2004 Assessment of the Electricity Markets in New England

By:

David B. Patton, Ph.D.
Pallas LeeVanSchaick
POTOMAC ECONOMICS, LTD.

Independent Market Monitoring Unit
ISO New England Inc.

June 2005

Table of Contents

I. Executive Summary

This report assesses the operational efficiency and competitiveness of New England's wholesale electricity markets during 2004. The current Standard Market Design ("SMD") wholesale electricity markets began operation on March 1, 2003 and include day-ahead and real-time energy markets, a regulation market, and a forward reserve market. These new markets were added to a pre-existing capacity market. The ISO is currently developing other ancillary services markets ("ASM"), including operating reserve markets, and other key enhancements to the initial SMD markets. The SMD markets are a considerable improvement over the previous market design, providing a much more efficient means to manage network congestion and set energy prices.

The SMD energy markets efficiently dispatch generation on the basis of supply offers to satisfy energy demand and operating reserve requirements, while preventing power flows on the network from exceeding transmission constraints. The markets establish locational marginal prices ("LMPs") that reflect the marginal system cost of serving load at each location on the network. When the market is functioning well, these prices ensure the efficient dispatch of generation in the short run, provide transparent price signals that facilitate efficient forward contracting, and are a primary component of the long-term incentives that guide generation and transmission investment and retirement decisions.

Based on the analyses presented in this report, we draw the following conclusions in three primary areas. The first area is the competitive performance of the electricity markets in New England. Overall, we conclude that the markets have performed competitively. We analyzed the overall market in New England, as well as a number of constrained areas within the market, and found little evidence that any suppliers were either economically or physically withholding resources to raise prices. However, this report confirms prior findings that in late 2004 one supplier began exercising local market power in the Boston area, and its conduct resulted in considerable increases in the operating reserve payments to the supplier. The market power resulted from the local reliability requirements for that area that compel the ISO to commit generation outside of the market processes. This issue was detected by the ISO's market monitor, who consulted with us regarding a proposal to modify the mitigation measure that would allow the Market Monitor to more effectively address this issue. This proposal has been approved by FERC and implemented by the ISO. In addition, we propose a change in the market rules that would further limit this conduct.

Second, we analyze the operation and short-term efficiency of the markets. In this area, we find that the markets have generally operated well. For the most part, prices in both the day-ahead and real-time energy markets have efficiently reflected underlying market fundamentals. For example, electricity prices in New England have closely tracked changes in underlying fuel prices as one would expect in a well-functioning market.

To maintain reliability in the constrained areas, however, the ISO has continued to commit substantial quantities of additional resources to supplement the market-based commitments in the day-ahead market. The need for these commitments is largely due to the limited quantities in these areas of resources that can start quickly, such as gas turbines. Therefore, the ISO must start up larger, slower-starting steam and combined-cycle generation to ensure reliability and manage voltage in constrained areas.

This additional online supply in constrained areas substantially reduces the congestion into these areas and generates significant supplemental charges to New England loads that are difficult for them to hedge (we refer to these charges as "uplift charges"). Therefore, reducing the need for these supplemental commitments should remain a high priority for the ISO. To that end, the ISO has made a number of changes in its market rules and worked with market participants to address the underlying reasons for the supplemental commitments. For example, the ISO has worked with participants to install equipment and make other operational changes to improve the ISO's ability to manage voltage in the Boston area without committing additional generation. Some of these improvements were implemented in late 2004 and contributed to a sharp reduction in supplemental commitments for voltage support in December 2004. A number of improvements have been implemented in 2005. Further enhancements are planned.

The ISO has also implemented a number of other significant changes, some of which were recommended in our prior assessment of the SMD markets.[1] For example, the commitment process in the day-ahead market was substantially improved in early 2005 to recognize transmission constraints and local reliability requirements in the constrained areas. This improvement has reduced the need for the ISO to rely on manual commitments by the day-ahead market operators. Due to the timing of this improvement, it is not reflected in the analyses in this report of 2004.

The third area in which we draw conclusions pertains to the long-run economic signals produced by the markets in New England. The issues and conclusions in this area are closely related to the short-run issues described above. In particular, the local reliability requirements that compel the ISO to make supplemental commitments in the constrained areas are not reflected in the market prices in New England. These supplemental commitments tend to alleviate congestion into the constrained areas and prevent the market prices from fully reflecting the economic value of energy produced and consumed in these areas. The fact that significant reliability requirements are not priced within the New England market framework causes the long-term economic signal in the key constrained areas to be understated, which limits the entry of new resources that are needed. These understated price signals have also resulted in a heavy reliance on reliability-must-run ("RMR") contracts to ensure that existing generation needed for reliability in these areas remains in operation. Reliability agreements are poor substitutes for efficient, transparent market prices.

We support four significant changes to the SMD markets, some of which we have recommended in past reports and the ISO is currently developing or implementing. These changes will address the efficiency of the long-term signals produced by the market and provide other economic and reliability benefits to the market.

First, the ISO has proposed a capacity market with location-specific requirements in the constrained areas. This locational installed capacity market ("LICAP") will provide a market for

---

1 *Six-Month Review of SMD Electricity Markets in New England*, Potomac Economics, February 2004.

capacity in the constrained areas to meet New England's local reliability requirements. This will help fill a significant gap in the ISO's current markets and should substantially reduce the need to employ reliability agreements or other types of supplemental arrangements to ensure that reliability requirements are satisfied.

The second major change in the ISO's markets is the development and implementation of its ASM project, which includes introducing real-time operating reserve markets and revising the forward reserve market to include locational requirements. The real-time operating reserve markets will be co-optimized with the real-time energy market and will enable the SMD markets to:

- reflect the economic relationship between operating reserves and energy that will lead to more efficient price signals for both products, particularly during supply shortages;
- more fully recognize the economic value of reliability requirements. It is these requirements that have resulted in supplemental commitment and out-of-merit dispatch;
- create incentives for units to provide operational flexibility to the system; and
- improve the efficiency of generator commitment and dispatch.

Implementing locational requirements in the forward reserve market should cause a larger share of the resources needed to meet local reliability requirements to be self-committed by the forward reserve suppliers and, as a consequence, it should reduce uplift costs in the constrained areas. It will also supplement the economic signals in the constrained areas by recognizing the need for reserves in those areas.

These first two changes in New England electricity markets, implementation of the LICAP market and the ancillary services markets, will mitigate the concerns raised in this report regarding the long-term economic signals produced by the markets. Hence, we recommend that the ISO make these improvements its highest priorities.

A third major change in the New England markets is a set of provisions being developed and tested that would better coordinate the physical interchange between New York and New England. We have been recommending the development of these provisions for the past three years in both New York and New England because they would facilitate a more seamless market in the Northeast, which would ensure that power is efficiently transmitted to the highest-value locations, achieve substantial economic savings for customers in the region, and improve reliability. These provisions also hold the promise of efficiently utilizing the controllable transmission lines between Connecticut and Long Island, two of the most congested locations in the Northeast. The ISO is working closely with the New York ISO ("NYISO") and the market participants to develop the provisions, and they all have recently engaged in a limited pilot project to test the feasibility of altering the physical interchange between the markets within an hour.

A fourth area of improvement is related to increasing the participation by demand-side resources in the market, which will increase market efficiency and reliability. Although the ISO has a variety of programs to facilitate demand-side participation in the New England markets, this participation has been limited. Therefore, we recommend that the ISO evaluate the economic provisions in its programs to determine whether they should be modified in the near term to increase participation. For example, participants in New York ISO's emergency demand response program are paid the higher of $500 per MWh or the LMP at their location when they curtail load in response to an ISO request. This program has generated more than 1000 MW of load curtailment during peak conditions. Importantly, the $500 per MWh payment sets the energy price in New York when the demand response is needed to avoid a shortage of reserves. As long as demand response is generally called under these circumstances (i.e., when needed to avoid a shortage) and the costs are allocated appropriately, it increases market efficiency. In the long term, the ISO should continue to expand the options for demand-response resources to participate in the real-time energy market, operating reserves markets, and capacity market.

Other potential improvements to the SMD markets involve modifications to certain operating procedures and rules that will increase the efficiency of the SMD markets. These recommended modifications are described in the following sections along with a summary of the report's findings and conclusions in each area.

Energy Price Trends

Energy prices have closely tracked movements in natural gas prices. 2004 began with a period of extraordinary volatility in the market for natural gas and corresponding volatility in electricity prices. After January, natural gas prices ranged between $5 and $10 per MMBTU, and movements in average electricity prices tracked the movements in fuel prices. This correlation between natural gas prices and electricity prices is consistent with a well-performing market given that: a) fuel costs constitute the vast majority of most generators' marginal costs, and b) natural gas-fired units are frequently on the margin, thus setting the market price in New England.

The energy prices in New England continued to exhibit very low levels of congestion, due in part to the continued need to commit supplemental generation in constrained areas. The additional supply online in the constrained areas associated with these commitments serves to reduce imports and congestion into these areas.

Prices during the summer were moderate due to the mild weather conditions. Although New England has implemented provisions to ensure that shortages of operating reserves are reflected in relatively high energy prices, no such shortages occurred in 2004.

Day-Ahead to Real-Time Price Convergence and Virtual Trading

In addition to energy prices' tracking movements in fuel prices, the energy prices in the day-ahead and real-time markets have converged more closely under SMD than have the corresponding prices in either PJM or New York. Measured at the New England Hub, New England exhibited a price premium in the day-ahead market of 3.0 percent in 2004, although the premium is only 1.9 percent of the highly volatile period during the Cold Snap in January 2004 is excluded. These results are generally consistent with historical patterns in New York and PJM prior to 2004. A day-ahead price premium is common in a two-settlement energy market. Buyers can reduce their risk by purchasing power at less volatile day-ahead prices, while some sellers may have additional outage risks when scheduled day-ahead. These factors tend to raise the day-ahead price relative to the expected real-time price, although this is generally mitigated by virtual traders who seek to arbitrage predictable price differences between the two markets.

In 2004, the average hourly price differences between the day-ahead and real-time markets were smaller in New England than in the other nodal energy markets. This reflects efficient scheduling behavior on the part of market participants in New England, as well as low overall price volatility during the year. In addition, we find signs of good overall consistency between the day-ahead and real-time market models, which contributes to convergence between the day-ahead and real-time markets. However, in 2004 there were significant differences in the Northeast Massachusetts/Boston area ("NEMA/Boston") import limits used in the day-ahead and real-time models. The import limit has generally been lower in the day-ahead market, resulting in higher levels of congestion in the day-ahead market than in real time. We recommend that the ISO investigate what factors have led to the systematic differences to determine whether improvements could be made to bring the limits into better alignment.

We find that virtual trading was initially more active after the day-ahead market opened in March 2003, but has since declined in contrast to its growth in other markets. Virtual trading is the practice of purchasing energy in the day-ahead market and selling it in the real-time market (virtual load) or selling energy in the day-ahead market and buying it back in the real-time market. These are purely financial transactions, and to trade the participant need not own or control any physical generation or load. Virtual bids and offers can be submitted for any location in New England.

During 2004, virtual schedules were assessed charges for supplemental generator commitments that averaged more than $1.32 per MWh outside load pockets and frequently more than $10 per MWh in Connecticut and NEMA/Boston. These charges reduce the incentive of virtual traders to arbitrage day-ahead and real-time prices, which has a negative impact on overall market efficiency. The ISO's Tariff was changed in March 2005 to allocate the costs of satisfying local reliability requirements to the network load in the constrained area, rather than to energy traded in the real-time market (including virtual trades). This is likely to improve price convergence and scheduling efficiency in Connecticut and NEMA/Boston. In addition, we recommend the ISO consider other changes in the allocation of these charges that would improve the incentives of virtual traders to arbitrage sustained differences between day-ahead and real-time prices.

Efficiency of New England Hub

The report evaluates congestion experienced in the day-ahead market at the New England hub. The hub is composed of 32 nodes in the geographic center of New England, and the hub price is an arithmetic average of the LMPs at these nodes. The New England hub price is calculated and posted by the ISO to facilitate trading in New England. When virtual purchases and sales are made in the day-ahead market at the hub, the injection or withdrawal is evenly distributed over the 32 nodes. In hours with relatively large net virtual load schedules, congestion has arisen between these nodes that would not reasonably be expected in real time. This intra-hub congestion reduces the value of the hub as a facilitator of trade because it can cause the hub price to be a poor reflection of the value of power in New England.

We show that while intra-hub congestion has declined markedly from 2003 to 2004, congestion-related price differences still occur at four of the 32 nodes. We also find that virtual scheduling at individual nodes within the hub has reduced intra-hub congestion. However, the volume of virtual trading at individual hub nodes has been small and not sufficient to completely eliminate these differences. It is likely that the allocation of significant uplift charges to virtual traders has inhibited them from further reducing congestion-related price differences at the hub. This supports our recommendation to reevaluate the allocation of these charges, but we find no need to make any changes in the definition of the hub.

Congestion and Local Reliability Requirements

New England experienced very little congestion during 2003 and 2004 under SMD. Congestion into historically-constrained areas, such as the NEMA/Boston area and Connecticut, has been notably mild. In fact, most of the price separation between net exporting regions and net importing regions has been due to transmission losses rather than transmission congestion. For instance in 2004, 74 percent of the difference between prices in Maine and Connecticut was due to losses and only 26 percent was due to congestion.

The report identifies operating requirements and procedures that tend to reduce the levels of congestion into chronically-constrained areas, including NEMA/Boston and several areas in Connecticut. To ensure reliability in these areas, the ISO operates with capacity requirements that are comparable to location-specific reserve requirements. Until the ISO implements locational reserve requirements or locational capacity requirements, these reliability requirements will not be reflected in the market outcomes. Hence, the SMD markets will not generally satisfy these requirements, which causes the ISO to make supplemental commitments and to dispatch generation out of merit order to maintain reliability in the constrained areas. This additional supply in the real-time market reduces real-time prices, mutes congestion into the constrained areas, and creates incentives for load to under-schedule in import-constrained areas. The ISO has taken some actions to reduce the supplemental commitments, and these are described below. In the longer term, the implementation of the ISO's LICAP market and ancillary services markets will more fully address these issues.

Financial Transmission Rights

In 2004, the Financial Transmission Rights ("FTRs") markets resulted in prices that more reasonably reflected expected congestion levels than they did in 2003. The improvement in the FTR pricing was likely due to higher volumes of offers in the market and experience gained by the market participants since the implementation of the SMD markets. FTR prices in the six-month auctions were consistent with congestion levels observed in the spot market during 2003, which generally exceeded the congestion incurred in 2004. However, summer peak demand was unusually low in 2004 and contributed to the reduction in congestion.

Market Operations

This section covers a wide variety of areas related to the operation of the SMD markets, including the market consequences of certain operating procedures and the scheduling actions of participants.

Price Corrections. Price corrections are frequently an indicator of implementation problems or software errors. While the rate of price corrections was low in 2003, it decreased 60 percent in 2004. Prices were corrected in less than one percent of the five-minute intervals since the implementation of SMD, and corrections have been substantially less frequent than the rate reported in New York.

Load Forecasting. Day-ahead load forecasting, which is an important determinant of efficient day-ahead commitment, has been very accurate. The average absolute forecast error for the peak load in New England has been 1.8 percent, compared to 2.1 percent in New York and 3.5 percent in PJM. The accuracy of the ISO's forecasts is important because the forecasts are a key input to the ISO's reliability assessments, forecasted transmission limits, and supplemental commitment decisions. They also provide information to market participants for their day-ahead scheduling and bidding.

Supplemental Commitment and Out-of Merit Dispatch. We find substantial quantities of supplemental commitment in both NEMA/Boston and Connecticut. Supplemental commitment for local reliability has increased 88 percent in NEMA/Boston from an hourly average of 324 MW in 2003 to an hourly average of 610 MW in 2004. Nearly all of the increase is due to additional commitment for voltage support prior to the day-ahead market. Supplemental commitment for local reliability has decreased 64 percent outside NEMA/Boston from an hourly average of 770 MW in 2003 to 275 MW in 2004. This reduction reflects fewer commitments for local 1st and 2nd contingency reliability requirements, particularly in Connecticut. While commitment for voltage support tends to decrease under peak demand conditions, commitments for local capacity requirements increase substantially on the highest-demand days. These commitments are necessary, in part, because these areas do not have a large quantity of quick-start resources that can help meet the capacity requirements of the local area while offline.

Supplemental commitments also frequently result in a significant quantity of out-of-merit dispatch, i.e., energy produced by resources whose energy offer prices are higher than the market energy price. This occurs because once they are committed, online resources must be dispatched at or above their minimum output parameter ("EcoMin"). These units cannot be shut down since their capacity is needed to satisfy the local capacity requirement. Since out-of-merit resources are treated as must-take resources (equivalent to assigning them a zero offer price) and are not eligible to set LMPs, they displace the marginal source of energy. This results in lower prices in constrained areas and in the broader New England market as well. Since most of the out-of-merit energy is produced from resources committed supplementally, the changes in out-of-merit energy generally mirror changes in supplemental commitments. Hence, the report finds that out-of-merit energy increased in NEMA/Boston from an average of 75 MW per hour in 2003 to 233 MW per hour in 2004 and decreased outside NEMA/Boston from an average of 291 MW per hour in 2003 to 217 MW per hour in 2004.

Supplemental commitments and out-of-merit energy dispatch create four issues in the New England market.

- They create inefficiencies because supplemental commitments are made with the objective of minimizing commitment costs (i.e., start-up, no-load, and energy costs at EcoMin), rather than minimizing the overall production costs.

- They tend to mute signals to invest in areas that would benefit the most from additional generation and transmission investment. They also stifle interest in registering potential demand response by diminishing the financial incentives for it.

- They can create incentives for generators frequently committed for reliability to avoid market-based commitment when they would be economic at the day-ahead LMP. This frequently induces the ISO to commit the resource in the Resource Adequacy Assessment ("RAA") process for local reliability where the generator is paid its bid price in the form of uplift. When the generator is not committed in the RAA, but expects to be economic at the real-time LMP, it simply commits itself after the RAA. The report finds that two generators in the NEMA/Boston area did this with regularity during the month of December, when they accounted for 88 percent of the unit-hours and 99 percent of the MWh of capacity self-committed after the RAA process.

- They cause a substantial amount of uplift costs that is difficult for participants to hedge and can be quite volatile, most of which are generated by commitments in Connecticut and NEMA/Boston. The uplift costs associated with these commitments are allocated in a variety of ways based on Tariff requirements. Some of these allocations can create inefficient incentives. The report discusses these allocations and recommends improvements.

The ISO has already implemented several changes that should reduce the need for supplemental commitments and improve the economic signals in the constrained areas. The most important change is the improvement in the commitment software and process to recognize transmission limits in the day-ahead market commitment. These limits include the first contingency limits and "proxy 2^{nd} contingency" limits that recognize the 2^{nd} contingency reliability requirements in Connecticut and Boston. Day-ahead market operators had previously accounted for these limits by manually adjusting the day-ahead market commitment. Other measures being pursued to minimize reliance on supplemental commitments in load pockets include:

- Coordinating with NSTAR and one of the suppliers in Boston to increase the capability of the transmission system to produce and absorb more reactive power in key locations – several improvements were made in 2004 and others should be completed in early 2005. These improvements will reduce the need for supplemental commitments for voltage support, the largest source of supplemental commitments in 2004.
- Developing a new Combined Cycle unit dispatch process to gain additional unit flexibility and non-spin capability in load pockets;
- Developing a new day-ahead commitment plan for units with reliability agreements;
- Identifying market enhancements to capture out-of-merit dispatch costs in reserve prices;
- Developing new ancillary services markets to provide better incentives for resources in the load pockets, particularly for new quick-start units; and
- Modifying the methodology for calculating references prices for units frequently committed for local reliability in constrained areas.

In addition, we recommend the following changes to further reduce the inefficiencies associated with supplemental commitments. We recommend that the ISO:

- Consider the merits of not allowing suppliers in load pockets to self-commit units after the RAA process unless they have suffered an outage on another unit or they provide comparable justification. This would reduce the quantity of supplemental commitments, improve the ISO's decision-making in the RAA process, and increase suppliers' incentives to offer resources competitively in the RAA since it would be their last opportunity to commit a unit;

- Allocate uplift for local 1st contingency commitments in the same manner as local reliability uplift is allocated. Currently, uplift for local 1st contingency commitments is assessed to market participants based on their scheduling behavior in the day-ahead and real-time market. Instead, it should be allocated to the physical load in the area that benefits from the commitment. This change would enhance incentives for virtual trading and price-responsive load scheduling in the day-ahead market. Additionally, it would recognize that commitments for local reliability protect all load in the area, regardless of whether the load settles in the day-ahead market or real-time market;

- Allocate uplift for voltage support commitments in the same manner as local reliability uplift is allocated. Currently, uplift for voltage support commitments is assessed to all New England load, although voltage support primarily benefits load in the local area. Assessing this uplift to the local area will provide appropriate incentives to upgrade the transmission system. This change is currently being considered by the NEPOOL Tariff Committee ; and

- Evaluate the assumptions underlying the calculation of the import limits to constrained areas to resolve the inconsistencies between the day-ahead and real-time limits. This would improve the efficiency of the day-ahead commitment and tighten convergence between day-ahead and real-time market outcomes.

Regulation Market

The regulation market includes an economic evaluation to determine a clearing price the evening prior to the market day. However, when the market was initially implemented with the start of SMD, participants whose offers were above the clearing price were able to self-schedule regulation in real time without affecting the clearing price. This flaw undermined the incentives for suppliers to offer regulation service at marginal cost and resulted in inefficient selection of resources for regulation. The ISO identified this concern and proposed that units not be allowed to self-schedule after the regulation clearing price has been determined. After this change was implemented in February 2004, regulation prices decreased by nearly 50 percent within two months and remained lower for the rest of 2004. Additional changes to the regulation market planned for 2005 should further improve its performance.

Competitive Assessment

This section of the report evaluates the market concentration and competitive performance of the markets operated by the ISO–New England in 2004. Under locational marginal pricing, there may be greater potential for certain participants to exercise market power in geographic markets that are smaller than the entire ISO–New England footprint. This evaluation characterizes the geographic areas and market conditions that are most susceptible to the exercise of market power by at least one large supplier. The following areas are examined:

- All of New England;
- Connecticut;
- Southwest Connecticut;
- Norwalk-Stamford , which is in southwest Connecticut;
- The Middletown portion of Connecticut; and
- NEMA/Boston.

