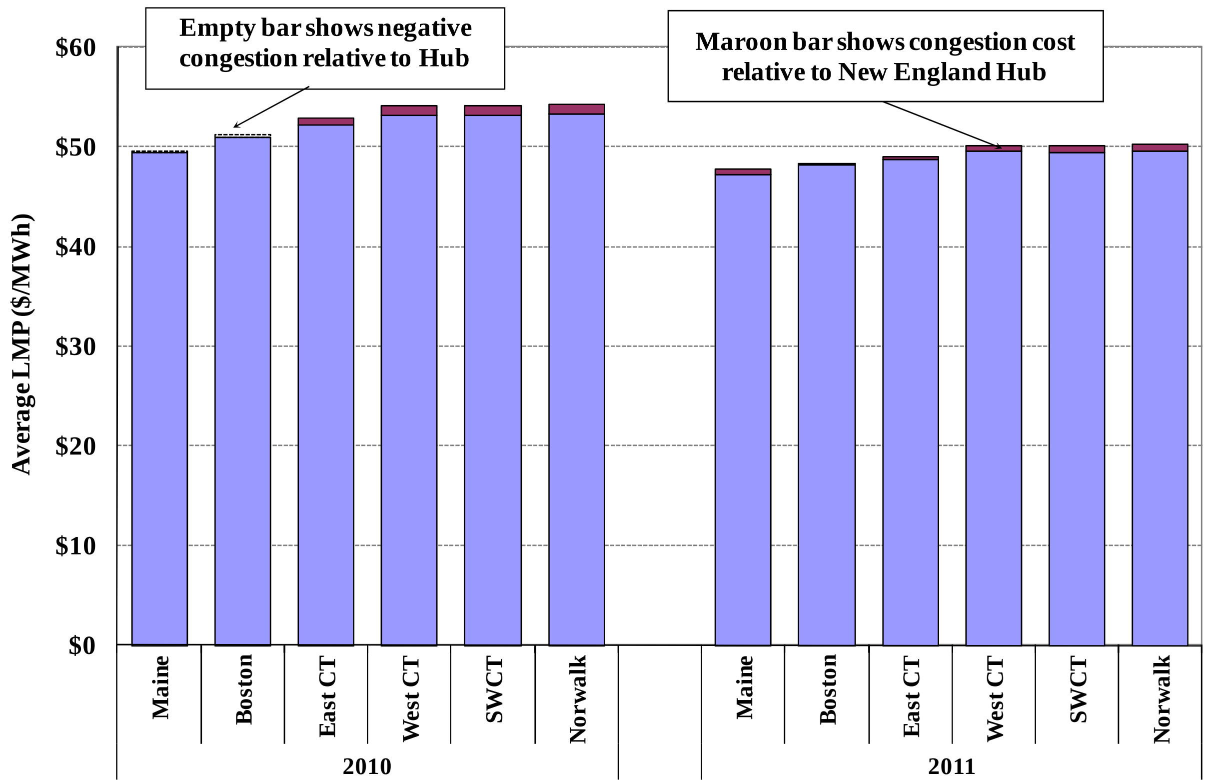
Historically, there have been significant transmission limitations between net-exporting and net-importing regions in New England. In particular, exports from Maine to the rest of New England have been limited by transmission constraints at times, while Connecticut and Boston were often unable to import enough power to satisfy demand without dispatching expensive local generation in the past.

We analyzed the differences in energy prices between several key locations during the study period. Figure 3 shows load-weighted average day-ahead LMPs in 2010 and 2011 for the Maine load zone, NEMA/Boston load zone, and four areas within Connecticut.[26] For each location, the load-weighted average LMP is indicated by the height of the solid bars. The maroon portion of

26 Connecticut is divided into: East Connecticut, the portion of West CT that excludes Southwest CT, the portion of Southwest CT that excludes Norwalk-Stamford, and Norwalk-Stamford.

the bars indicates positive congestion to the location from the New England Hub, while negative congestion is indicated by the empty bars. Thus, areas that are import-constrained (e.g., areas within Connecticut) exhibit positive congestion from the Hub.

Figure 3: Average Day-Ahead Prices by Location
2010 – 2011



Congestion levels remained very low in 2011. The average differential in the congestion component of the LMP between West Connecticut and Maine was negligible in 2011, down slightly (from 2 percent of the LMP) from 2010. The average differential in the loss component of the LMP between these locations was much more significant at 5 percent of the average LMP.

The figure shows that all areas exhibited very low levels of congestion in both 2010 and 2011. Significant transmission upgrades were placed in service in Boston, Connecticut, and Southeast Massachusetts from 2007 to 2009, which greatly increased the transfer capability into these areas and eliminated most of the historical congestion into these regions. As a result, the LMPs no longer provide significant incentives for locating new resources in net-importing regions, such as Connecticut and Boston.

D. Day-Ahead Market Performance: Energy Price 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 risk. 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.

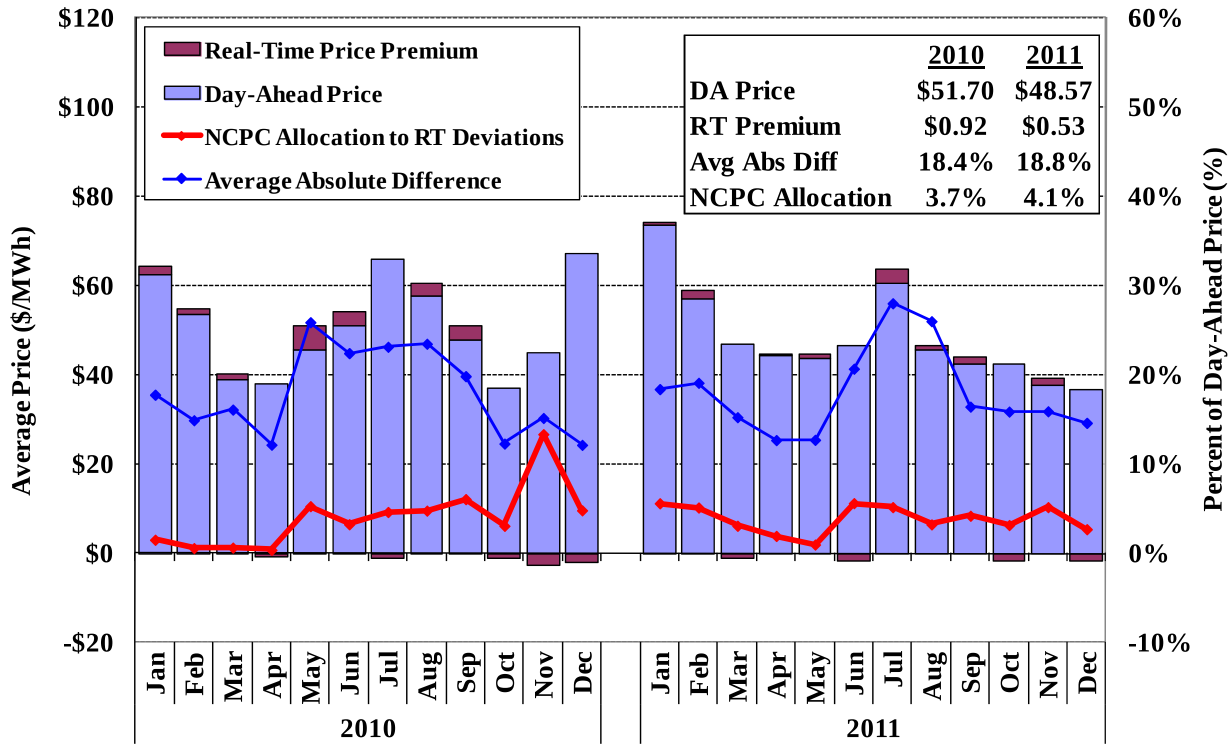
In well-functioning day-ahead and real-time markets, we expect that day-ahead and real-time prices will not systematically diverge by a substantial amount. If day-ahead prices were predictably higher or lower than real-time prices, participants should adjust their purchases and sales in the day-ahead market to bring the prices into convergence. However, day-ahead prices can 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.

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 4 summarizes day-ahead prices and the convergence between day-ahead and real-time prices at the New England Hub in each month of 2010 and 2011.[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.

Figure 4: Convergence of Day-Ahead and Real-Time Prices at New England Hub
2010 – 2011



	2010	2011
DA Price	$51.70	$48.57
RT Premium	$0.92	$0.53
Avg Abs Diff	18.4%	18.8%
NCPC Allocation	3.7%	4.1%

---

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

Figure 4 shows that the majority of months (15 out of 24 months in 2010 and 2011) exhibited a real-time premium, although the average real-time premium fell from nearly 2 percent of the average day-ahead price in 2010 to 1 percent in 2011. The reduction was partly attributable to increased system surplus capacity in 2011 since higher operating capacity margins generally lead to reduced price volatility and lower average real-time prices.[28]

We do not believe this result is efficient because day-ahead premiums 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 (such as when high-cost peaking resources are committed to satisfy the real-time demand). 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 satisfy the real-time system needs.

One reason for the real-time premiums in the past two years is that the average allocation of NCPC charges increased significantly in 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. Virtual trading patterns and NCPC allocations to virtual transactions are evaluated in the next subsection.

The second measure of price convergence evaluated in the figure above is the average absolute difference between day-ahead and real-time prices. This measure is calculated by averaging the absolute value of the hourly differences between day-ahead and real-time prices. As a percentage of the average day-ahead price in each year, the absolute differences averaged approximately 19 percent in 2011, consistent with 2010. The average absolute difference was particularly elevated during the peak summer months as one would expect due to the hotter conditions and higher associated load levels.

---

28 See Section VI.D.

E. Uplift Allocation and Virtual Trading

1. The Role of Virtual Trading

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 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.

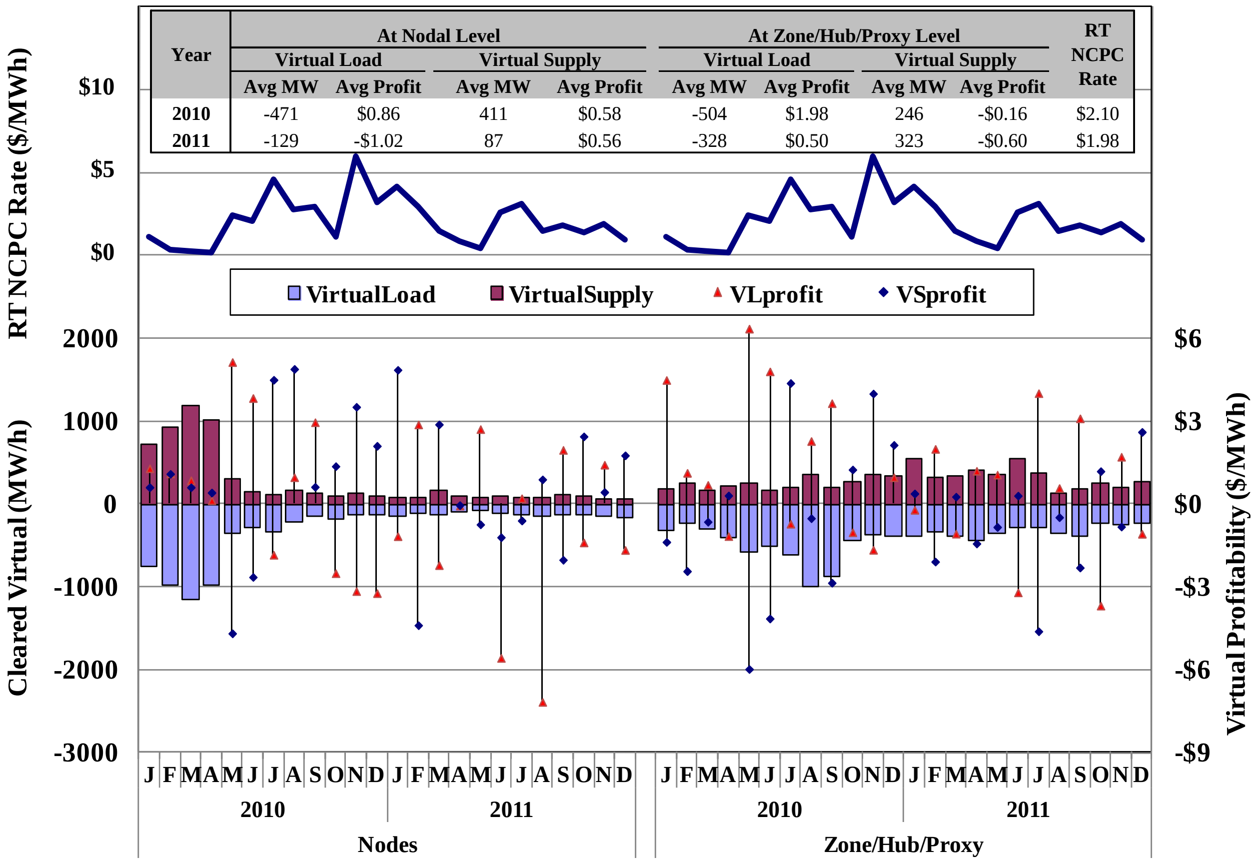
Virtual transactions that are scheduled in the day-ahead market are settled against real-time energy prices. Virtual demand bids are profitable when the real-time energy price is higher than the day-ahead price; likewise, virtual supply offers are profitable when the day-ahead energy price is higher 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.

ISO-NE allows virtual traders to schedule transactions at every pricing location, including individual nodes and more aggregated locations, such as the New England Hub and load zones. This provides flexibility for traders to arbitrage the price differences at various locations between day-ahead and real-time, leading day-ahead prices to converge with real-time prices.

2. Virtual Trading Levels and Profits

Figure 5 shows the average volume of virtual supply and demand that cleared the market in each month of 2010 and 2011 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 positions in the day-ahead market compared to the price at which these positions were covered in the 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.[29]

Figure 5: Virtual Transaction Volumes and Profitability
2010 – 2011



Year	At Nodal Level: Virtual Load Avg MW	At Nodal Level: Virtual Load Avg Profit	At Nodal Level: Virtual Supply Avg MW	At Nodal Level: Virtual Supply Avg Profit	At Zone/Hub/Proxy Level: Virtual Load Avg MW	At Zone/Hub/Proxy Level: Virtual Load Avg Profit	At Zone/Hub/Proxy Level: Virtual Supply Avg MW	At Zone/Hub/Proxy Level: Virtual Supply Avg Profit	RT NCPC Rate
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

The figure shows that scheduled virtual transactions decreased from 2010 to 2011. Overall, scheduled virtual load fell 53 percent and scheduled virtual supply fell 38 percent from 2010 to 2011. The decrease was due primarily to the decline in the volumes of virtual trading at the nodal level where scheduled virtual load and virtual supply fell 73 percent and 79 percent, 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.

29 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.

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.

The real-time NCPC rate increased substantially from $0.46 per MWh in the first four months of 2010 to $3.60 in the last eight months of 2010 and $1.98 in 2011. Several factors contributed to the increase in NCPC charges. First, supplemental commitment for system reliability rose substantially from May 2010 to December 2010 when a large flexible pump storage resource was unavailable, contributing to elevated NCPC rates in these months. Second, virtual trading at the nodal level declined substantially after the modeling inconsistency was remedied in May 2010. This caused the NCPC costs to be allocated to a smaller quantity of real-time deviations (which includes virtual transactions), thereby increasing the average NCPC charge rate.

The relatively high levels of NCPC charges that have been allocated to real-time deviations since May 2010 provide disincentives for firms to schedule virtual transactions. Hence, it is likely that some virtual traders have chosen not to schedule virtual trades (or increased the margin between their bid or offer and the expected real-time price) that would have improved the consistency between day-ahead and real-time prices.

Figure 5 shows that virtual traders earned profits of $14 million in 2010 but incurred losses of $1 million in 2011.[30] The average profit of cleared virtual supply at individual nodes was $0.56 per MWh in 2011, compared to a loss of $0.60 per MWh at more aggregated locations. Virtual load was considerably more profitable than virtual supply over the past two years, which was largely due to the prevailing real-time price premiums.[31] Virtual trades that are profitable generally contribute to better convergence between day-ahead and real-time prices. Unprofitable virtual trades can raise a number of potential concerns so we evaluate them in the next sub-section.

30 Virtual transactions would net a loss on average after paying NCPC charges in both 2010 and 2011 (e.g., gross profitability of all cleared virtual transactions was $0.98 per MWh in 2010 compared to an average real-time NCPC charge rate of $2.10 per MWh in 2010).

31 Gross 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) in each year.

ISO-NE currently allocates nearly all real-time "Economic" NCPC charges to deviations between the day-ahead and real-time schedules.[32] 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.

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 and 2011, 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 are working 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.

32 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.

3. 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 2011, we found that the firms with the most significant virtual losses 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.

F. Conclusion

Energy prices decreased 6 to 7 percent in 2011, driven primarily by lower natural gas prices, reduced load levels and increased net imports from neighboring areas. Very little transmission congestion occurred as the transmission investments made in recent years continued to provide substantial capability into historically constrained areas.

Differences between day-ahead and real-time prices were relatively small in 2011, but the sustained real-time premiums raised a potential concern that the market was 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

A key function of LMP markets is to set efficient energy prices that reflect the economic value of binding transmission constraints. These prices govern the short-term dispatch of generation and establish long-term economic signals that govern investment in new generation and transmission assets. Hence, a primary focus of this report is to evaluate locational marginal prices and associated congestion costs.

A. Congestion Costs and the Role of FTRs

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.[33]

Financial Transmission Rights (FTRs) can be used to hedge the congestion costs of serving load in congested areas or as speculative investments for purchasers who forecast higher congestion revenues between two points than the cost of the associated FTR. An FTR entitles a participant 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 the bilateral contract.

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

33 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.

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 decreased approximately 50 percent from 2010 to 2011 because congestion in the day-ahead market decreased. This decrease was largely due to lower natural gas prices and slightly lower load levels. Lower natural gas prices reduce the cost of redispatching resources to manage constraints, while lower load levels reduce the need to import power to import-constrained areas. Payments to FTR holders are funded by the congestion revenue collected by ISO-NE. In 2011, 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 2011, 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.

B. 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.

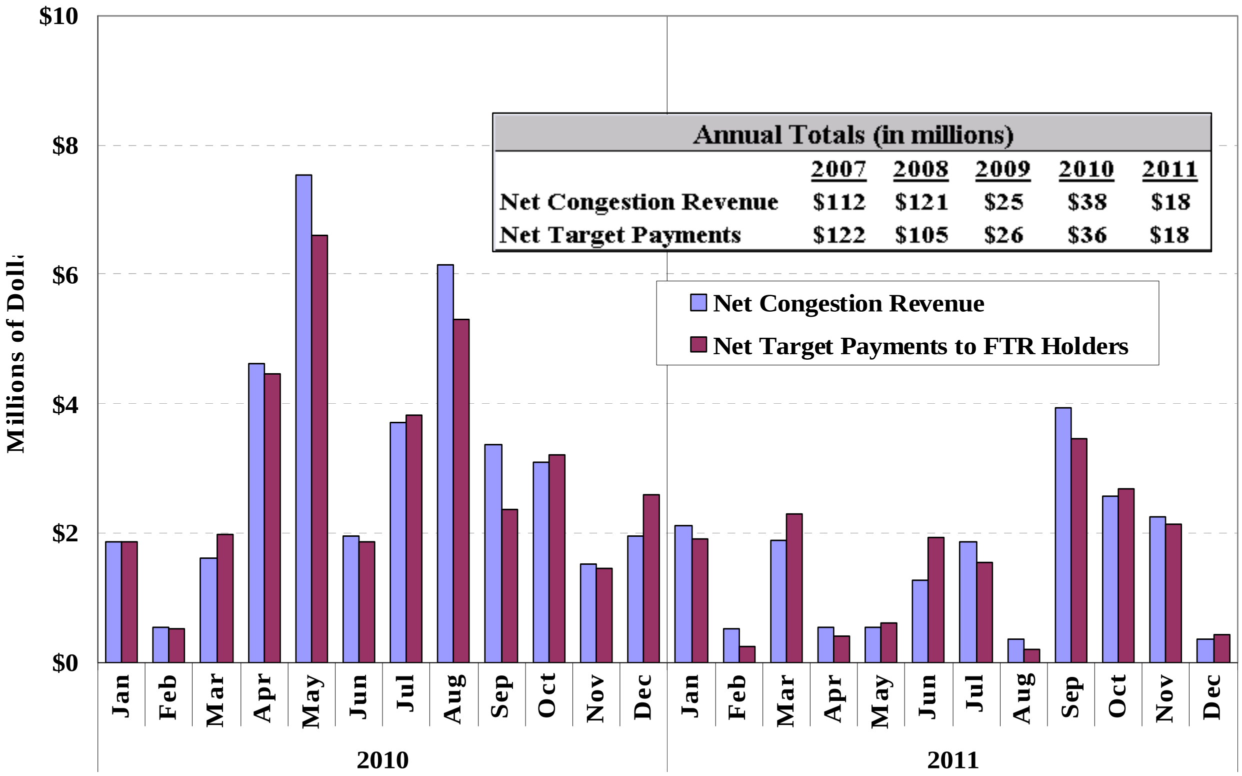
Day-ahead congestion revenue is equal to the megawatts scheduled to flow across a constrained interface times the day-ahead shadow price (i.e., the marginal economic value) of the interface. Real-time congestion revenue is equal to the change in scheduled flows (relative to the day-ahead market) across a constrained interface times the real-time shadow price of the interface. 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 6 compares the net congestion revenue collected by the ISO-NE with the net target payments to FTR holders in each month of 2010 and 2011. 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 capacity 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 6: Congestion Revenue and Target Payments to FTR Holders
2010 – 2011



The net congestion revenue decreased from approximately $38 million in 2010 to $18 million in 2011, a 52 percent decrease. Likewise, the net target payments to FTR holders decreased from $36 million in 2010 to $18 million in 2011. The decrease in congestion in 2011 was primarily due to two factors. First, load levels decreased, particularly in the summer, reducing the frequency of congestion into import-constrained areas. Second, natural gas prices decreased, reducing redispatch costs and associated congestion-related price differences.

As shown in the inset table, the net congestion revenues after 2009 were substantially less than in prior years due 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 subsection B below.

The figure also shows that net congestion revenues were more than the net target payments to FTR holders in most months during 2010 (8 months) and 2011 (7 months). As a result, the total net congestion revenues for the 12 months in both 2010 and 2011 were sufficient to fund 100 percent of the net target payments.

Congestion revenues exceed the net target payments to FTR holders when the amount of FTRs purchased along a congested transmission path is lower than the actual transfer capability in the day-ahead and real-time markets. For example, assume 1,000 MW of FTRs are sold into a constrained area because that is the normal limit into the area. If the interface is increased to 1,100 MW in the day-ahead market and the interface is congested, ISO-NE will collect 110 percent of the congestion revenue it needs to satisfy the target payments to the holders of the FTRs into the constrained area.

C. 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 7 shows day-ahead and real-time congestion prices and FTR prices for each of the eight ISO-NE load zones and five sub-areas of interest in 2011. 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. Aside from the eight load zones, the figure shows prices for Boston and four areas within Connecticut. Connecticut is divided into: East Connecticut, the portion of West Connecticut that excludes Southwest Connecticut, the portion of Southwest Connecticut that excludes Norwalk-Stamford, and Norwalk-Stamford. 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. The table compares the average day-ahead and real-time congestion prices and FTR prices from the New England Hub to Connecticut in 2010 and 2011.

Figure 7: FTR Auction Prices vs. Day-Ahead and Real-Time Congestion
Average Difference from New England Hub in Peak Hours, 2011

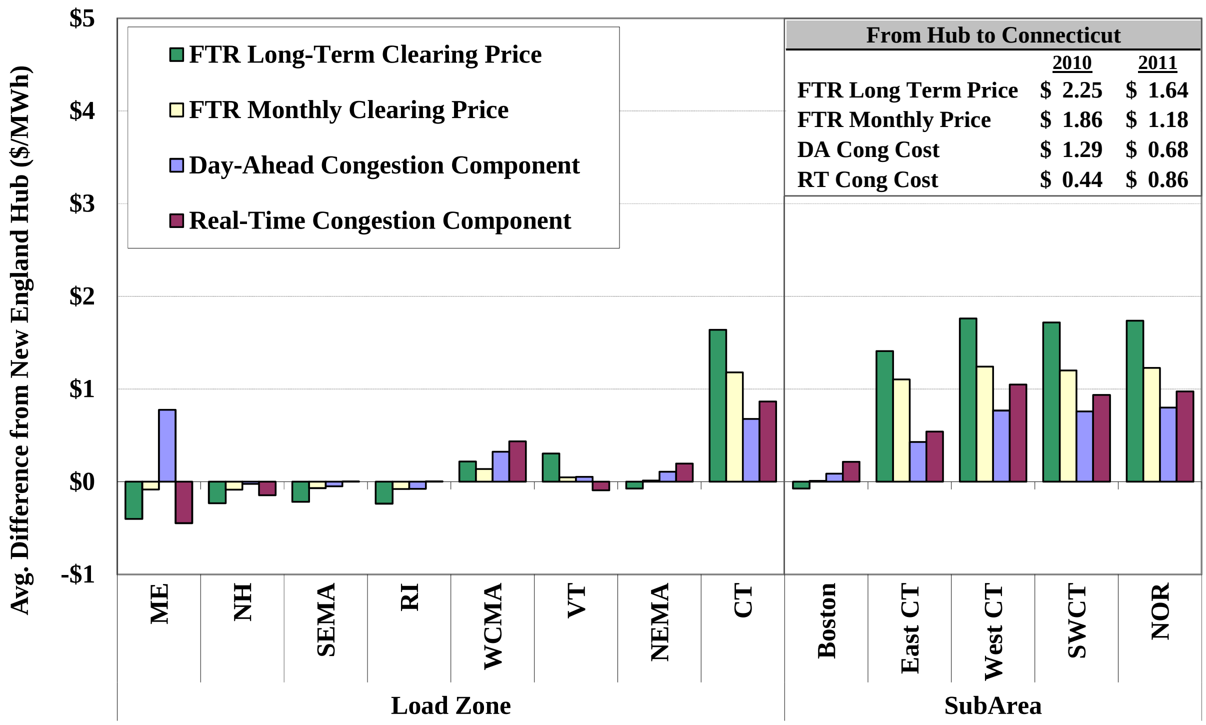


Figure 7 shows that in most areas during 2011, monthly FTR prices were more consistent with congestion prices in the day-ahead market than were annual FTR prices. For example, the annual FTR prices from the New England Hub to the various areas in Connecticut were $0.94 to $0.99 per MWh higher than the corresponding day-ahead congestion values, while the monthly FTR prices were only $0.43 to $0.67 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 inset table shows that monthly FTR auction prices exceeded the day-ahead and real-time congestion prices from the New England Hub to Connecticut by a significant margin in both 2010 and 2011. This suggests that participants' expectations in the monthly auctions were higher than the congestion in the day-ahead and real-time markets in both years. The over-estimates of congestion in the monthly FTR auctions were likely influenced by the congestion that occurred in prior periods. Congestion has decreased substantially over the past three years due to significant transmission upgrades that were completed between 2007 and 2009.