The first part of our assessment evaluates each geographic market using a pivotal supplier analysis to determine the demand conditions under which a supplier may have market power. We find that the largest suppliers in Norwalk-Stamford, Middletown, and NEMA/Boston are pivotal in a large number of hours. However, these areas contain large amounts of nuclear capacity and capacity covered under RMR agreements which significantly mitigate the incentives to exercise market power. After taking nuclear capacity and RMR agreements into account, the areas where market power was the greatest concern were: (i) NEMA/Boston, where a supplier is pivotal in 16 percent of all hours and (ii) all of New England, where a supplier is pivotal in 2 percent of all hours. However, this analysis suggests that once the RMR agreements expire, market power will be a significant concern within Connecticut as well.

The second part of this assessment examines market participant behavior to determine whether it was consistent with the profitable exercise of market power. We measure potential economic and physical withholding for all resources and find little evidence of behavior that is consistent with the exercise of market power by large suppliers. Congestion was relatively low during 2004 due to mild summer weather and substantial supplemental commitment. Strategic withholding is likely to be a bigger concern in the future under higher demand conditions with less excess commitment in local areas. Thus, we recommend that the ISO continue to monitor structural and behavioral market power indicators.

While there is no substantial evidence that any suppliers exercised market power by withholding capacity to raise clearing prices, a supplier with resources needed for specific local reliability requirements can still exercise market power if it does not face competition from other suppliers that can also meet those requirements. A supplier can exercise market power by inflating the guarantee payments the ISO must make to utilize the supplier's resources. Although such suppliers face pay-as-bid incentives, if they increase their offer prices by more than they would if they faced competition from other suppliers, one may conclude that they are exercising local market power.

Based on our review of the commitment patterns, offers, and uplift payments made in the Boston area, we conclude that in late 2004 a significant exercise of local market power began that has continued into 2005. Although this was detected when it began occurring, it could not be effectively mitigated using the existing mitigation measures for economic withholding because the reference prices used for the resources in question were inflated. To correct this, the ISO filed for a change in the reference price calculation methodology that should substantially improve the effectiveness of the mitigation measures in addressing this conduct. This change has been approved by FERC and implemented by the ISO. In addition, adoption of our recommendation to remove the flexibility for units in constrained areas to self-schedule after the RAA process would further mitigate this form of local market power.

Conclusion

This report concludes that the SMD markets have been operated well and performed competitively in 2004. Locational pricing and congestion management under the SMD energy markets are substantial improvements over the prior markets in New England. The report identifies five major changes to the SMD markets that will increase their efficiency and improve the long-term economic signals they provide to govern investment and retirement decisions.

The report also discusses a number of the changes that have been made since the beginning of 2004 to improve the performance of the market, and recommends a number of additional changes to specific market rules and procedures. Although these recommendations are important, the implementation of the LICAP market and operating reserves and regulation markets that are jointly optimized with the energy markets are the most important potential improvements to the current SMD markets and should remain the highest priorities.

II. Prices and Market Outcomes

In this section, we review trends in prices that relate to the performance of the New England wholesale market during 2004. This includes analyses of overall price trends, price convergence between the day-ahead and real-time markets, and transmission congestion.

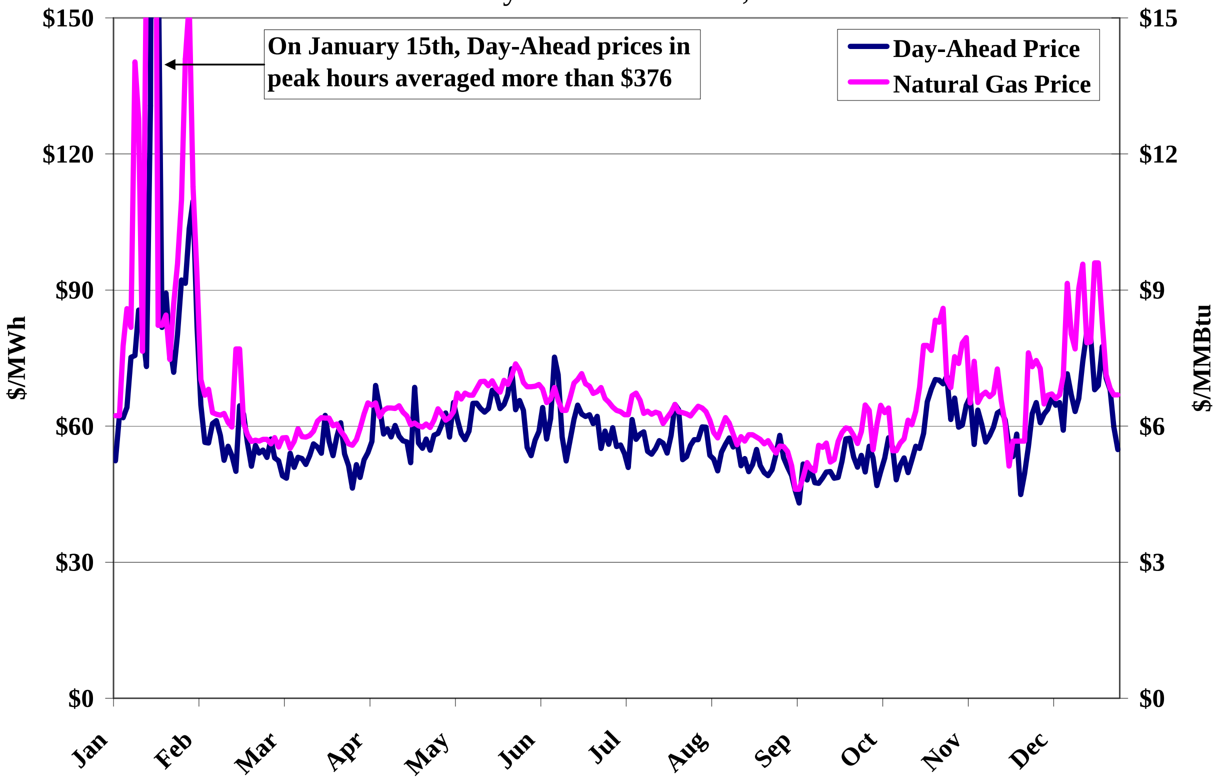
A. Price Trends

Our first analysis examines trends in day-ahead prices at the New England Hub location during 2004. The New England Hub represents an average of prices at 32 individual pricing nodes located at the geographic center of New England. This hub price has been developed and published by the ISO to disseminate price information that will facilitate bilateral contracting.

Figure 1 shows the daily load-weighted average price at the New England Hub in the day-ahead market for each weekday in 2004. Most of the high day-ahead prices were experienced in January, late October, and December when natural gas prices were particularly high—more than $8/MMbtu. Natural gas prices should be a key driver of electricity prices if the market is operating competitively since natural gas-fired generating units set electricity prices in a large share of the hours. Natural gas-fired generation represents 38 percent of all supply in New England, but is the marginal source of supply in a much higher share of the hours because the natural gas units have marginal costs that are higher than base load nuclear, coal, and hydroelectric generating units. These lower cost resources are frequently fully utilized, causing natural gas-fired resources to be dispatched and set the market clearing price. As expected, therefore, electricity prices have been highly correlated with natural gas prices.

This figure also shows that prices did not rise substantially during the summer load conditions. During the summer, the daily average day-ahead price at the Hub never reached $80/MWh, and it was less than $60/MWh on August 30, the day when the annual peak load occurred. A number of factors likely contributed to the moderate summer prices, including the relatively high capacity margins resulting from new investment in the region and relatively mild weather conditions.

Figure 1
Daily Average Day-Ahead Prices at New England Hub
Weekdays 6 AM to 10 PM, 2004



Note: Daily average prices are load-weighted.

B. Price Convergence

In this subsection, we evaluate the convergence between day-ahead and real-time prices in various locations in New England. Price convergence is important because it indicates whether the markets exhibit efficient intertemporal arbitrage. Such arbitrage between the day-ahead and real-time markets is essential for ensuring that the commitment of generation through the day-ahead market is efficient. For example, if prices are consistently lower in the day-ahead market than the market fundamentals would dictate, generation will tend to be under-committed by the day-ahead market.

The day-ahead market allows participants to make forward purchases and sales of power for delivery in the real time. The existing multi-settlement markets in the Northeast (i.e., New York

ISO and PJM) have historically exhibited a small premium in the day-ahead market.[2] This can be explained by the relative risks faced by participants in the day-ahead and real-time markets. Loads can insure against volatility in the real-time market by purchasing power in the day-ahead market. Generators selling in the day-ahead market are exposed to some risk by committing financially day-ahead because an outage after the day-ahead market could compel the generator to purchase replacement power at relatively high prices. If participants are risk-averse, these factors will generate a premium on average in the day-ahead market. However, a predictable day-ahead price premium encourages virtual traders to schedule virtual supply (i.e. to sell short at the day-ahead price and buy back at the real-time price). This response puts downward pressure on day-ahead prices and should limit the size of the average day-ahead premium.[3]

We evaluated weighted-average day-ahead and real-time prices at nine locations, including the New England Hub. Two measures of convergence are calculated, the average price difference and the average of the absolute value of the hourly price differences between the two markets. The average price difference shows whether prices over the entire period were higher in the day-ahead or real-time markets. The second measure shows the size of the hourly differences between the day-ahead and real-time prices. This can be an important aspect of the price convergence between the two markets. For example, if the day-ahead price is $100 higher in half of the hours and $100 lower in the other half of the hours, the average difference would be zero while the average of the absolute value of the hourly difference would be $100 per MWh. One could conclude that the markets exhibit poor convergence in this example, notwithstanding the average difference of zero. These values for these two measures are shown in Table 1.

2 This was not the case in PJM in 2004, which is being investigated by the market monitor in PJM. New York City also has exhibited real-time price premiums that have been linked to modeling differences between the day-ahead and real-time markets.

3 Under some conditions, rational traders can cause the real-time price to be higher than the day-ahead price. See Bessembinder, H., and Lemmon, M, *Equilibrium Pricing and Optimal Hedging in Electricity Forward Markets*. Journal of Finance 57 (June 2002): 1347-82.

Table 1
Average Day-Ahead and Real-Time Price Differences
January to December 2004

	Average Clearing Price			Average of Hourly
	Day-Ahead	**Real-Time**	**Difference**	**Absolute Price Difference**
New England Hub	$53.72	$52.14	$1.58	$8.24
Maine	$48.60	$47.78	$0.82	$7.09
New Hampshire	$52.08	$50.73	$1.36	$7.81
Vermont	$53.94	$52.33	$1.61	$8.25
WC Mass	$53.86	$52.34	$1.52	$8.25
Rhode Island	$52.81	$51.23	$1.59	$8.07
SE Mass	$52.33	$50.73	$1.60	$7.95
NE Mass/Boston	$53.46	$51.47	$2.00	$8.57
Connecticut	$54.62	$52.82	$1.80	$8.88

Based on these results, we conclude that there was relatively efficient convergence between day-ahead and real-time price. In each location, average day-ahead prices were $0.82 to $2.00/MWh higher than average real-time prices. A slight premium in the day-ahead market is consistent with the results of other multi-settlement markets in the Northeast and with expectations based on the discussion of the risk factors above. However, virtual trading activity tends to moderate the size of the day-ahead price premium.

Table 1 also shows the absolute average hourly difference between day-ahead and real-time prices, which ranged from $7.09 to $8.88 per MWh. This is lower than comparable results in the New York ISO and PJM markets as shown later in this section. This supports the conclusion that the market participants in New England have effectively arbitraged the day-ahead and real-time markets. These results also indicate that the day-ahead and real-time market models have been consistent (i.e., consistent transmission limits and other constraints), which has been a significant issue in other markets. However, these conclusions are tempered by the fact that the lack of price spikes and transmission congestion in 2004 has made prices in New England less volatile than historical prices. Tight price convergence is expected under conditions of low price volatility.

The ISO–New England instituted day-ahead and real-time LMP markets in March 2003. Over time, market participants gain experience that should improve their ability to forecast real-time conditions. Improved foresight should lead to better convergence between day-ahead and real-time prices. Table 2 compares day-ahead and real-time price statistics from the summers of 2003 and 2004 to assess whether price convergence has improved since the start of SMD.

Table 2
Average Day-Ahead and Real-Time Price Differences
March to December, 2003 & 2004

	Real-Time Clearing Price		Day-Ahead - Real-Time Price Difference		Hourly Absolute Price Difference	
	2003	2004	2003	2004	2003	2004
New England Hub	$48.55	$50.25	$0.37	$0.96	$7.90	$7.29
Maine	$44.16	$46.28	$0.71	$0.09	$7.06	$6.22
New Hampshire	$47.32	$48.91	$0.50	$0.70	$7.62	$6.89
Vermont	$48.71	$50.49	$0.89	$1.01	$8.18	$7.28
WC Mass	$48.64	$50.49	$0.30	$0.90	$7.78	$7.32
Rhode Island	$47.66	$49.38	$0.40	$0.98	$7.95	$7.14
SE Mass	$47.50	$48.90	$0.18	$1.01	$7.65	$7.02
NE Mass/Boston	$48.05	$49.70	$0.75	$1.43	$7.96	$7.74
Connecticut	$50.17	$51.07	$0.28	$1.02	$9.07	$7.93

For the New England Hub and all eight zones, Table 2 shows that real-time prices were slightly higher in 2004 than in 2003 by approximately $2/MWh. The difference in some locations was smaller, such as in Connecticut where it was only $0.90/MWh higher in 2004. In 2004, the day-ahead premium rose significantly at the Hub, West-Central Mass, Rhode Island, South-East Mass, NEMA/Boston, and Connecticut, while the premium decreased in Maine.[4]

Although a day-ahead price premium can be expected, the size of the premium is likely affected by certain aspects of settlement rules. There are several types of costs incurred by the ISO–New England in the process of operating the system that are "uplifted" to load. Since allocations of uplift to day-ahead scheduled load are smaller than allocations to load that is not scheduled day-

[4] The Table excludes January and February in both years. This will eliminate the highly volatile prices that occurred during the January 2004 "Cold Snap."

ahead, uplift charges are higher for load purchased at the real-time price. Thus, the day-ahead price premium is at least partly the result of charging less uplift to energy purchased at the day-ahead price. These uplift charges are addressed in greater detail in Section IV.E. of this report.

Table 2 shows a decrease in the average hourly absolute price difference in all locations. In particular, the average difference for Connecticut decreased by $1.14/MWh. NEMA/Boston showed the smallest decrease in price differences from 2003 with a decrease in absolute difference of $0.22/MWh. The overall improvement in price convergence likely reflects that market participants have learned more about factors that influence prices.

We further analyzed day-ahead and real-time price convergence by comparing on a more local basis the average price differences and the average absolute hourly price differences in New England to those in New York and PJM. Table 3 shows this comparison for three locations within each market.

Table 3
Day-Ahead and Real-Time Price Convergence
New England Compared to Adjacent Regions, January to December, 2004

	Average Price Difference - Day Ahead minus Real Time	Average Absolute Hourly Price Difference
In New England:		
Maine	$0.82	$7.12
New England Hub	$1.58	$8.27
Connecticut	$1.81	$8.91
In New York:		
Zone A (West)	$0.72	$9.12
Zone G (Hudson Valley)	$2.01	$10.63
Zone J (New York City)	-$1.47	$13.09
In PJM:		
Western Hub	-$0.74	$10.30
New Jersey Hub	-$1.77	$12.84
Delmarva Peninsula	-$0.81	$11.36

The three locations shown in Table 3 were intended to include an export-constrained area, an import-constrained area, and a central market location. According to the first measure of price convergence, New England had a consistent day-ahead price premium in 2004, which is consistent with the experience of New York and PJM prior to 2004. However, in 2004, New York City and PJM had significant real-time price premiums. In New York, the range in price premiums between zones is largely the result of modeling inconsistencies between the day-ahead and real-time markets. In PJM, there were several substantial changes to the market including the integration of the ComEd and AEP control areas. In prior years, the day-ahead prices exceeded the real-time prices and the market monitor for PJM has stated that he is currently reviewing the causes of low day-ahead prices relative to real-time in 2004. Based on the second measure of price convergence, the analysis indicates that New England prices have converged more closely than the prices in either New York or PJM.

Last, we analyzed the differences in real-time prices between several key locations during the study period. In a market with locational marginal pricing, differences in prices occur between locations to reflect variations in transmission costs due to losses and congestion. Losses result whenever power flows across the transmission network. These are larger when power is transferred over long distances and at lower voltages.

Transmission congestion arises in both the day-ahead and real-time markets when transmission capability is not sufficient to allow the lowest-cost resources to be fully dispatched and their output transmitted to the locations where it should be consumed. When congestion arises, the SMD markets establish a spot price for energy at each location on the network that reflects the marginal system cost of meeting load at each location. The marginal system cost can vary substantially over the system, reflecting the fact that higher-cost units must be dispatched in place of lower-cost units to serve incremental load at some locations without overloading any transmission facilities. This will result in higher spot prices at these “constrained locations” than we would see in the absence of congestion.

Table 4 quantifies the relative prices of congestion and losses for Maine (which tends to be a generation pocket), the New England Hub (which is centrally located), and NEMA and Connecticut (which tend to be load pockets).

Table 4
Average Price of Real-Time Transmission Losses and Congestion by Location
January to December, 2004

	Average Clearing Price ($/MWh)				As a Percent of Reference	
	Total	**Losses**	**Congestion**	**Energy**	**Losses**	**Congestion**
New England Hub	$52.18	$0.72	-$0.08	$51.54	1.4%	-0.1%
Maine	$47.83	-$2.88	-$0.83	$51.54	-5.6%	-1.6%
NE Mass/Boston	$51.52	-$0.04	$0.02	$51.54	-0.1%	0.0%
Connecticut	$52.86	$0.83	$0.50	$51.54	1.6%	1.0%

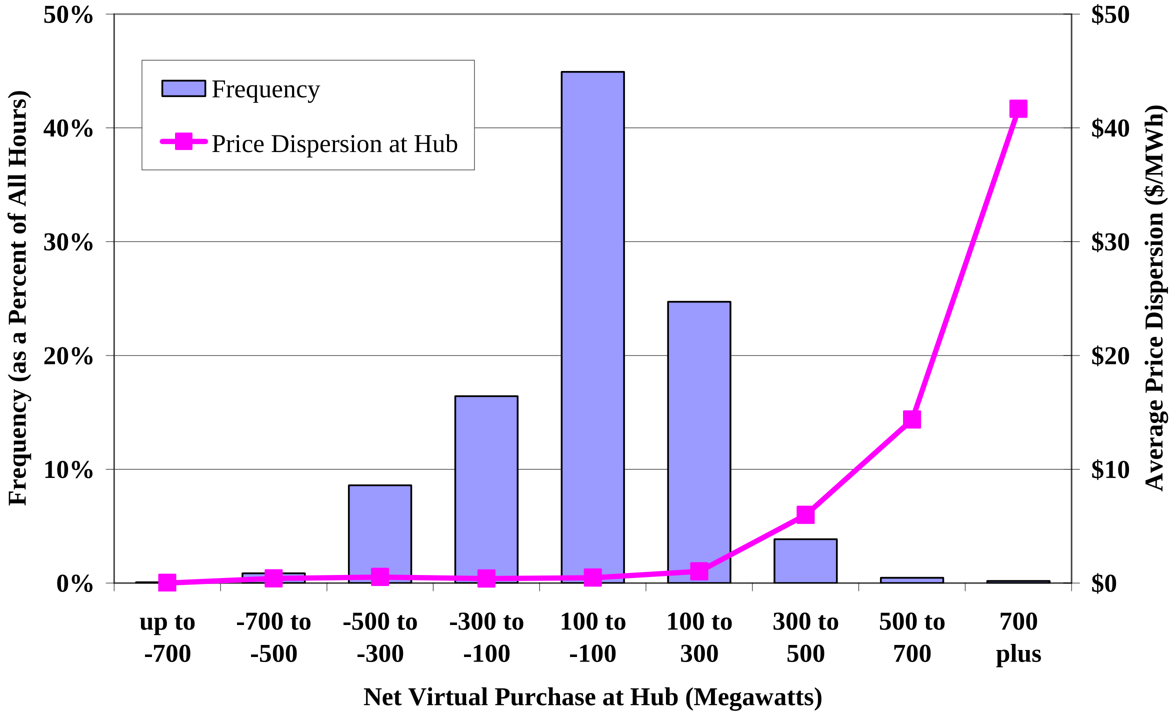
The average congestion and losses prices shown in Table 4 are measured relative to the price at the reference point, which is the weighted average price in New England. Thus, in Maine, prices are lower than the New England average by $2.88/MWh due to losses and $0.83/MWh due to congestion. In Connecticut, where significant import limits were expected, the price associated with losses has been larger in magnitude than the price of congestion. The congestion component of price for NEMA/Boston was only ten cents higher than for the New England hub. This reflects the low levels of congestion that have prevailed since the implementation of SMD relative to what was expected. The next section evaluates congestion patterns in detail. In addition, we note that the average price at the New England Hub is actually higher than for NEMASS/Boston due to the higher marginal losses at that location.

C. Day-Ahead Pricing at the New England Hub

This subsection evaluates congestion experienced in the day-ahead market at the New England Hub. When virtual purchases and sales are made in the day-ahead market at the hub, the injection or withdrawal is evenly distributed over the 32 nodes that comprise the hub. In some periods with relatively large net virtual purchases, the distribution of these virtual transactions can result in significant network flows on the facilities that interconnect the hub nodes. These flows can create congestion in the day-ahead market that would not occur in real time.