Given the volatile nature of congestion patterns and the variations in congestion patterns resulting from transmission upgrades in recent years, we found that FTRs were reasonably valued in the FTR auctions. The FTR market responded to changes in patterns of day-ahead congestion during 2011. Although this response has been slow, we conclude that the FTR markets performed reasonably well in 2011.

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.[37] 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). 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 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 market, the forward reserve market, and the regulation market.

[37] These requirements are intended to satisfy the reliability standards promulgated by the North American Electric Reliability Corporation (NERC) and the Northeast Power Coordinating Council, Inc. (NPCC).

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 must hold an amount of 10-minute reserves (i.e., TMSR plus TMNSR) equal to the largest generation contingency on the system, which averaged 1,325 MW in 2011. Based on system conditions, the operator determines how much of the 10-minute reserve requirement to hold as spinning reserves. ISO-NE held an average of 39 percent of the 10-minute reserve requirement in the form of spinning reserves during intervals with binding TMSR constraints in 2011.[38]

The ISO must hold an amount of 30-minute reserves (i.e., TMSR plus TMNSR plus TMOR) equal to the largest generation contingency on the system plus half of the second-largest contingency on the system. The 30-minute reserve requirement averaged approximately 2,010 MW in 2011. 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 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.

38 The TMSR requirement is binding when a non-zero cost is incurred by the market to satisfy the requirement. This occurred in 3.8 percent of the intervals in 2011.

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 be no lower than $10 per MWh. The marginal system cost used to schedule the reserves and set reserve clearing prices is equal to the marginal redispatch cost of the resource. When excess reserves are available without incurring any costs, reserve clearing prices will be $0 per MWh.

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 because a megawatt of TMSR would help satisfy all three constraints. Likewise, the TMNSR clearing price would be $12 per MWh 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 meets a significant share of its local area reserve requirements with imported reserves. For example, imported reserves satisfied 34 percent of the Connecticut requirement during constrained intervals in 2011.

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:

- $100 per MWh for the system-level 30-minute reserve constraint;
- $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.[39]

These values are differentiated to reflect values of the reserves and the reliability implications of shortages in the various classes of reserves. Furthermore, 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.[40] For example, shortages of both 30-minute and 10-minute reserves would produce a clearing price of

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

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

$950 per MWh for the system-level 10-minute reserves ($100 plus $850 per MWh) and energy prices likely exceeding $1,000 ($950 plus the marginal price 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 close to the ISO-NE market's energy offer cap of $1,000 per MWh during sustained periods of significant operating reserve shortages. The use of RCPFs to set efficient prices during operating reserve shortages has been endorsed by FERC.[41]

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 $150 per MWh, the real-time market would not schedule sufficient reserves to meet the requirement and the reserve clearing price would be set to $100 per MWh. This is efficient as long as 30-minute reserves truly have a reliability value of $100 per MWh. However, if the operator intervenes at this point to maintain the required level of reserves at a cost greater than $100 per MWh, this out-of-market action will undermine the efficiency of the market because (a) it artificially lowers energy and reserve prices, and (b) it is more costly than the value of the reserves being maintained. Accordingly, ISO-NE and NEPOOL filed on March 22, 2012 to adjust the RCPF for the system-level 30-minute reserve requirement from $100 per MWh to $500 per MWh, effective June 1, 2012. If approved by FERC, this adjustment will lead to more efficient real-time pricing and scheduling decisions for external transactions and internal resources.

3. Market Outcomes

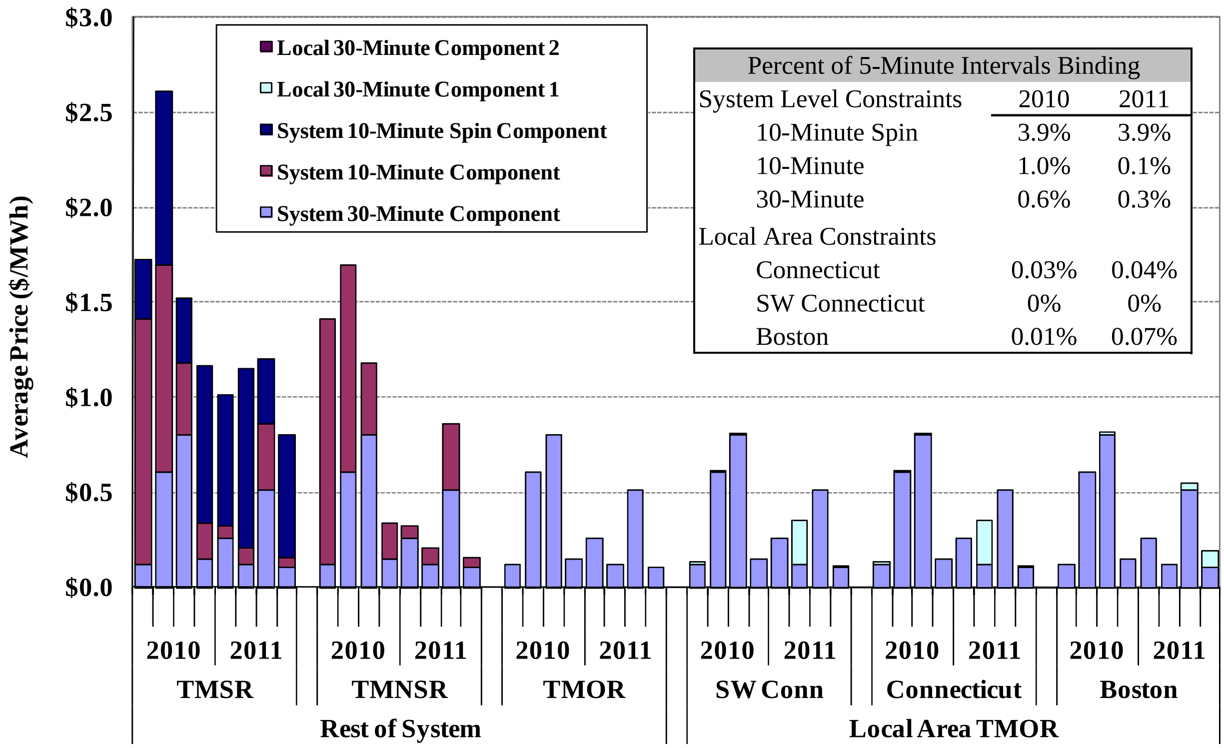
Figure 8 summarizes average reserve clearing prices in each quarter of 2010 and 2011. The left side of the figure shows prices outside the local reserve zones for all three service types. The right side of the figure shows prices in the three local reserve zones for TMOR only. Each price is broken into components associated with the underlying requirements. For example, the

---

41 *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).

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 8: Quarterly Average Reserve Clearing Prices by Product and Location
2010 – 2011



Percent of 5-Minute Intervals Binding		
System Level Constraints	2010	2011
10-Minute Spin	3.9%	3.9%
10-Minute	1.0%	0.1%
30-Minute	0.6%	0.3%
Local Area Constraints		
Connecticut	0.03%	0.04%
SW Connecticut	0%	0%
Boston	0.01%	0.07%

The figure shows that reserve constraints were binding infrequently in New England in 2010 and 2011. The most frequent binding constraint was the system-level 10-minute spinning reserve requirement, which was binding in roughly 4 percent of market intervals in both years. The other reserve requirements were binding in less than 1 percent of all intervals.

Overall, the clearing prices for operating reserves fell from 2010 to 2011. Outside the local constrained areas, the average prices for each reserve product decreased:

- TMSR prices fell from $1.76 per MWh in 2010 to $1.04 per MWh in 2011;

- TMNSR prices fell from $1.16 per MWh in 2010 to $0.39 per MWh in 2011; and
- TMOR price fell from $0.42 per MWh in 2010 to $0.25 per MWh in 2011.

The lower reserve clearing prices were primarily caused by increases in the amount of available (i.e., on-line or fast-start) capacity from 2010 to 2011. On average, the minimum daily surplus capacity increased from 750 MW in 2010 to 930 MW in 2011.[42] The amount of available capacity increased primarily because a large flexible pump storage resource returned to service following a seven-month outage in December 2010. The resource's return to service significantly increased the overall ramping capability of the system.

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 constraints have bound very infrequent since transmission upgrades were made in Boston in the spring of 2007 and in Connecticut between 2007 and 2009.

Average reserve clearing prices are relatively low because reserve clearing prices are $0 in the vast majority of real-time intervals. This reflects that there is 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 just 0.1 percent of intervals in 2011, indicating that the requirement can be met at no cost with surplus capacity in 99.9 percent of intervals.

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

42 Surplus capacity is the amount of capacity that is online or capable of starting quickly in excess of the amount required to satisfy load and reserve requirements. The minimum daily surplus capacity is the lowest quantity of surplus capacity that was available in any interval on a particular day.

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 $1.07 per MWh. This is composed of $1.04 per MWh for TMSR outside the local areas and the Boston 30-minute reserve component of $0.03 per MWh.

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 several reserve products in the Forward Reserve Market auction. There are two categories of forward reserve capacity: TMNSR and TMOR. The forward reserve market has five geographic zones: Boston, Southwest Connecticut, Connecticut, Rest of System (i.e., areas outside Connecticut and Boston), and the entire system (i.e., all of New England). With two exceptions, the reserve products sold in the forward reserve market are consistent with reserves in the real-time market. First, the forward reserve market has no TMSR requirement. Second, the forward reserve market has a minimum requirement for reserves in Rest of System, while the real-time market has no corresponding requirement. However, the minimum requirement for reserves in Rest of System was eliminated beginning in the Summer 2011 Procurement Period. [44]

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

44 See the FERC order at http://www.iso-ne.com/regulatory/ferc/orders/2010/dec/ er11-2038-000_12-17_10_ltr_order_accept_frm_enchance.pdf.

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.[45]

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 2011/12 Procurement Period (October 2011 to May 2012) was held in August 2011. 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.[46]
- For Rest of System (i.e., areas outside Connecticut and Boston), the minimum purchase requirement is 600 MW. This is multiplied by a factor that accounts for the typical availability of resources deployed for TMOR to determine the TMOR requirement for each auction.[47] However, this requirement was eliminated beginning in the Summer 2011 Procurement Period.
- 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

45 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.

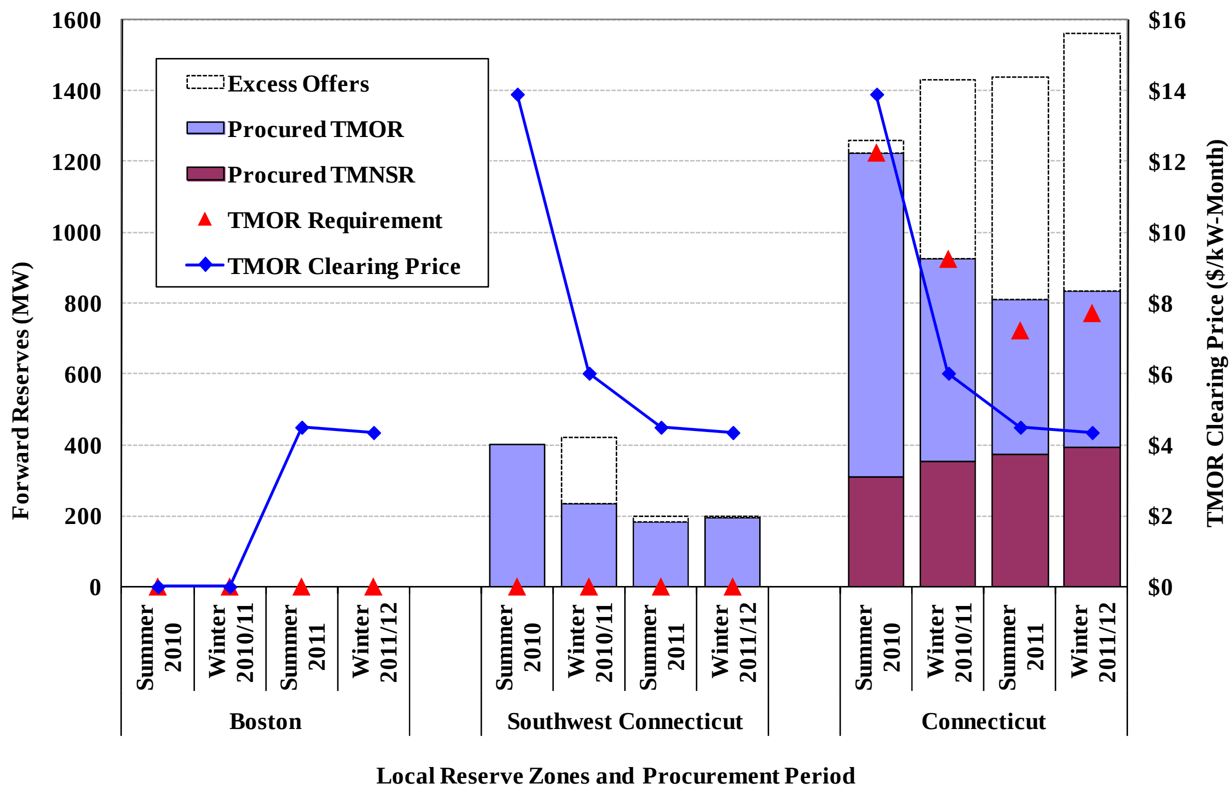
46 Usually, the forecasted largest contingency is the HQ Phase II Interconnection and the forecasted second largest contingency is the combination of the Mystic 8 and Mystic 9 generating units.

47 This factor is referred to as the "R-factor" which reflects the estimated availability of resources deployed to provide system-wide contingency reserves. The R-factor has been 1.33 for years and changed to 1.2 for the Winter 2010/11 auction, which corresponds to the availability discount factor for peaking resources that was filed by the ISO in February 2010.

priced according to the sum of the values of the underlying products, although this is limited by the $14 per kW-month price cap.[48]

The following two figures summarize the quantities purchased in the last four forward reserve auctions towards each requirement. Figure 9 shows auction outcomes for the three local reserve zones, and Figure 10 shows auction outcomes for the system-level and Rest of System requirements. For each local reserve zone in each procurement period, Figure 9 shows the TMOR clearing price, the quantity of TMOR and TMNSR procured, the forward reserve requirement, and the quantity of excess offers that was cleared in the auctions.

Figure 9: Summary of Forward Reserve Auction for Local Areas
Procurement for June 2010 to May 2012



48 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.

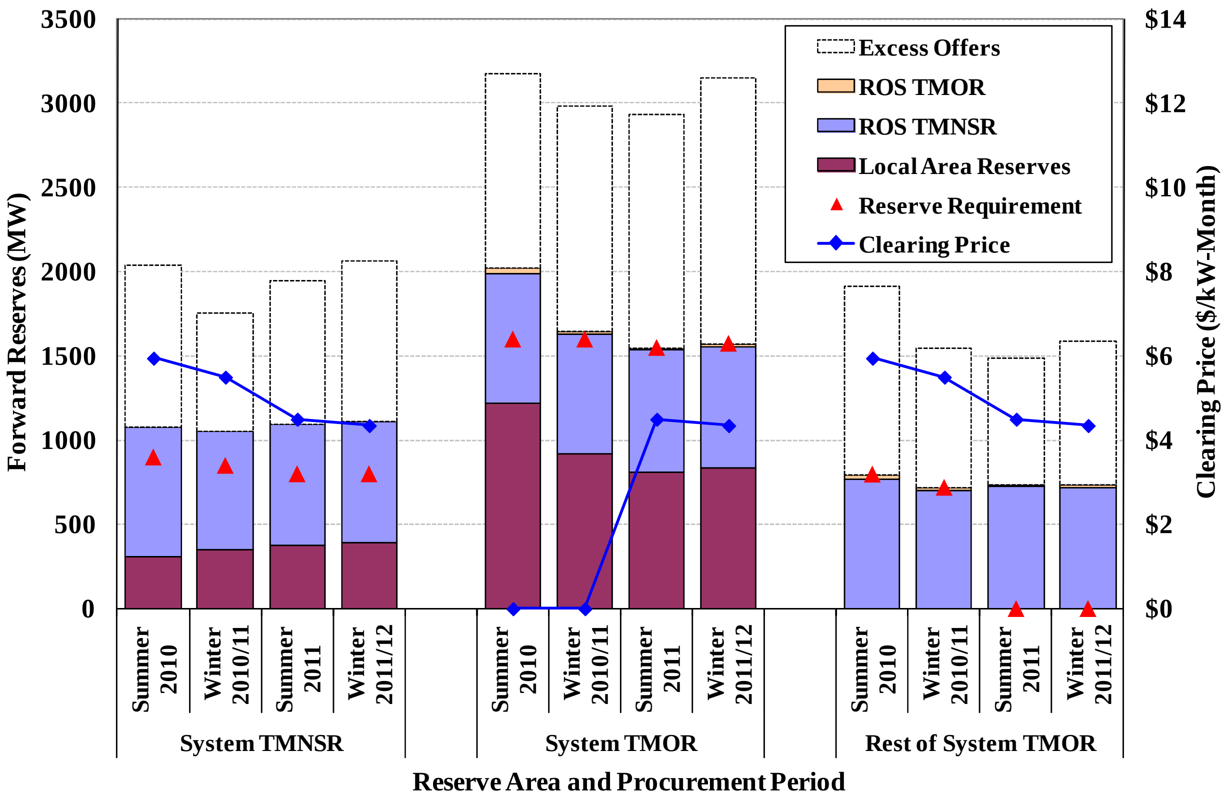
The figure shows that no resources were offered or procured in the Boston reserve zone in the last four auctions because there were no reserve requirements for internal Boston resources in these auctions. A substantial amount of transmission capability was added into the Boston area in 2007, leading ISO-NE to assume 830 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 because of External Reserve Support) in the last four auctions. Accordingly, the TMOR prices in Boston cleared at the same levels with the system TMOR prices.

In Connecticut, TMOR prices cleared below the $14 per kW-month cap for the first time in the Summer 2010 Procurement Period, since that was the first procurement period in which the Connecticut requirement was satisfied. The Connecticut reserve requirement fell over the last four procurement periods from 1,225 MW in the Summer 2010 period to 772 MW in the Winter 2011/2012 period. The reduction was due to the increases in assumed “External Reserve Support” to Connecticut, which is largely attributable to the new transmission capability added into the area. The offer quantities have increased over the last four procurement periods from nearly 1,260 MW in the Summer 2010 period to 1,560 MW in the Winter 2011/2012 period. This increase was due to the sales from new fast-start resources and increased participation by existing fast-start capacity. As a result, the clearing prices of TMOR fell from $13.90 per kW-month in the Summer 2010 period to $4.35 per kW-month in the Winter 2011/2012 period. In the last two auctions, TMOR clearing prices in Connecticut were the same as prices at the system level because the Connecticut reserve requirements were not binding.

The forward reserves procured for Southwest Connecticut are shown both separately and as a subset of the total procurement for Connecticut. The reserve requirement for internal resources in Southwest Connecticut was 0 MW in all four auctions. This is due primarily to transmission upgrades into Southwest Connecticut that were brought into service in early 2009. Starting in the Summer 2009 auction, nearly the entire requirement for Southwest Connecticut was satisfied by External Reserve Support. As a result, the local requirement was satisfied in all four auctions and TMOR prices cleared at the same levels with the TMOR prices in Connecticut.

Figure 10 shows the same analysis for the system-level and Rest of System requirements. For each procurement period, Figure 10 shows the TMOR clearing price, the quantity of TMOR and TMNSR procured, the forward reserve requirement, and the quantity of excess offers that was sold in the auctions.

Figure 10: Summary of Forward Reserve Auction for Outside Local Areas
Procurement for June 2010 to May 2012



Outside of the local reserve areas, the forward reserve requirements were satisfied in each auction. In the first two auctions shown in Figure 10, the Rest of System TMOR requirement was binding, while the system-level TMOR and TMNSR requirements were not binding because they were satisfied by the purchases for other requirements (i.e., no additional costs had to be incurred or purchases made to satisfy the system-wide TMOR and TMNSR requirements). Hence, TMNSR and Rest of System TMOR sold at the same price.

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). 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.

The amount of TMOR offers accepted outside Connecticut was very limited, accounting for only 1 percent of total procured forward reserves in the four auctions. The procurement from Connecticut accounted for 56 percent, and the TMNSR procured from Rest of System accounted for the remaining 43 percent.

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.[49] 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 2011. Suppliers that do not meet their forward reserve obligations incur a Failure to Reserve Penalty.[50]

49 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 $5.55 per MMbtu and the forward reserve heat rate was 15.31 MMbtu per MWh for December 2011. Hence, it resulted in a Forward Reserve Threshold Price of approximately $85 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.

50 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.

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.[51] Suppliers can reduce this risk by 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.

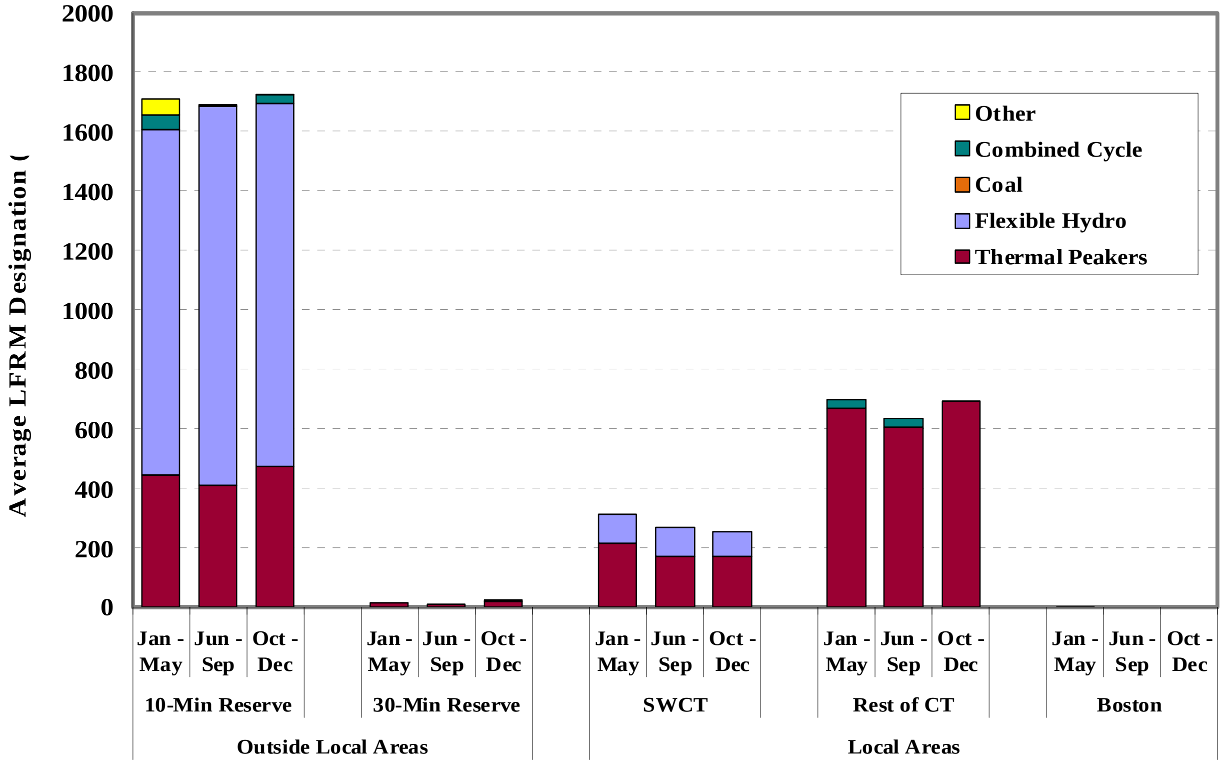
Third, suppliers may forego profitable energy sales as a result of offering incremental energy at the Forward Reserve Threshold Price. For instance, suppose the Forward Reserve Threshold Price is $100 per MWh and a supplier assigns a generator that has incremental costs of $60 per MWh to provide forward 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.

51 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.

The following analysis evaluates how market participants satisfied their forward reserve obligations in 2011 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
2011



Approximately 98 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 (1.7 percent) of the forward reserves in 2011. 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, the 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 62 MW in 2011.

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 examines the pattern of supply offers from regulation providers. The end of this subsection summarizes our conclusions related to the regulation market.

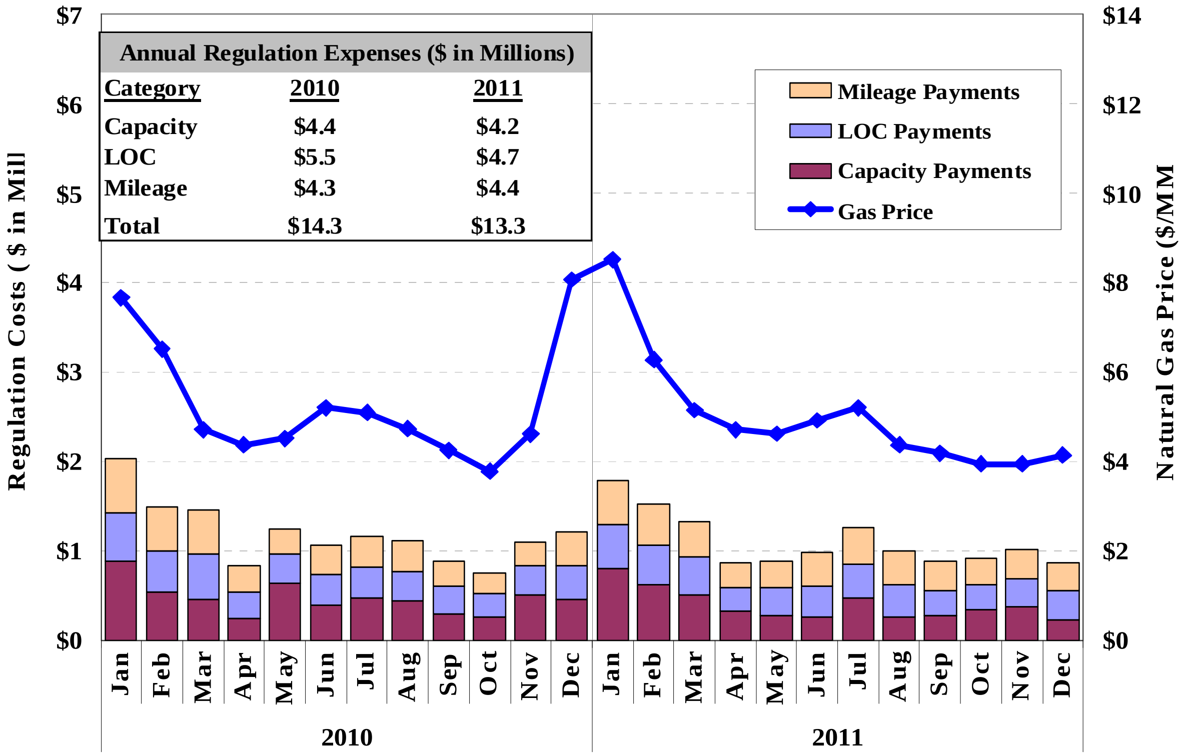
1. Regulation Market Expenses

Resources providing regulation service receive the following payments:[52]

- 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 2010 and 2011. The figure also shows the monthly average natural gas price.