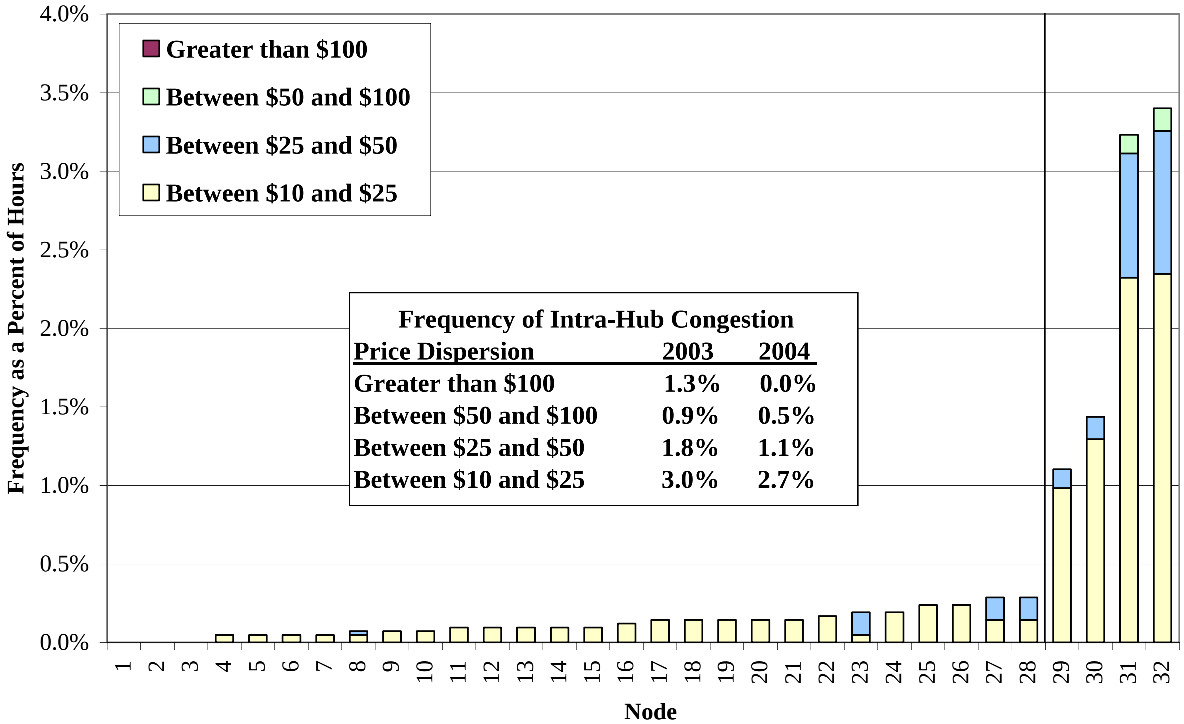
To evaluate the extent to which this has occurred, Figure 2 shows the relationship between virtual scheduling and congestion within nodes that comprise the hub during 2004. The bars in the figure show how frequently various levels of net virtual purchases were made at the New England Hub. For each level of net virtual purchasing, the line shows the difference between the highest and lowest price at the 32 hub nodes (i.e., the price dispersion). The dispersion in prices between the hub nodes is a measure of the intra-hub congestion. Hence, we characterize periods of high average price dispersion within the hub as those exhibiting high intra-hub congestion.

Figure 2
Quantity of Virtual Purchase vs. Price Dispersion at New England Hub
January to December, 2004



The columns in this figure show the frequency with which different levels of net virtual purchases have occurred; the magnitude of the virtual purchases increases from left to right. This figure shows that the highest levels of congestion between nodes at the New England Hub coincide with relatively high volumes of net virtual purchases. However, large net virtual purchases were relatively uncommon during 2004, with purchases exceeding 300 MW occurring in fewer than 5 percent of the hours. While it would be preferable for the hub to not exhibit congestion, these results do not necessarily indicate a sustained problem with the hub. To the extent that the intra-hub congestion is realized at a limited number of nodes, participants should have the ability to submit virtual bids and offers at physical nodes in order to arbitrage the prices. Although virtual bids and offers cannot be placed at hub nodes, the majority of hub nodes have a physical node counterpart. These counterparts may be used by participants to improve convergence of prices at the hub nodes. Figure 3 evaluates whether this is the case by showing the distribution of the significant congestion values across the 32 nodes in the hub.

Figure 3
Congestion between the NE Hub Nodes
Weekdays, 6am to 10pm, 2004



This figure shows how frequently there are substantial differences between the average congestion price at the hub and the congestion prices at each of the hub nodes. This analysis indicates that the most significant congestion within the hub has been focused on a limited number of nodes. For example, the vast majority of the congestion differences that are greater than $10 have occurred at four nodes. All four of the most congested nodes have physical counterparts. The table shown in the figure indicates that congestion between nodes at the New England Hub was substantially reduced in 2004. The figure above suggests that participants should be able to submit a limited quantity of virtual bids and offers at these nodes to arbitrage the prices within the hub as they become familiar with these congestion patterns.

To assess whether market participants have scheduled virtual trades at nodes within the Hub that have reduced congestion, Table 5 summarizes the impact of nodal virtual trading on congestion by showing the total quantity of nodal virtual transactions scheduled at nodes within the New England Hub during 2003 and 2004.

Table 5
Nodal Virtual Trading and Congestion within New England Hub
2003 & 2004

Category	**2003**	**2004**
Virtuals Transactions Scheduled at Individual Nodes within Hub (MWh)		
Total Virtual Transactions Scheduled	38707	20378
Virtual Transactions that Reduced Congestion	666	2203
Virtual Transactions that Increased Congestion	622	568
Avg. Congestion Between NE Hub and Location of the Virtual Transaction		
When the Virtual Transactions Reduced Congestion	$12.81	$27.55
When the Virtual Transactions Increased Congestion	$3.38	$2.82

Table 5 shows that in both years the total virtual transactions scheduled was much larger than the subset that are scheduled in periods when they increase or reduce congestion. This is due to the fact that there is rarely congestion within the New England Hub. However, when congestion did occur within the Hub, the virtual transactions have predominantly helped to reduce congestion. This is particularly true in 2004, which may be an indication that participants have learned to arbitrage the intra-hub congestion more effectively.

The table also shows the average value of the congestion between the New England Hub and the individual nodes where virtual transactions are scheduled during the congested periods. This analysis shows that the value of the congestion is generally substantially larger during periods when virtual transactions reduce it. This is consistent with participants submitting virtual transactions to arbitrage the intra-hub congestion when it becomes significant. The fact that the average value of the congestion is relatively low when the virtual transactions increase congestion limits potential concerns that participants' scheduling of virtual transactions has created intra-hub congestion and thereby distorted the day-ahead prices there.

Based on this analysis, we conclude that market participants have improved in their ability to arbitrage intra-hub congestion in the day-ahead market by submitting virtual bids and offers at the hub nodes. Hence, redefining the hub should not lead to significant improvements.

III. Transmission Congestion and Financial Transmission Rights

Setting efficient energy prices that reflect the economic consequences of all binding transmission constraints is one of the most important functions of the SMD markets. These prices guide the short-term dispatch of the generators to manage the congestion as efficiently as possible and establish long-term economic signals that govern investment in new generation and transmission assets. Hence, evaluating the locational marginal prices and associated congestion costs is a primary component of this report.

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 generated and consumed on the network. The price difference 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 embodied in the congestion component of the LMP at each location.[5]

A participant may hedge congestion charges in the day-ahead market by holding Firm Transmission Rights (FTRs). An FTR entitles a participant to payments corresponding to the congestion-induced difference in prices between two locations in a defined direction. For example, a participant that holds 150 MW of FTRs from point A to zone B is entitled to 150 times the locational energy price at zone 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 its bilateral contract if it owns an FTR between the same receipt and delivery points as the bilateral contract.

Energy purchased and sold in the real-time market includes only the deviations from the day-ahead schedules. Hence, a participant that purchases more energy in the day-ahead market than

[5] 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, assuming no transmission losses.

it consumes in real time will sell the excess energy into the real-time market. Similarly, a participant that sells more energy in the day-ahead market than it produces in real time will be a purchaser in the real-time market.

Likewise, settlement of congestion costs in the real-time market is related only to deviations from the day-ahead schedules. Participants with day-ahead schedules do not pay real-time congestion charges related to their scheduled quantities. Because the real-time spot market is a balancing mechanism for day-ahead contracts, net congestion charges should be zero in real time as long as the transmission limits and external loop flows have not changed from those assumed in the day-ahead market. In other words, any congestion charge to a real-time purchase would be offset by a payment to a real-time sale. This would not be the case if the transmission limits or other modeling inputs in the day-ahead market were different than the inputs to the real-time market. Inconsistencies in limits or other modeling inputs can compel the ISO to incur substantial costs to reduce the flow on constrained facilities in real time, which would be recovered through uplift charges. This has not been a problem under the ISO's operation of the SMD markets.

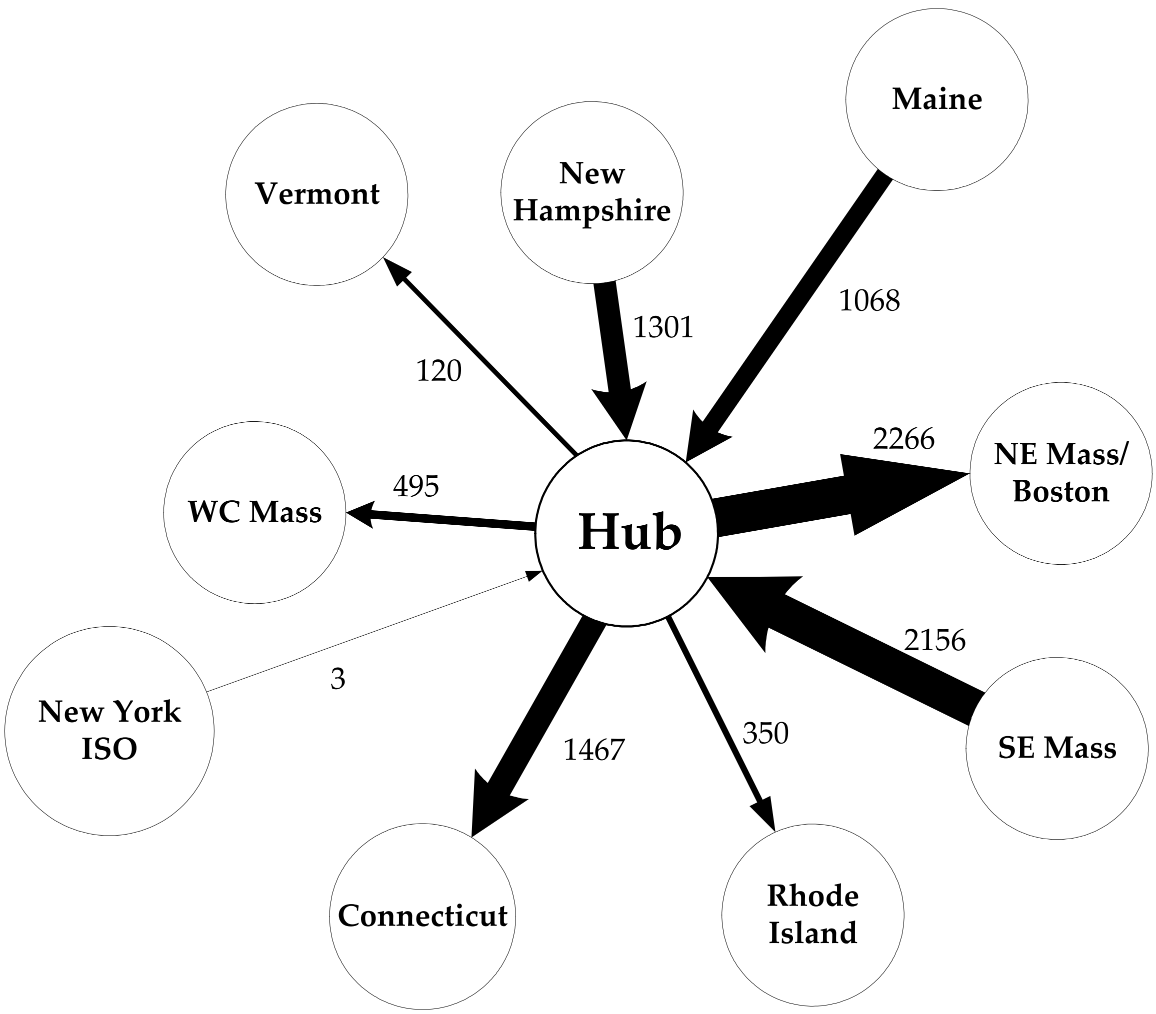
A. FTR Purchases

To begin our evaluation of congestion management in New England, we first assess the pattern of FTR purchases. As discussed above, an FTR is purchased between a designated source and sink. An FTR entitles the purchaser to receive the difference in the prices at the FTR's source and sink points, excluding losses, times the FTR quantity. 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. In well-functioning markets, the FTR prices should be highly correlated with the actual congestion on the system. In addition, the pattern of FTR purchases should correspond to the attendant power flows associated with the location of loads and generation.

In 2004, the ISO auctioned FTRs with one-month and six-month terms. The longer-term FTRs allow market participants greater certainty by locking-in congestion hedges further in advance.

Currently, the ISO releases 50 percent of the transmission capability of the system in the long term auction, while any remaining capability is made available in the one-month auction. Our first analysis in this subsection calculates the net purchases of inter-zonal FTRs for each of the eight New England Zones and is shown in Figure 4. The net purchases from the six-month auction are combined with the quantities from the one-month auction in the figure.

Figure 4
Net FTR Purchases between New England Zones



To simplify Figure 4, we show all of the FTR purchases relative to the New England Hub rather than showing the actual sources and sinks. Since FTRs have the properties of geometric vectors, an FTR between any two zones is equivalent to an FTR from the first zone to the hub plus an FTR from the Hub to the second zone. If a zone was a net source for FTRs (more FTRs exit the zone than enter the zone), then the arrow in Figure 4 is directed from the zone to the New England Hub (e.g., Maine). If the zone is a net sink, then the arrow points from the New England Hub to the zone (e.g., Connecticut).

The patterns shown in the figure are generally consistent with expectations. Maine, South East Massachusetts, and New Hampshire zones have been net sources for FTRs, consistent with the fact that these zones tend to exhibit net exports of power. NEMA and Connecticut, and to a lesser extent Vermont, have been net sinks for the FTRs. This is also generally consistent with historical power flows into these areas. Rhode Island and West-Central Massachusetts were slight net importers of FTRs in 2004, which is notable because these zones were initially net exporters after the implementation of SMD.

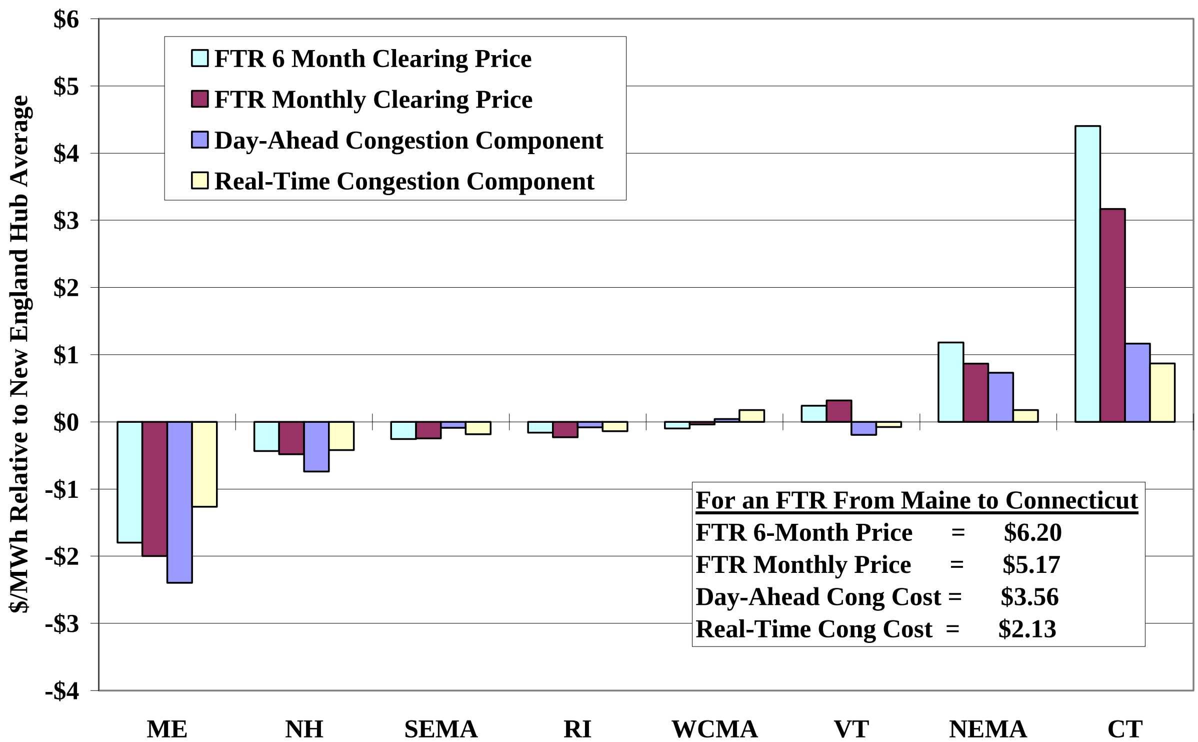
Another notable aspect of the pattern of FTR purchases is that the net quantity of FTRs into Connecticut was only 1467 MW, while the total import capability typically ranges from 2000 MW to 2500 MW. Although the quantity of FTRs sold was less than the import capability *into* Connecticut, other constraints within Connecticut are binding, such as those into Southwest Connecticut, which limit the quantity of FTR purchases. The following sections evaluate the levels of congestion and FTR prices into each zone in New England.

B. Congestion Patterns and FTR Prices

To evaluate the congestion experienced under SMD, we analyzed day-ahead and real-time congestion costs relative to revenues earned by holders of FTRs in the various zones. In a well-functioning system, these values should be highly correlated. We made this comparison for the ISO's eight zones and the New England Hub.

Figure 5 shows day-ahead and real-time congestion costs compared to FTR prices during 2004 for each of the eight New England zones. The congestion costs shown are the average for on-peak hours and are calculated relative to the New England hub. Hence, if the congestion component in the figure indicates $4 per MWh, this is interpreted to mean the congestion cost to transfer a MW of power to the New England Hub is $4. The congestion cost between any two points shown in the figure is the congestion price at the sink location less the congestion price at the source location. The analysis is limited to the on-peak hours since the load and the power flows on the system are greatest in these hours.

Figure 5
FTR Auction Prices vs. Day-Ahead and Real-Time Congestion
Zonal Averages Shown Relative to New England Hub Price
2004 -- Weekdays 6 AM to 10 PM



The 6-month FTR auction clearing price is the average purchase price from the two semi-annual auctions, reported in dollars per MWh by location. Likewise, the monthly FTR auction clearing price is the purchase price for the 12 monthly FTR auctions in dollars per MWh. Like the congestion costs, the purchase price for an FTR between two locations is the difference between the prices at the destination and origin points for the FTR. For example, a $1.00/MWh FTR price for Maine and $0.50/MWh FTR price for Connecticut would indicate a total price for an FTR from Maine to Connecticut of $1.50/MWh.

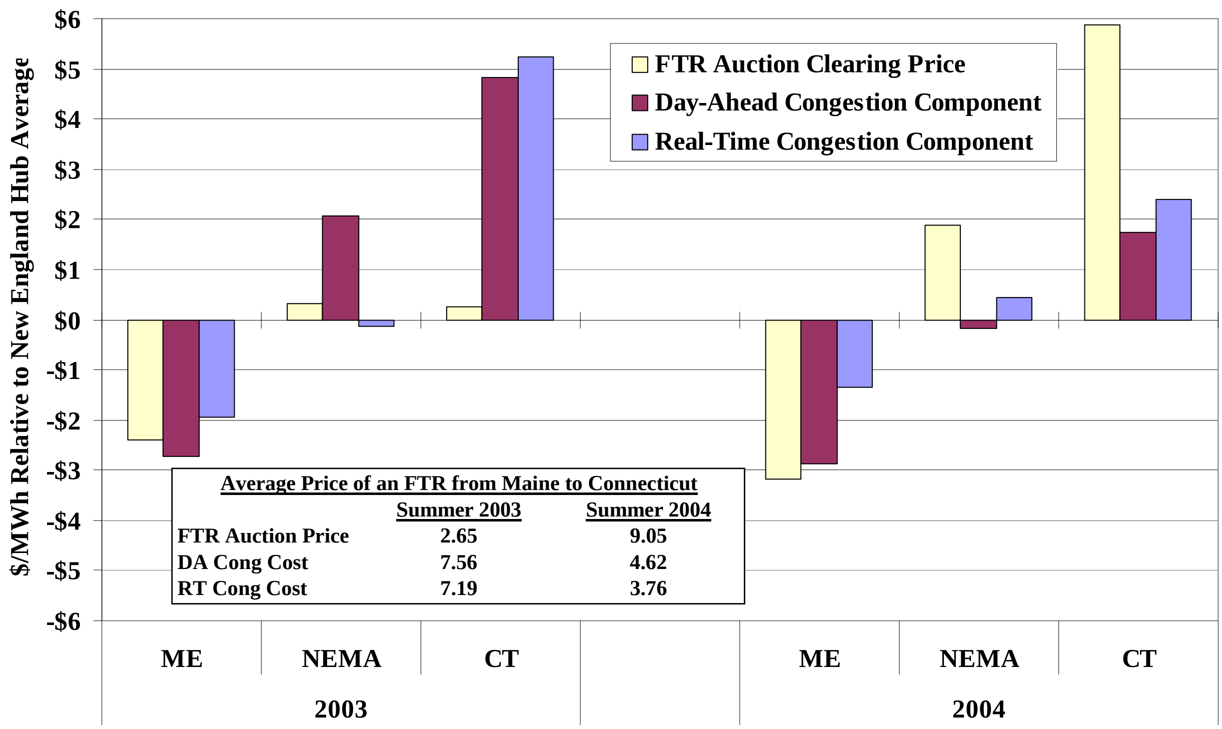
The zones listed along the horizontal axis are generally ordered in accordance with their historical congestion patterns relative to the hub. Hence, the zones listed toward the left tend to face congestion as they export power to zones toward the right. This should result in negative congestion and negative FTR values for zones on the left of the horizontal axis and positive values for zones on the right. We generally expect that congestion costs would be correlated with FTR revenues.

During 2004, congestion costs and FTR payments were generally well-correlated, although the magnitudes of FTR prices and congestion costs differed. As Figure 5 shows, the monthly FTR prices were lower than the semi-annual FTR prices (more negative in the exporting areas), while day-ahead and real-time congestion was generally smaller in magnitude than monthly FTR prices. The exceptions were the two most significant export-constrained areas, Maine and New Hampshire, where the day-ahead congestion was larger in magnitude (more negative) than the monthly FTR prices. The fact that the semi-annual FTR prices were consistently higher than the monthly FTR prices in the import-constrained areas suggests that market participants forecasted higher levels of congestion, but revised their expectations downward as it became apparent that less congestion was occurring. One reason that congestion may have been less than expected is the mild load conditions that prevailed during 2004, which generally results in lower congestion.

For Maine, the six-month and one-month FTR prices as well as day-ahead and real-time congestion prices correspond well. FTR prices and congestion prices into New Hampshire have been negative and relatively low, consistent with its historical experience. Rhode Island, West-Central Massachusetts, South-East Massachusetts, and Vermont experienced virtually no congestion which was consistent with the FTR auction prices. Only NEMA/Boston and Connecticut were significantly more expensive than New England Hub due to congestion. Figure 5 indicates that the average cost of an FTR from Maine to Connecticut was $6.20/MWh in the six-month auctions, nearly 75 percent higher than the average cost of congestion in the day-ahead market. The NEMA zone experienced low levels of congestion, contrary to historical experience. These low levels of congestion can be attributed to transmission upgrades, the installation of a considerable quantity of new generation, mild load conditions during 2004, and supplemental commitment and out-of-merit dispatch, particularly to address voltage requirements.

The greatest stress on the transmission system typically occurs during the summer when cooling demand is at its peak. The increased stress generally results in higher congestion costs and greater financial risks for market participants, making FTRs most valuable during the summer. Figure 6 shows FTR prices and congestion costs for the three most congested zones during the summers of 2003 and 2004.

Figure 6
FTR Auction Prices vs. Day-Ahead and Real-Time Congestion
Maine, NEMA/Boston, and Connecticut Relative to New England Hub Average Price
June to August, 2003 & 2004 – Weekdays 6 AM to 10 PM



Average Price of an FTR from Maine to Connecticut

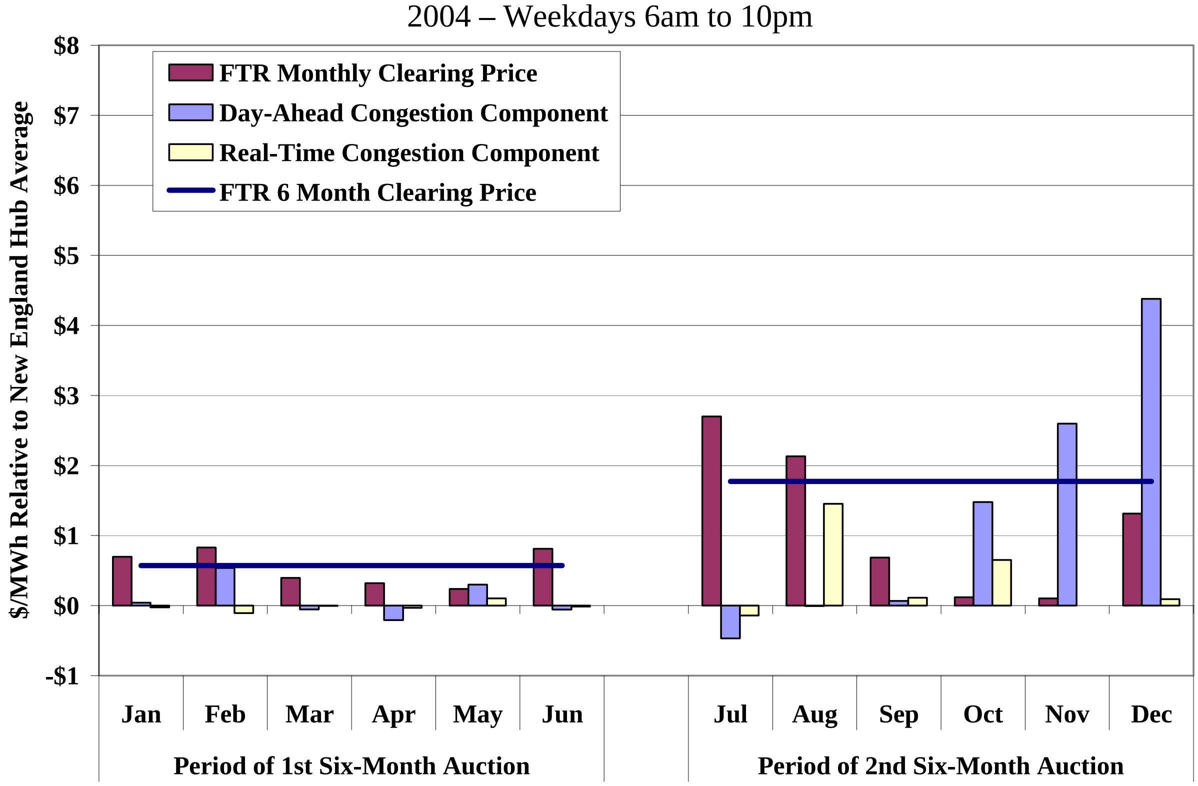
	Summer 2003	Summer 2004
FTR Auction Price	2.65	9.05
DA Cong Cost	7.56	4.62
RT Cong Cost	7.19	3.76

Figure 6 averages prices for the months of June, July, and August. Therefore, the FTR prices include only results from the monthly auctions. The figure shows that the average prices of FTRs into NEMA and Connecticut were far less than actual day-ahead congestion costs in the summer of 2003, while FTR prices were significantly higher than congestion costs in 2004. It is likely that the lower day-ahead and real-time congestion during the summer in 2004 is largely attributable to the milder weather and lower peak loads in 2004.

Regarding the higher FTR prices, increased liquidity in the FTR markets in 2004 significantly improved how well the FTR prices reflected reasonable expectations of congestion. It is notable that the prices of FTRs in 2004 were very close to the actual day-ahead congestion experienced in 2003. This suggests that market participants may have based their estimates of congestion 2004, in part, on market outcomes from 2003.

NEMA and Connecticut warrant closer examination as they are areas that have been subject to significant constraints historically and represent a substantial share of the total load in New England. Figure 7 shows monthly comparisons of FTR prices with congestion costs for NEMA. We examine the data for Connecticut below.

Figure 7
FTR Auction Prices vs. Day-Ahead and Real-Time Congestion
NE Mass Averages Shown Relative to New England Hub Average Price
2004 – Weekdays 6am to 10pm



For the first six months in 2004, Figure 7 shows that the average six-month FTR auction price was comparable to the one-month FTR prices. However, actual congestion costs fell below the monthly FTR prices in every month except May when day-ahead congestion costs were comparable. In the second six months, the long term FTR price was higher than the average of the monthly prices. The monthly FTR prices between July and December exhibit a pattern typical of historical congestion levels, with high levels in the hot and cold months and low levels in the fall. The pattern of actual congestion differed significantly from expectations in the last six months. Day-ahead congestion costs grew progressively from July to December and were not correlated with Real-time congestion costs. To evaluate this pattern, we show the average difference on a daily basis between the day-ahead limit and the real-time limit into NEMASS/Boston.

Figure 8
Difference between Real-Time and Day-Ahead Limit into NEMASS/Boston
Daily Average -- 2004

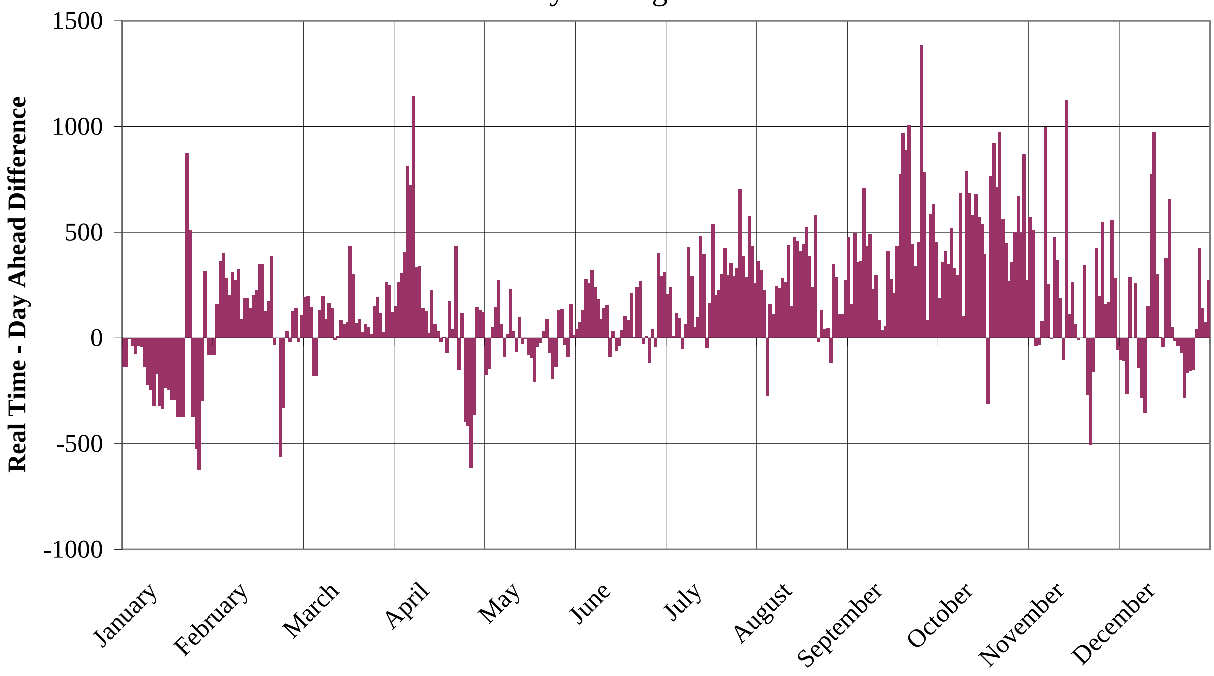


Figure 8 shows that the day-ahead limit has been consistently lower than the real-time limit. These limits are calculated to reflect the second-contingency requirements in the area (known as the proxy second-contingency limits). The methodology is the same for calculating the day-ahead and real-time limits. However, the real-time limit is based on actual operating conditions, while the day-ahead limit is based on a forecast of the next day's operating conditions. Many of the factors used to calculate the limit can change between the day-ahead forecast and real-time including:

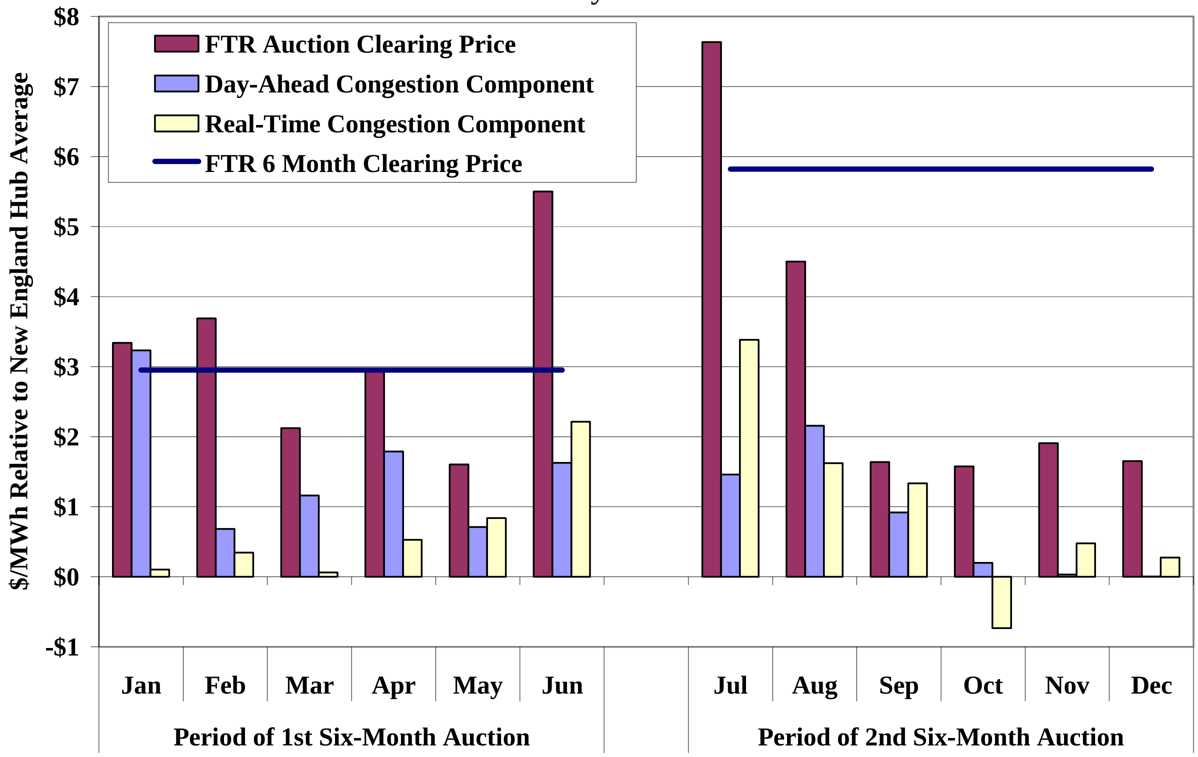
- the thermal limits of individual elements that make up the interface and reactive power flows;
- outages of key transmission lines and generators;
- the size of the largest generator contingency;
- the quantity of 30-minute reserves available on on-line and off-line quick start resources in the load pocket; and
- the amount of load that can be shed in the event of a contingency.

Naturally, there will always be differences between forecasted and actual conditions, which will lead to some differences between day-ahead and real-time limits. However, these differences should be random and result in a difference in the limits between the day-ahead and real-time market that is close to zero on average. Since the real-time limits have been higher than the day-ahead limits on a sustained basis in 2004, we recommend that the ISO investigate the factors that have led to these systematic differences to determine whether improvements could be made to bring the day-ahead and real-time limits into better alignment.

In addition, the convergence issues between the day-ahead and real-time prices in NEMA were not self-correcting due to the large operating reserve charges associated with the local reliability commitments that were allocated to real-time deviations in 2004. These charges resulted in large and volatile costs for any market participants engaging in virtual transactions to arbitrage these differences. However, in early 2005 the ISO addressed this issue by modifying its allocation of the local reliability-related operating reserve costs to all physical load in the constrained area. This eliminates the bulk of the charges to virtual transactions and allows them to act on their incentive to arbitrage the day-ahead and real-time prices in the constrained areas.

Figure 9 shows monthly comparisons of FTR prices with congestion costs for Connecticut during peak hours in 2004.

Figure 9
FTR Auction Prices vs. Day-Ahead and Real-Time Congestion
Connecticut Averages Shown Relative to New England Hub Average
2004 – Weekdays 6 AM to 10 PM



As with NEMA/Boston, the average six-month FTR auction price was comparable to the one-month FTR prices during the first six months in 2004. Although, actual congestion costs were lower than monthly FTR prices in every month. In the second six month block, the six-month FTR was priced higher than the average of the monthly prices. The monthly FTR prices were relatively high during the summer months and then decreased considerably in fall and winter. Congestion costs exhibit the same general pattern as FTR prices, but the magnitudes of day-ahead and real-time congestion were considerably lower. Like the results in other areas, the fact that the actual congestion was less than the FTR prices into Connecticut was likely due, in part, to the relatively mild load levels in 2004.

Overall, we find that the FTR markets have more closely corresponded to reasonable expectations of congestion in 2004 than in 2003. We believe this is due to experience gained by market participants under the SMD markets and increased liquidity of the FTR markets. In addition, we have reviewed the FTR market processes and do not find any structural or methodological impediments to efficient FTR pricing. To better understand the congestion patterns in New England, we analyze generator commitment and dispatch patterns in the next section. In particular, we examine commitment and dispatch that occur outside of the market processes.

IV. Market Operations

In this section, we evaluate a number of issues related to the operation of the SMD markets. These issues include: the accuracy of the ISO's load forecasts, frequency of price corrections, supplemental commitment of generating resources, and out-of-merit dispatch. These issues are important because they can substantially affect the efficiency of the New England market's price signals.

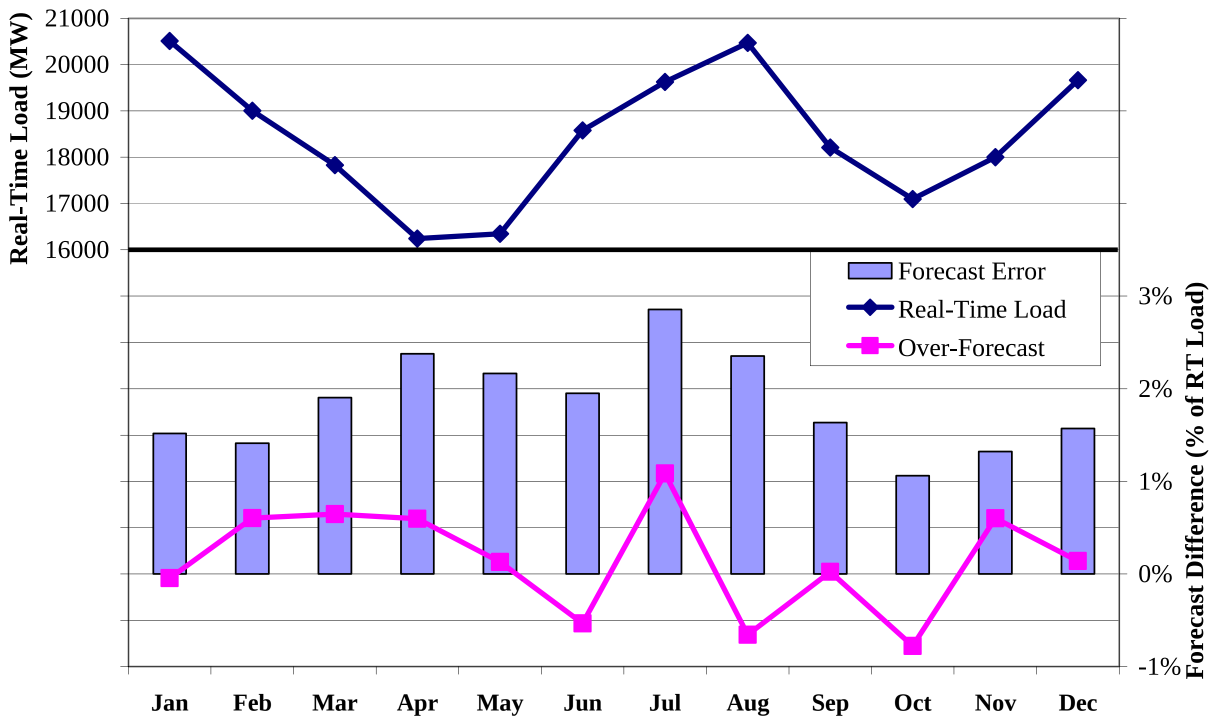
A. Accuracy of ISO Load Forecasts

The accuracy of ISO load forecasts is important for efficient market operations. Inaccurate load forecasts can cause the ISO to commit too much or too little capacity. As we explain in more detail below, excessive supplemental commitments can distort real-time prices and increase uplift costs. Therefore, it is desirable that day-ahead forecast accurately predict actual loads.

Figure 10 shows daily peak load on weekdays as well as two measures of forecast error averaged on a monthly basis during 2004. The figure shows a characteristic pattern of high loads during the winter and summer and mild load during the spring and fall. While the annual peak load of over 24 GW occurred during August, the figure shows that the average daily peak was actually highest in January. Forecasted demand tracked actual load very closely. The average difference between the forecast load and actual load in 2004 was only 0.1 percent, with the forecast being slightly higher on average. On a monthly basis, this average over-forecast was generally close to zero, but ranged as high as 1.1 percent in July and as low as -0.8 percent in October. The lack of consistency in this monthly statistic is an indication that the forecast has not been systematically biased for any sustained period.

To measure the average forecast error associated with the daily peak demand, we also calculated the average of the absolute value of the difference between the forecasted peak demand and the 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 difference between the forecast load and actual load would be zero. Our analysis shows that the forecast error as a percent of the actual peak demand averaged 1.8 percent. On a monthly basis, the average forecast error for the daily peak ranged from as high as 2.8 percent in July to 1.1 percent in October. Generally, the forecast error was highest during the summer when load fluctuates the most and uncertainty associated with the weather is the highest. The average forecast error for the other operating markets in the same timeframe ranged from 2.1 percent in New York to 3.5 percent in PJM.

Figure 10
Average Daily Peak Forecast Load and Actual Load
Weekdays



Note: Over-forecast is the percentage by which the day-ahead forecast exceeded real-time load. A negative percentage value indicates an under-forecast.

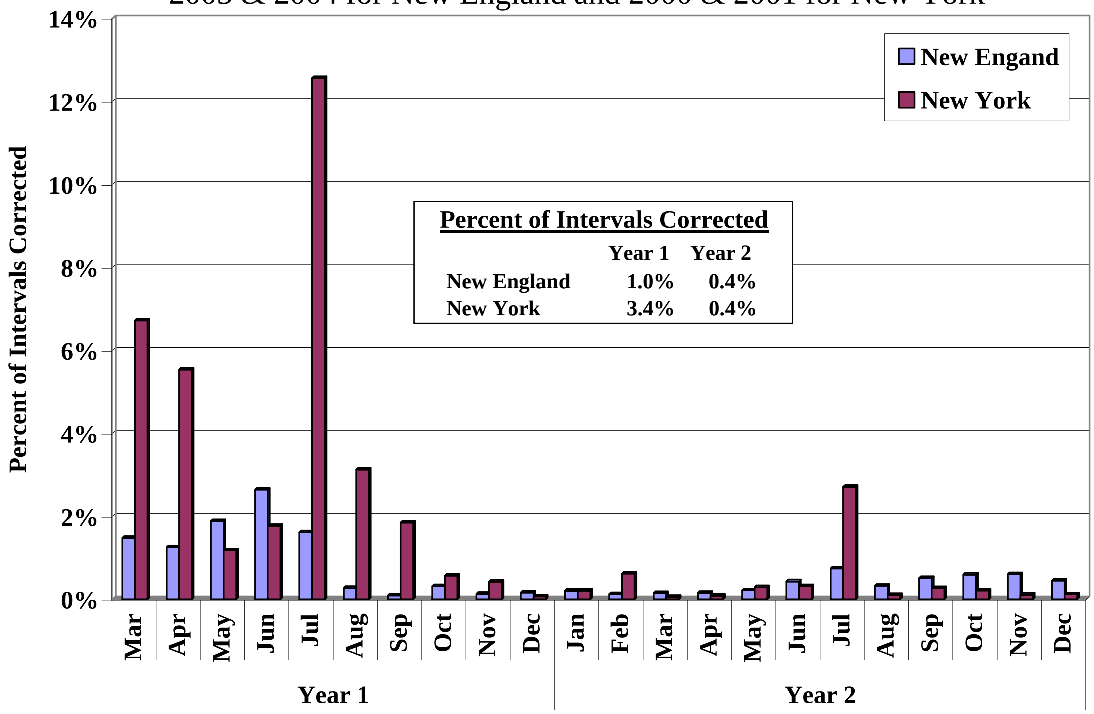
Because the forecast error levels are small in magnitude and less than values for neighboring markets in the region, we find that the ISO's load forecasting has been reasonably accurate. This is important because it provides a foundation for efficient commitment of resources in New England.