Figure 12: Regulation Market Expenses
2010 – 2011



52 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 7 percent from $14.3 million in 2010 to $13.3 million in 2011, which was due to:

- A 9 percent reduction in average regulation requirement in 2011 from 2010;
- An increase in low-cost regulation offers after new entry in Connecticut; and
- A large pump storage resource returning to service after a lengthy outage in 2010.

The figure also shows that variations in monthly regulation market expenses were 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 2010 and 2011. 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:

- Estimated Capacity Payment- 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.[53]

On October 20, 2011, FERC issued Order 755 on Frequency Regulation Compensation, which will require ISO-NE and other ISOs to make certain modifications to their market designs.[54] 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 expect to file modifications by April 30, 2012 to comply with the Order.[55]

This part of the section evaluates the offer patterns of regulation suppliers in 2011. 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

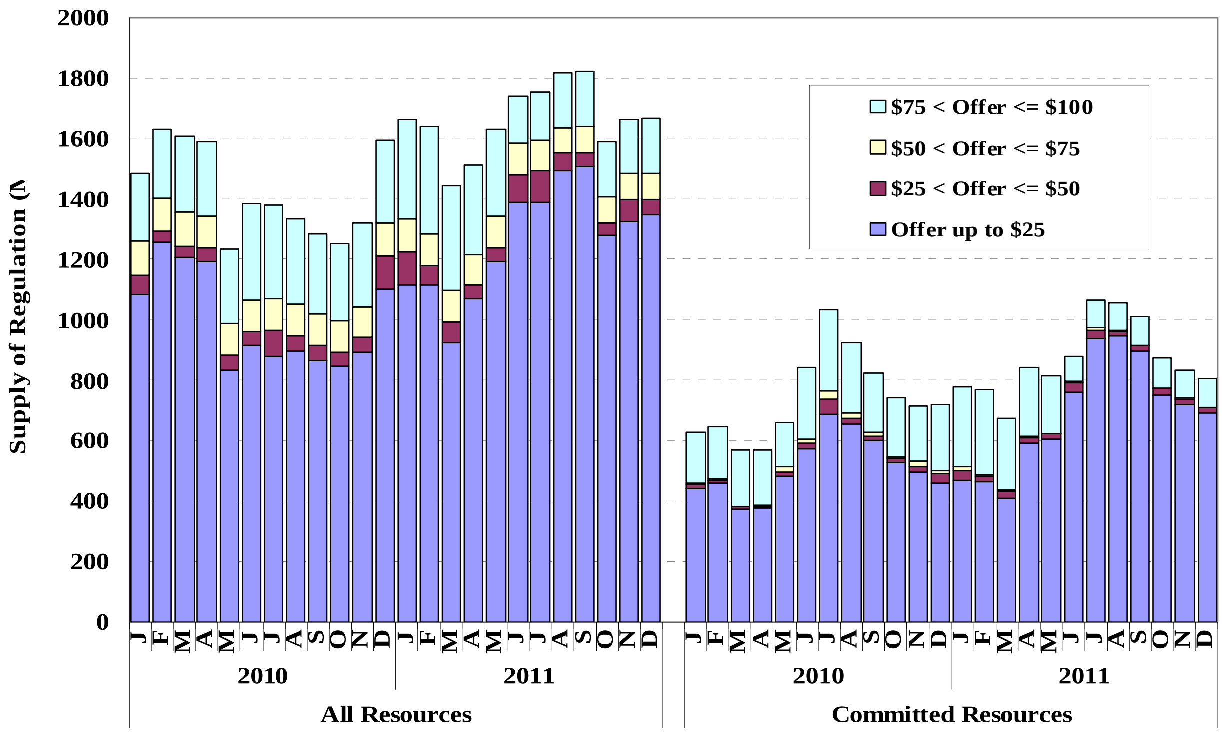
53 However, if the RCP rises from one iteration to the next, the model will use the previous iteration to rank resources. For additional details, see Section 3.2.5 of ISO-NE Manual M-11 on Market Operations.

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

55 For a description of the proposed modifications, see Participants Committee Meeting Materials for April 13, 2012, Agenda Item #8.

regulation offers from all resources and from online resources. Figure 13 shows monthly averages of the quantity of regulation offered into the market in 2010 and 2011 for two categories of offers. 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 colors on the bars in the chart show the average quantities offered by offer price range.

Figure 13: Monthly Average Supply of Regulation
2010 – 2011



In Figure 13, the left panel 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 1,200 MW and 1,800 MW in most months of 2010 and 2011. In 2010, however, the average quantity of total regulation offers was approximately 270 MW lower in the months of May to November than in other months. This was primarily because a large flexible generator was unavailable during the period.

The portion of regulation offers in each price range was also relatively consistent over the past two years, except that the low-cost portion of regulation offers decreased from May to November

2010 due to the unavailability of a large flexible resource. On average, 76 percent of the total regulation offers were below $25 per MWh, 10 percent were between $25 and $75 per MWh, and 14 percent were more than $75 per MWh in 2011.

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. On average, approximately 52 percent of the regulation offered in the day-ahead market was available to the hourly real-time selection process in 2011. 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.

The average quantity regulation offers from online resources increased roughly 17 percent from 2010 to 2011 as a large flexible pump storage resource returned to service in December 2010 that generally supplies a significant portion of ISO-NE's regulation service. In addition, natural gas prices fell while coal prices rose significantly in 2011. The narrowing spread between coal prices and natural gas prices combined with the better fuel efficiency of most gas-fired combined cycle units led them to be committed more frequently. This increased the supply of regulation from online units because gas combined cycle units provide most of the regulation capability in New England.

During both 2010 and 2011, significantly more regulation capability was offered into the market than was actually procured 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.

D. Conclusions

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 TMSR clearing price fell from roughly $1.75 per MWh in 2010 to $1.04 per MWh in 2011. This decrease was partly due to the increase in the average surplus capacity on the system in real time, which is evaluated and explained in Section VI.D of this report. 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.

In the forward reserve market, the TMOR clearing price in Connecticut fell 49 percent from the 2010/11 Capability Period to the 2011/12 Capability Period. Given that forward reserve payments are adjusted based on capacity prices, the effective net price of forward reserves in Connecticut actually fell 71 percent, to $1.28 per kW-month, in the 2011/12 Capability Period. Connecticut prices fell because:

- The Connecticut reserve requirement fell from an average of 1,025 MW in the 2010/11 Capability Period to 756 MW in the 2011/12 Capability Period. The local reserve requirement decreased in 2011 because the recent transmission upgrades in Connecticut reduced the local forward reserve requirements in the area; and
- Offer quantities increased due to the introduction of new peaking capacity and increased participation by existing peaking capacity.

TMOR clearing prices outside Connecticut fell 22 percent because of the decrease in forward capacity prices and the elimination of the Rest of System TMOR requirement, which reduced the amount of reserves that must be purchased from outside Connecticut and Boston. As expected, we found that 98 percent of the resources assigned to satisfy forward reserve obligations in 2011 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 (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, so the obligations of forward reserve suppliers are not consistent with the day-to-day operational needs of the system. Consequently, excess reserves are available on some days in local areas and system wide, while insufficient reserves are available on other days and require out-of-market commitments for reliability.

Overall, the regulation market performed competitively in 2011. On average, approximately 870 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." ISO-NE plans to comply with the Order by: (i) altering the criteria for minimizing the total expected cost of procuring regulation to incorporate a separate movement offer, and (ii) paying regulation providers according to a Vickrey-style auction rather than by setting market clearing prices. 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 2011. 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 helps lower the costs of meeting reliability standards 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 several aspects of transaction scheduling between ISO-NE and adjacent control areas. Section A summarizes scheduling between New England and adjacent areas in 2011. 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 2010 and 2011. Figure 14 shows the average net imports across the three interfaces with Quebec and New Brunswick by month, for peak and off-peak periods.[56] The net imports across the two interfaces linking Quebec to New England are combined.

Figure 14: Average Net Imports from Canadian Interfaces
2010 – 2011

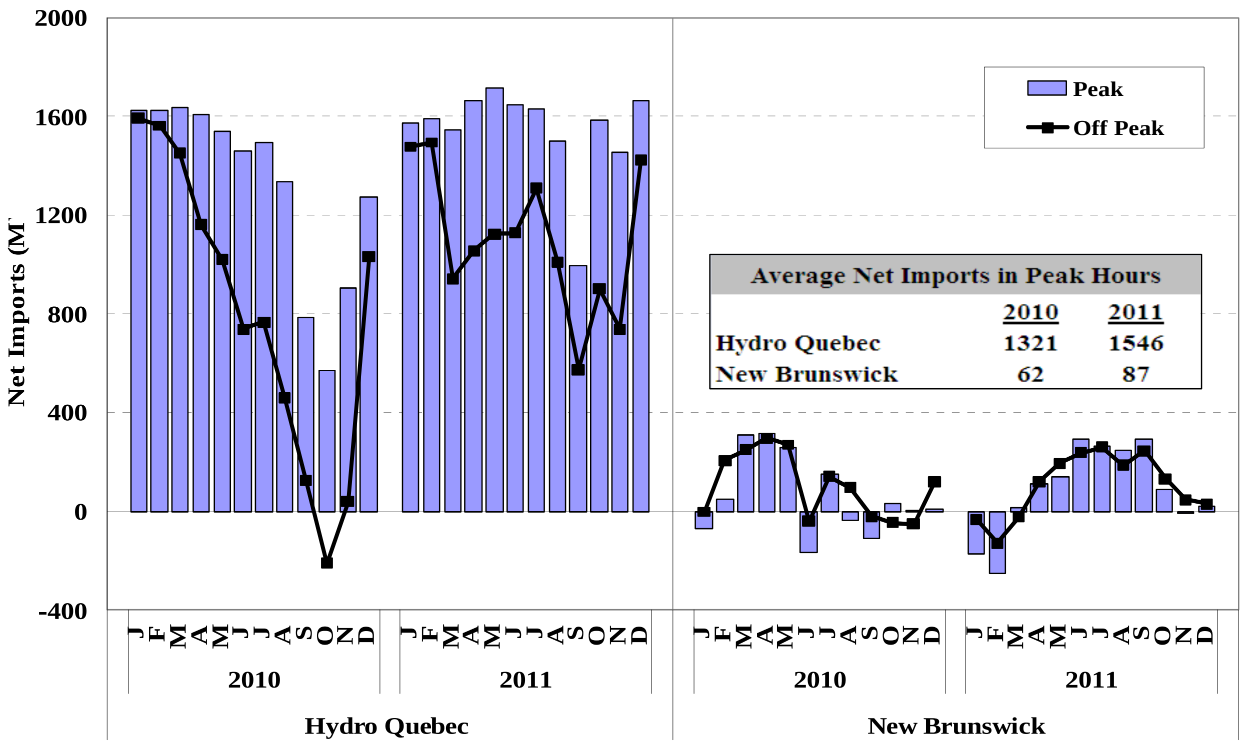


Figure 14 shows that power is generally imported from Quebec and New Brunswick. Across the two interfaces with Quebec, average net imports were higher during peak hours than during off-peak hours by roughly 510 MW in 2010 and by 450 MW in 2011. This reflects the tendency for hydro resources in Quebec to store water during low demand periods in order to make more

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

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 2011, when average net imports 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. Net imports over the New Brunswick interface were much lower than over the Quebec interfaces and did not vary significantly from peak to off-peak hours.

Figure 15 shows average net imports across the three interfaces with New York by month in 2010 and 2011 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
2010 – 2011

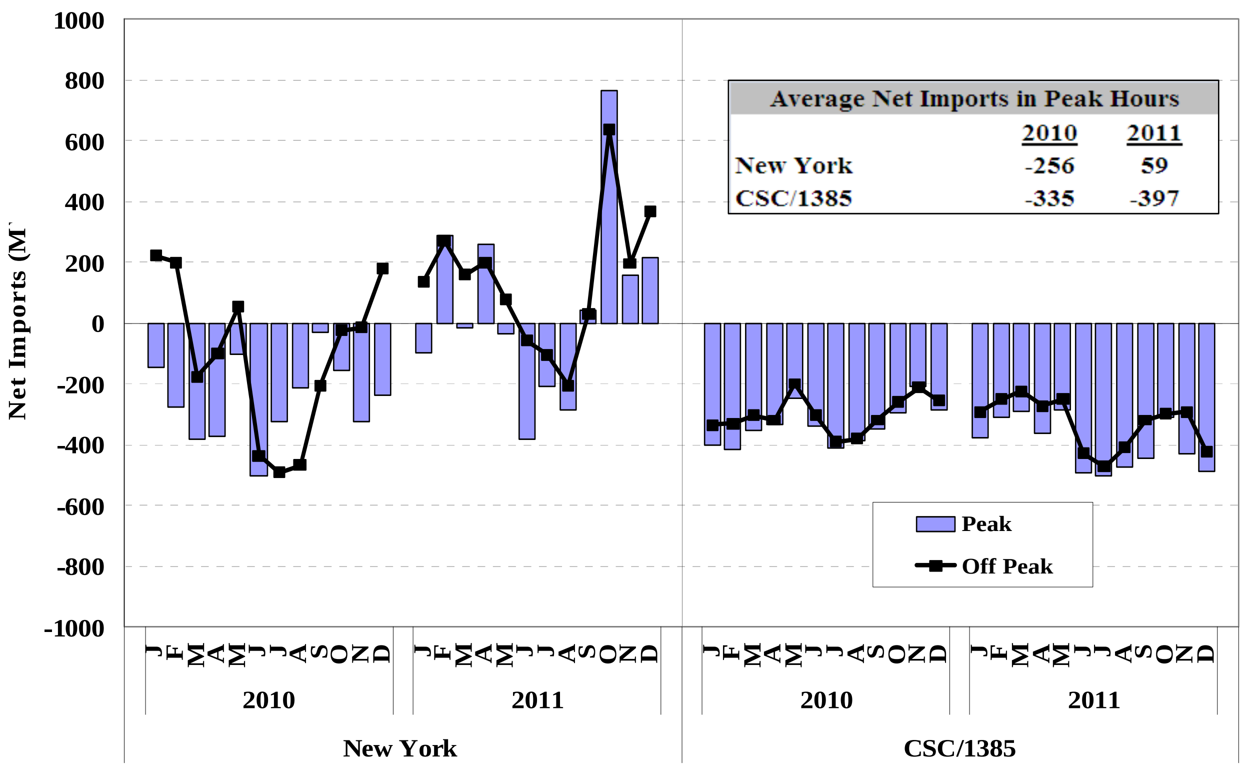


Figure 15 shows that the direction and the level of flows varied considerably across the primary interface with New York (i.e., the Roseton interface) during the past two years, reflecting the variations in relative prices in the two markets. On average, ISO-NE was a net exporter to New

York across the primary interface in 2010, and a net importer in 2011. Also, ISO-NE tends to import more power from (or export less power to) New York in the winter months. This is partly because ISO-NE is more reliant on natural gas generation, which is typically most expensive in the winter months.

The figure also shows that flows were relatively consistent from New England to Long Island across the Cross-Sound Cable and the 1385 Line, averaging approximately 320 MW in 2010 and 360 MW in 2011. The Cross-Sound Cable and the 1385 Line have transfer capabilities of 330 MW and 200 MW, respectively. Both lines are usually fully utilized to export power to Long Island when they are in service. Accordingly, the average level of exports across the 1385 Line increased in June 2011 after the completion of upgrades that increased the normal transfer capability from 100 MW to 200 MW.

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 when small amounts of additional imports can substantially reduce prices.

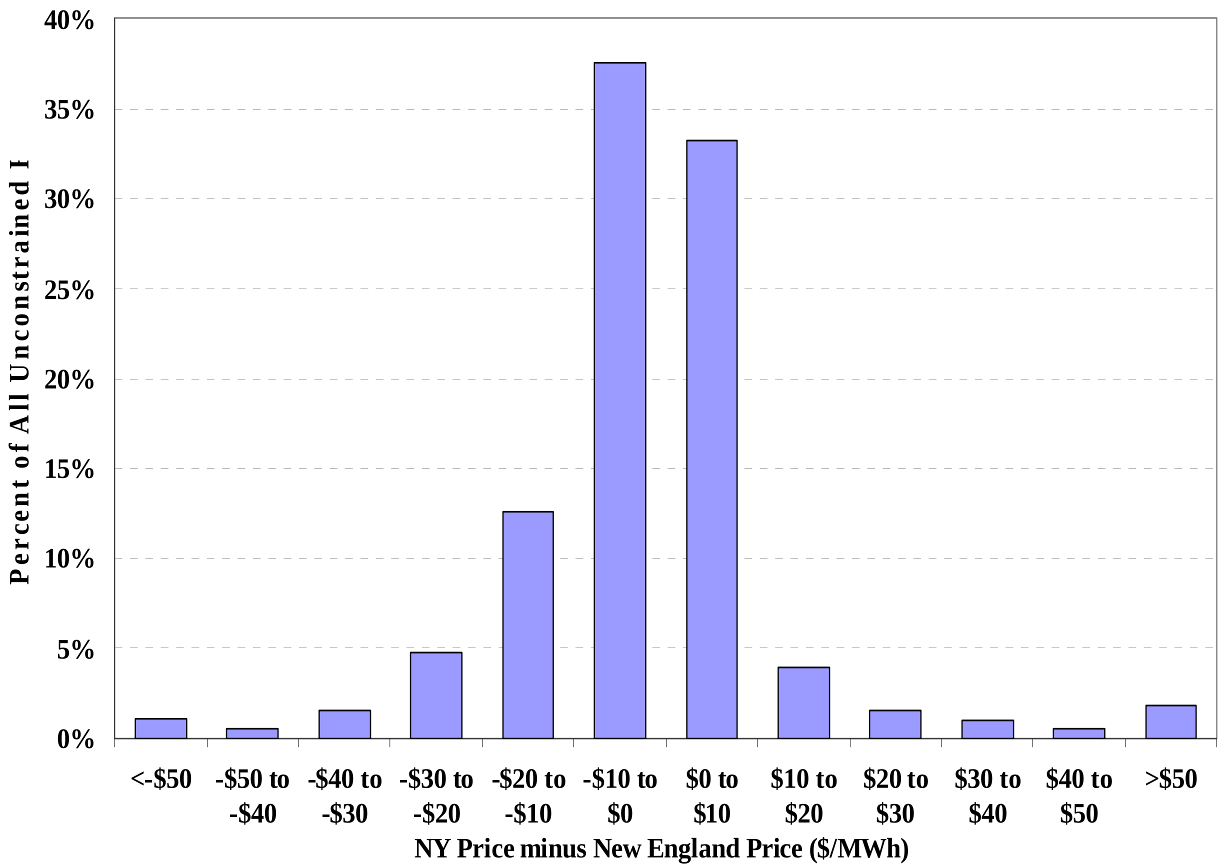
Several factors prevent real-time price differences between New England and New York from being fully arbitraged. First, market participants may not be able to predict which side of the interface will have a higher real-time price at the time when transaction bids and offers must be submitted. Second, differences in the procedures and timing of scheduling in each market serve as barriers to full arbitrage. Third, 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. Last, risks associated with curtailment and congestion can reduce the incentives of participants 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.

The following figures focus on the efficiency of scheduling across the primary interface between New England and New York. The Cross-Sound Cable is not evaluated in the following figures because it is scheduled under separate rules.[57] The 1385 Line is also not included because it was usually fully scheduled from Connecticut to Long Island in 2011. Hence, it was normally transmission constrained and, therefore, generally not responsive to price variations. Figure 16 shows the distribution of real-time price differences across the primary interface between New England and New York in hours when the interface was not constrained.[58]

57 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.

58 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 Upstate New York
Unconstrained Hours, 2011



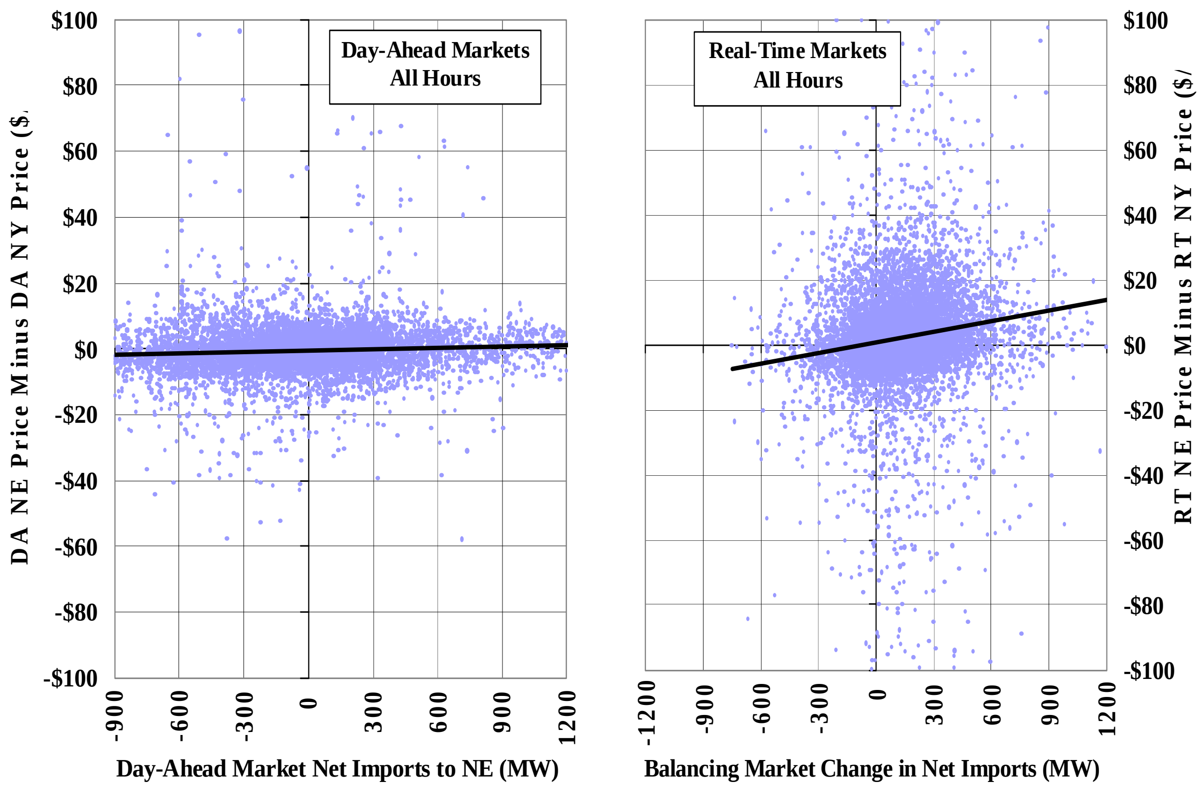
While the factors described above prevent complete arbitrage of price differences between regions, trading should help keep prices in the neighboring regions from diverging excessively. Nonetheless, Figure 16 shows that approximately 30 percent of the unconstrained hours have real-time price differences of greater than $10 per MWh. In 4 percent of the hours, the price difference is greater than $40/MWh.

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.

Although market participants have not fully arbitraged the interface between New York and New England, the next analysis evaluates whether the incremental changes in participants' schedules have been consistent with the relative prices in the two regions and have, therefore, improved price convergence and efficiency.

Figure 17 shows the net scheduled flow across the interface versus the difference in prices between New England and upstate New York for each hour in 2011. 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, 2011



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 market and in the real-time market. However, the correlation in the day-ahead market is extremely weak, which indicates the difficulty participants have in scheduling transactions efficiently. The correlation is stronger in the 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, although this response is highly variable.

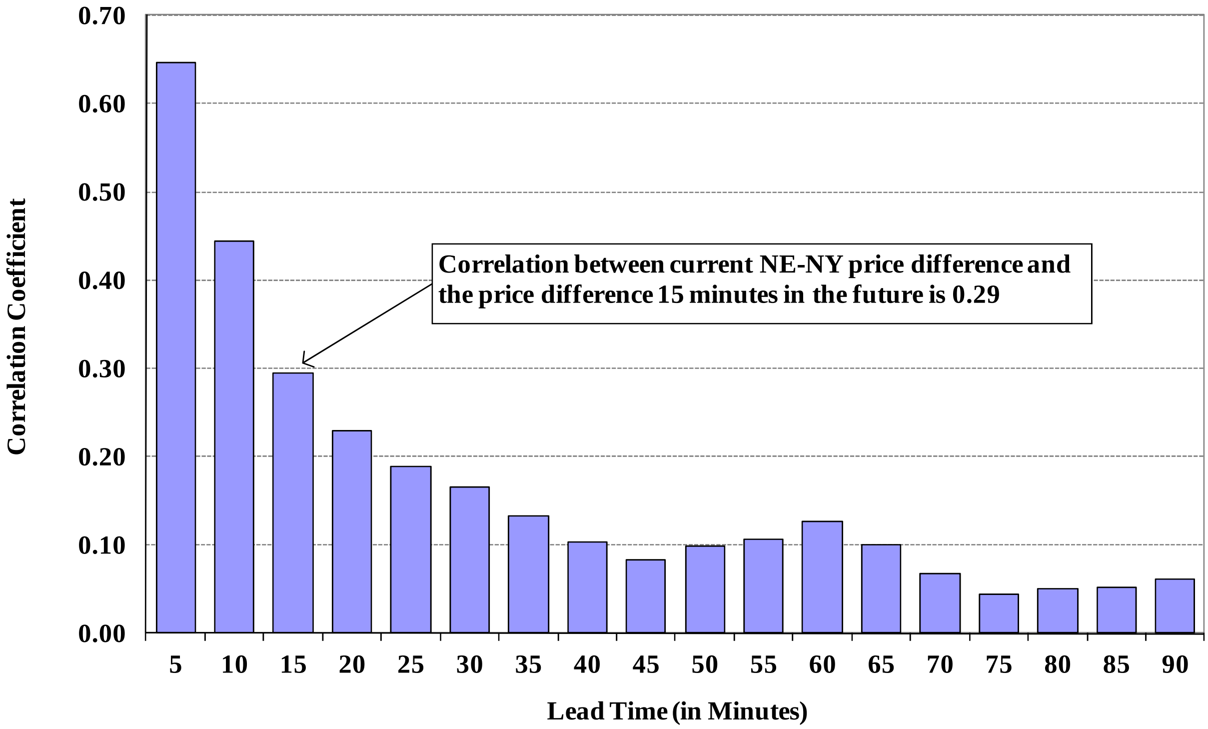
The difficulty of predicting changes in market conditions in real-time is reflected in the wide dispersion of points most notable on the right side of Figure 17. More than 41 percent of the points in the real-time market panel are in unprofitable quadrants – upper left and lower right – indicating hours 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 shows 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 2011 were $2.3 million in the day-ahead market and $7.6 million in the real-time market (not accounting for transaction costs).[59] 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.