B. Price Corrections

This subsection evaluates the rate of price corrections that have occurred during 2004. 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 these corrections since substantial and frequent corrections can harm the integrity of the market.

Price corrections tend to be more frequent during the transition to new markets or the implementation of significant software changes. Therefore, we expect that price corrections should be less frequent in 2004 than initially after the start of SMD. Figure 11 below shows the percent of prices that were corrected in New England during the first 22 months of SMD compared to the same months in 2000 and 2001 in New York, which were the NYISO's first two years of operation.[6]

Figure 11
Price Corrections in First Two Years of Market Operations
2003 & 2004 for New England and 2000 & 2001 for New York



6 The NYISO markets were implemented in November 1999.

The figure shows that New England required significantly less price correction in its first year of operating its nodal energy markets. In 2004, New England improved its frequency of price corrections considerably over 2003. The New York ISO made a number of improvements to its market rules and the software in its second year that reduced the corrections to levels comparable to New England in 2004. Similar data were unavailable for the PJM market. These results support the conclusion that ISO did a very good job in developing and implementing the SMD markets.

C. Commitment for Local Congestion and Reliability

In New England, there are several load pockets that import a significant portion of their total consumption. In order to ensure that these areas can be served reliably, additional capacity within the load pocket must be committed. Specifically, additional capacity may be required to:

- Make certain sufficient capacity is on-line to resolve local first contingency limits (i.e. ordinary transmission interface limits).
- Ensure that reserves are sufficient in local constrained areas to respond to a second contingency;
- Support the voltage of the transmission system in specific locations; and
- Manage low-voltage constraints on the distribution system that are not modeled in the day-ahead market software (known as Special Constraint Resources ("SCRs")).

The New England market commits resources in the day-ahead market based on multi-part offers. Offers include a cost to start a unit that is offline, a "no-load" cost reflecting the fixed hourly cost of keeping a unit online, and an energy offer curve reflecting the offer price for the unit's incremental output. In order for a unit to be committed in the day-ahead market, demand bids from load serving entities and virtual traders must express a willingness to pay enough for the energy from the unit that it is economic to incur the start-up, no-load, and incremental offer of the unit. However, demand bidders should not be willing to pay substantially higher prices day-ahead than they anticipate in the real-time market the following day. Thus, day-ahead market-based commitment is always limited by expectations of real-time prices.

As stated above, there is a requirement to ensure sufficient capacity is on-line in load pockets to manage voltage and provide reserves. However, New England does not currently operate spot markets for reserves or voltage support. Thus, the real-time market will not adequately reflect the value of on-line capacity in locational prices for energy, reserves, or voltage support. As in any forward financial market, the day-ahead market prices will tend to converge with the real-time prices. Hence, the day-ahead prices will also not reflect the value of additional capacity in local area prices.

Market-based commitment will generally not be sufficient to meet local reliability requirements. However, the ISO has attempted to increase the extent to which the market-based commitments will satisfy the local reliability requirements by modeling a lower transfer limit into the constrained area to reflect the 2nd contingency reliability requirements. This lower limit is referred to as the "proxy 2nd contingency limit". Nonetheless, supplemental commitments are frequently needed to meet local requirements. Supplemental commitments may occur in either the day-ahead market process or later in the Reliability Adequacy Assessment ("RAA") process. There are two ways in which supplemental commitments are made:

- The commitment software recognizes a need for capacity (but not energy) in a local area and commits the resources with the lowest commitment costs that satisfy the need.
- The operator recognizes a constraint not modeled in the software, especially voltage support, and manually commits resources to manage the constraint. This may not be the lowest-cost way to manage the constraint.

Although it is preferable for the commitment software to make supplemental commitments rather than to do so manually, neither method adequately reflects the cost of maintaining reliability in clearing prices. Furthermore, since these units must be dispatched at or above their economic minimum generation level ("EcoMin"), these commitments will tend to reduce energy prices by displacing energy that would have been produced by other units committed through the market.

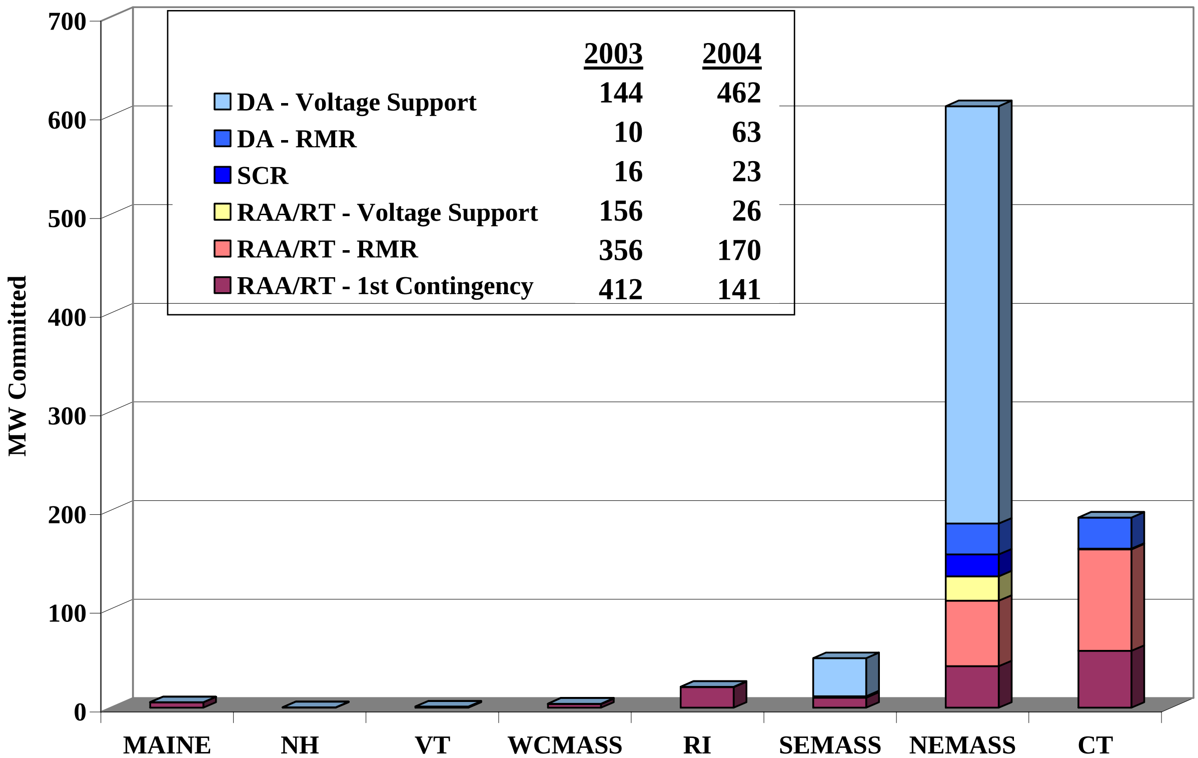
Thus, supplemental commitment tends to mute locational price signals associated with resolving transmission congestion. Because congestion price signals can be muted in the real-time, day-ahead demand bidders may not bid enough to commit the minimum set of resources necessary to manage normal transmission constraints. These are called "1st Contingency" commitments, because normal transmission flows must be maintained at levels that would allow the network to withstand the single largest contingency. Hence, supplemental commitment may be required simply to manage normal transmission congestion. This section gives a detailed summary of supplemental commitment during 2004 for voltage support, 2nd contingency requirements, and 1st contingency transmission limits.

1. Generation Committed for Local Needs

Because committing units to provide reserves and/or voltage support in a local area increases the on-line supply in the congested area, it can diminish locational price signals in the areas. Hence, it is important to monitor the extent to which these actions occur and the locations where they occur. Therefore, we calculate the average quantity of commitments made to satisfy local requirements during the daily peak in each zone in New England. These commitments during 2004 are shown in Figure 12. The commitments that are made to withstand the 2nd largest contingency in a local area are referred to as reliability must run or "RMR" commitments. These are distinct from the reliability agreements discussed earlier in this report. Reliability agreements provide capacity payments to units that must remain in operation over a specified timeframe (e.g., the next year) to maintain reliability. RMR commitments are made as part of the operation of the system on a given day to ensure that sufficient resources are online in an area to withstand the first and second largest contingencies. RMR commitments can occur for units that do or do not have a reliability agreement with the ISO.

For the purposes of this analysis, the MW commitment level is the entire capacity of the committed unit, regardless of the energy it produced in real time. However, the commitments' effect on prices depends on the energy produced from these units, particularly the energy produced out of merit order.

Figure 12
Commitment for Local Reliability by Zone
2004 – Daily Peak Hour



Note: Capacity committed day-ahead for RMR or voltage support that would have been economically committed in the day-ahead market is excluded. The 2003 averages exclude January and February.

Figure 12 shows an average of 885 MW committed for local reliability during 2004. Nearly half of this capacity was committed in the day-ahead market to maintain voltage in the NEMA area. Significant quantities were also committed to satisfy RMR and first contingency requirements in NEMA and Connecticut, primarily during the RAA process. Local reliability commitments were down 19 percent in 2004 relative to the previous year due to substantial reductions in commitments for RMR (generally for 2nd contingencies) and 1st contingencies. NEMA/Boston experienced a large increase in day-ahead commitments for voltage support. The decrease in RMR and 1st contingency commitments in NEMA/Boston is expected since capacity brought on-line for voltage support also provides local reserves and relief for local congestion. This shift in supplemental commitments from the real-time market to the day-ahead market is consistent with a recommendation we had made in our prior report on the operation of the SMD markets in New

England.[7] In this report, we had recommended that the ISO pre-commit units they know are needed prior to the day-ahead market.

The day-ahead commitment software is designed to commit the set of resources that minimizes overall production costs. Since some local constraints are not represented in the day-ahead market, the operators must make supplemental commitments. To the extent that the commitment of a particular unit to satisfy a local requirement is known, it is most efficient to commit the unit before the day-ahead market software runs. This allows the software to determine the lowest-cost solution, taking into account the manual commitment. When an additional resource is committed supplementally, it may no longer be efficient to commit one or more units that were committed in the day-ahead market. This tends to make some units committed economically through the day-ahead market run out of economic merit in real-time. Therefore, it is most efficient for the day-ahead software to determine the lowest-cost set of offers taking into account units that must run for local reliability to the extent that they are known when the day-ahead market runs.

There was a large shift of supplemental commitments from the RAA process to the day-ahead market in 2004, because the ISO began pre-committing units needed for voltage support in the day-ahead market. Approximately 59 percent of the supplemental commitment occurred during the day-ahead market in 2004, compared with just 14 percent in 2003. The additional commitments for voltage support in the day-ahead market reduced the need for additional supplemental commitments after the day-ahead market for other local reliability needs. Ultimately, this results in a more efficient commitment by reducing the potential for excess generation commitments that are associated with the supplemental commitments that occur after the day-ahead market.

Because the market effects of local reliability commitments are likely to be the greatest in the highest-load hours, Figure 13 shows the average quantities in the peak hours on the five highest-load days during the study period. This figure shows that the supplemental commitments made

[7] *Six-Month Review of SMD Electricity Markets in New England*, Potomac Economics, February 2004.

to meet local RMR and first contingency requirements were substantially higher while commitments for voltage were considerably lower than the average for all days.

Figure 13
Average Supplemental Commitment by Zone
Top 5 Peak Load Days

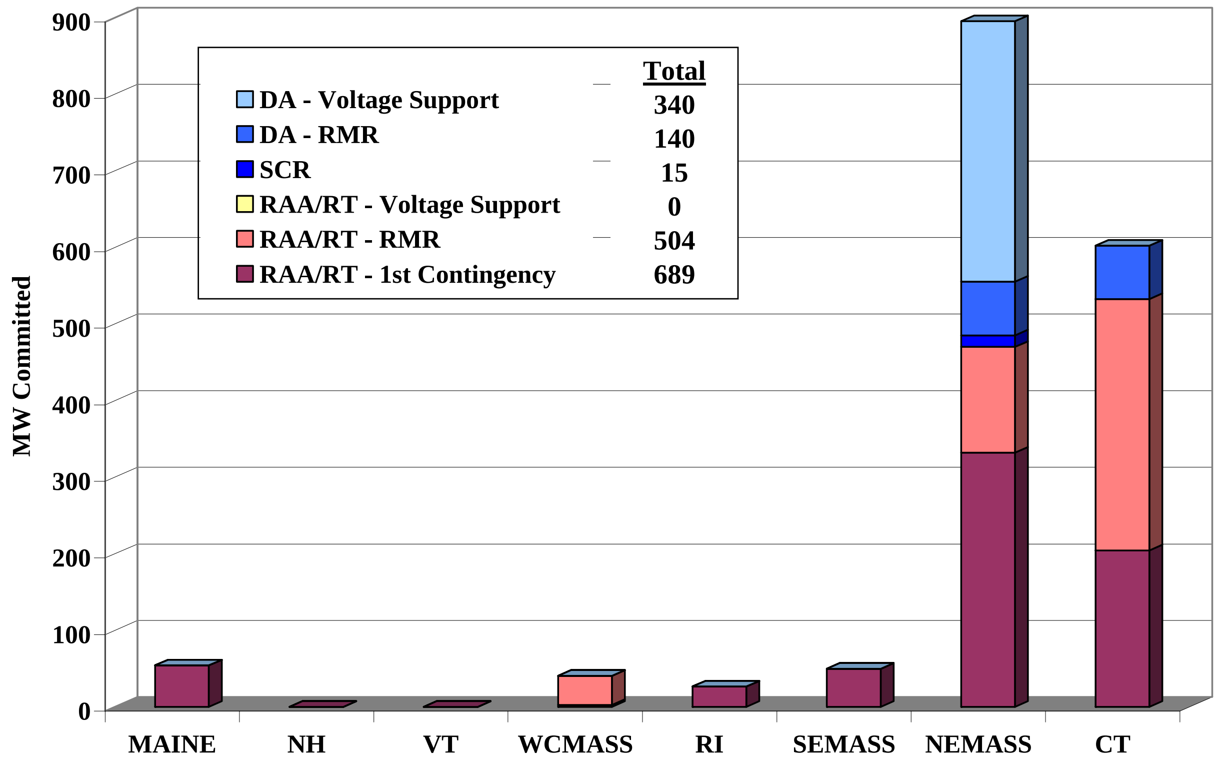


Figure 13 shows that commitments for local reliability averaged 1688 MW on these five days. Approximately 80 percent of this capacity was committed to satisfy RMR requirements (i.e., local reserves needed to respond to 2nd contingencies) and local congestion management (i.e. 1st contingencies). While 603 MW was committed for local reliability in Connecticut, nearly 900 MW was committed in NEMA/Boston. The latter figure is particularly large given that it constitutes 25 percent of the installed summer capability in NEMA/Boston.

While 59 percent of all supplemental commitments for local reliability in 2004 were made in the day-ahead market, only 28 percent of those shown in Figure 13 occurred day-ahead. As discussed above, committing additional resources in the RAA process causes some units committed day-ahead to run out of economic merit in real-time. Real-time prices are depressed when substantial amounts of energy are produced out of merit.

Given the predominance of the supplemental commitments in Connecticut and NEMA during 2004, we analyzed the patterns in these two zones. Figure 14 shows supplemental commitments for local needs in NEMA/Boston on a monthly basis during 2004.

Figure 14
Commitment for Local Reliability in NEMA/Boston
2004 – Daily Peak Hours

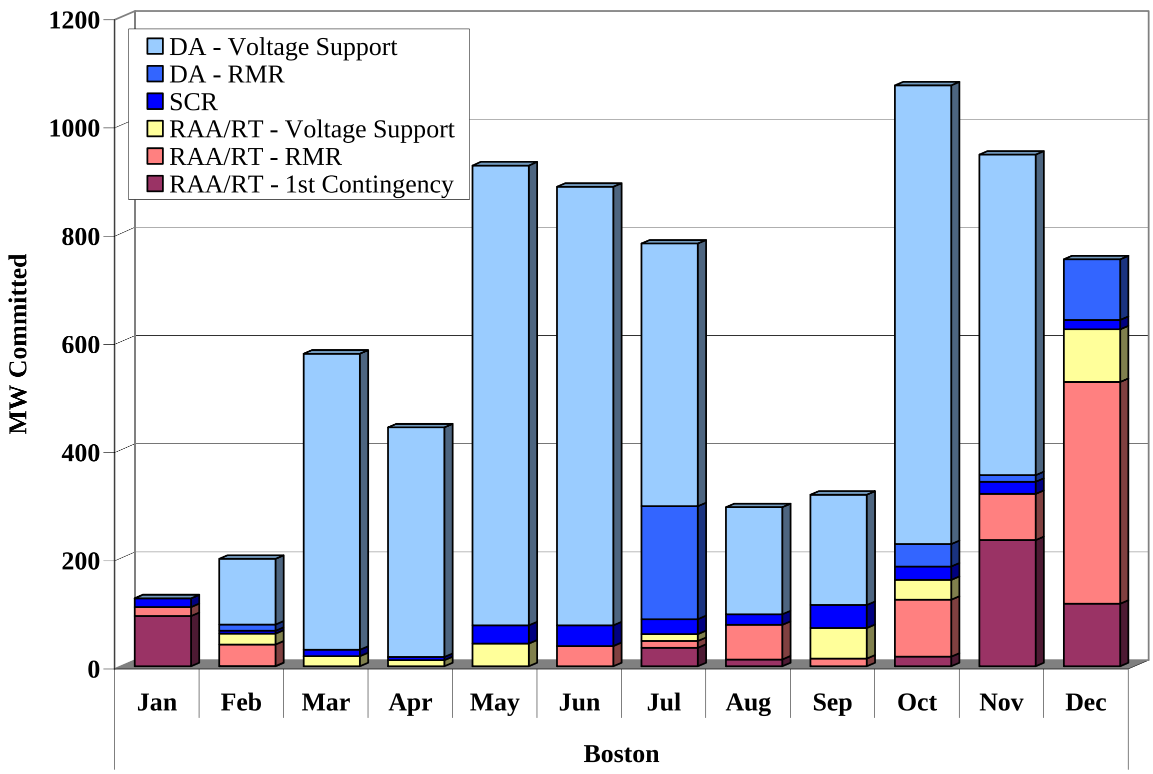


Figure 14 shows that a significant amount of supplemental commitment in the NEMA/Boston zone was flagged for voltage support, particularly day-ahead. Day-ahead commitments for voltage support were high from March through July and again during October and November. Voltage support commitments were relatively low during the highest load months of January, August, and December. Generally, less commitment is necessary for voltage during high load periods when substantial quantities of capacity are already on-line to serve demand.

Day-ahead commitments for voltage declined sharply in December for daily peak hours, although some were still made during off-peak hours. The reduced commitments for voltage support were replaced with additional RMR units. This reflects the fact that when units are brought on for voltage support, they also help satisfy reserve requirements and relieve transmission congestion. So, when fewer units are committed for voltage support, more must be committed for other local reliability reasons.

To address the problem of supplemental commitment for voltage support in Boston, the ISO is pursuing or has implemented the following five initiatives. A number of these initiatives were completed in the 4th quarter of 2004 and were largely responsible for the sizable reduction in supplemental commitments for voltage support in Boston in December 2004.

- Work with the owners of Mystic 8 and Mystic 9 to increase their ability to produce reactive power by a total of 100 MVar – Completed 4th quarter, 2004;
- Work with NSTAR to return a shunt reactor to service with the capability of absorbing 80 MVar – Completed 4th quarter, 2004;
- Work with NSTAR to quickly repair a load tap changer in the Woburn 345/115 kV transformer that will enable three shunt reactors to be more effective in absorbing reactive power – Completed 4th quarter, 2004;
- Revise the ISO's Boston area operating guide based on these three upgrades and train operations staff on new procedures – Proposed 2nd quarter, 2005; and
- NSTAR will install a new 150 MVar shunt reactor to absorb reactive power – Proposed 2nd quarter, 2005.

Figure 15 shows supplemental commitments for local needs in Connecticut on a monthly basis during 2004.

Figure 15
Commitment for Local Reliability in Connecticut
2004 – Daily Peak Hours

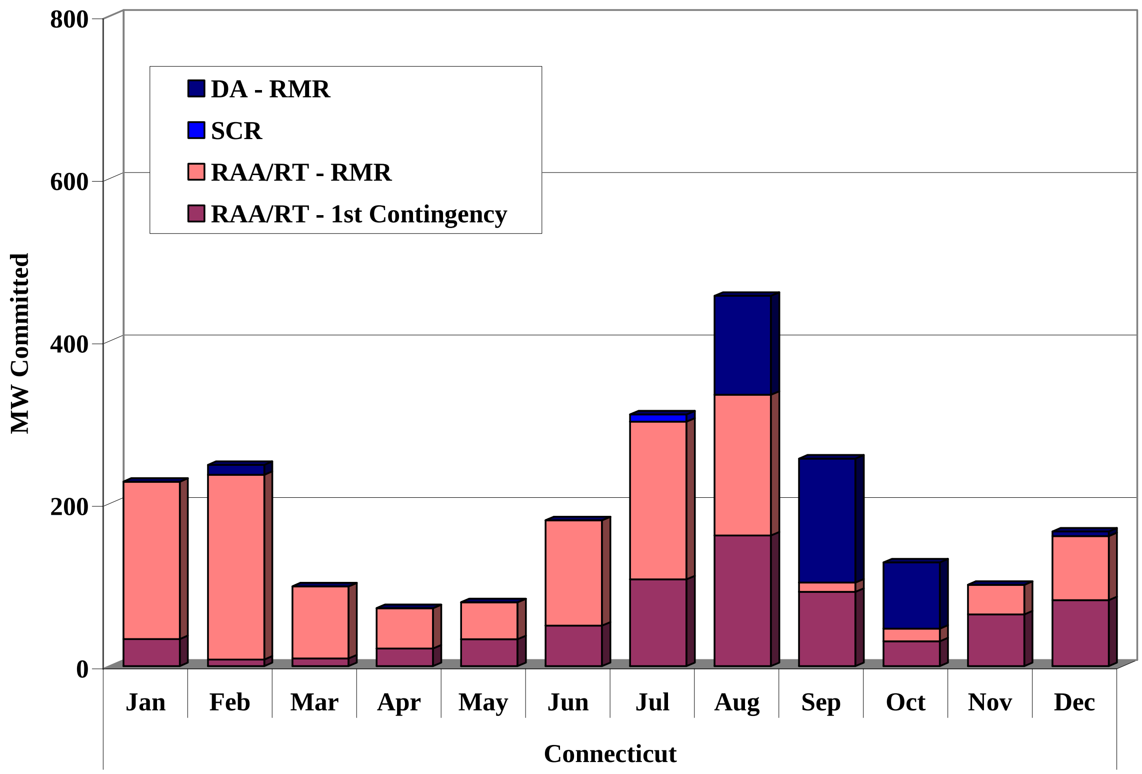


Figure 15 shows that out-of-merit commitment in Connecticut increased and decreased with load over the year, peaking during the summer and winter. This is because local 1st contingency and RMR requirements rise with demand levels. No units were committed for voltage support and SCR commitments only occurred in July. The figure also shows that most of the RMR commitments during August, September, and October were made in the day-ahead market. Over the entire year, however, RMR commitments were ordinarily made through the RAA process.

The RMR commitments in Connecticut are made to address second contingency reliability requirements. The ISO must have sufficient reserves available in each area to respond to the largest generation and/or transmission contingency after the first contingency has occurred. The reserves required in each area vary hourly depending on the availability of quick-start resources, the flow on the interface into the area, the size of the second contingency, and other factors. Due to the limited quantity of quick-start resources in these areas, a large portion of these reserves must be held by on-line resources. If additional quick-start resources are added in these areas over the longer term, the frequency and quantity of supplemental commitment would be substantially reduced.

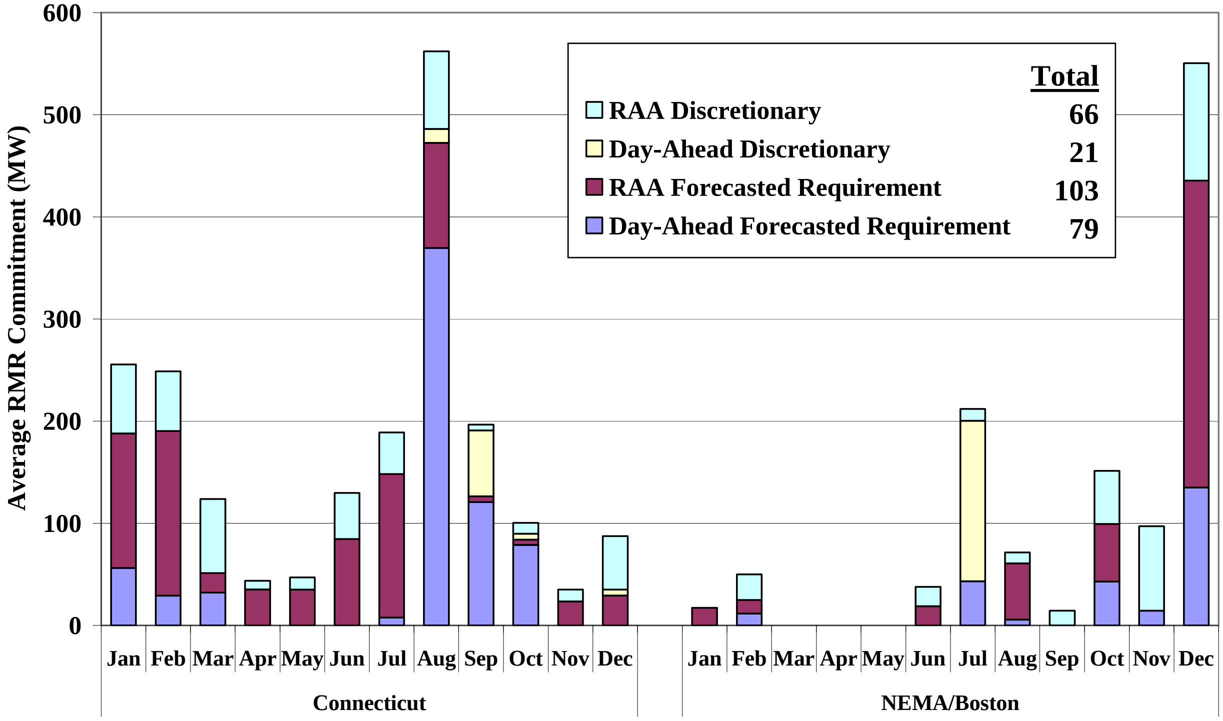
The incidence of supplemental commitment in NEMA and Connecticut, together with the unusually low congestion that has prevailed, suggests that supplemental commitments have contributed to the lack of congestion price differences in these zones. In a subsequent section, we will examine how much energy runs out-of-merit as a result of these supplemental commitments.

2. Evaluation of RMR Commitments

Supplemental commitments in constrained areas can significantly affect the market outcomes. Therefore, it is important that they only be done when truly needed. This subsection evaluates the performance of the ISO in making RMR commitments for NEMA and Connecticut through the day-ahead and RAA processes. Based on ISO operations data for 2004, we divided the quantities of RMR commitments made by the ISO operators between those needed to meet the forecasted RMR requirements versus additional discretionary commitments. Discretionary commitments are those that did not appear to be necessary to meet the local reliability requirements in NEMA or Connecticut.[8] Discretionary commitments may be made for a variety of reasons, including concern by the operators regarding the forecasted peak load in the constrained area or the status of a key resource in the area. The results of this analysis are shown in Figure 16.

8 If only a portion of an RMR resource is needed to meet the forecasted RMR requirements, the entire unit is classified as satisfying the RMR requirement, rather than discretionary.

Figure 16
Reliability Commitments in Constrained Areas
Day-Ahead and Reliability Adequacy Assessment – 2004



The results in Figure 16 show that during 2004, RMR commitments averaged 100 MW day-ahead and 170 MW in the RAA.[9] Generally, the ISO made limited quantities of discretionary RMR commitments in the months studied, which averaged less than 90 MW in both areas. In 2004, roughly 39 percent of day-ahead and 21 percent of real-time RMR commitments were discretionary. To the extent that they are not necessary to maintain reliability, the ISO should continue to reduce these commitments because they can inefficiently mute the transmission congestion into the constrained areas.

The ISO has undertaken the following projects that are expected to reduce the need for RMR commitments and rely more on the market to reflect the value of resources in load pockets. First, the commitment software used by the ISO was modified in early 2005 to minimize the total commitment and dispatch costs associated with satisfying the day-ahead load, subject to the

9 Figure 12 reports only 63 MW of day-ahead RMR commitments, because an average of 37 MW were economic at day-ahead market prices and were thus efficient to commit.

transmission constraints that limit the flow into and out of various areas in New England. Prior to 2005, the commitment software did not recognize import constraints and local capacity requirements, However, the day-ahead market operators would make manual adjustments to the commitment to resolve clear inefficiencies in the commitment related to transmission constraints (e.g., insufficient commitments in Boston or excess commitments in Maine). This software enhancement should improve the efficiency of commitment for local reliability and reduce the need for supplemental commitments.

Second, the ISO–New England Market Monitor identified a problem with the methodology for calculating references prices under the Tariff.[10] Units that are frequently committed for local reliability generally receive a large share of their compensation through guarantee payments corresponding to their offer that result in operating reserve charges to loads. Hence, these suppliers are faced with “pay-as-bid” incentives and do not have an incentive to offer their units at marginal cost. This violates an assumption underlying the reference level calculation methodology, which is that reference levels should be based on periods where the supplier has an incentive to offer at marginal costs. Higher reference levels associated with these incentives can result in inflated guarantee payments and higher operating reserve charges. In addition, the higher offers facilitated by the higher reference levels make it less likely the resources will be committed through the market and more likely they will be committed for local reliability. To address this problem, the ISO–New England made a filing to FERC proposing to change its reference level methodology for units frequently committed and dispatched out of merit order.

The other measures proposed by the ISO to minimize reliance on RMR commitments in load pockets include:

10 Reference prices are used to monitor and, when warranted, to mitigate attempts to exercise market power. A reference price serves as a competitive benchmark for performing the tests that determine whether mitigation may be warranted. The reference price is intended to reflect a generator’s marginal cost, including legitimate risk and opportunity costs. This is an appropriate benchmark because, absent market power, a supplier will maximize its profit by continuing to increase its output until the cost of producing additional output (i.e., its marginal cost) is higher than the market clearing price.

- Develop new Combined Cycle unit dispatch process to gain additional unit flexibility and non-spin capability in load pockets;
- Develop new Day ahead commitment plan for RMR units;
- Identify market enhancements to capture out-of-merit dispatch costs in reserve prices;
- Develop new ASM markets to provide better incentives to for resources in the load pockets, particular for new quick-start units.

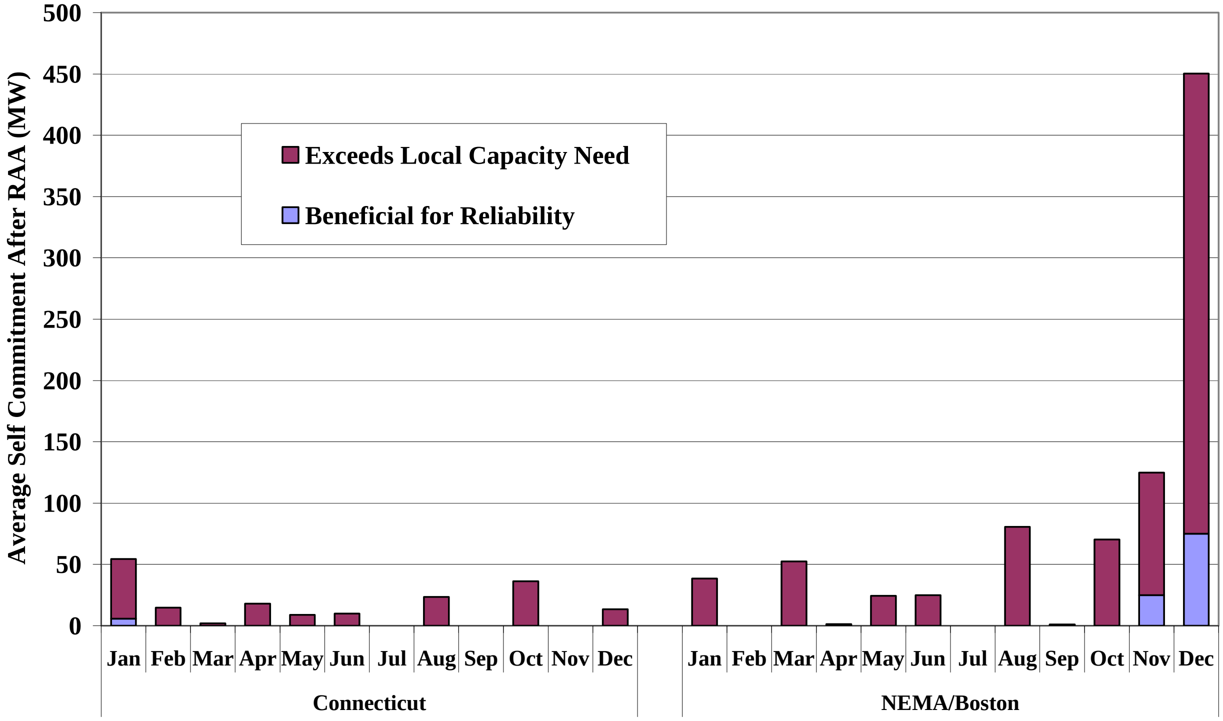
3. Self-Commitment after the RAA

In local areas that are frequently constrained, market-based commitments are not sufficient to secure reliability. Hence, the ISO regularly supplements market-based commitments with additional commitments. The previous section indicates that commitments are made for local 2nd contingency coverage in NEMA/Boston and Connecticut on a regular basis. In the RAA, this is done by forecasting the minimum necessary on-line capacity and then committing additional generators as needed to meet the requirement.

Before making a supplemental commitment, the ISO counts capacity committed prior to this evaluation in the following categories: (i) day-ahead, (ii) after the day-ahead for voltage support, (iii) after the day-ahead for SCR, (iv) self committed in the re-offer period, and (v) committed in the RAA for local 1st contingencies. If the ISO is still short of the local capacity requirement after these commitments, it will commit additional RMR resources. However, if a generator commits itself after the RAA, it can lead to surplus capacity in the load pocket. Moreover, the ISO may need to pay uplift for an RMR-committed unit that would not have been necessary if the ISO had been aware of all self-commitments when it conducted the RAA.

The following figure summarizes the extent to which self-commitment after the RAA has helped meet any remaining local 2nd contingency requirement versus how often it has led to excess capacity in local areas.

Figure 17
Self Commitment after the Resource Adequacy Assessment
Connecticut and NEMA/Boston – 2004



In Connecticut, an average of 15 MW was self committed after the Resource Adequacy Assessment. In NEMA/Boston, the average self commitment after the RAA was also relatively small before November. In November and December in particular, self commitment after the RAA became significant. Only a small quantity of the self commitments was necessary to meet the local capacity requirement. While it can be efficient to have more than the minimum capacity required in each local area, most of the self commitments in NEMA/Boston during November and December occurred after the ISO had already committed units for 2nd contingency coverage. If the ISO knew in advance that these units would be self scheduled, it would have needed to commit fewer units for 2nd contingency coverage.

In some cases, the ISO can de-commit a resource that had been committed in the RAA if a self-schedule occurs later that eliminates the need for the commitment. However, the figure shows that this has not been fully effective and the de-commitment is not without cost. The committed generator may have incurred cost to procure fuel in response to the commitment instruction and can, therefore, be harmed if the commitment instructions are not reliable. In addition, these costs can affect generators' incentives to offer their resources in the RAA process.

Self commitment after the RAA can lead to inefficient market outcomes in the constrained areas. When excess capacity is committed in the load pockets because the ISO has committed units in the RAA for 2nd contingencies prior to a self-commitment, the real-time prices will generally be depressed, congestion into the area will be muted, and operating reserve charges will increase. Since it is difficult for load serving entities to predict when units will choose to self schedule, the result will likely be worse convergence between day-ahead and real-time prices. Figure 7 in Section III indicates that day-ahead congestion into NEMA/Boston rose to significant levels from October to December, while there was almost no real-time congestion during the same period.

This convergence issue was more significant in 2004 than in would be now because the large operating reserve charges associated with the RMR commitments were allocated to real-time deviations, resulting in large and volatile costs for any market participants engaging in virtual transactions to arbitrage these differences. However, in early 2005 the ISO addressed this issue by modifying its allocation of the RMR-related operating reserve costs to all physical load in the constrained area. This eliminates the bulk of the charges to virtual transactions and allows them to act on their incentive to arbitrage the day-ahead and real-time prices in the constrained areas.

The pattern of self commitment after the RAA has likely contributed to poor price convergence as load serving entities are not able to perfectly predict how much load they need to purchase day-ahead.

While there are some legitimate reasons for self commitment after the RAA, the rise in this activity is also consistent with incentive problems that result from frequent supplemental commitment. In NEMA/Boston, local reliability requirements are generally satisfied outside the market process, and these units are paid their offer when the clearing price is not sufficient for them to recover their as-bid cost. Even under perfect competition, units with pay-as-bid incentives rationally offer above costs. Generators frequently committed for local reliability usually have some degree of local market power, and thereby have a greater incentive to offer above marginal cost. If such units submit high-priced offers in the RAA process and are not committed, they would potentially forego the opportunity to sell energy profitably in the real-time market because they will be offline. However, they do not incur this cost because they have flexibility to self-commit the units after the RAA process if they are not selected. Hence, the market rules make this a low-risk strategy.

If the rise in self commitment after the RAA is caused by inefficient incentives, the units that were frequently self-committed in December should be the same units that are frequently committed for local reliability (because they should self-schedule when not selected in the RAA process). The following figure shows the pattern of commitment of the two units that frequently self schedule after the RAA. The two units account for 88 percent of the unit-hours and 99 percent of the MWh self committed after the RAA throughout New England in December.

Figure 18
Two Units Most Frequently Self Committed After RAA
Frequency and Reason for Commitment
September to December, 2004

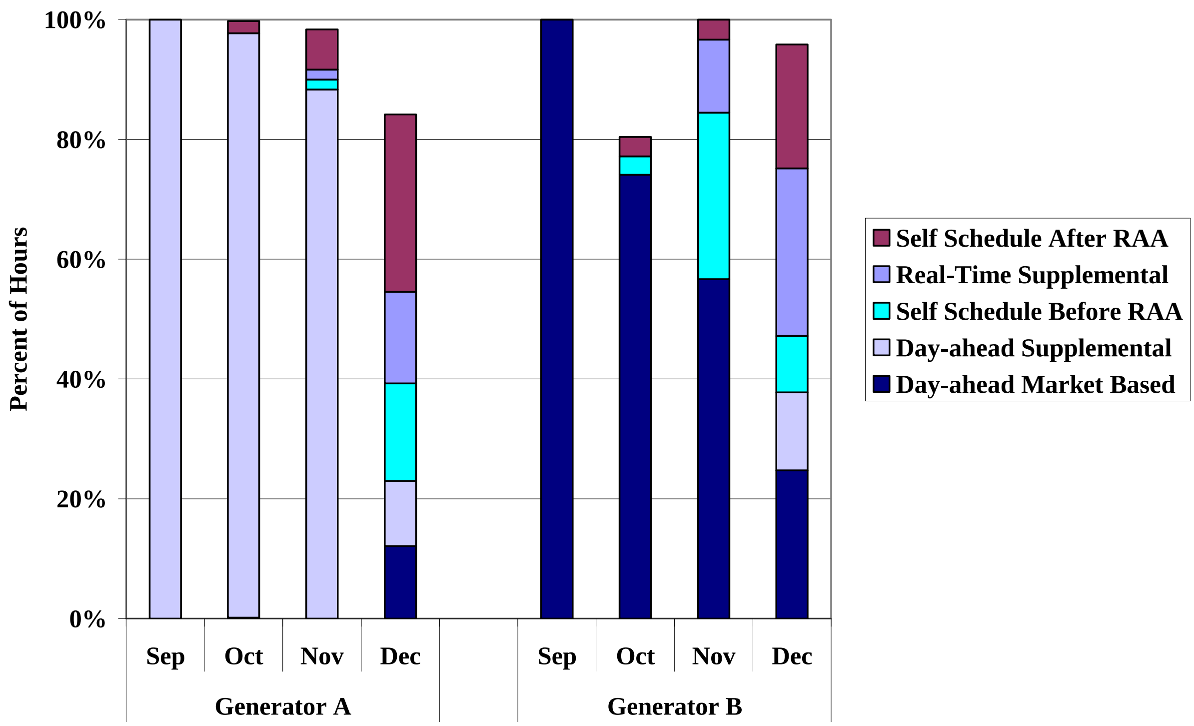


Figure 18 shows that both generators were committed nearly 100 percent of the time, which is justified based on their operating costs. However, Generator A was rarely committed economically through the day-ahead market while the frequency with which Generator B was committed in the day-ahead market decreased sharply from September to December. Generator A was committed in the day-ahead for local reliability (primarily for voltage support) in more than 90 percent of the hours from September to November. By December, commitment of these units for local reliability had decreased substantially, and both were frequently being self-scheduled when they were not committed by the ISO. This indicates that the owner deemed them to be economic at the expected real-time prices. This self-scheduling generally occurred after the RAA process.

The RAA ensures that sufficient capacity is on-line to meet the local 2nd contingency requirements by committing additional generation for RMR reasons. However, some of these RMR commitments made in the RAA process become unnecessary after additional units self commit. This excess capacity depresses real-time prices and results in additional uplift costs. Furthermore, it is evident that generators frequently committed for local reliability have an incentive to wait until after the RAA process to inform the ISO of their decision. To address this incentive problem, we recommend that generators in load pockets be prohibited from self committing after the RAA, unless it is for a legitimate reason (e.g., replacing a unit that is forced out of service).

4. Local Commitment Conclusions

The analysis in this section indicates that the ISO has been committing generating resources consistent with its market-wide and locational reliability requirements under the current market processes and procedures. However, these procedures are resulting in substantial supplemental commitment after the day-ahead market, which can affect real-time prices and increase uplift costs. Work is underway that will address these concerns, including:

- Projects that are planned to reduce the need for supplemental commitments in the constrained areas (e.g., the voltage support improvements in Boston);
- Introducing locational requirements in the forward reserve market;

- Developing full ancillary services markets with locational requirements;

In addition, we recommend:

- Prohibiting suppliers in constrained areas from self-committing generation after the RAA process;
- Improving the consistency of the 2^{nd} contingency proxy limits in the day-ahead and real-time market;

D. Out-of-Merit Dispatch

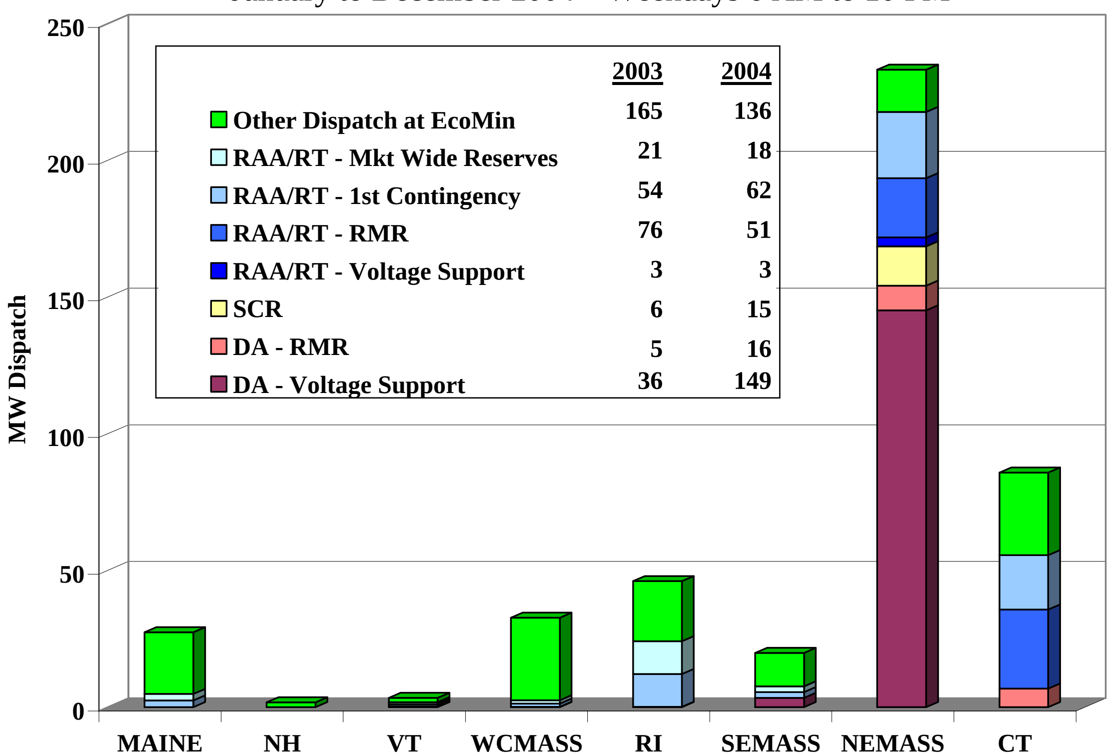
Out-of-merit dispatch occurs in real time when energy is produced from an output range on a unit whose incremental energy offer is greater than the LMP at its location. In general, resources may be dispatched out of merit because either 1) they would not be economic under the current market conditions, but are needed to meet an operational or reliability requirement; or 2) they are economic under the current market conditions, but are ineligible to set the clearing price. In either case, the out-of-merit generation is treated as "must-take" in the market – equivalent to a resource with an offer price of zero.