The next analysis examines the correlation between the lead times for scheduling transactions and the predictability of price differences between adjacent markets. Figure 18 **Error! Reference source not found.**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 0.29 in 2011.

59 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, 2010



Not surprisingly, Figure 18**Error! Reference source not found.** 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 less than 0.3 at 15 minutes ahead of real time, which is the shortest scheduling lead time currently used by any RTO. 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. [60]

Through their respective stakeholder processes, ISO-NE and NYISO decided to move forward with the CTS proposal to improve coordination between markets.[61] Accordingly, a market design project for CTS is currently under way and may be effective as soon as late 2013.[62] 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 (CTS) process is being implemented 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

---

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

61 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.

62 See the 2012 ISO-NE Wholesale Markets Project Plan, page 19.

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 the efficient procurement of 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. Whenever possible, operations should be performed in a manner that results in efficient real-time price signals. This is because efficient real-time price signals 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 2011. 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;
- Frequency of price corrections; and
- Efficiency of real-time ex post prices.

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. This 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 ramp to their maximum output level more quickly than most baseload units, and better enable the system operator to respond rapidly to unexpected changes in load. Such operating conditions can result in especially tight market conditions, making it 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 2011 (because fast-start peaking units are relatively inflexible once they are started, they frequently do 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.[63] 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.[64] 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.[65]

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

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

64 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.

65 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.

UDS schedules fast-start generators using the following criteria:

- Offline fast-start generators- UDS considers commitment costs by adding the amortized start-up and "no-load" offers to the incremental offer.[66] 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 the 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 is even lower than the incremental offer of the fast-start generator.

The following example illustrates the challenges for efficient pricing when fast-start generators are economically deployed by the real-time market. Suppose UDS needs to schedule an additional 15 MW in an import-constrained area and the most efficient way is to start up a fast-start generator with an incremental offer price of $75 per MWh, a no-load offer price of $300/hour, a start-up offer price of $500, a minimum output level of 18 MW, and a maximum

66 For example, suppose a 20 MW fast-start unit has an incremental offer of $75 per MWh, a no-load offer of $300/hour, and a start-up offer of $500 (which UDS amortizes over one hour). The average total offer of the unit is $115 per MWh = ($75 per MWh + $300/hour ÷ 20 MW + $500/hour ÷ 20 MW).

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. 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. As a result, the owner of the fast-start unit would receive an NCPC payment to make up the difference between the total offer and the real-time market revenue, resulting in additional uplift charges to the market.

2. Evaluation of Fast-Start Deployments by UDS in 2011

The following two analyses assess the efficiency of real-time pricing during periods when fast-start units were deployed in merit order in 2011. The first analysis summarizes the quantities of fast-start capacity deployed economically by UDS on average each day. It also examines the extent to which the real-time LMP revenues that such units receive are consistent with their total offers. The second analysis evaluates how real-time prices would be affected if the average total offers were fully reflected in real-time LMPs.

The first analysis shown in Figure 19 summarizes the consistency of the average total offer (including no-load and start-up costs amortized for 1 hour) of fast-start generators that were deployed economically by UDS with the average real-time LMP 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 such occurrences in four categories that exclude the start-up component of the offer. These categories are shown according to the size of the difference between the average total offer and the average real-time LMP. This comparison is shown separately for hydro and thermal peaking units in each month.

Figure 19: Comparison of Real-Time LMPs to Offers of Fast-Start Generators
First Hour Following Start-Up by UDS, 2011

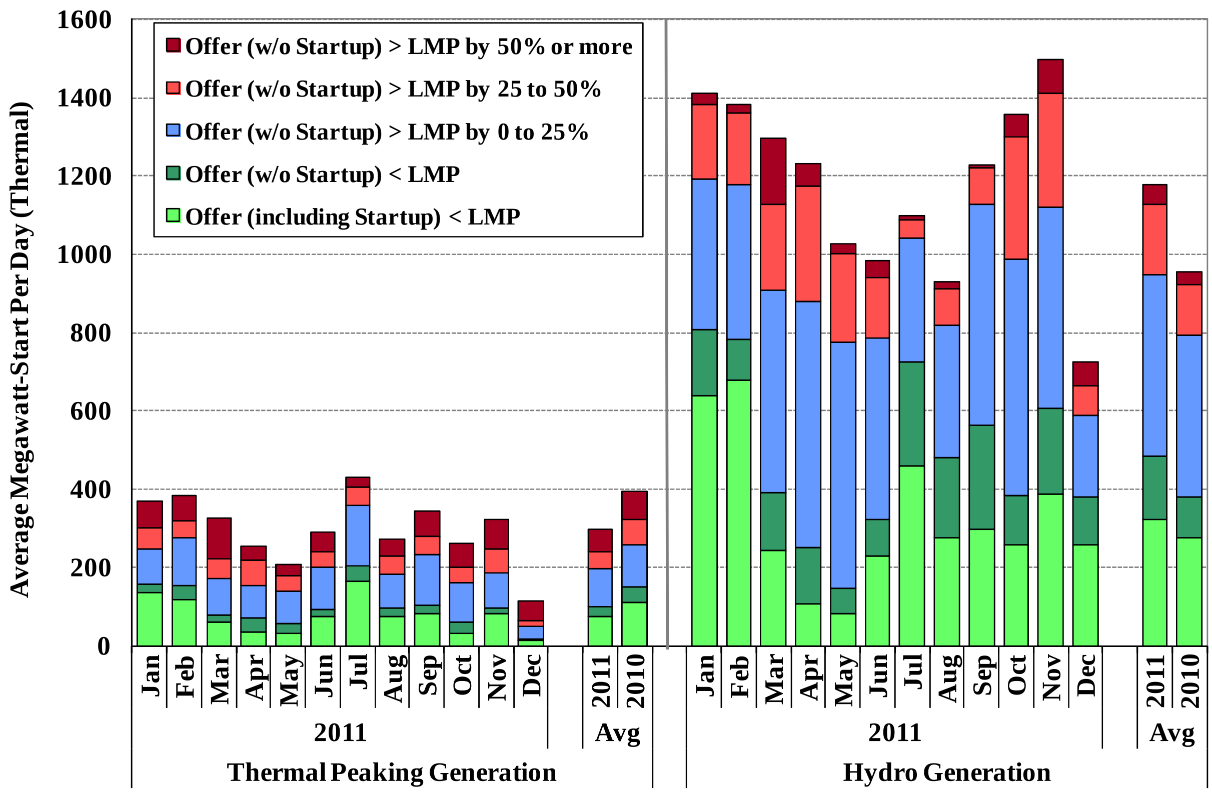


Figure 19 shows that flexible hydro generation accounted for the majority of fast-start generation that was started by UDS in 2010 and 2011. The operation of flexible hydro generation rose 23 percent from 2010 to 2011 because a large pump storage hydro generator was in service all of 2011 that was out of service from May to December 2010.[67] The operation of thermal peaking resources fell in 2011 because the average real-time capacity surpluses were higher in 2011.

The overall efficiency of real-time prices when fast-start resources are committed by UDS did not change significantly from 2010 to 2011. The average total offer (including start-up costs) was higher than the real-time LMP in 73 percent of starts in 2011, which is comparable to 2010. 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.

Nearly 300 MW of thermal peaking generation were started each day in 2011. The average amount of thermal peaking capacity that was started by UDS decreased 25 percent from 2010 to 2011, which was partly due to the increase in surplus capacity in real-time in 2011. As detailed in Section VI of the report, average surplus capacity during daily peak load hours increased 17 percent from 2010 to 2011. Roughly 75 percent of the committed thermal peaking capacity exhibited a total offer (including start-up costs) greater than the average real-time LMP. Even when start-up costs are excluded, 66 percent of the thermal peaking generation started exhibited offers that exceeded the average LMP over the minimum run time. Although thermal peaking generators are deployed in a relatively limited number of hours, they are frequently the marginal source of supply to the system 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.

The prior analysis in Figure 19 shows that the full costs of the thermal peaking units are frequently higher than real-time LMPs. This indicates that fast-start units that are committed in economic merit order usually rely on NCPC payments to recoup their full offer costs. More importantly, it indicates that real-time prices do not accurately reflect the marginal cost of

67 The Energy Information Agency's EIA-923 database reports the monthly electricity output of individual generators at http://www.eia.doe.gov/cneaf/electricity/page/eia906_920.html.

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.[68]

The analysis summarizes the portion of the fast-start units' costs that were not fully reflected in real-time LMPs in 2011. 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 2011.[69] The figure excludes fast-start units that were started in import-constrained areas since the LMP of the fast-start unit during such events would be representative of only a limited area of New England.[70] 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.

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). Ramp demands are highest on the system during these periods so fast-start generation is sometimes needed to meet these demands.

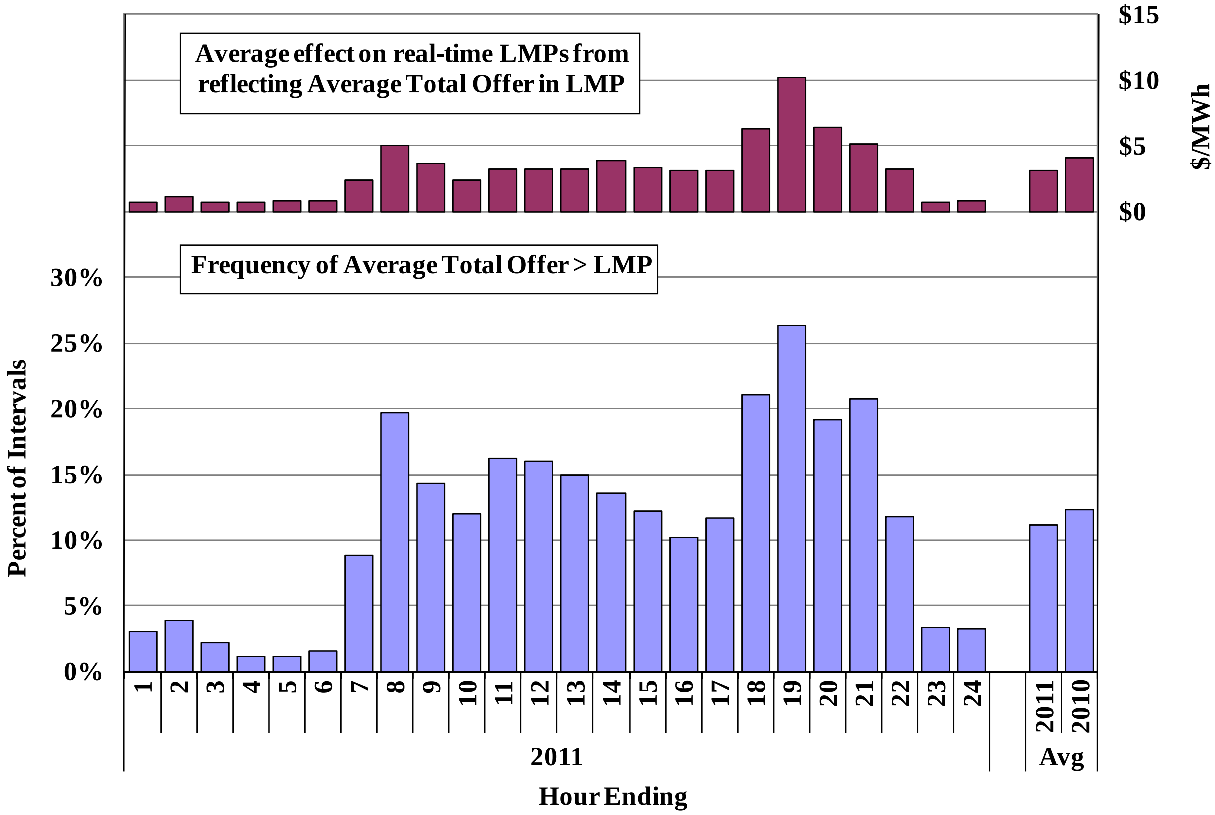
---

68 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-cost gas turbine started in the same hour would not affect prices because they are inframarginal.

69 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.

70 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.

Figure 20: Difference Between Real-Time LMPs and Offers of Fast-Start Generators
First Hour Following Start-Up by UDS, 2011



Fast-start units were started economically by UDS when their average total offer exceeded the real-time LMP over the minimum run time in 11 percent of all hours in 2011 and 12 percent of all hours in 2010. If the average total offers were fully reflected in the energy price in these hours, the average real-time LMP would increase approximately $3.10 per MWh in 2011.[71] In 2011, the price effect would be largest in hour-ending 19 when the average LMP would rise by $10.25 per MWh.

71 This is roughly $1.0 per MWh lower than in 2010, which is likely due to the lower natural gas prices and less frequent starts of thermal peaking units in 2011.

However, 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 would be smaller than the effects reported in Figure 20, since the figure does not consider the market response to more efficient real-time prices. Importantly, these responses would substantially improve efficiency because higher-cost peaking generation would be displaced by lower-cost intermediate generation and net imports. In addition, the changes in the market's economic signals that would result by causing peaking resources to set prices more reliably when they are marginal would improve economic efficiency over the long term by facilitating more efficient contracting and investment.

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 to consider using a similar mechanism to improve the efficiency of real-time pricing when fast-start resources are the marginal sources of supply.[72]

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

72 The MISO is currently working on the implementation of this project, which is known as "ELMP." For the latest timeline for the project, see "https://www.midwestiso.org/Events/Pages/ELMPTT20111216.aspx."

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.[73, 74]

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 system-level 30-minute reserves and determined that the RCPF should be raised to $500 per MWh. The ISO found that the real-time market sometimes experienced a shortage of reserves because the marginal cost of scheduling the available reserves exceeded the RCPF. In such cases, the ISO would take additional actions to maintain the required level of reserves. This included curtailing exports to neighboring areas, manually dispatching online generators that had available capacity that was not providing 30-minute reserves; or manually committing slow-start generators to bring additional capacity online.

The ISO will implement the $500 per MWh RCPF for system-level 30-minute reserves on June 1, 2012. This should ensure that the available capacity of online and fast-start units are fully scheduled to maintain adequate reserves whenever possible, reducing the need for manual actions to maintain reserves. Furthermore, the new RCPF should provide more efficient price

73 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).

74 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 binding constraint shadow prices.

signals during reserve shortages, which will provide better incentives for resources to be available 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 five 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:

- Real-Time 30-Minute Demand Response Program. 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.
- Real-Time 2-Hour Demand Response Program. 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.
- Real-Time Profiled Response Program. 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.
- Real-Time Price Response Program. These resources may reduce load (but are not required to do so) when they receive notice on the previous day. If they reduce their load, they receive 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.[75] The fourth program is a price-based program that provides a mechanism for loads to respond when the wholesale price is 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.[76]

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.

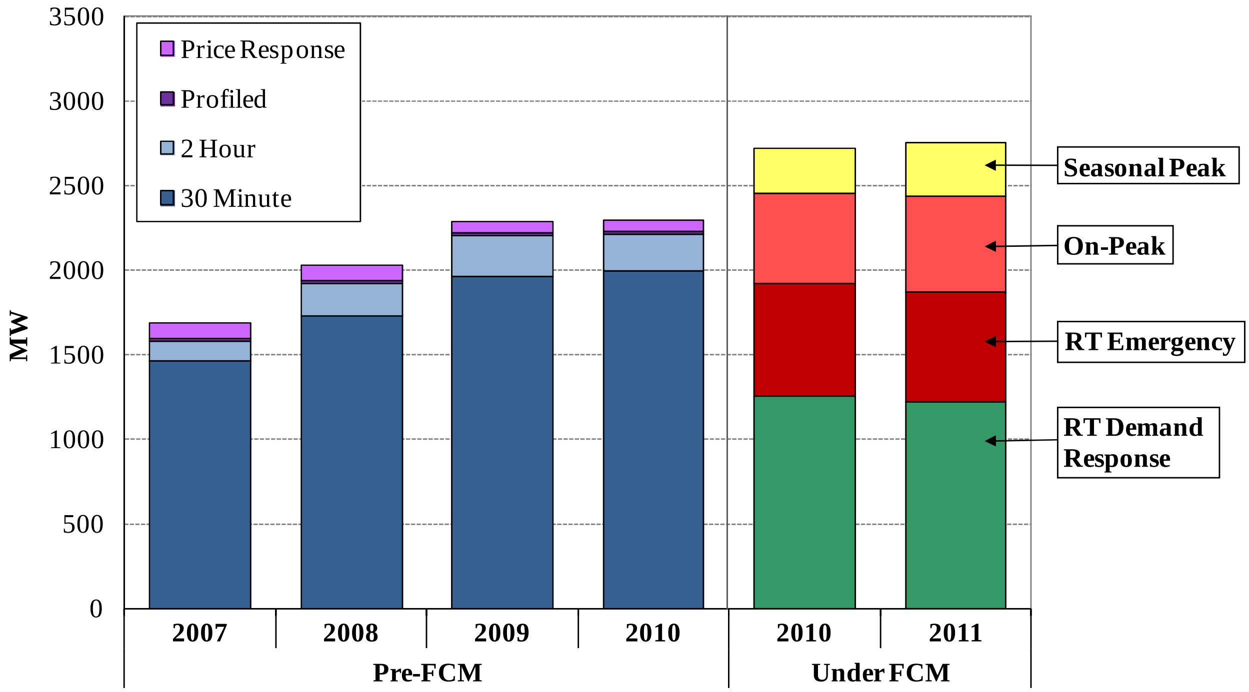
The first two are *active* (i.e., dispatchable) demand resources that operate based on real-time system conditions via dispatch by the ISO. They are defined at the Dispatch Zone level and

75 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.

76 Resources in the Real-Time Price Response Program do not receive capacity payment. This program is set to expire on May 31, 2012. Beginning June 1, 2012, the Transition Period Price Responsive Demand Program will pay demand response resources that curtail the real-time LMP rather than the higher of the LMP or $100 per MWh.

reduce energy demand during OP-4 conditions.[77] 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.[78] Demand response participation has surged in New England in recent years. Figure 21 shows the quantity of resources enrolled in each of the real-time demand response programs from 2007 to 2011. 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 21: Real-Time Demand Response Program Enrollments
2007 – 2011



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

77 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.

78 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 2,755 MW of demand resources were enrolled by the end of 2011, with 68 percent (or 1,877 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 VIII.

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. [79,80]

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 the last three years, there have been only two occasions when emergency demand response resources were activated. On July 22, 2011 the ISO activated a total of 642 MW of Real Time Demand Response resources for five and half hours (12:15 to 17:45). On December 19, 2011 the

79 "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.

80 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.

ISO activated a total of 504 MW of Real Time Demand Response resources for two and half hours (7:26 to 10:00). Both events were caused by a system-wide capacity deficiency that was driven by unexpected high load and generation outages.[81] Table 1 summarizes prices and the levels of supply and demand during the two events.

Table 1: Resource Availability During Activation of RT Demand Response
2011

Date	Hour Ending	NE Hub LMP DAM ($/MWh)	NE Hub LMP RTM ($/MWh)	RT Avail Capacity (MW)	Net Imports NY (MW)	Net Imports Canada (MW)	DAM Sched Load (GW)	RT Load (GW)	DAM Not Commit (GW)	Post-DAM Commit (GW)
22-Jul	12	$183.36	$265.31	30.4	192	1827	23.4	27.3	3.8	2.1
	13	$201.06	$422.29	30.4	97	1852	23.9	27.7	3.8	2.1
	14	$224.16	$558.55	30.4	-169	1902	24.1	27.7	3.8	2.3
	15	$234.43	$474.29	30.4	-268	1854	24.4	27.8	3.8	2.3
	16	$231.60	$455.18	30.4	-283	1904	24.3	27.7	3.8	2.6
	17	$231.16	$176.56	30.4	-13	2013	24.2	27.5	3.8	3.0
	18	$209.00	$108.06	30.4	-347	2067	23.7	27.2	3.8	3.0
	19	$188.29	$67.34	30.4	98	2067	23.3	26.6	4.3	3.5
19-Dec	7	$40.20	$65.50	30.7	358	1308	15.0	16.4	11.7	0.3
	8	$44.97	$277.59	30.7	135	1438	15.9	17.6	11.0	0.3
	9	$39.44	$31.97	30.7	213	1521	15.9	17.3	11.0	0.3
	10	$39.81	$30.95	30.7	408	1537	15.8	17.2	11.0	0.3

The table specifically shows:

- The day-ahead and real-time LMPs at the New England Hub;
- The amount of available offered capacity in real-time;
- The net imports from other control areas in real-time;
- The day-ahead net scheduled load and real-time load;
- The amount of slow-start capacity that was offered in the day-ahead market, but not scheduled; and
- The amount of slow-start capacity that was committed after the day-ahead market in the RAA process or as a self-commitment.

81 On both days, the Real Time Demand Response resources were activated via OP-4 Action 2, while the Real-Time Emergency Generation resources were not activated.

The table shows that there was a relatively deep shortage of 30-minute reserves for four hours on July 22 when real-time demand response resources were activated. Accordingly, the reserve clearing price for 30-minute reserves was $100 per MWh (the RCPF) during the four hours. As discussed in Section V.B, it is likely that the low RCPF for 30-minute reserves contributed to the severity of the shortage in at least two ways.

First, the table shows that there was a large amount of available import capability from New York that could have been used to bring more power into New England if market participants had expected real-time prices to be higher in the event of a 30-minute reserve shortage, or seen higher prices after the shortage began.

Second, the amount of energy scheduled in the day-ahead market was 3 to 4 GW lower than the amount of real-time load. Participants will have the incentive to purchase more energy in the day-ahead market if they expect higher real-time prices in the event of a 30-minute reserve shortage. Increased purchases would likely lead to more resources would have been scheduled in the day-ahead market.

Hence, in addition to causing the price signals during shortage conditions to understate the value of the shortage, the low RCPF for 30-minute reserves actually increases the frequency and severity of the shortages. The ISO is addressing these issues by raising the RCPF for system-level 30-minute reserves to $500 per MWh, which should lead to more efficient pricing and greater availability of resources during peak load conditions.

This change should improve the ISO’s pricing during shortages, however demand response activations can have even larger effects on prices when it causes the system to no longer be short of operating reserves (i.e., when the response is larger than the magnitude of the shortage). This occurred during the last two hours of the real-time demand response activation on July 22. In hour-ending 17, there would have been a shortage without the real-time demand response activation, so the demand response resources were effectively the marginal resources during this hour. Accordingly, economic efficiency would have dictated that the marginal costs of the demand response resources (i.e., the value of foregone consumption) set the real-time prices of energy and reserves.

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.[82] 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 the work the ISO plans to allow peaking resources to set prices will likely be applicable to demand response resources. 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 offered at marginal cost). Because they are consistent with the optimized dispatch, they are an efficient reflection of the prevailing market conditions.

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

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.[83] 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 2011:

- The current implementation of ex post pricing results in a small (0.3 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.

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 22 summarizes differences

83 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.

between ex ante and ex post prices in 2011 at a location close to the New England Hub.[84] 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 2011. 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 2011.

Figure 22: Average Difference Between Five-Minute Ex Post and Ex Ante Prices
2011

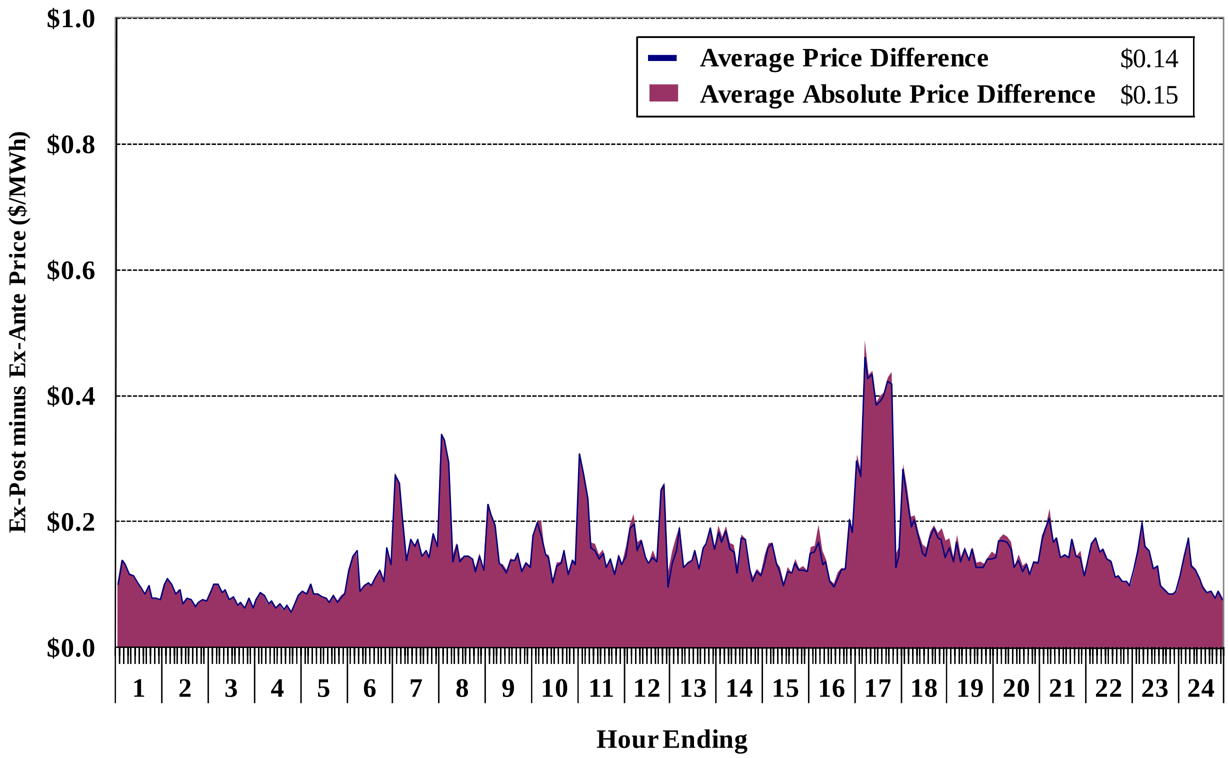


Figure 22 shows that the average ex post price is greater than the average ex ante price for every interval. This persistent bias is the result of the interaction between the following two factors. First, loss factors change slightly due to the time lag between the calculation of the ex ante and

84 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.

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.[85]

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.[86]

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 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

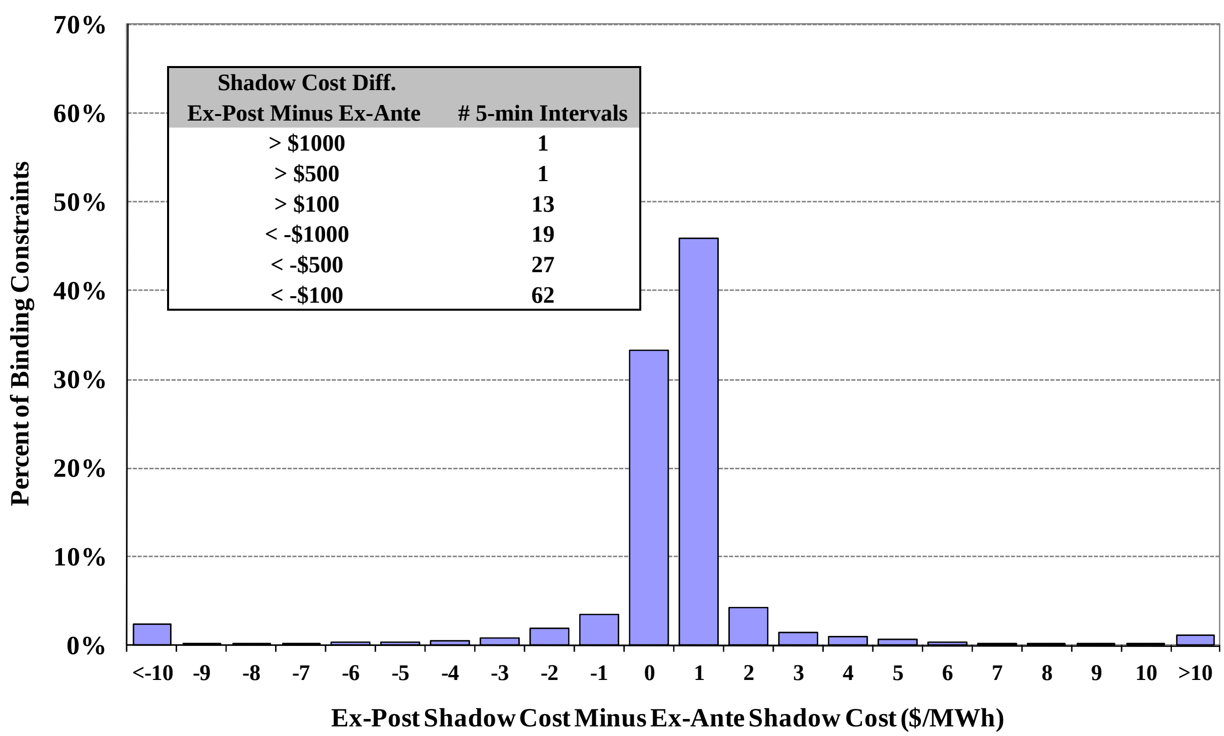
85 Uninstructed deviation penalties are penalties applied to suppliers that are not within a specified range of the dispatch instruction sent by ISO-NE.

86 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.

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 23 summarizes differences in constraint shadow prices between ex post and ex ante calculations in 2011. 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 23: Difference in Constraint Shadow Costs Between Ex Post and Ex Ante
All Binding Constraints, 2011



The average difference was not significant in 2011. Nearly 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 13 intervals during which ex post shadow prices were at least $100 per MWh higher than ex ante prices, and 62 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:

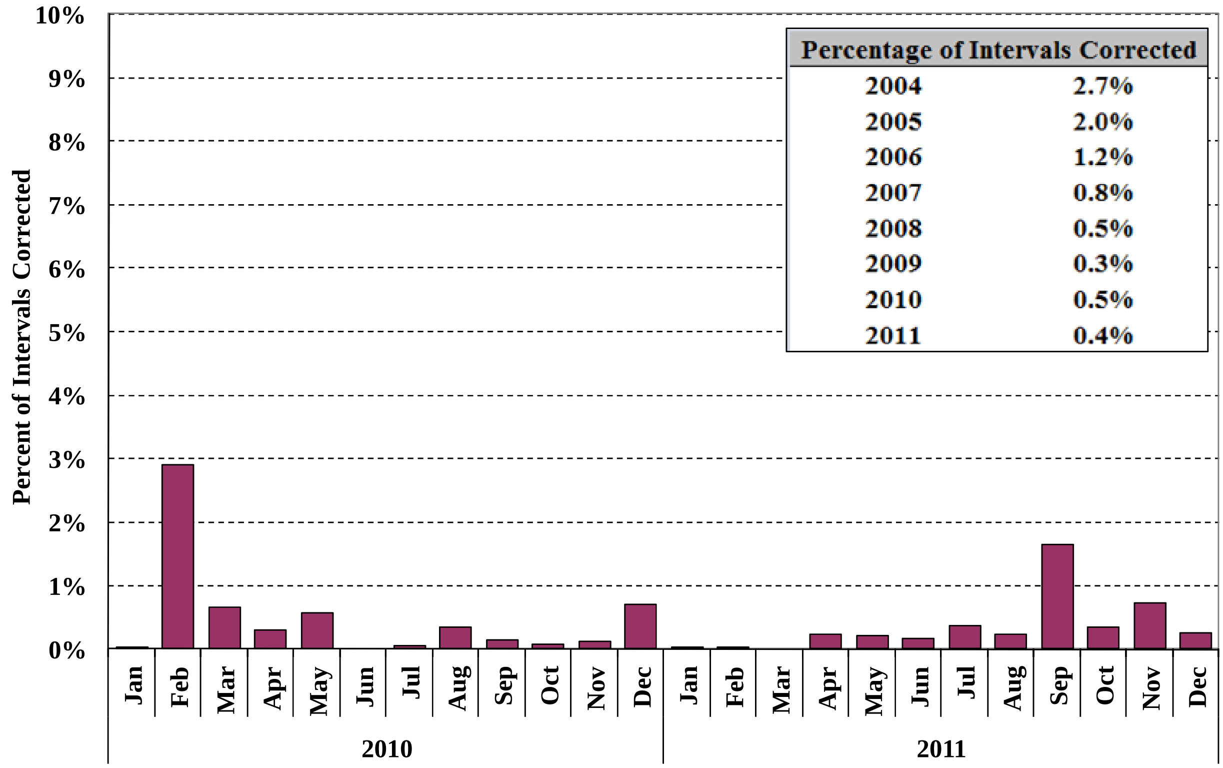
- Are slightly biased in the upward direction in uncongested areas;
- Introduce small potential inefficiencies when they are not consistent with dispatch instructions; and
- Sometimes distort the value of congestion into constrained areas.

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 2011. 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 24 shows the rate of real-time price corrections in New England in each month of 2010 and 2011. The inset table shows the annual rate of price corrections in the past eight years.

Figure 24: Rate of Real-Time Price Corrections
2010 – 2011



The figure shows that real-time price corrections were infrequent in both 2010 and 2011. The rate was less than one percent in all but two months during 2010 and 2011. September exhibited the highest rate of price correction of any month in 2011. This was caused by price corrections for 12 hours on one day in September 2011 due to software errors, which affected the LMPs at only three pricing nodes. 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 60 percent of the intervals that experienced price corrections in 2010 and 2011 were due to issues with the real-time software's Dead Bus Logic, which affects the LMPs at very few pricing nodes.[87] Hence, during many of the real-time intervals with price corrections, the effect of the price correction on the market was very limited.

---

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

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 suppliers to offer competitively, 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 2011. 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 in areas that rely on fast-start generators to manage local congestion.
  - ✓ *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, requiring the ISO to curtail exports and take other manual actions outside the market. This has led to inefficiently low real-time prices that did not properly reflect the cost of maintaining reliability. We recommended in the 2010 Assessment evaluate and update the 30-minute reserve RCPF. ISO-NE performed this evaluation and filed to increase the RCPF to $500 per MWh, effective June 1, 2012.
- The new RCPF level will provide market participants better incentives to schedule in the day-ahead market and schedule with 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.
  - ✓ *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.

- Finally, 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 2011, 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 in a manner that is economically efficient. 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.

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).

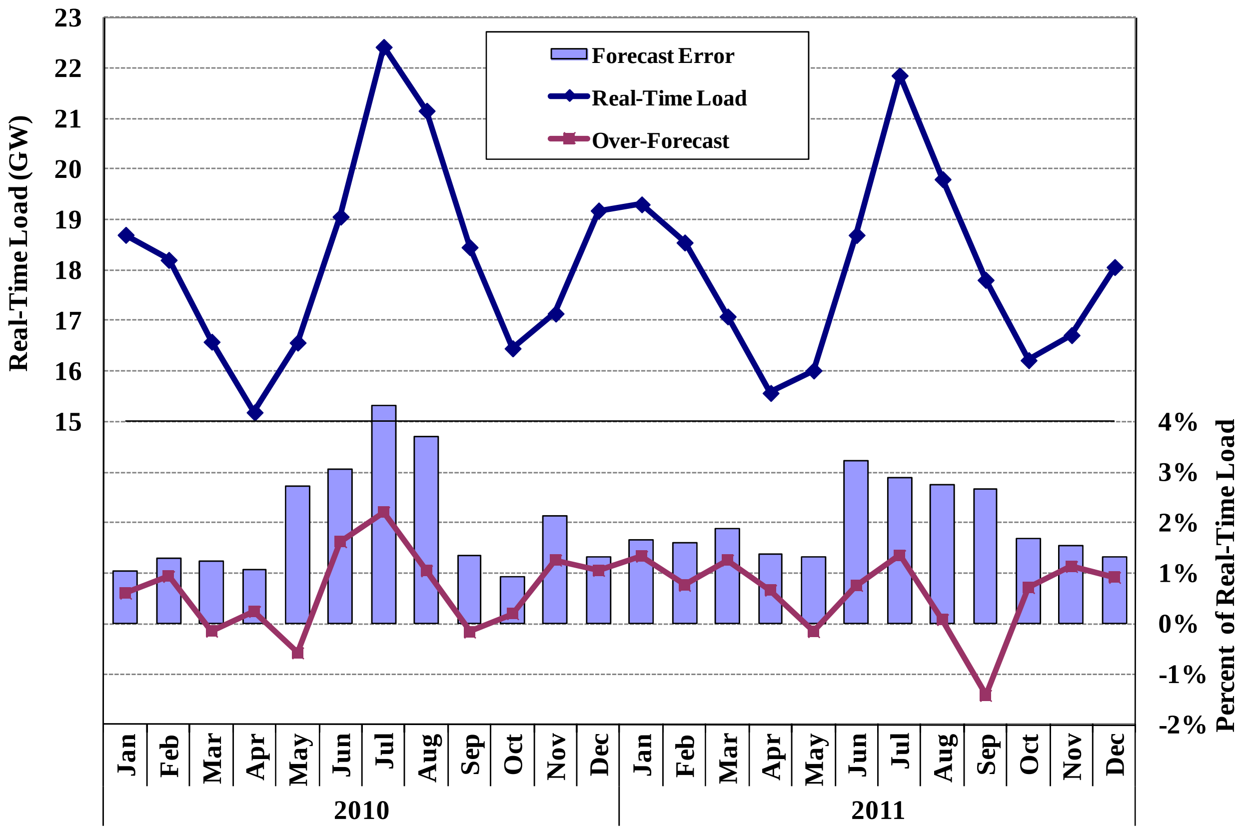
Accurate load forecasts promote efficient scheduling and unit commitment. Inaccurate load forecasts can cause the day-ahead market and/or the ISO to commit too much or too little capacity, which can affect prices and uplift. Therefore, it is desirable for the day-ahead forecast to accurately predict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

Figure 25 summarizes daily peak loads and two measures of forecast error on a monthly basis during 2010 and 2011. 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. [88] 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.

---

88 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.

Figure 25: Average Daily Peak Forecast Load and Actual Load
Weekdays, 2010 – 2011



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 2010 to 2011. The annual peak load of 27.7 GW occurred on July 22, 2011, up approximately 3 percent from the peak load of 27.1 GW in 2010. July 22, 2011 was the hottest day since 1960, and it exhibited the second–highest all-time peak load level. [89] However, the average load declined nearly 1 percent, from 15.0 GW in 2010 to 14.9 GW in 2011. In addition, the frequency of peak load conditions exceeding 20 GW decreased from 531 hours in 2010 to 312 hours in 2011. The decline in load levels was particularly notable in the summer months when average load fell 2.5 percent from the previous year, which was primarily due to milder summer weather in 2011.

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

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

years: 0.7 percent in 2010 and 0.6 percent in 2011. The ISO regularly evaluates the performance of its load forecasting models to ensure there are no factors that bias the forecast unjustifiably.[90]

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 2.0 percent in 2011, consistent with 2010. The forecast error tends to increase during the summer months. In 2011, the forecast error averaged 3.0 percent in the summer months (i.e., June to August) and just 1.7 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.

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)).

90 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.

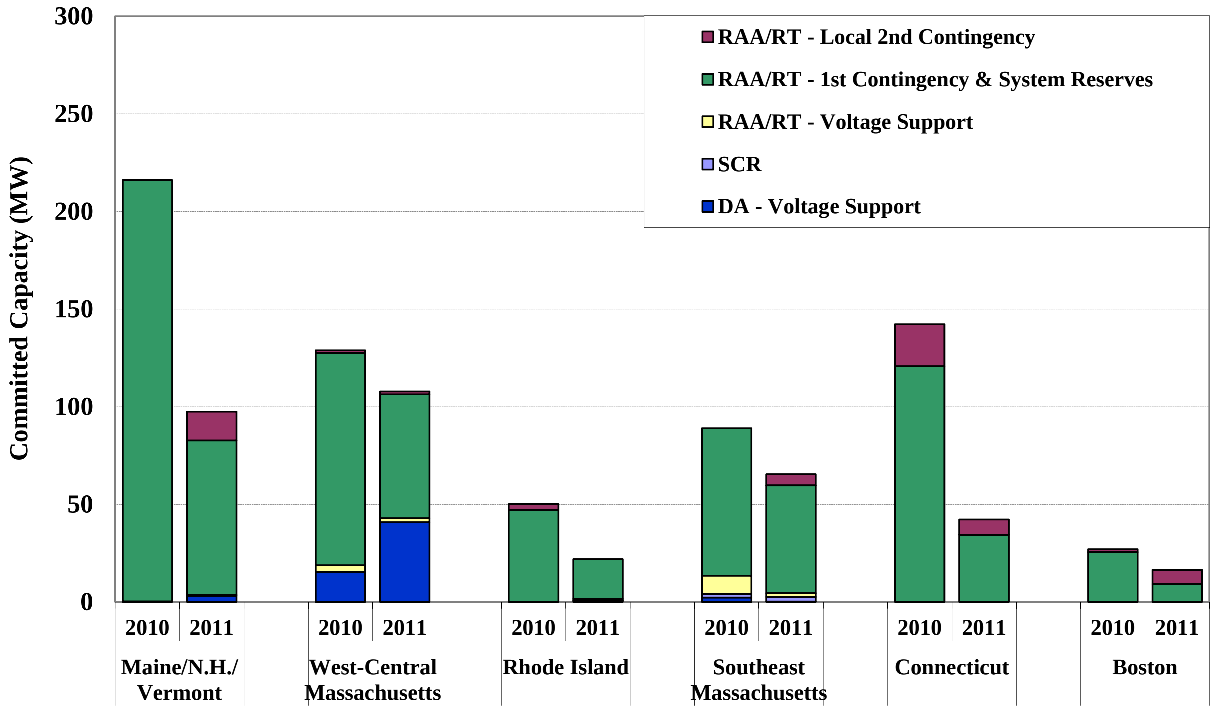
In the day-ahead market, generators are scheduled based on the bids and offers submitted by buyers and sellers. A generator is committed when demand bids from load serving entities and virtual traders are high enough for the unit to be economic given its start-up, no-load, and incremental offer components. 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 three years and discusses several initiatives by the ISO to reduce the frequency and effects of supplemental commitment.

Figure 26 shows the average amount of capacity committed to satisfy local and system-level requirements in the daily peak load hour in each zone from 2010 to 2011.[91] 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.

Figure 26: Commitment for Reliability by Zone
Daily Peak Hour, 2010 – 2011



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

91 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.

Total supplemental commitments fell significantly from an average of 650 MW in 2010 to 350 MW in 2011.[92] 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 made for system-wide reserves. The amount of supplemental commitment for system-wide reserves (and local first contingencies) fell from an average of 590 MW in 2010 to 260 MW in 2011. These commitments accounted for 91 percent and 75 percent of total reliability commitments in 2010 and 2011, respectively. We evaluate the need for these commitments and their effects on real-time energy prices later in this section.

One reason that the need to commit resources after the Day-Ahead market for system-wide reserves was greater in 2010 than in 2011 was that a large flexible pump storage hydro resource was out of service during most of 2010. The return of this unit in December 2010 increased the amount of fast-start capability available to satisfy system capacity requirements in the RAA process, thereby reducing the need to supplementally commit non-fast start resources. This pattern highlights the benefits provided by fast-start resources, underscoring the importance of setting price signals in the day-ahead and real-time markets that properly reflect the value that such resources provide to the system.

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. 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

92 The reduction in supplemental commitment was even more significant from 2007 (when the quantity averaged 1,670 MW). This reduction 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.

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.

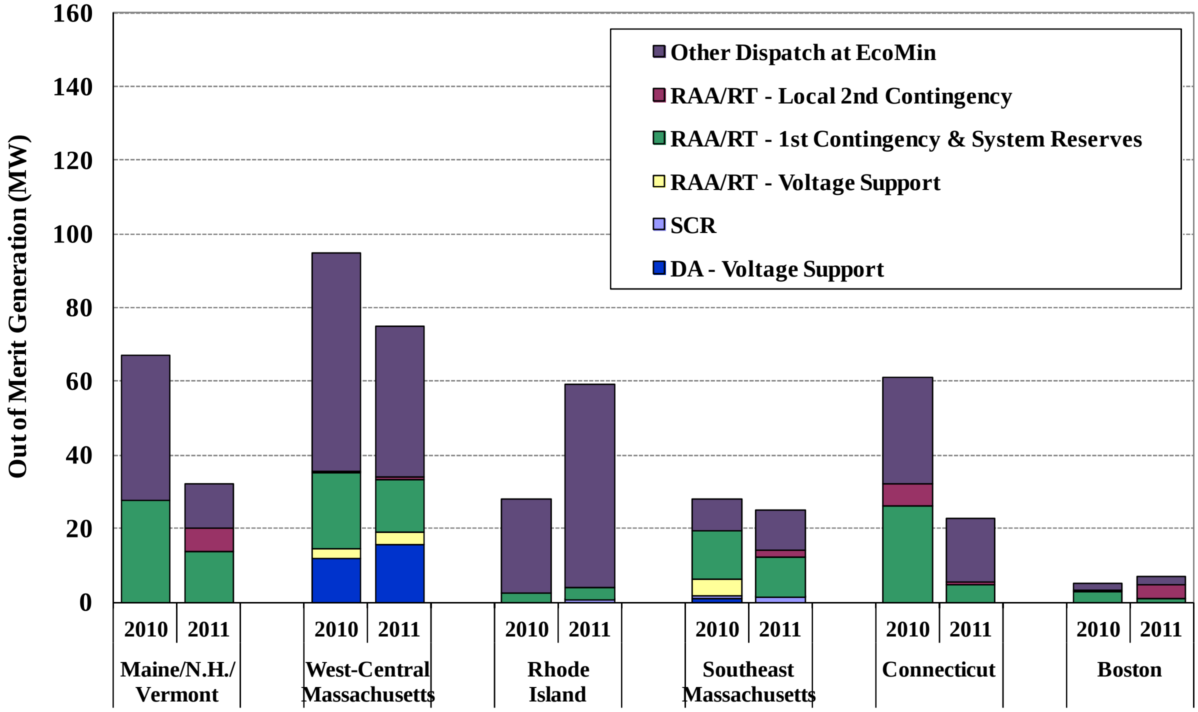
Out-of-merit generation occurs for several reasons. First, a unit may run at its EcoMin to satisfy its minimum run time after having run in-merit in previous hours or in anticipation of running in an upcoming hour. This is efficient because the software is minimizing cost over the total run-time of the 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.[93]

Figure 27 summarizes the average out-of-merit generation by location during peak hours (weekdays 6 AM to 10 PM, excluding holidays) in 2010 and 2011. The figure shows five categories of out-of-merit generation on units that are committed (and occasionally dispatched) for reliability reasons.[94] The figure also shows an "other dispatch" category that includes generation from units that were economically committed but are running at their EcoMin.

93 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.

94 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.

Figure 27: Average Hourly Out-of-Merit Generation
Weekdays 6 AM to 10 PM, 2010 – 2011



In most regions, Figure 27 shows that the majority of the out-of-merit generation was attributable to non-local reliability units being dispatched at EcoMin in 2010 and 2011. However, this was not the case in Maine and Southeast Massachusetts where most of the out-of-merit dispatch was from units committed through the RAA process for reliability.

The average quantity of out-of-merit generation from units committed for reliability fell from 121 MW in 2010 to 83 MW in 2011. The decline in out-of-market generation from units committed for reliability tracked the reduction in supplemental commitments and was caused by the same underlying factors. The reduced commitment for reliability in Maine and Connecticut, in particular, led to proportionate reductions 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 declined (however, to less extent) in recent years, from an average of 221 MW in 2008 to 139 MW in 2011. The overall decrease since 2008 was consistent with the reduction in reliability commitments as a result of the transmission upgrades in Connecticut, Boston, and Southeast Massachusetts that were completed in 2009.

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.

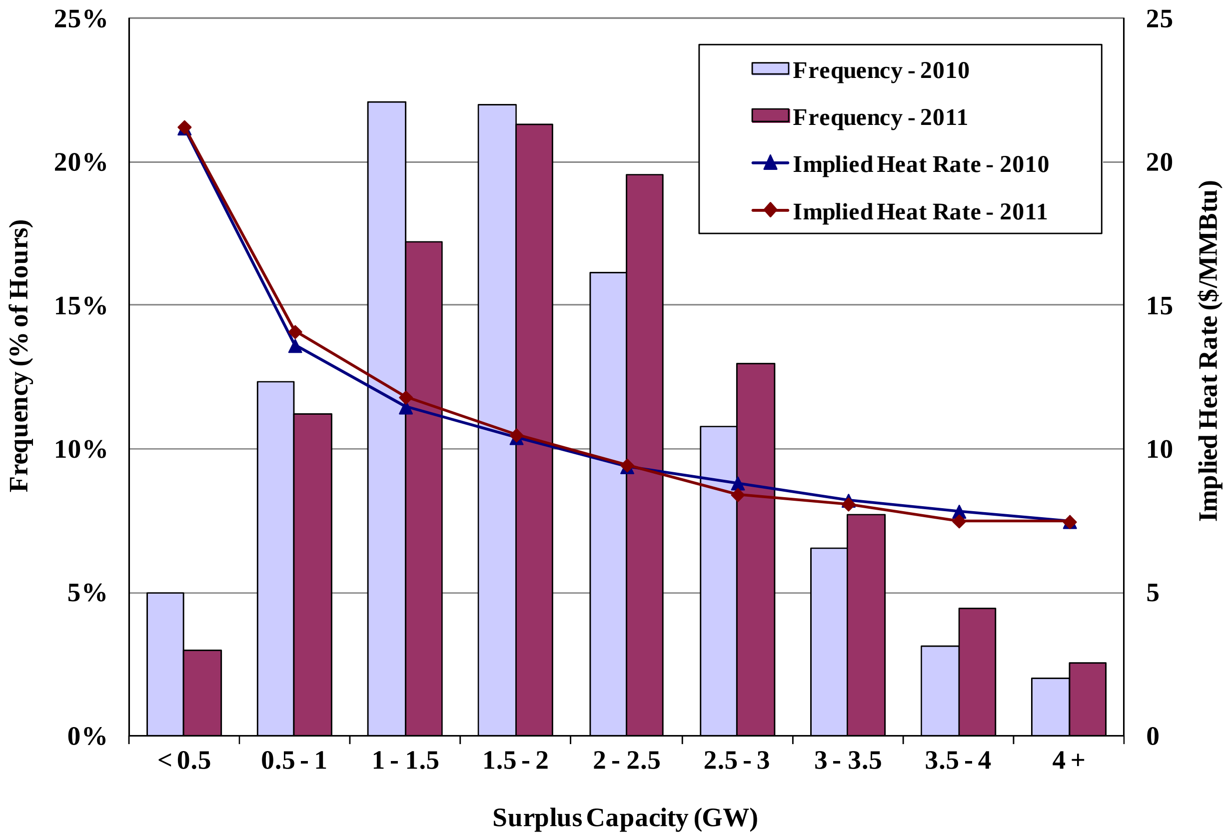
Surplus capacity is also important because it constitutes the resources that are available to respond to unexpected changes in real-time operating conditions. Accordingly, the quantity of surplus capacity exhibits a strong negative correlation with real-time energy prices. This section evaluates the 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 28 summarizes the relationship of surplus capacity to real-time energy prices at ISO-NE Hub in each peak hour of 2010 and 2011. 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 1.0 to 1.5 GW of surplus capacity in approximately 22 percent of the peak hours in 2010 and 17 percent in 2011. The lines show the average real-time implied marginal heat rate at ISO-NE Hub in the hours that correspond to each range of surplus capacity. For example, in hours when there was 1.0 GW to 1.5 GW of surplus capacity, the average real-time implied marginal heat rate was 11.5 MMbtu per MWh in 2010 and 11.8 MMbtu per MWh in 2011. The implied marginal heat rate is shown in order to normalize real-time energy prices for changes in natural gas prices during 2010 and 2011.[95]

95 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.