Out-of-merit generation tends to reduce energy prices by causing lower-cost resources to set the energy price. In a very simple example, assume the two resources closest to the margin are a $60 per MWh resource and a $65 per MWh resource, with the market clearing price set at $65 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-take resource with a $0 offer. Assuming the energy produced by the $100 resource displaces all of the energy from the $65 resource, the energy price will decrease to $60 per MWh.

A unit may be dispatched out of merit for three main reasons. First, a unit may run at its EcoMin to satisfy its minimum run time after having run in merit for several previous hours or in anticipation of running in an upcoming hour. Such a unit may also be at its EcoMin when providing reserves. The real-time market software cannot dispatch a unit below its EcoMin so it will dispatch the unit at its EcoMin if the unit must remain online even when its incremental energy offer is above the market price. 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 committed for reliability after the day-ahead market are committed without regard to their incremental energy offer and are, therefore, more likely than units committed competitively in the day-ahead market to have incremental offers higher than the LMP.

Third, a unit may be out of merit in real time to satisfy reliability requirements in real time. Similar to the supplemental commitments, operators may request certain units to be run at higher levels than their energy offers would justify. This can be necessary for a number of reasons, including (a) voltage support on transmission or distribution facilities; (b) managing congestion on local distribution facilities; or (c) providing local reserves to protect against second contingencies. Figure 19 summarizes by zone the average out-of-merit dispatch for weekday hours (6 AM to 10 PM) during 2004, and it includes a table comparing 2003 and 2004.

Figure 19
Average Hourly Out-of-Merit Dispatch by Zone
January to December 2004 – Weekdays 6 AM to 10 PM



Note: Capacity committed day-ahead for RMR or voltage support that would have been economically committed in the day-ahead market are excluded. The 2003 averages exclude January and February.

As expected, the level of out-of-merit dispatch is much lower than the level of supplemental commitment. In addition, Figure 19 shows that virtually all of the out-of-merit dispatch outside of the constrained areas is attributable to economically committed units dispatched at EcoMin. However in Boston and Connecticut, most of the out-of-merit dispatch is from units committed for local reliability. The average quantity of out-of-merit dispatch was comparable between 2003 and 2004, with the exception of energy from units committed day-ahead for voltage support in Boston. Out-of-merit dispatch from voltage units in Boston was three times larger in 2004 than in 2003.

There are two factors related to the commitment process that contribute to the quantities of resources in the "Other Dispatch at EcoMin" category. First, the excess commitments shown in the prior section will generally increase the supply on the system and cause higher-cost resources to reduce their output to EcoMin. Second, because the day-ahead market commitment model did not recognize first-contingency transmission constraints in export-constrained areas or "generation pockets" (e.g., Maine), the commitment process may result in more units being committed than can be dispatched given the network constraints. This can cause the output from units in these areas to be reduced to EcoMin. Starting in 2005, the ISO has begun to reflect transmission congestion in the commitment model which helps reduce the inefficient commitment of resources in export constrained areas. As a result, this will tend to reduce the amount of out-of-merit dispatch at EcoMin.

The primary causes of out-of-merit dispatch in the constrained areas, including NEMA and Connecticut, are resources committed to satisfy first-contingency, second-contingency (RMR resources), and voltage requirements. In Connecticut, the decreased supplemental commitment for local reliability has led to a significant reduction in out-of-merit dispatch. In Boston, the increase in voltage support commitment day-ahead has led to significant increases in out-of-merit dispatch, so that more than half of all out-of-merit dispatch occurred in Boston in 2004. This is particularly notable because only 12 percent of installed capacity in New England is located in Boston.

While the overall frequency of commitment in Boston for voltage support increased in 2004, the portion of this capacity which is dispatched out-of-merit has also increased. In 2003, the ISO was usually able to commit units for voltage support with low EcoMin values. However in 2004, the available units have typically had high minimum operating levels, which has led to more out-of-merit energy.

Figure 20 shows the monthly pattern of out-of-merit dispatch quantities in NEMA. This figure shows that the incidence of out-of-merit dispatch has been highly correlated with the pattern of supplemental commitment shown in Figure 14 from a previous sub-section.

Figure 20
Average Hourly Out-of-Merit Dispatch in NEMA/Boston
January to December, 2004 – Weekdays 6 AM to 10 PM

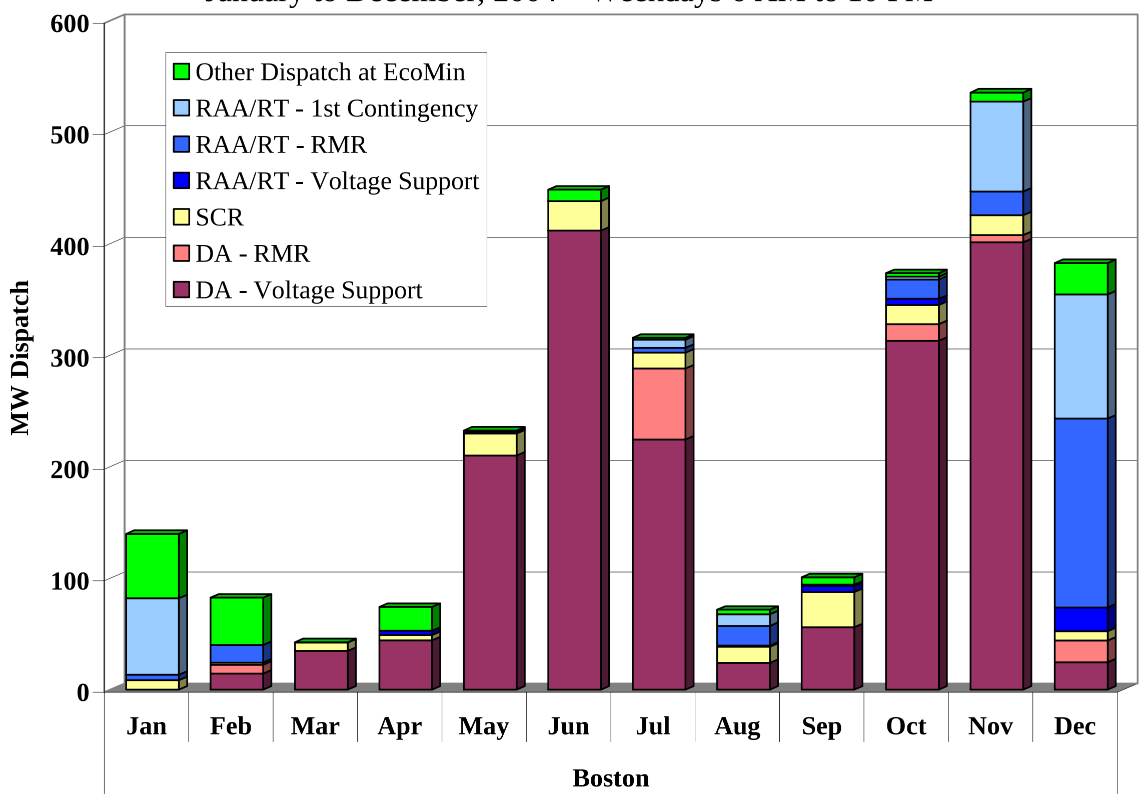
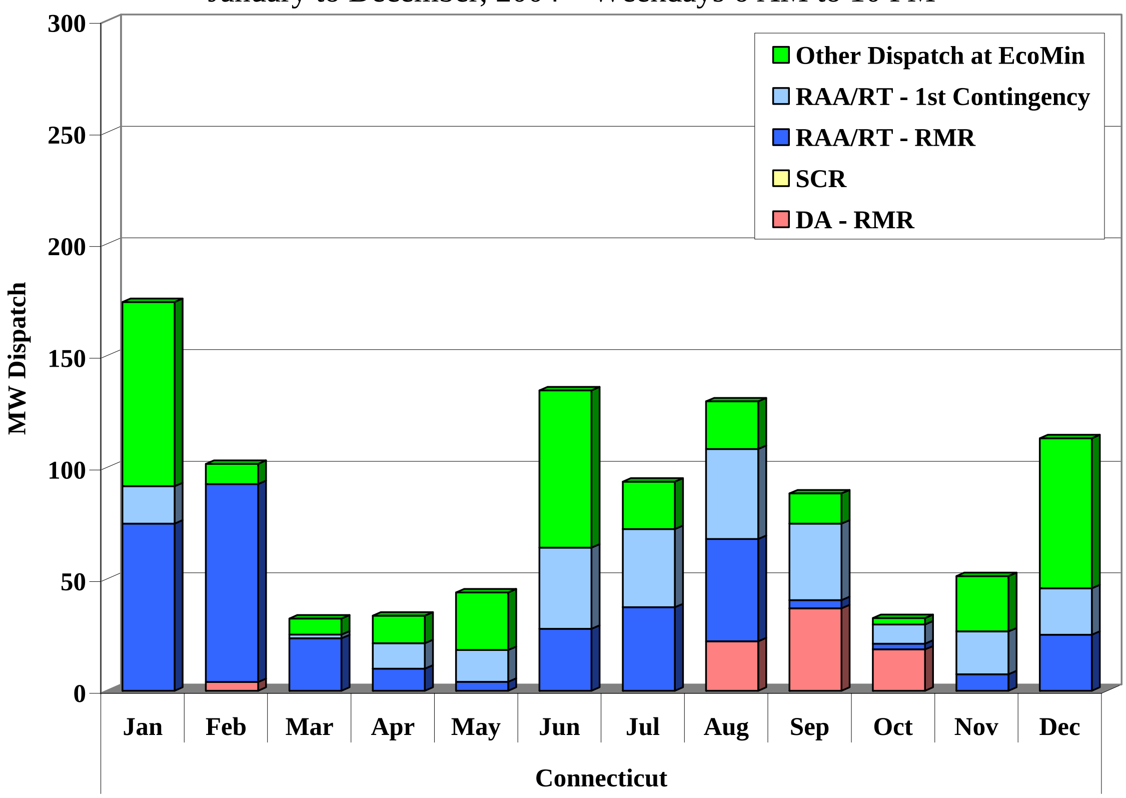


Figure 20 shows that out-of-merit dispatch from voltage support resources averaged more than 200 MW during five months of 2004. During June and November, the average out-of-merit dispatch from voltage support units exceeded 400 MW. These quantities are very large relative to the amount of energy generated in Boston, which was 1,526 MW in 2004. The average generation from Boston units was 1,410 MW during November, so that 38 percent of this output was dispatched out-of-merit. In December, the out-of-merit dispatch from 1st contingency and 2nd contingency units was high because of extreme winter load levels and a reduction in commitment for voltage support.

Figure 21 shows the monthly pattern of out-of-merit dispatch quantities in Connecticut which are substantially lower than 2003 due to a reduction in supplemental commitment for local reliability. To the extent out-of-merit dispatch occurs in Connecticut, it is more frequent in the summer and winter peak load months because high load conditions can increase the incidence of system conditions needing to be resolved by out-of-merit dispatch and commitment.

Figure 21
Average Hourly Out-of-Merit Dispatch in Connecticut
January to December, 2004 – Weekdays 6 AM to 10 PM



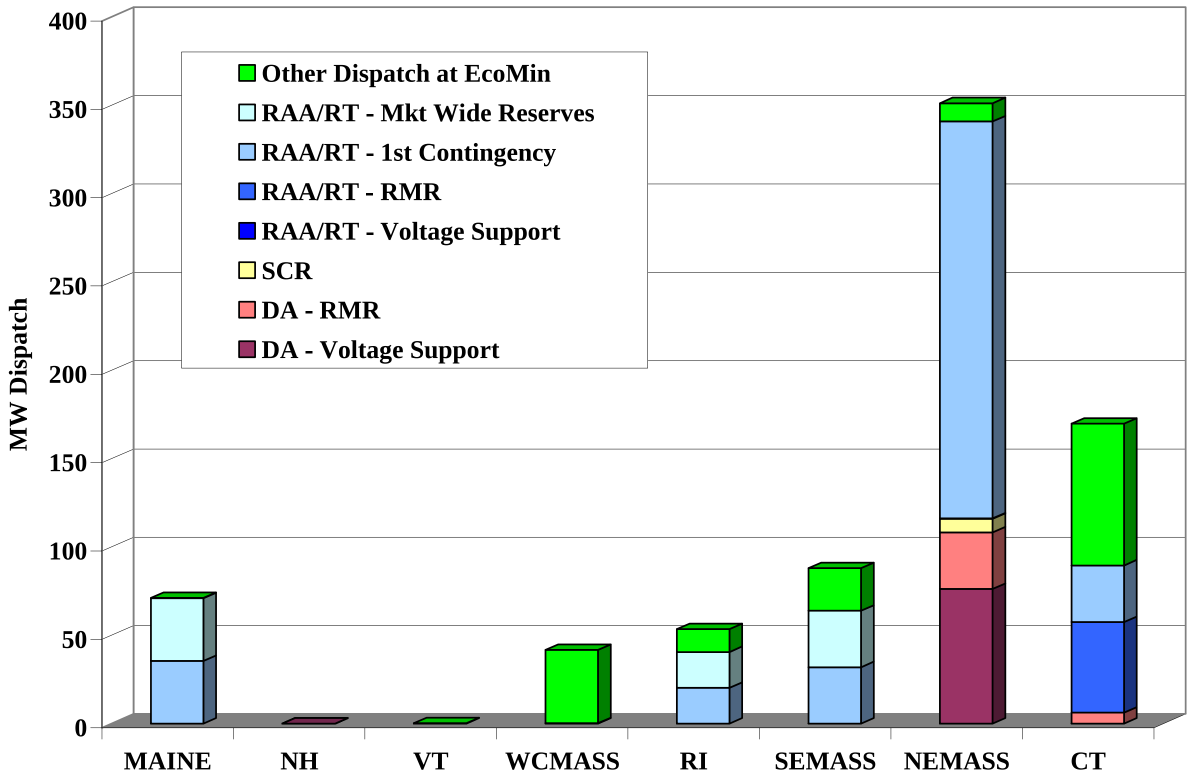
Although some resources may need to be dispatched out of merit in any system, this should be minimized because it can undermine the efficiency of the locational energy prices. Furthermore, owners of units that are frequently called out-of-merit order will have an incentive to offer in excess of marginal costs, which can also affect locational price signals when they are taken in merit order. When units are offered above marginal costs, it reduces the likelihood that they will be committed economically through the day-ahead market, thereby contributing to the need for supplemental commitments. Hence, it is a pattern that can be self-reinforcing.

These results are consistent with those reported by the ISO in its 2003 report reviewing its experience with Peaking Unit Safe Harbor ("PUSH") offer rules that allow "peaking resources" to submit offer prices that would cover the resources' fixed costs.[11] The PUSH report found that although one or more peaking units produced energy in almost two-thirds of the intervals during the summer months of 2003, these units were rarely in merit. Only 8 percent of the output produced by the PUSH units was in merit.

Prices tend to be more sensitive to out-of-merit dispatch during peak-demand periods when the market is clearing at a steep portion of the supply curve (i.e., where supply is relatively inelastic). Because prices are more sensitive under these conditions, out-of-merit dispatch will have a larger effect on prices. Therefore, we examined the highest-demand days to determine the nature of out-of-merit actions at those times. This analysis is shown in Figure 22.

11 *A Review of PUSH Implementation and Results*, ISO-NE, December 2003 ("PUSH Report"). For purposes of the PUSH provisions, peaking resources were defined as those that have capacity factors less than 10 percent.

Figure 22
Average Hourly Out-of-Merit Dispatch by Zone
Top 5 Peak Days



Like the comparable figure in the prior subsection on supplemental commitment, this figure shows that on the highest-demand days the out-of-merit dispatch for local 1^{st} and 2^{nd} contingencies increases significantly in the constrained areas while most other categories of out-of-merit dispatch decrease. This emphasizes the importance of recognizing these reliability requirements within the SMD market framework, which will allow prices in these areas to reflect these requirements. The most important market improvement in this regard is the implementation of operating reserve markets that include the local reserve requirements in the constrained areas.

E. Uplift Costs

In some cases, locational prices are not sufficient to support the costs of resources required to serve load and meet all applicable reliability requirements. There are several ways in which these costs are guaranteed to the owners of these resources and recovered from loads through uplift charges. These payments vary depending on the reason the unit needs the payment (e.g., commitment, dispatch, and/or fixed costs). Similarly, the costs associated with these payments are allocated differently depending on the reason for the action and whether it occurred before or after the day-ahead market.

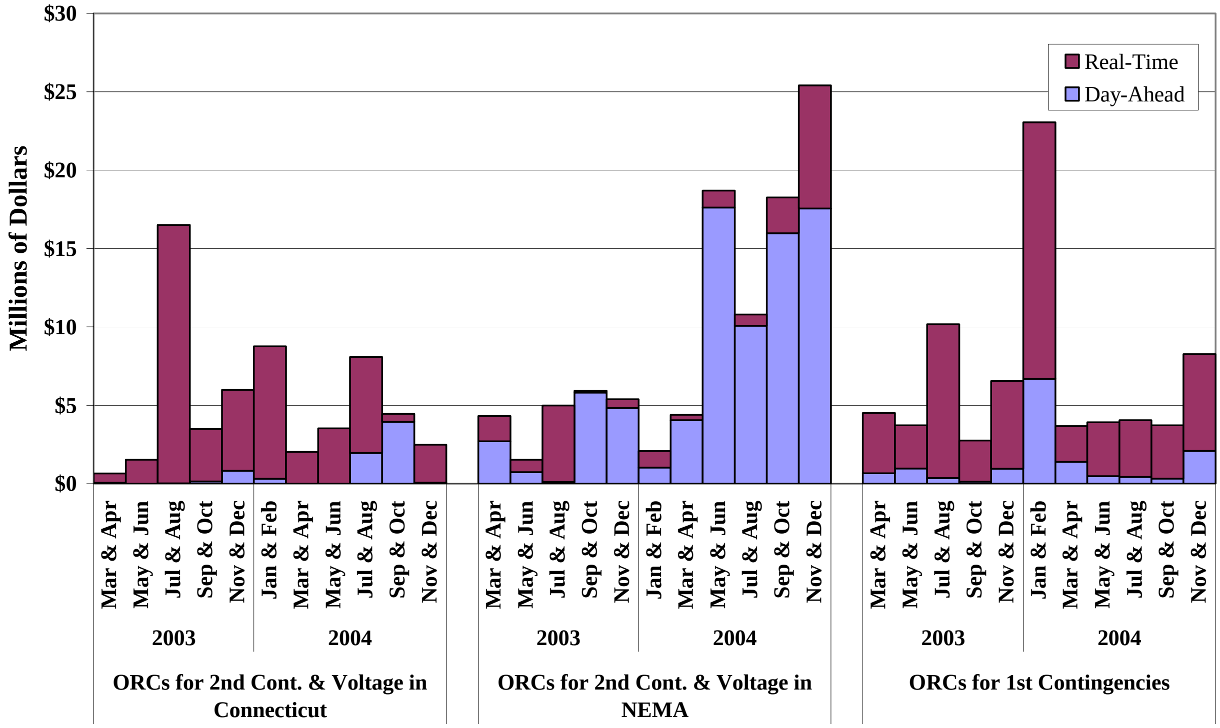
The day-ahead uplift payments arise during the day-ahead market process (not including the RAA and other processes that occur after the market closes). Units may be designated in the day-ahead market for voltage support, as reliability must-run resources, or as special constraint resources. To the extent that these units do not recover their commitment costs in the day-ahead market they will receive uplift payments, referred to as "Operating Reserve Credits" ("ORCs"). Units that are committed economically that do not recover their as-bid production costs through the day-ahead energy market receive Day-Ahead Economic ORCs. These units tend to have high commitment costs relative to their incremental energy costs and can be economic to commit even when they cannot recover their full commitment costs through the day-ahead market.

Units committed after the day-ahead market closes can also be committed for RMR, SCR, and voltage support. In addition, payments to units committed in the RAA process for other reasons are called Real-Time Economic ORCs. While some of these units are committed to meet market-wide forecasted energy and operating reserve requirements, the majority are committed to resolve local 1st contingency requirements. Similarly, units are designated as "RMR" resources when they are committed for local 2nd contingency coverage.

There are several units in highly congested areas that are required for reliability reasons, but cannot earn enough from the energy or capacity markets to pay fixed costs. Furthermore, given expectations of future market conditions, no new investment is expected to replace these resources. To maintain local reliability, these resources are covered under reliability agreements that ensure fixed cost recovery. To the extent that energy and capacity revenues are not sufficient to recover fixed costs for these resources, uplift payments make up the difference. The uplift payments resulting from supplemental commitment and out-of-merit dispatch in 2003 and 2004 under SMD are shown in Figure 23. These payments are divided between (a) ORCs for 2nd

contingency and voltage support commitments in Connecticut; (b) ORCs for 2nd contingency and voltage support commitments in NEMA; and (c) ORCs for 1st contingency commitments. The uplift for SCR commitments and reliability agreements are not shown.