Figure 28: Surplus Capacity and Implied Marginal Heat Rates
Based on Real-Time LMPs at the Hub in Peak Hours, 2010-2011 [96]



The figure shows a strong correlation between the quantity of surplus capacity and the implied marginal heat rate in real time. In 2011, the average implied marginal heat rate ranged from approximately 21 MMbtu per MWh in hours with less than 0.5 GW of surplus capacity to 7.5 MMbtu per MWh in hours with more than 4 GW of surplus capacity.

Overall, the average implied heat rate during peak hours fell from 11.0 MMbtu per MWh in 2010 to 10.6 MMbtu per MWh in 2011. This is primarily attributable to the increase in the amount of surplus capacity from 2010 to 2011. The share of hours with less than 0.5 GW of surplus capacity decreased from 5 percent in 2010 to 3 percent in 2011. This is significant because these hours exhibited substantially higher price levels with average implied marginal heat rates above 20 MMbtu per MWh. Likewise, the share of hours with less than 1.5 GW of surplus capacity decreased from approximately 40 percent in 2010 to 31 percent in 2011. The amount of surplus

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

capacity increased primarily because a large pump storage resource that was out of service for most of 2010 returned to service in December 2010.

Despite the increase from 2010, the surplus capacity levels were still lower in 2011 than in recent years prior to 2010. Before the completion of transmission upgrades in 2009 in Boston, Connecticut, and Southeast Massachusetts, the ISO routinely committed over 1 GW of capacity after the day-ahead market for local reliability in these areas, which led to substantially higher levels of surplus capacity. Although the recent operation with lower levels of surplus capacity has led to higher real-time price volatility, it generally improves the efficiency and overall performance of the market as real-time prices more completely reflect the true needs of the system. These effects also generally lead to lower NCPC costs.

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-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 29 shows the first six months of 2011, and Figure 30 shows the last six months of 2011.

Figure 29: Daily Supplemental Commitments and Surplus Capacity
January to June, 2011

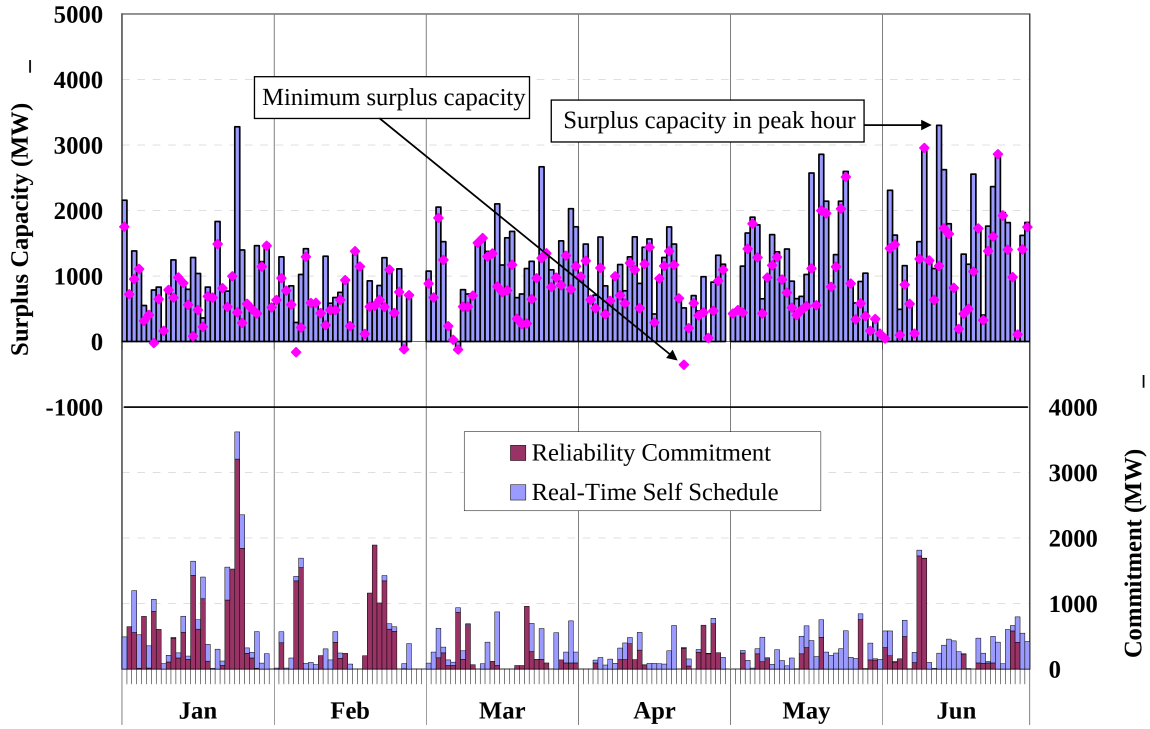
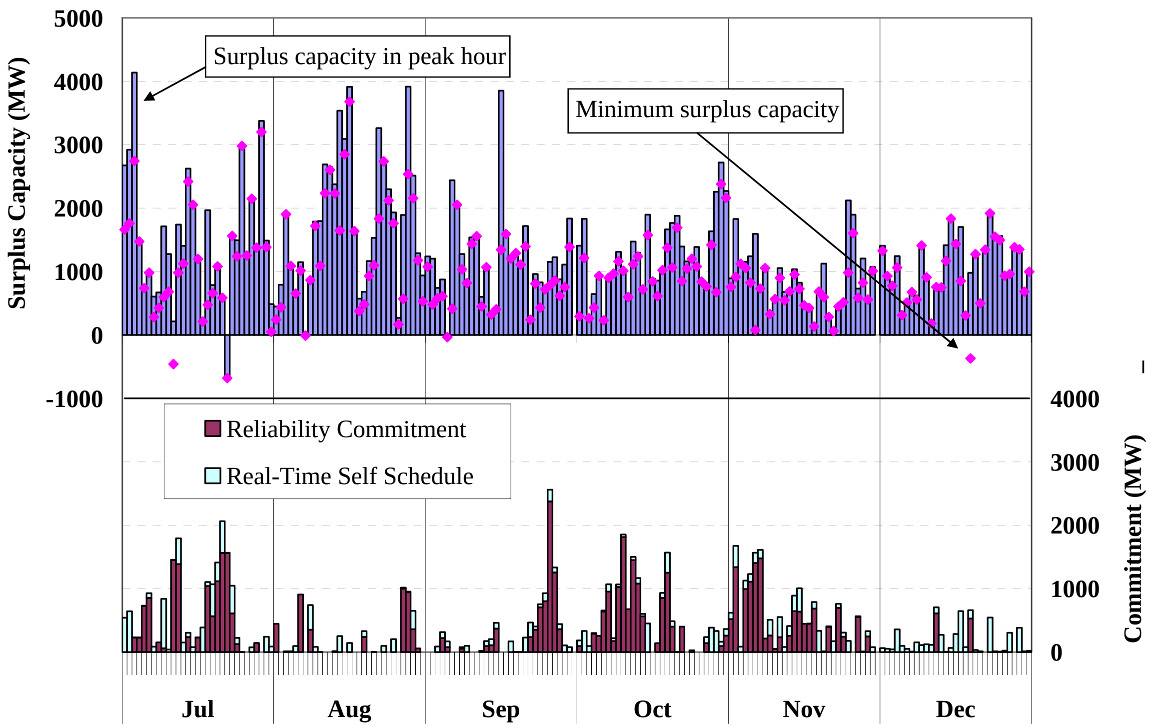


Figure 30: Daily Supplemental Commitments and Surplus Capacity
July to December, 2011



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 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 42 percent of the supplemental resource commitments in 2011 were needed to maintain system level reserves in retrospect.[97] The fact 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.

97 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 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".

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; [98] and
- The assumed level of imports and exports.

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. However, we recognize that conditions can sometimes arise that would justify an increase in the desired capacity surplus.

With regard to the import and export assumptions, we believe that improvements are possible based on our review. In general, the assumptions regarding imports and exports are based on the day-ahead scheduled transactions. 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; and
- The value of the day-ahead exports may not justify the costs of the supplemental commitments made to support them.

This is particularly true when exports 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 (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

98 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)."

(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 imports and exports.

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 has used reliability agreements[99] and supplemental commitment to ensure reliability, particularly in local import-constrained areas. Reliability agreements give the owners of uneconomic generating facilities supplemental payments to keep them in service. Supplemental commitments bring uneconomic capacity online at times when market clearing prices are insufficient. 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 2010 and 2011. The main categories of uplift are:

- Reliability Agreements/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.[100] The last set of Reliability Agreements expired on June 1, 2010 when the first FCM period started. 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 2011, 97 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.[101] The remaining uplift associated with day-ahead rather than real-time commitments was allocated to day-ahead load schedules in the local zone.

---

99 Reliability agreements expired on June 1, 2010 when the first Forward Capacity Commitment Period began.

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

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

- 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.

- Economic and First Contingency Protection Resources – In 2011, 93 percent of this uplift was allocated to Real-Time Deviations throughout New England.[102] 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).[103]

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 reliability agreements and supplemental commitment in 2010 and 2011. The year-over-year change in each category of uplift is shown as well.

102 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.

103 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.

Table 2: Allocation of Uplift for Out-of-Market Energy and Reserves Costs
2010 – 2011

	Millions of Dollars		% Change
Category of Uplift	**2010**	**2011**	**2010-2011**
Reliability Agreements/ FCM Reliability Credits	**$24.0**	**$1.4**	**-94%**
Local Second Contingencies	**$3.9**	**$6.0**	**53%**
Special Case Resources	**$1.6**	**$3.4**	**113%**
Voltage Support	**$5.2**	**$5.9**	**13%**
Economic*			
Quick-Start Resources	**$12.0**	**$11.9**	**-1%**
Other Resources	**$72.7**	**$42.5**	**-42%**
Total	**$119.4**	**$71.1**	**-40%**

* The category of Economic includes uplift for commitments made for system-wide reserve requirements and first contingency requirements.

Uplift charges fell significantly from $119 million in 2010 to $71 million in 2011, primarily for two reasons. First, out-of-market capacity payments (including reliability agreements and FCM reliability credits) fell from $24 million in 2010 to $1.4 million in 2011 as the last of the reliability agreements expired on June 1, 2010 when the first Forward Capacity Commitment Period began. The reliability agreement costs accounted for 92 percent of the out-of-market capacity payments in 2010. The remaining out-of-market capacity payments were paid to two units in Connecticut in 2010 and 2011 under FCM because their de-list requests were rejected in the first Forward Capacity Commitment Period for reliability purposes.

Second, the “Economic” category of uplift fell from $85 million in 2010 to $54 million in 2011. The reduction in supplemental commitment for system-wide reserves (which is quantified further in Subsection B above) led to concomitant incident reductions in the NCPC payments to such resources. These reductions are largely attributable to the return to service of a large pump storage hydro resource in December 2010. These savings highlight the value of flexible fast-start resources in maintaining system reliability at relatively low cost.

The "Economic" NCPC payments associated with fast-start resources were approximately $12 million in both 2010 and 2011. These payments result primarily from instances when fast-start resources are dispatched in merit order by the real-time dispatch model, but do not set the LMP once they are online. In other words, they are started economically and can set the price for the interval in which the model starts them, but they must remain on in subsequent intervals when they are frequently no longer economic and are running at their minimum output level. Hence, fast-start resources frequently do not recover sufficient LMP revenue to recoup their as-offered costs.[104] 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).[105]

G. Conclusions and Recommendations

We conclude that the ISO's operations to maintain adequate reserve levels in 2011 were reasonably accurate and consistent with the ISO's procedures. In the past two years, the amount of capacity committed for local reliability has been much lower than in previous years due to significant transmission upgrades in Connecticut, Boston, and Southeast Massachusetts. A by-product of the reduction in commitments for local reliability is an increase in the amount of supplemental commitment for system-wide reliability. Local reliability commitments generally increase the total resources available for system reliability, so the reduced local reliability commitment led to more days when additional resources had to be committed in the RAA process to satisfy the system requirements.

In 2011, supplemental commitments for system-wide reliability fell 56 percent as a large pump storage resource returned from an extended outage in December 2010 that provides flexible fast-start capability resource. However, NCPC payments for supplemental commitments of slow-

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

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

start resources remained substantial at $43 million in 2011. This highlights the value of fast start resources as low cost sources of operating reserves and the importance of day-ahead and real-time prices that signal the benefits provided by such resources. Such price signals encourage efficient investment in new resources that have flexible characteristics.

We regularly review patterns of supplemental commitment and the resulting out-of-merit generation because they can raise the following market issues:

- Dampening of economic signals to invest in areas that would benefit the most from additional investment in generation, transmission and demand response resources.
- Large and volatile uplift charges that can be difficult for participants to hedge, and which may discourage participation in the ISO-NE market.
- Incentives for generators frequently committed for reliability to avoid market-based commitment in order to seek additional payments through the reliability commitment process.

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. 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. 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. The ISO is planning to introduce hourly day-ahead energy offers and intraday reoffers in 2015.[106]

We also recommend several 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.

106 See 2012 Wholesale Markets Project Plan, page 20.

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 (FCA1) was held in February 2008, facilitating the procurement of installed capacity for the period from June 2010 to May 2011. Six auctions have been held to date, which have satisfied ISO-NE's planning requirements through May 2016.[107] 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 of the report provides background on the FCM rules and evaluates the outcomes of the first five auctions. This section also discusses the need for certain specific market reforms. A summary of our conclusions and recommendations is at the end of the section.

A. Proposed Changes to ISO-NE's Capacity Market

In recent years, the ISO and stakeholders have discussed reforms to the FCM. The Commission's most recent order addressing a broad set of market design topics was issued in April 2011.[108] The order included several significant directives for the ISO to work with stakeholders to:

- Model eight capacity zones corresponding to its eight Load Zones;
- Strengthen the supply-side market power mitigation rules;

107 The latest auction, FCA 6, was held on April, 2, 2012 for the Capacity Commitment Period of June 1, 2015 to May 31, 2016. However, the auction results are not included in this report.

108 See Order on Paper Hearing and Order on Rehearing, Docket ER10-787-000, et al. (Issued April 13, 2011).

- Extend the price floor in the auction through at least FCA6 and until appropriate buyer-side market power mitigation measures can be implemented; and
- Develop buyer-side market power mitigation rules in order to address the shortcomings of the proposed Alternative Price Rule.[109]

After lengthy discussions with stakeholders, consensus was not achieved. As a result, the ISO filed in January 2012 to make short-term changes for FCA7 (to be held in June 2012) while discussions continue on long-term changes. The short-term changes include modeling four (out of eight) capacity zones and extending the price floor of $3.15 per kW-month through FCA7.

As the External Market Monitoring Unit (EMMU), we filed a protest in February 2012 that made several recommendations, including that the Commission:

- Impose a deadline for the ISO to file the changes required by the Commission’s April 2011 Order by November 2012 so the changes can be implemented in time for FCA8 in June 2013;
- 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).

FERC issued an Order on March 30, 2012 requiring the ISO to file changes in the FCM by December 3, 2012 that will apply to FCA8.

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,

109 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.[110]

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)[111] 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[112] in each zone, less retirement and export bids, are compared to the zone's Local Sourcing Requirement (LSR).[113] If the amount of capacity is greater than the LSR, the zone will not be modeled as a separate import-constrained zone in the auction. 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.
- 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

110 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.

111 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.

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

113 The LSR is the minimum amount of capacity that is needed in the load zone. Since FCA1, 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 FCA4, 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).

commitment period from one to five years at the time of qualification. 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 Monitoring Unit (IMMU) 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

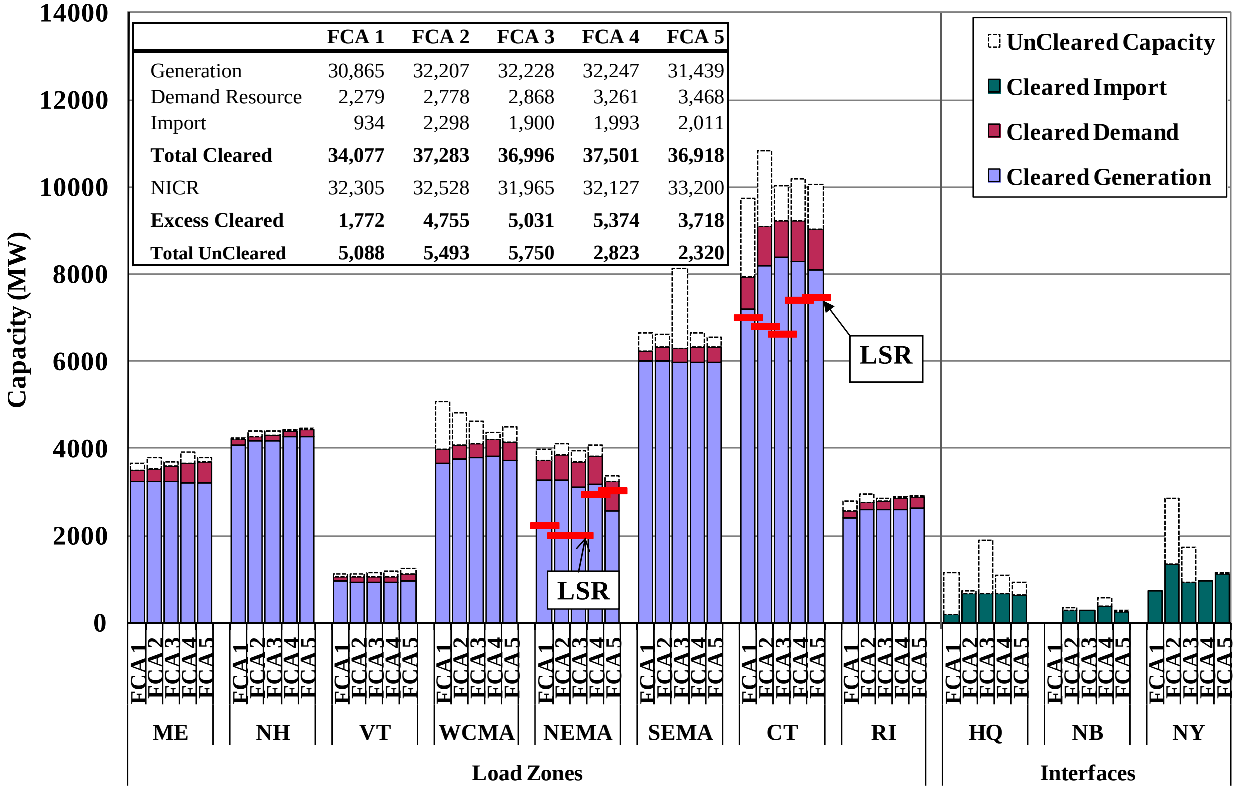
Five FCAs were held before the end of 2011: the first in February 2008 for the commitment period of 2010/2011 (FCA1), the second in December 2008 for the commitment period of 2011/2012 (FCA2), the third in October 2009 for the commitment period of 2012/2013 (FCA3), the fourth in August 2010 for the commitment period of 2013/2014 (FCA4), and the fifth in June 2011 for the commitment period of 2014/2015 (FCA5). In each auction, there was a substantial surplus of capacity over the NICR. Accordingly, each auction cleared at the floor price: $4.50 per kW-month in FCA1, $3.60 in FCA2, $2.95 in FCA3 and FCA4, and $3.21 in FCA5.

No import-constrained zones were deemed necessary because the amount of existing capacity exceeded the LSR in each area. Maine was modeled as an export-constrained zone in all five 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 five FCAs, the de-list bids of existing suppliers, and the procurement of new capacity.

1. Summary of Capacity Auction Results

Figure 31 summarizes the procurements in the first five 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.[114] The amounts of cleared resources are shown relative to the LSRs for Connecticut and NEMA and relative to the NICR for all of New England.

Figure 31: FCM Auction Clearing Summary by Location
FCA1 – FCA5



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 exceeded the NEMA LSR by 1.5 GW, 1.8 GW, 1.7 GW, 0.9 GW, and 0.2 GW in the five FCAs. It exceeded the Connecticut LSR by 1.2 GW, 2.3 GW, 2.6 GW, 1.8 GW, and 1.6 GW in the five FCAs.

From FCA1 to FCA2, the amount of excess capacity rose significantly in Connecticut 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

114 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.

Public Utility Control (DPUC). From FCA3 to FCA4, the amount of excess capacity fell in Connecticut and in NEMA as a result of changes in the method calculating the LSRs. Until FCA4, 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).[115] 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 FCA4. From FCA4 to FCA5, the amount of excess capacity fell in NEMA due to the planned retirement of the last two units at the Salem Harbor plant totaling nearly 600 MW.

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 FCA1 to 5.4 GW in FCA4. Substantial excess capacity cleared in the first five auctions as a result of the price floor. Under the original Settlement Agreement, the price floor was supposed to be eliminated after FCA3, but it has been extended at least through FCA6, and the ISO has proposed it be extended through FCA7 as well.[116] The price will likely clear at the floor in the next auction given the amount of excess supply and the vertical demand curve implicit in the FCM design.

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 FCA1 to 10 percent in FCA5. Roughly 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

115 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.

116 See the ISO's January 2012 filing.

Brunswick, and NYISO also accounted for a significant portion of the procured capacity, increasing from 934 MW in FCA1 to approximately 2,000 MW in the last four FCAs. [117]

In each auction, a substantial amount of qualified resources did not clear. New proposed resources accounted for more than 80 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 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 mitigation rules.

Figure 32**Error! Reference source not found.** evaluates several categories of accepted de-list bids in the first five FCAs. The figure shows four categories of de-list bids: permanent, static, export or administrative, and dynamic.[118] Accepted de-list bids are also separated according to

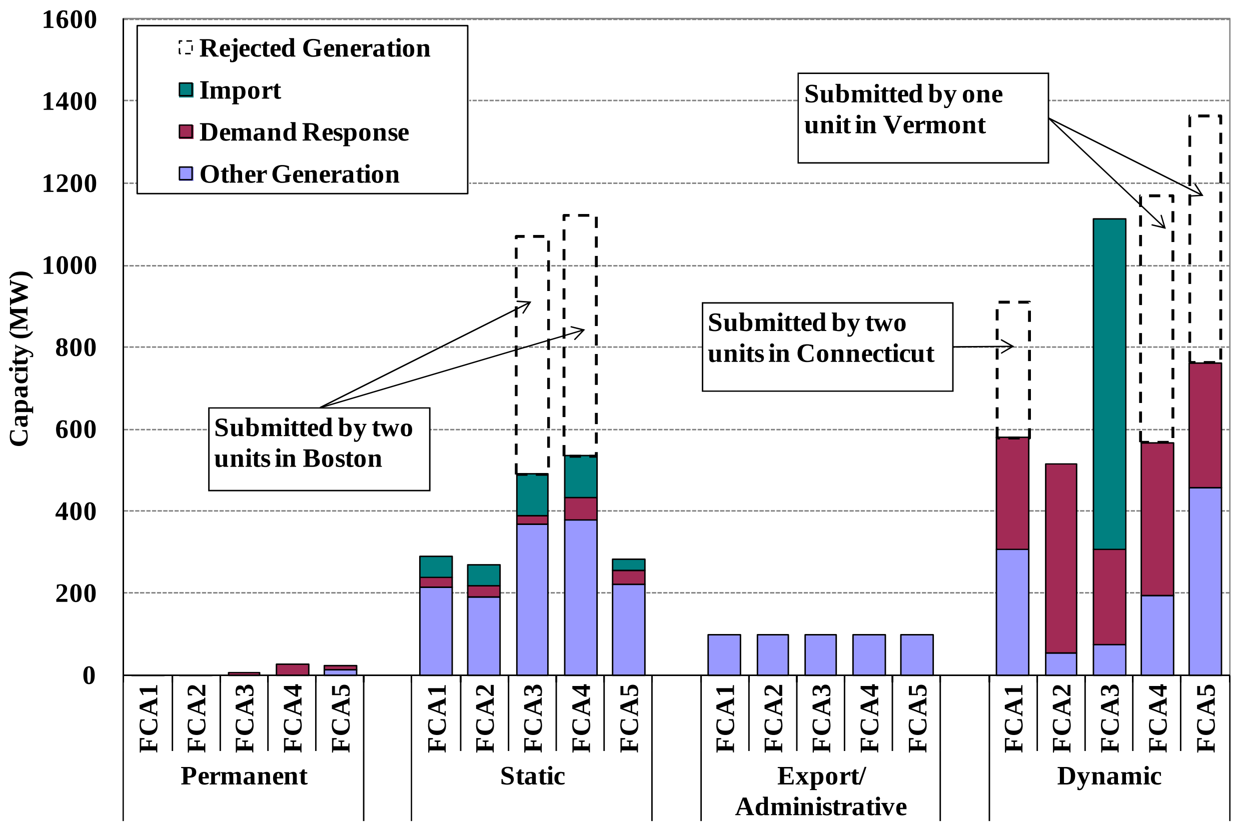
---

117 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.

118 Each category of de-list bid is defined in Tariff Section 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 IMMU 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 IMMU 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.

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 32: Summary of Accepted De-list Bids by Type
FCA1 – FCA5



Approximately 2.7 GW of generation resources have 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 some or all of the commitment period. Forty-four percent of this capacity was able to de-list without delay, while the ISO initially prevented 34 percent of the resources attempting to delist from doing so for reliability reasons. The remaining 22 percent of delist bids continue to be rejected by the ISO for reliability reasons. The fact that a large share of the supply attempting to de-list has been 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 delist bids in all but one of the FCAs have demonstrated that this is not the case:

- 330 MW of generation de-list bids in Connecticut in FCA1;
- 585 MW of generation de-list bids in Boston in FCA3 and FCA4; and
- 604 MW of generation de-list bids in Vermont in FCA4 and FCA5.

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 four 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 highlighted the following three market design concerns.

First, the Connecticut and Boston LSRs were much lower than the capacity requirements that were implied by the Transmission Security Analysis, which was the basis for rejecting de-list bids. As a result, de-list bids were rejected to protect Connecticut and Boston area reliability even though the Connecticut LSR was satisfied by nearly 1,200 MW in FCA1 and the Boston LSR was satisfied by over 1,600 MW in FCA3. In principle, markets should be always designed to satisfy the reliability needs of the system, which allows market prices to accurately reflect these needs. Accordingly, the ISO filed and implemented tariff changes in time for FCA4 to modify the LSR criteria to be consistent with Transmission Security Analysis used to determine whether a de-list bid should be rejected for zone-level reliability.

Second, 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. Hence, it will improve the capacity market signals when the ISO implements the Commission's directive to model eight capacity zones in each FCA.

Third, the Boston LSR was raised in FCA4 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.

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 33 shows the amounts of new capacity that were procured in the first five 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 FCA1 or it cleared in FCA1 through FCA5.[119]

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.[120]

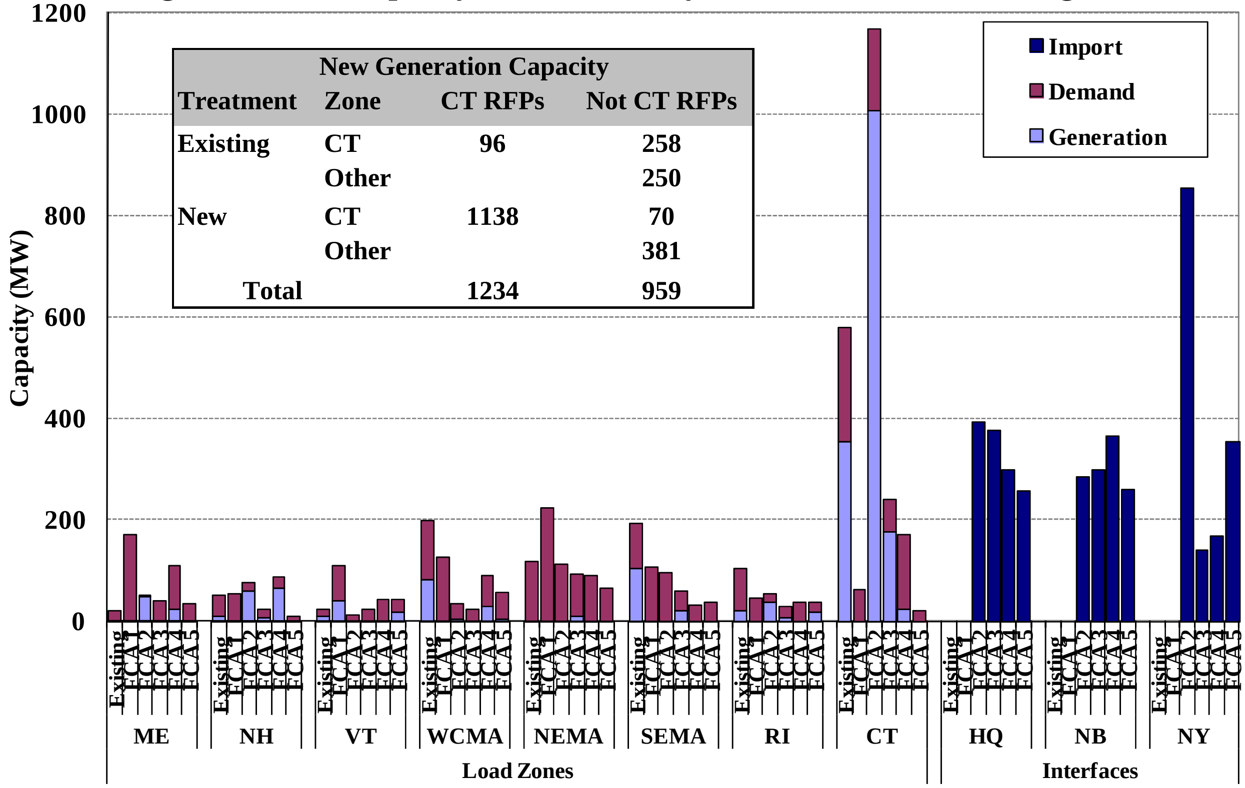
In the first five FCAs, 8.0 GW of new capacity was procured from generation, demand response resources, and imports.[121] The discussion following the figure reviews and evaluates the procurements of new capacity by resource type that are shown in Figure 33.

119 Resources expected to be in-service prior to the first Capacity Commitment Period could elect to be treated as existing resources in FCA1. Accordingly, they are able to submit de-list bids rather than supply offers.

120 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.

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

Figure 33: New Capacity Procurement by Location in FCAs 1 through 5



Import Capacity

A large quantity of new capacity was sold by importers in FCA2 through FCA5, indicating that the suppliers expected the revenues from providing capacity to New England during the Capacity Commitment Period 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 no longer used. Similarly, the amount of capacity that de-lists in order to export may increase in the future if the floor price is removed.

Demand Response Capacity

Demand response resources have sold substantial amounts of capacity under FCM, indicating 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.[122] 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 (2,193 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 five 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 generation received additional payments under the RFPs of the Connecticut DPUC and more than 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

122 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.

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 451 MW of new generating resources cleared in the five 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 five 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 more than 10 percent. 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 FCA should 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.

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.[123] 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

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

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

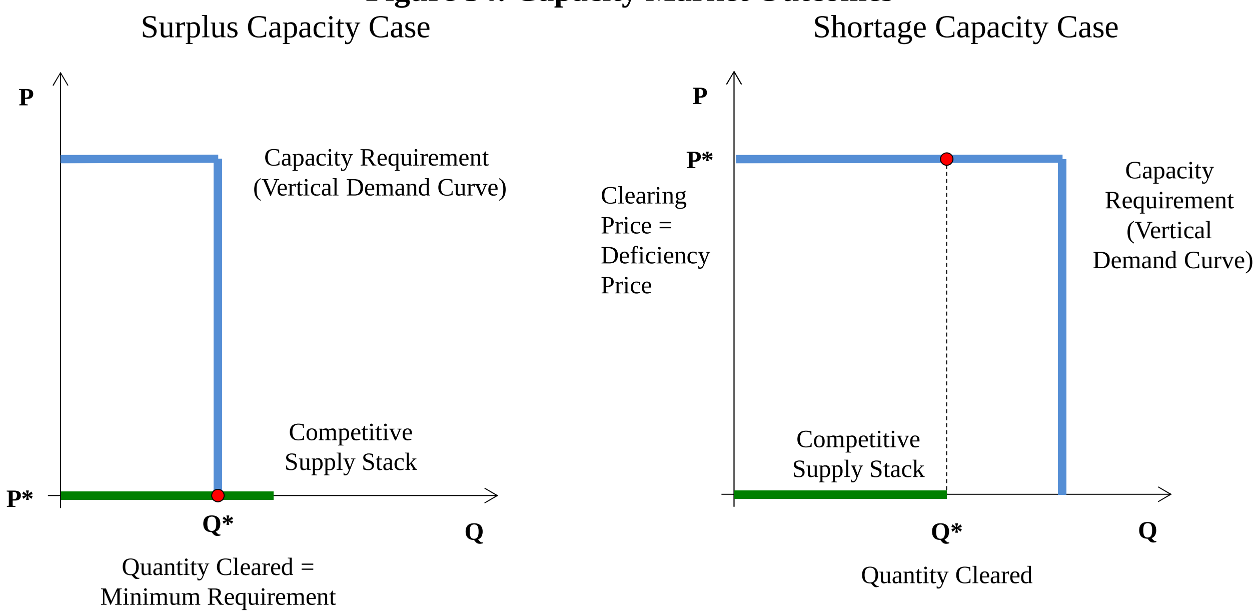
In workably competitive capacity markets, the competitive offer for capacity (i.e., the marginal cost of selling capacity) is generally close to zero.[124] A supplier's offer represents the lowest price it would be willing to accept to sell capacity. This is determined by whether there are costs the 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 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 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 34 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).

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

Figure 34: 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.

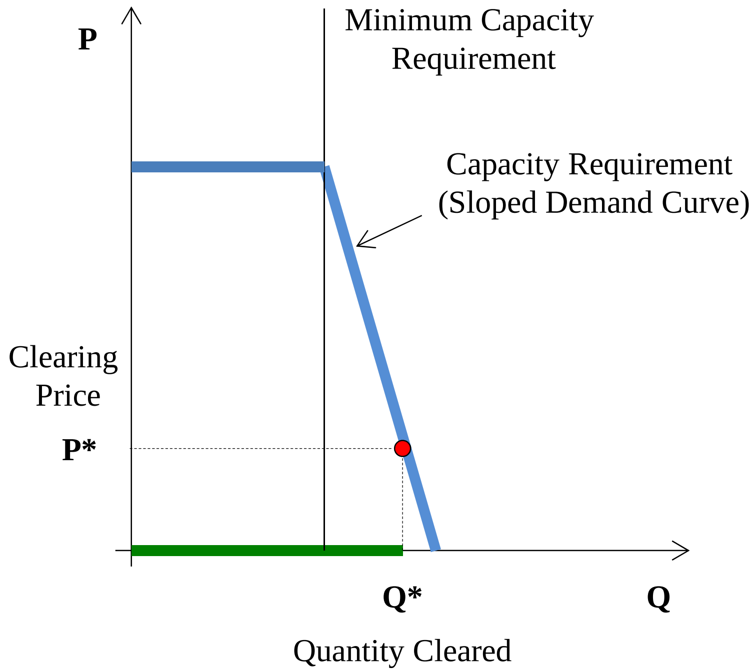
Second, since prices produced by such a construct do not accurately reflect the true marginal value of capacity, the market will not provide efficient long-term economic signals to govern investment and retirement decisions.

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 35 illustrates the sloped demand curve and the difference in how prices would be determined.

Figure 35: 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 recommended 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.

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

The purpose of a capacity market is to provide a mechanism whereby firms have efficient incentives to make large capital investments in new resources based on their expectations of market conditions over a period of decades. Greater certainty from stable and predictable market revenues lowers 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 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.[125]

- 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.[126]

---

125 See Tariff Section III.13.1.1.2.2.3(b).

126 See Tariff Section III.13.1.1.2.2.4.

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.

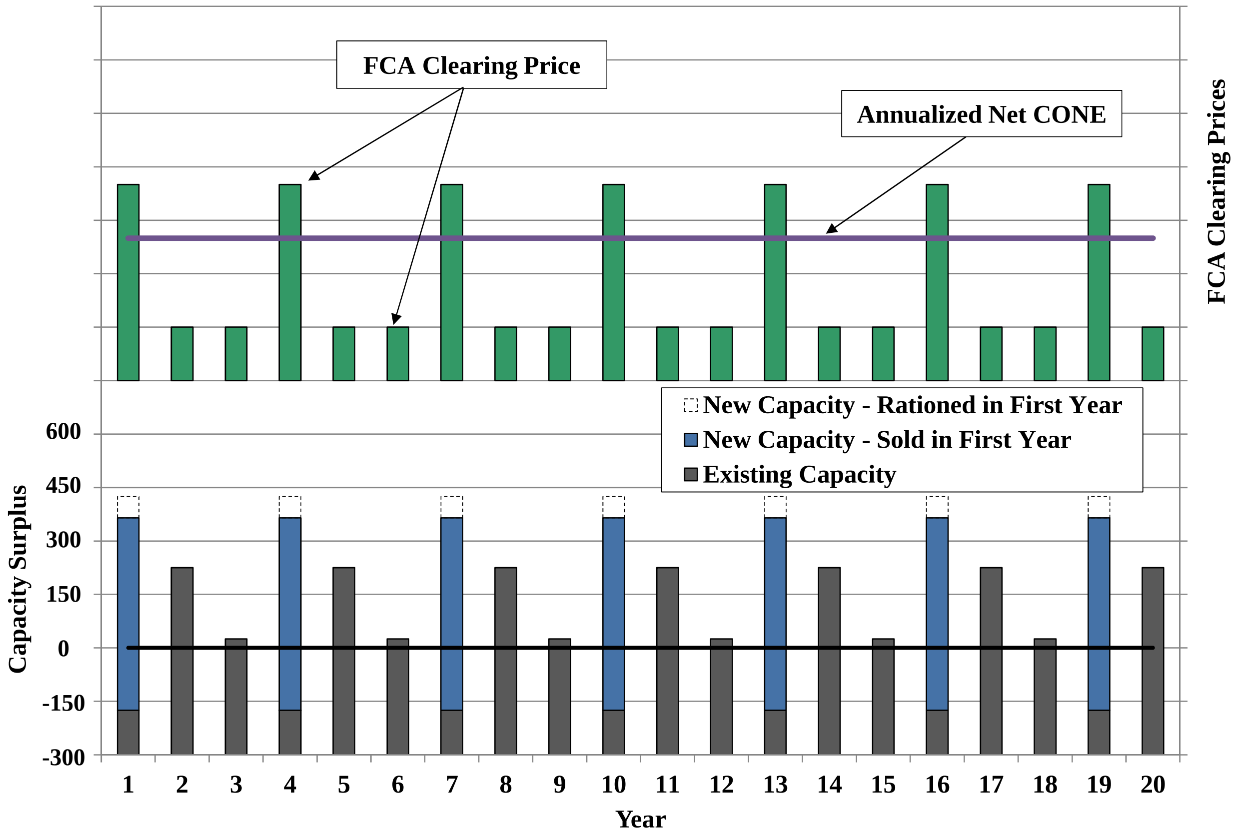
The first part of this subsection discusses the anticipated cycle of investment under the current FCM rules and how this will lead to volatile capacity clearing prices, which may in turn provide disincentives for new entry and for the capital expenditures necessary to keep existing resources in service. 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.

1. 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 36 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 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 36: 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 36.

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 36 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 36 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.[127]

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.

---

127 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 percent lower than the annualized net CONE of \$2.75.

2. 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 36.[128] 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 difference in 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).

128 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 it 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. 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 2015. The use of out-of-market payments by the ISO to retain existing resources has been virtually eliminated. This has significantly improved incentives to capacity suppliers compared with the current reliance on reliability agreements to retain existing capacity.

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. The first of these issues was addressed in time for FCA4.

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 will be partially resolved for FCA7, although the Commission in its April 2011 Order directed the ISO model all eight capacity zones in the long-run. This is consistent with our recommendation that the ISO model a full set of capacity zones and capacity transfer rights that more fully reflect the ISO's locational requirements.

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 FCA4.

The Commission's April 2011 Order also directed the ISO to work with stakeholders to develop improvements to FCM in a number of areas. The market power mitigation changes are particularly important because improved locational modeling of ISO-NE's capacity requirements will cause attempts to withhold or enter uneconomically to have larger price effects. Therefore, we recommend that the ISO implement effective buyer-side and seller-side mitigation measures in time for FCA8 in June 2013.

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.

VIII. Competitive Assessment

This section evaluates the competitive performance of the ISO-NE wholesale markets in 2011. 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 inadequate generation resources or insufficient transmission capability. 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.[129] We address four 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; and
- Potential physical withholding.

Although there are some areas and conditions under which suppliers have market power, the ISO has market power mitigation measures that are employed to prevent suppliers from exercising market power. 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 withhold at levels that would not warrant mitigation. A summary of our conclusions regarding the overall competitiveness of the wholesale market is included at the end of this section.

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

129 See, e.g., Section VIII of the "2010 Assessment of Electricity Markets in New England", Potomac Economics.

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.[130] 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 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.

130 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.

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 meet demand in real time. 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 other 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:[131]

- 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.

The findings from the structural market power analyses in this section are used to focus the analyses of potential economic and physical withholding in Sections C and D.

131 Lower SEMA was evaluated in prior reports, but is excluded from this report because the transmission constraints into the area was virtually eliminated since July 2009 when network upgrades were completed.

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 feasible withhold. Smaller "run-of-river" hydroelectric facilities are generally more limited in their ability to change output level.
Fossil-Fired	Fossil-fired units generally have relatively wide dispatch ranges and marginal production costs that closer to the prevailing LMP. Hence, they are generally the easiest and least costly resources to withhold.

Figure 37 shows import capability and two categories of installed summer capability for each region: nuclear units and all other generators in 2010 and 2011.[132] These supplies are shown as a percentage of 2010 and 2011 peak loads, respectively, although a substantial quantity of additional capacity (typically around 2,000 MW) is also necessary to maintain operating reserves

132 The import capability shown for each load pocket is the transfer capability during the peak load hour, reduced to account for local reserve requirements.

in New England. The figure shows that while imports can be used to satisfy 12 percent of the load in the New England area under peak conditions in both years, 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 37: Supply Resources versus Summer Peak Load in Each Region
2010 - 2011

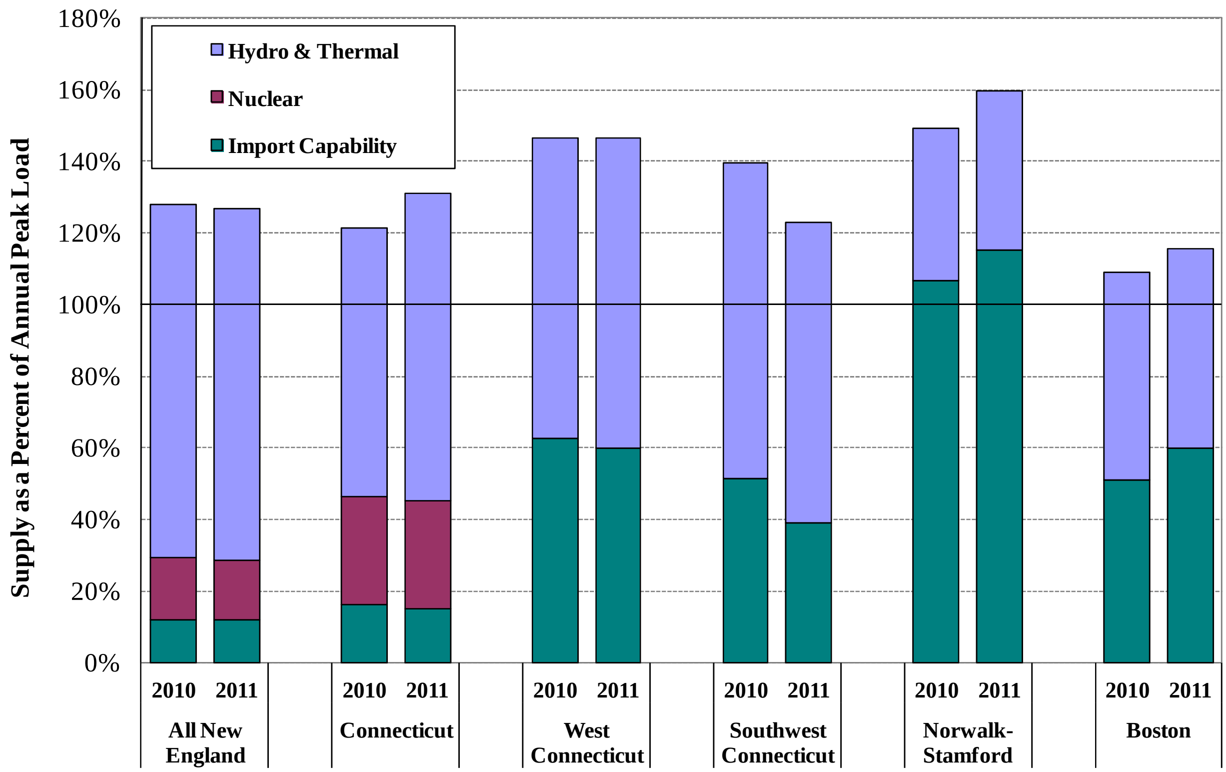


Figure 37 shows that the internal supply as a share of peak load decreased slightly from 2010 to 2011 in Boston, Southwest Connecticut, and Norwalk-Stamford. This is because the summer peak load levels rose 3 percent overall from 2010 to 2011, while there were very few changes to the supply of internal resources in these regions. However, the internal supply as a share of peak load increased from 2010 to 2011 in western Connecticut and Connecticut as a total of 820 MW of new resources entered in Connecticut in 2011 (four 50 MW gas turbines in western Connecticut and one 620 MW combined cycle unit in northern Connecticut).

The amount of import capability into each region did not change significantly from 2010 to 2011.[133] The modest 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.

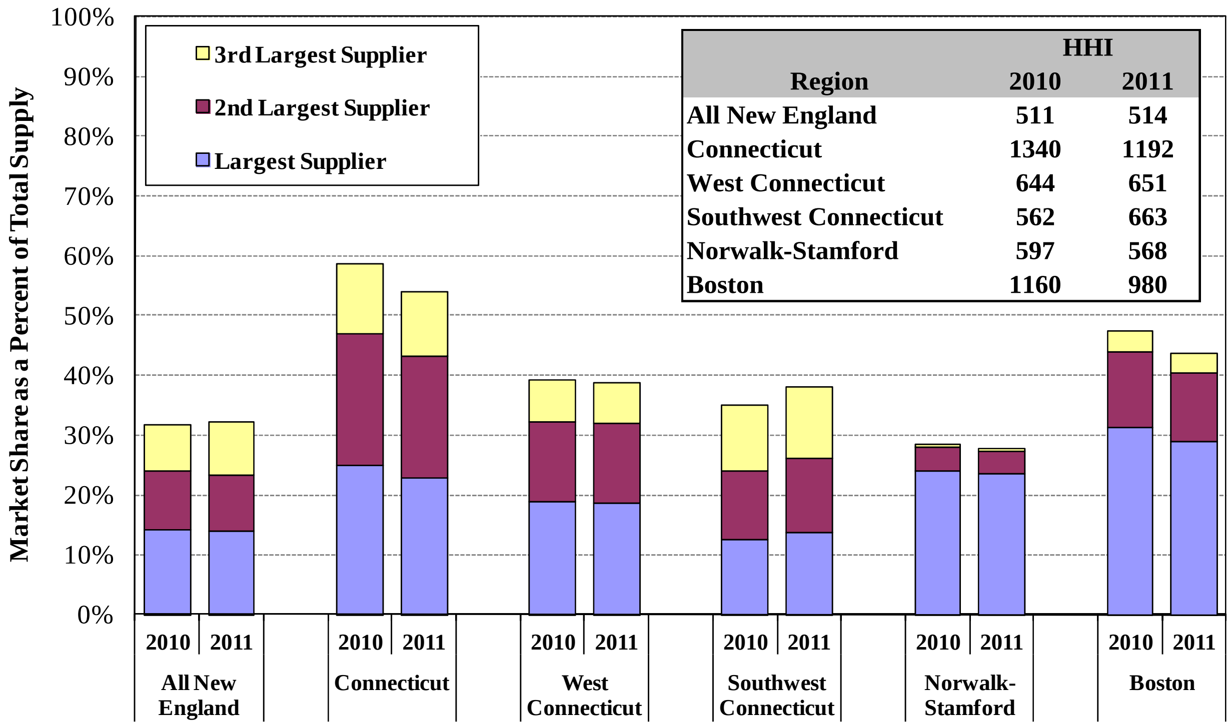
Therefore, supply conditions were generally consistent in most areas from 2010 to 2011. Figure 37 also shows the margin between peak load and the total available supply from imports and native resources. The total supply exceeded peak load in each region, ranging from 16 percent in Boston to 60 percent 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 38 shows the market shares of the largest three suppliers in the annual peak load hours in 2010 (on July 6) and in 2011 (on July 22). 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.

133 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 38: Installed Capacity Market Shares for Three Largest Suppliers
July 6, 2010 and July 22, 2011



Region	HHI 2010	HHI 2011
All New England	511	514
Connecticut	1340	1192
West Connecticut	644	651
Southwest Connecticut	562	663
Norwalk-Stamford	597	568
Boston	1160	980

Figure 38 indicates a substantial variation in market concentration across New England. In all New England, the largest supplier had a 14 percent market share in 2011. In the load pockets, the largest suppliers had market shares ranging from 14 percent in Southwest Connecticut to 29 percent in Boston in 2011. Likewise, there is variation in the number of suppliers that have significant market shares. For instance, Norwalk-Stamford had three native suppliers with very different market shares in 2011, while Southwest Connecticut had three native suppliers with comparable market shares.

The figure shows that market shares did not change significantly from 2010 to 2011 even though the ownership of some generating assets has changed. Most notably, one large firm exited the market in early 2011 when another firm acquired its entire fleet in New England (nearly 3,000 MW). The new firm also had a purchase agreement with the new 620 MW combined cycle unit in Connecticut, making it the largest supplier in Boston and the second largest supplier in New

England in 2011.[134] 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 small differences in market shares on the peak load day between 2010 and 2011 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 2011.[135] The HHI for Norwalk-Stamford is 568, 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 2011, with 1192 and 980, 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.

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.[136] The HHI's usefulness is limited by the fact that it reflects only the supply-side, ignoring demand-side factors that affect the

134 See Order Authorizing Disposition of Jurisdictional Facilities, 133 FERC ¶ 61,248, December 22, 2010.

135 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.

136 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. It is also important to note the HHI analysis employed by the antitrust agencies is not intended to determine whether a supplier has market power. For example, see Severin Borenstein, James B. Bushnell, and Christopher R. Knittel, "Market Power in Electricity Markets: Beyond Concentration Measures," Energy Journal 20(4), 1999, pp. 65-88.

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. 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 output of some of its resources is needed to meet demand 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.[137] This section of the report identifies the frequency with which one or more suppliers were pivotal in areas within New England during the study period.

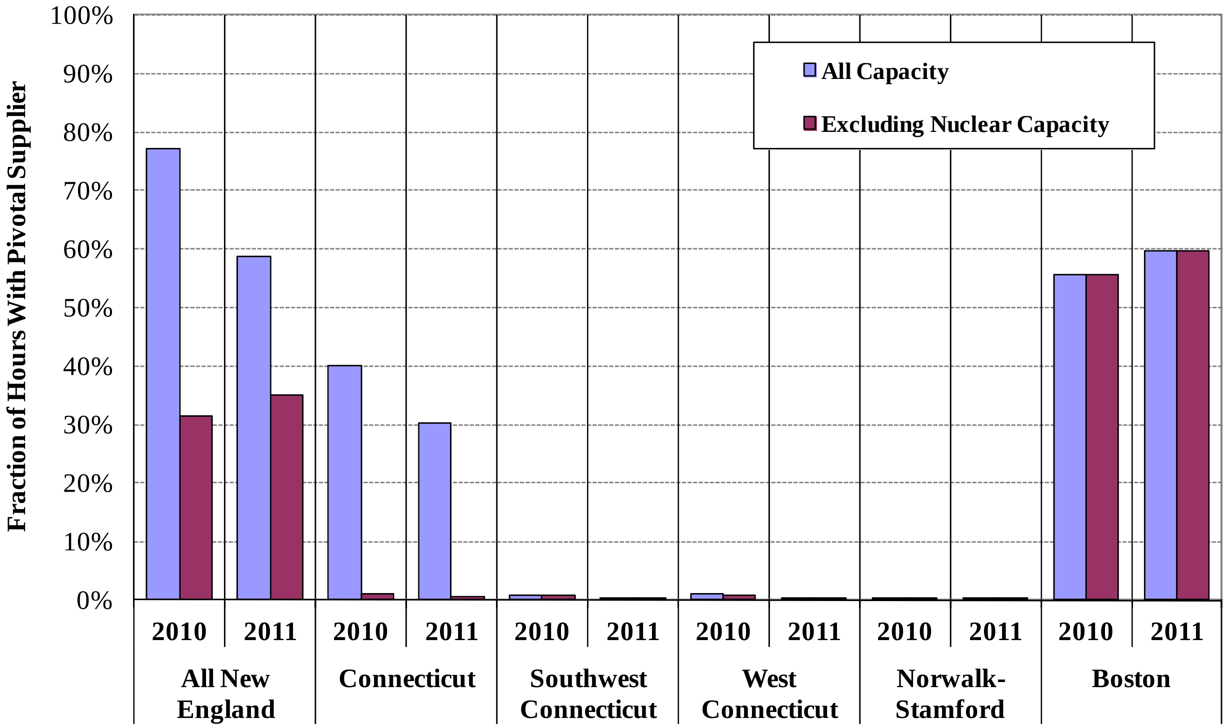
Even small suppliers can be pivotal for brief periods. For example, all suppliers are pivotal during periods of shortage. This does not mean that all suppliers should be deemed to have market power. As described above, suppliers must have both the *ability* and *incentive* to raise prices to have market 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 39 shows the portion of hours where at least one supplier was pivotal in each region during 2010 and 2011. The figure also shows the impact of excluding nuclear units. As discussed above,

137 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.

owners of nuclear units are less likely to engage in economic or physical withholding of these units.

Figure 39: Frequency of One or More Pivotal Suppliers by Type of Withheld Capacity
2010-2011



Including all categories of capacity, the pivotal supplier analysis raises potential concerns regarding three of the six areas shown in Figure 39. 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 60 percent of the internal capacity. In Boston, none of the largest supplier's capacity was nuclear capacity.

Although Connecticut had a pivotal supplier in 30 percent of hours in 2011 and 40 percent of the hours in 2010, 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 59 percent to 35 percent of hours in 2011, and from 77 percent to 31 percent of hours in 2010). 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 Connecticut and all of New England, the pivotal frequency declined from the prior year, from 40 percent to 30 percent in Connecticut, and from 77 percent to 59 percent in all of New England.[138] These decreases were primarily driven by the substantial increases in surplus capacity due to the combined effect of lower load levels, increased imports, and increased internal supply (the return of a large resource from outage and the new entry in Connecticut) . These factors greatly increased the supply margin in these areas in 2011.

The pivotal supplier summary in Figure 39 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 in each geographic market.

As described above, market power tends to be more prevalent as the level of demand grows. 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

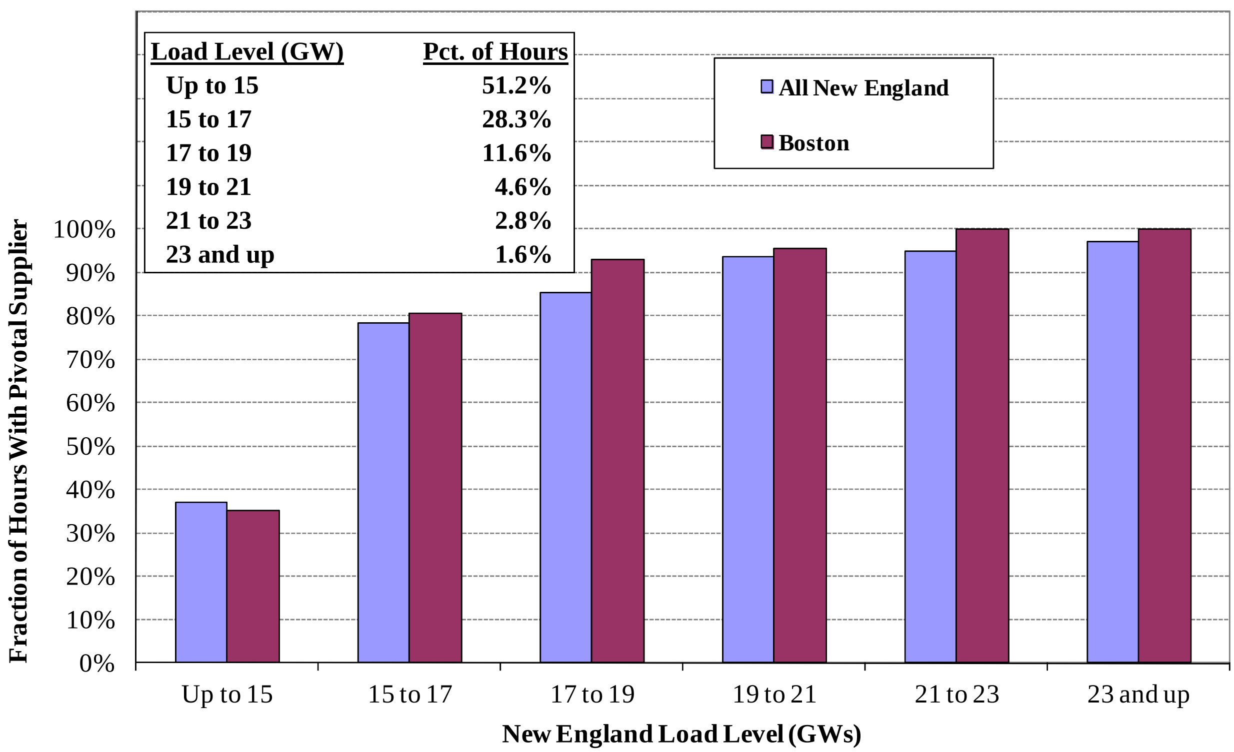
138 We refined the assumptions of available capacity in our pivotal supplier analysis and report different numbers for 2010 in this report than in the previous report.

withholding is most likely to be profitable, Figure 40 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 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 2011. 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 80 percent of hours when the load exceeded 15 GW in New England. In all of New England, the largest supplier was pivotal in 78 percent of the hours when load exceeded 15 GW. The pivotal frequency fell to around 35 percent in Boston and all of New England during hours when load was below 15 GW in New England.

Figure 40: Frequency of One or More Pivotal Suppliers by Load Level
2011



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 exceeds 15 GW during 2011. 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 Monitoring Unit (IMMU) 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 IMMU 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 hydro 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.

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 41 shows output gap results for Boston by load level. Output gap statistics are shown 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 41: Average Output Gap by Load Level and Type of Supplier
Boston, 2011

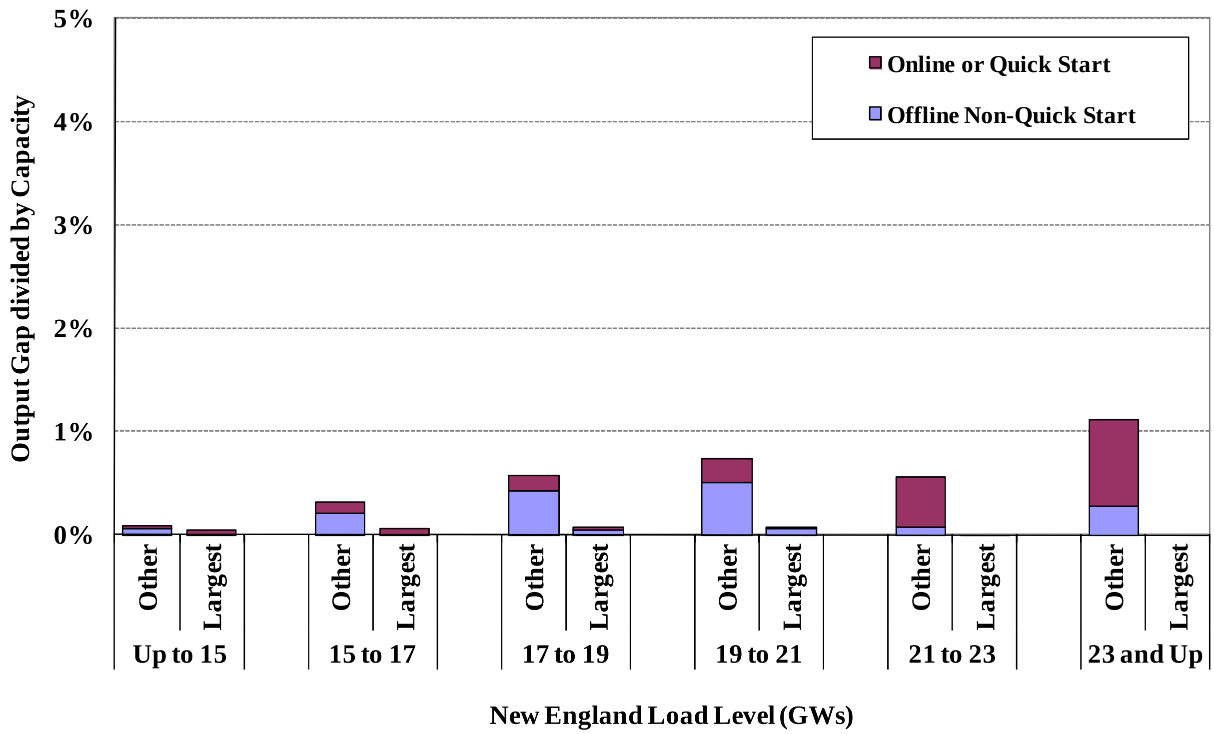
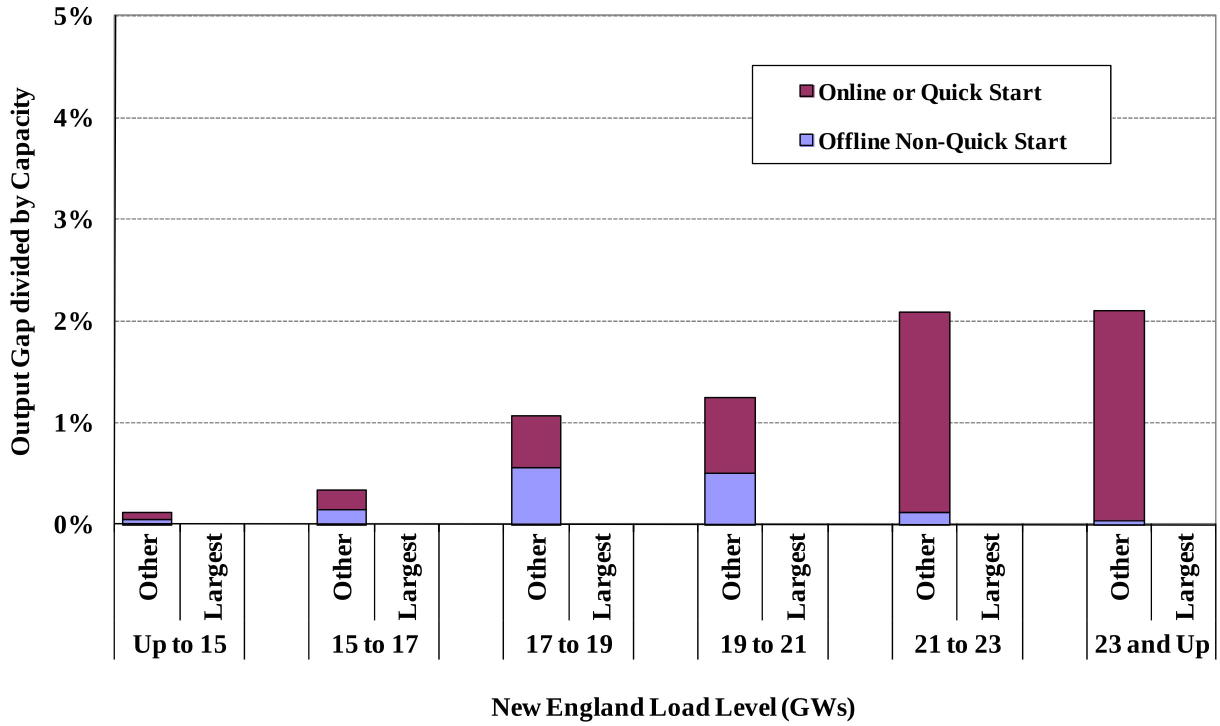


Figure 41 shows that the overall amount of output gap for the largest supplier in Boston was very small as a share of its total capacity in 2011, almost zero at all load levels. The output gap for the other suppliers was higher than that for the largest supplier, but the highest level was only about one percent when load exceeded 23 GW. 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 2011 in Connecticut. Figure 42 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 42: Average Output Gap by Load Level and Type of Supplier
Connecticut, 2011

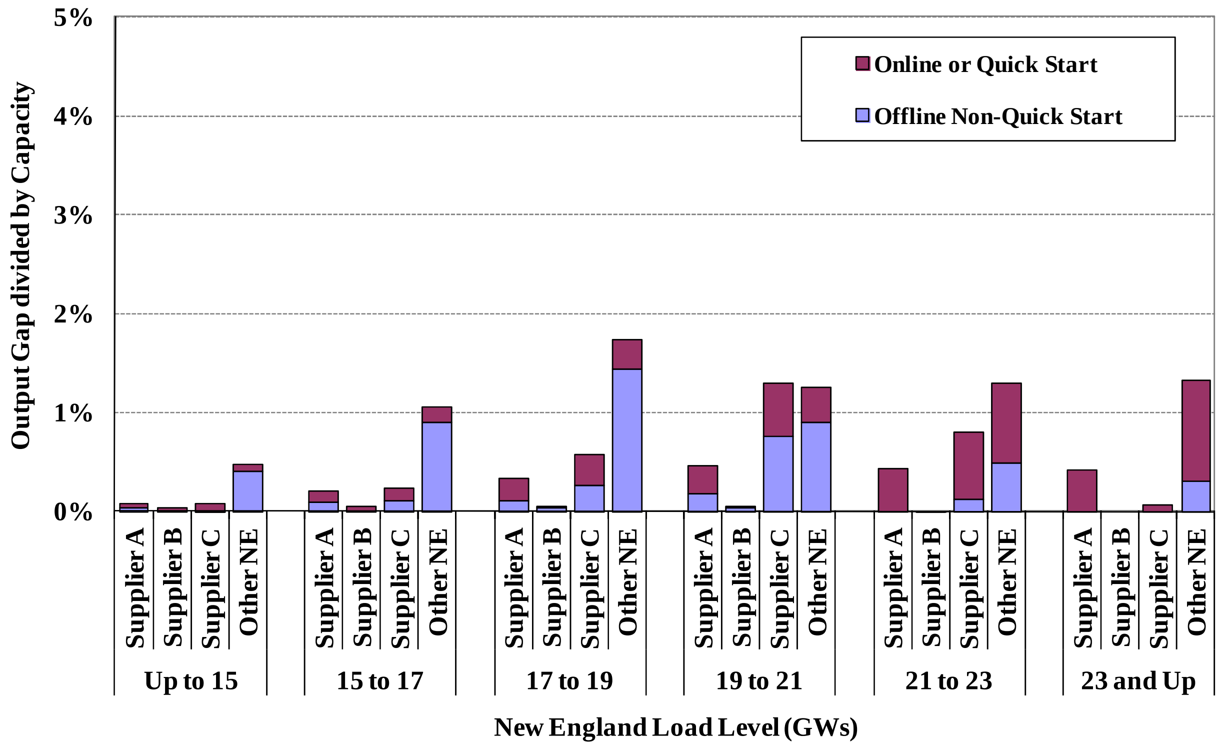


The pivotal supplier analysis indicated that the largest supplier in Connecticut was pivotal in about 30 percent of all hours when all capacity is considered, although the largest supplier owns exclusively nuclear capacity and had no output gap in 2011. Figure 42 also shows that the total output gap of all other suppliers was very low (< 2 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 43 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 59 percent of the hours during 2011. Suppliers B and C are the second and third largest suppliers in New England and were pivotal during 36 percent and 20 percent of the hours, respectively. All other suppliers are shown as a group for reference.

Figure 43: Average Output Gap by Load Level and Type of Supplier
All New England, 2011



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 under all load conditions. It is especially notable that the output gap levels for the three largest suppliers were lower than 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 relatively low and the largest suppliers' output gap amounts are lower than 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 2011.

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 44 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 2011, especially when load was less than 15 GW. This was primarily driven by two 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 44: Forced Outages and Deratings by Load Level and Supplier
Boston, 2011

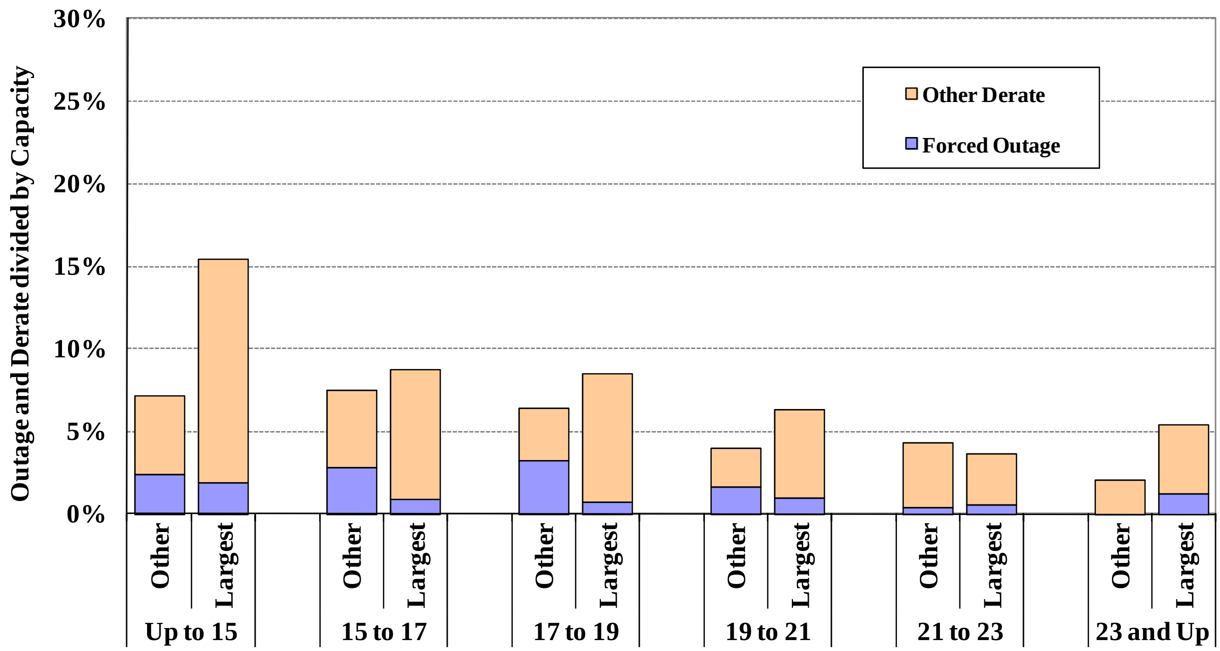


Figure 44 shows a pattern of deratings and outages consistent with expectations in a competitive market. Although levels of outages and deratings for the largest supplier were high at low load levels, they were comparable to other suppliers when load exceeded 15 GW (when withholding is most likely to 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 45 summarizes declarations of forced outages and other deratings in Connecticut by load level in 2011. The figure shows these statistics for the largest supplier in the area and compares them with statistics for other suppliers.

Figure 45: Forced Outages and Deratings by Load Level and Supplier
Connecticut, 2011

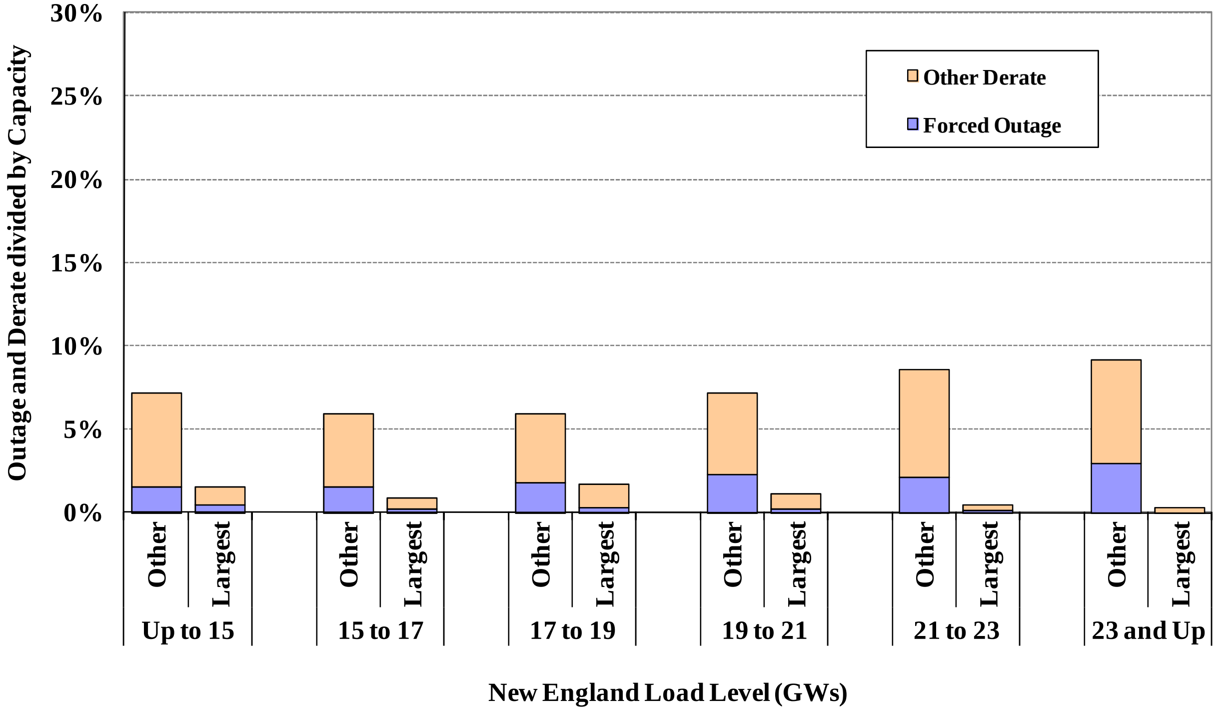
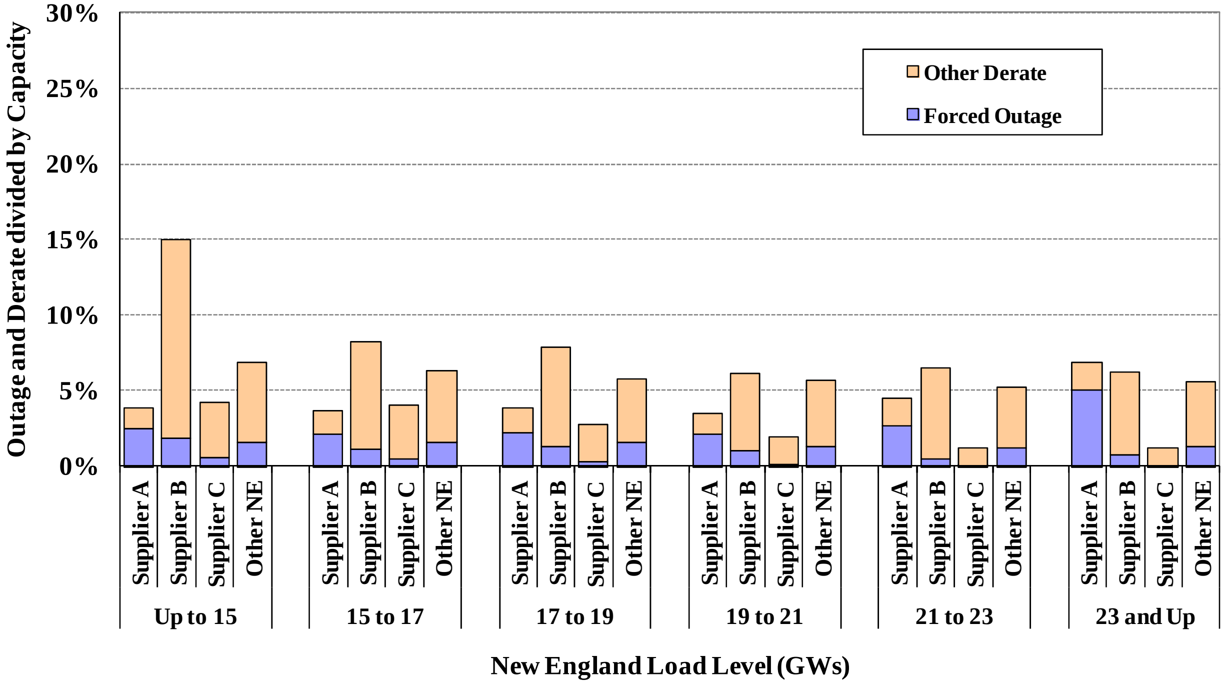


Figure 45 shows that the physical derating and forced outage quantities for the largest supplier and all other suppliers in Connecticut were moderate under all load conditions in 2011 and especially low during high load conditions. Hence, these deratings and outages do not raise concerns about physical withholding in Connecticut.

3. Potential Physical Withholding in All New England

Having analyzed the two major constrained areas in New England, Figure 46 summarizes the physical withholding analysis for all of New England by load level in 2011. 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 59 percent of the hours during 2011. Suppliers B and C are the second and third largest suppliers in New England and were pivotal during about 36 percent and 20 percent of the hours, respectively. All other suppliers are shown as a group for comparison purposes.

Figure 46: Forced Outages and Deratings by Load Level and Supplier
All New England, 2011



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 19 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. 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 2011. The pivotal supplier analysis suggests that market power concerns exist in several areas in New England. However, the abuse of this market power is limited 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.[139] The ISO substantially strengthened the market power mitigation measures in April 2012 when it implemented Automated Mitigation Procedures (AMP). 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. Nonetheless, we continue to monitor market outcomes closely for potential economic and physical withholding together with the IMMU.

139 See Market Rule 1 Appendix A – Market Monitoring, Reporting and Market Power Mitigation for details.

Document Content(s)

ISONE 2011 EMMU Report_Final_June 2012.PDF.............................1-183