Figure 23
Uplift for Supplemental Commitment and OOM Dispatch
March 2003 to December 2004



Aggregate uplift costs for 2nd contingency commitments in Connecticut decreased in 2004, particularly during the summer. Furthermore, a larger share of the uplift in Connecticut moved to the day-ahead market. NEMA/Boston experienced a substantial increase in uplift costs, primarily due to day-ahead commitments for voltage support. Uplift for voltage support in NEMA/Boston increased from $14 million for the ten months shown in 2003 to $64 million in the same ten months of 2004. The uplift costs in NEMA/Boston have risen significantly in Boston since the summer of 2004 and have begun to shift from the day-ahead market to the real-time market. Not counting January and February, ORCs for 1st contingency commitments decreased slightly in 2004 relative to 2003, although, approximately $18 million of this was incurred from January 14th to January 19th, during the "Cold Snap."[12]

Since the start of SMD, the majority of the uplift payments for supplemental commitments have been to address local reliability requirements in Connecticut and NEMA/Boston. While ORCs for 2nd contingency coverage are charged to the local area, voltage support charges are assessed to the entire market. Furthermore, the uplift costs associated with commitments made to satisfy 1st contingency requirements in Connecticut and NEMA/Boston are allocated to the entire market. Since supplemental commitment for voltage support and 1st contingencies makes additional reserves available in local areas, these often take the place of 2nd contingency commitments.

The costs of the various sources of uplift are allocated in different ways. The following table shows how the different types of uplift costs are allocated.

Table 6: Allocation of Uplift Costs

	Day-Ahead Market	RAA and Real-Time Dispatch
ORCs for Second Contingencies:		
Prior to March 1, 2005	**Allocated to load scheduled in the day-ahead market in the zone where the RMR unit is located.**	**Allocated to real-time deviations in the zone where the RMR unit is located.**
After March 1, 2005	**Allocated to physical load in the zone where the unit is located.**	
ORCs for First Contingency	**Allocated to load scheduled in the day-ahead market in all of New England.**	**Allocated to real-time deviations in all of New England.**
ORCs for VAR and Voltage Support	**Allocated to all physical load in New England**	
ORCs for SCR (local distribution)	**Allocated to transmission operator that requested the SCR**	
Reliability Agreement Costs	**The charges are assessed to the physical load in the zone where the RMR unit is located.**	

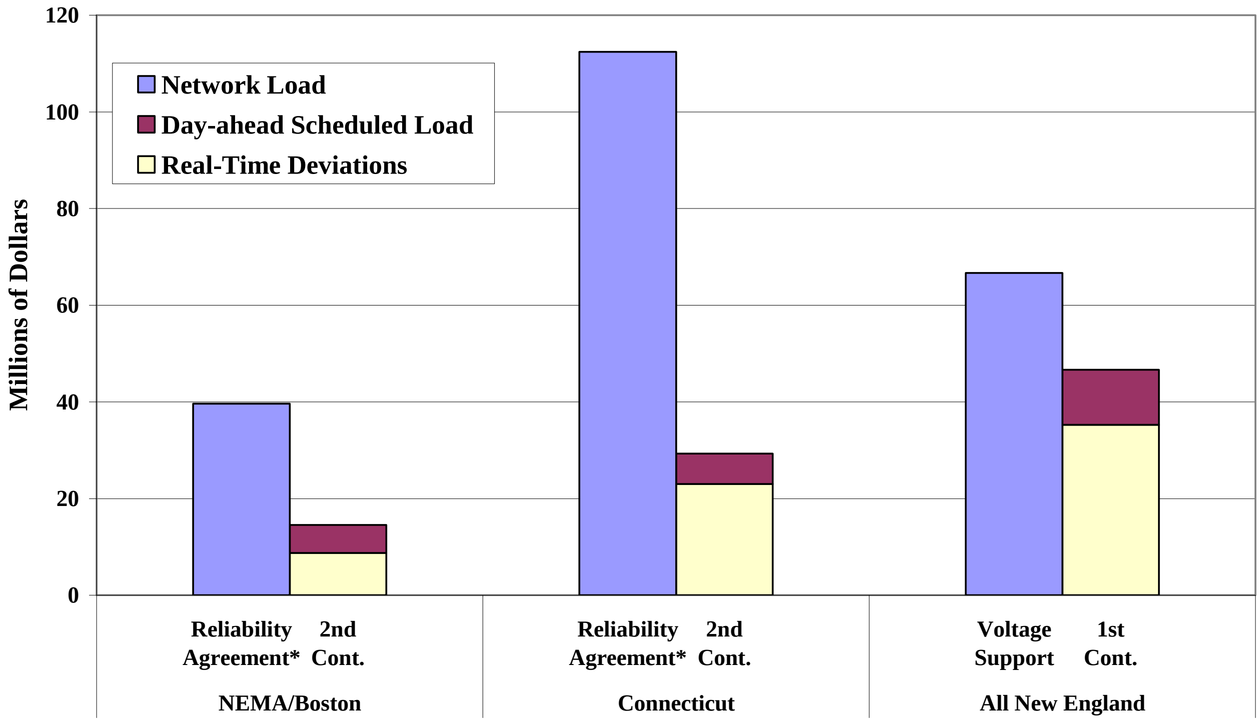
12 *"Report on Electricity Supply Conditions in New England during the January 14-16, 2004 "Cold Snap"*, ISO-NE, Inc.

Prior to March 1, 2005, day-ahead ORCs for 2nd contingencies were allocated to purchases and sales in the day-ahead market, while real-time ORCs (incurred after the day-ahead market) for 2nd contingencies were allocated to real-time purchases and sales (i.e., real-time deviations). In March 2005, the RMR cost allocation was modified to be allocated to all physical load in the zone to address disincentives that the prior allocation created for engaging in virtual trading in the constrained areas. Uplift costs for voltage support and payments under reliability agreements are allocated to physical load -- voltage support is allocated market-wide while reliability agreement costs are allocated zonally.

Most of the supplemental commitments for local 2nd contingencies occur after the day-ahead market and are thus considered real-time ORCs. In 2004, these costs were allocated only to the real-time deviations, which represented a very small portion of the entire load in the zone. This allocation methodology can have a substantial effect on certain types of conduct, including virtual trading and price-sensitive day-ahead demand. Virtual trades are sales or purchases of energy in the day-ahead market that are settled in real time because there is no corresponding physical load or generation. Hence, the entire quantity of the virtual load or generation is a real-time deviation. Likewise, load-serving entities (LSEs) that bid price-sensitively in the day-ahead market, opting to purchase in real time if they deem that prices are unjustifiably high in the day-ahead market, will be allocated a large share of the uplift costs because their entire real-time purchase will be a real-time deviation.

Figure 24 summarizes the total costs of uplift associated with reliability agreements and supplemental commitment (excepting SCR uplift) that are allocated to various areas within New England.

Figure 24
Allocation of Uplift for Out-of-Market Energy and Reserves Costs
2004



Note: While information is not available on the breakdown of payments under reliability agreements by zone, this analysis assumes that the ratio of payments to fixed cost guarantees is the same for NEMA/Boston and Connecticut.

The total cost incurred in these areas was $309 million in 2004. While approximately $54 million was assessed to NEMA/Boston and $142 million was assessed to Connecticut, $113 million was charged to all of New England. 97 percent of the uplift for voltage support was incurred from committing units in NEMA/Boston, but since these costs are shared by all network load, only 27 percent of the charges are assessed there.

While it is undesirable to generate large ORC payments, some level is unavoidable at this point in the market's development. However, the current allocation methods can have adverse incentive effects and revisions would serve the interest of economic efficiency. First, and foremost, the practice of allocating the cost of real-time ORC payments to real-time deviations from day-ahead schedules creates artificial disincentives for virtual trading and price-sensitive load purchases in the day-ahead market.

Virtual trading and price-sensitive load purchases facilitate price convergence between the day-ahead and real-time markets. To the extent that the allocation methodology creates a disincentive for these actions, larger and more volatile prices differences between the day-ahead and real-time markets will prevail, leading to higher overall costs of serving load. Virtual trading also mitigates market power in the day-ahead market. If a participant attempts to raise day-ahead prices by withholding resources, this will be undermined by virtual sales and loads that choose to buy in the real-time market. However, the effectiveness of this process depends on the real-time market being a viable option for participants without them incurring substantial additional costs.

To address these concerns, the ISO–New England has changed its Tariff to assess ORCs for local 2nd contingency commitments to physical load in the zone rather than day-ahead schedules and real-time deviations. This change was made effective, March 1, 2005, and will likely lead to additional virtual trading in Connecticut and NEMA/Boston which are the only zones with significant ORCs for local 2nd contingency commitments. This change was justified not only by the economic efficiency considerations described above, but also by the fact that most of the commitments that result in uplift costs are made to protect the reliability of all load.

The one type of operating reserve credit cost that arguably could be allocated to real-time deviations is the cost of the Real-time ORCs associated with commitments necessary to satisfy the forecasted load in the RAA. The RAA is used to ensure that sufficient generation is committed to meet forecast load in real time. To the extent loads are under-scheduled in the day-ahead market, these costs may appropriately be allocated to real-time purchases. This differs from allocating the uplift costs to all real-time deviations, which would include loads that were over-scheduled in the day-ahead market. However, these make up a relatively small share of the category labeled above as 1st contingency. Actual commitments for local 1st contingencies, like those for local 2nd contingencies, are incurred to protect physical load in specific areas, and therefore, should be allocated to the physical load in the zone that benefits from the supplemental commitment.

F. Market Operations -- Conclusions

In general, we conclude that the markets operated well during 2004. Price corrections have been rare, and load forecasting has been accurate. However, substantial quantities of supplemental commitments continue to occur in both NEMA/Boston and Connecticut. These commitments are necessary, in part, because these areas do not have a large quantity of quick-start resources that can help meet the capacity requirements of the local area while offline.

Supplemental commitments and out-of-merit energy dispatch create four issues in the New England market.

- They create inefficiencies because supplemental commitments are made with the objective of minimizing commitment costs (i.e., start-up, no-load, and energy costs at EcoMin), rather than minimizing the overall production costs.

- They tend to mute signals to invest in areas that would benefit the most from additional generation and transmission investment. They also stifle interest in registering potential demand response by diminishing the financial incentives for it.

- They can create incentives for generators frequently committed for reliability to avoid market-based commitment when they would be economic at the day-ahead LMP. This frequently induces the ISO to commit the resource in the Resource Adequacy Assessment ("RAA") process for local reliability where the generator is paid its bid price in the form of uplift. When the generator is not committed in the RAA, but expects to be economic at the real-time LMP, it simply commits itself after the RAA. The report finds that a very small number of generators in the NEMA/Boston area did this with regularity during the month of December, when they accounted for 88 percent of the unit-hours and 99 percent of the MWh of capacity self-committed after the RAA process.

- They cause a substantial amount of uplift costs that is difficult for participants to hedge and can be quite volatile, most of which are generated by commitments in Connecticut and NEMA/Boston. The uplift costs associated with these commitments are allocated in a variety of ways based on Tariff requirements. Some of these allocations can create inefficient incentives. The report discusses these allocations and recommends some improvements.

The ISO has already implemented several changes that should reduce the need for supplemental commitments and improve the economic signals in the constrained areas. The most important change is the improvement to the commitment software and process to recognize transmission limits in the day-ahead market commitment. These limits include the first contingency limits and "proxy 2nd contingency" limits that recognize the 2nd contingency reliability requirements in Connecticut and Boston. Day-ahead market administrators had previously accounted for these limits by manually adjusting the day-ahead market commitment. Other measures being pursued to minimize reliance on supplemental commitments in load pockets include:

- Coordinating with NSTAR and one of the suppliers in Boston to increase the capability of the transmission system to produce and absorb more reactive power in key locations – several improvements were made in 2004 and others should be completed in early 2005. These improvements will reduce the need for supplemental commitments for voltage support, the largest source of supplemental commitments in 2004.
- Developing a new Combined Cycle unit dispatch process to gain additional unit flexibility and non-spin capability in load pockets;
- Developing a new day-ahead commitment plan for RMR units;
- Identifying market enhancements to capture out-of-merit dispatch costs in reserve prices;
- Developing new ASM markets to provide better incentives for resources in the load pockets, particular for new quick-start units; and
- Modifying the methodology for calculating references prices for units frequently committed for local reliability in constrained areas.

- Developing a new day-ahead commitment plan for units with reliability agreements;
- Identifying market enhancements to capture out-of-merit dispatch costs in reserve prices;
- Developing new ancillary services markets to provide better incentives for resources in the load pockets, particularly for new quick-start units; and
- Modifying the methodology for calculating references prices for units frequently committed for local reliability in constrained areas.

In addition, we recommend the following changes to further reduce the inefficiencies associated with supplemental commitments. We recommend that the ISO:

- Consider the merits of not allowing suppliers in load pockets to self-commit units after the RAA process unless they have suffered an outage on another unit or they provide comparable justification. This would reduce the quantity of supplemental commitments, improve the ISO's decision-making in the RAA process, and increase suppliers' incentives to offer resources competitively in the RAA since it would be their last opportunity to commit a unit;

- Allocate uplift for local 1st contingency commitments in the same manner as local reliability uplift is allocated. Currently, uplift for local 1st contingency commitments is assessed to market participants based on their scheduling behavior in the day-ahead and real-time market. Instead, it should be allocated to the physical load in the local area that benefits from the commitment. This change would reduce disincentives for virtual trading and price-responsive load scheduling in the day-ahead market. Additionally, it would recognize that commitments for local reliability protect all load in the area, regardless of whether the load settles in the day-ahead market or real-time market;

- Allocate uplift for voltage support commitments in the same manner as local reliability uplift is allocated. Currently, uplift for voltage support commitments is assessed to all New England load, although voltage support primarily benefits load in the local area. Assessing this uplift to the local area will provide appropriate incentives to upgrade the transmission system. This change is currently being considered by the NEPOOL Tariff Committee; and

- Evaluate the underlying assumptions in the calculation of the import limits to constrained areas to resolve the inconsistencies between the day-ahead and real-time limits. This would improve the efficiency of the day-ahead commitment and tighten convergence between day-ahead and real-time market outcomes.

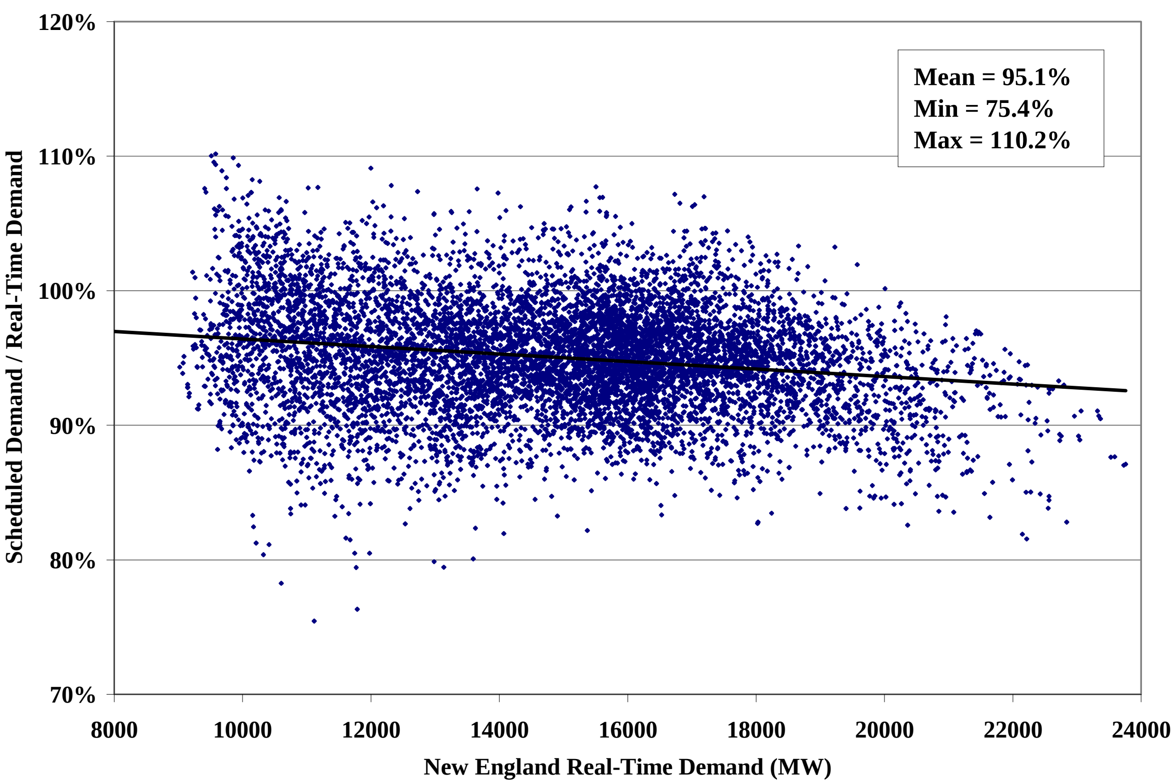
V. Demand Scheduling in the Day-Ahead Market

In this section, we examine the load-scheduling pattern in the day-ahead market to determine whether it has been consistent with efficient market operations. We also analyze virtual trading – both virtual supply and virtual demand.

A. Load Scheduling

Demand bidding can have important effects on market efficiency. Under-scheduling demand in the day-ahead market can lower day-ahead prices and contribute to the need to commit supplemental resources, which can distort real-time prices as explained in the prior section. Figure 25 shows the hourly ratio of demand scheduled in the day-ahead market to the actual real-time demand in New England. The scheduled day-ahead load includes the physical demand scheduled plus the net virtual load scheduled (virtual load minus virtual sales).

Figure 25
Ratio of Scheduled Demand to Real-Time Demand in New England
2004



This figure shows that the average ratio of the day-ahead demand to the real-time demand is 95.1 percent -- i.e., demand is under-scheduled by 4.9 percent, on average. While there is a substantial range of hourly values, the vast majority of hours are scheduled at between 90 percent and 100 percent of actual demand. The figure also reveals that there is a tendency for under-scheduling to increase as the actual level of demand increases. To understand the causes and effects of the under-scheduling, we analyze the load scheduling by zone. Figure 26 shows the weighted average ratio of demand scheduled in the day-ahead market to real-time demand for each zone in 2003 and 2004.

Figure 26
Average Ratio of Scheduled Demand to Real-Time Demand
March 2003 to December 2004

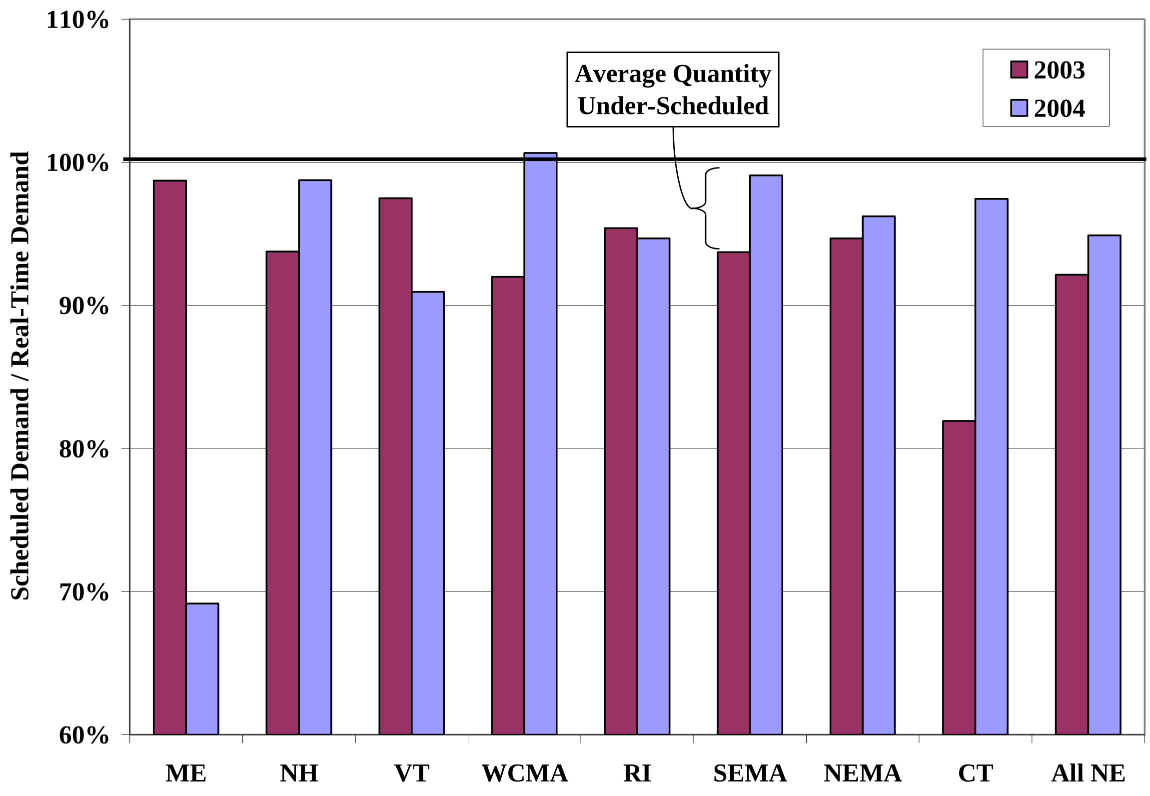


Figure 26 shows that the percentage of demand scheduled in the day-ahead market rose from 92 percent in 2003 to 95 percent in 2004. While this is a small percentage increase, it implies that the quantity of under-scheduling has decreased substantially, from 8 percent to 5 percent. While this was driven by increased scheduling in five zones, there were significant decreases in demand scheduling in Maine and Vermont. Indeed, Maine and Vermont were the only zones that did not schedule 95 percent or more of their demand in the day-ahead market.

In 2004, the fraction of Connecticut demand scheduled day-ahead was much higher than in the previous year. In 2003, the under-scheduling in Connecticut was consistent with the effects of the supplemental commitment after the day-ahead market and out-of-merit dispatch described in the prior section. Non-market-based commitment and dispatch tends to depress real-time prices. This creates a premium in the day-ahead market and participants will naturally act on these economic incentives to reduce their day-ahead schedules. This can take the form of reduced schedules by LSEs in the area, reduced virtual loads, or increased virtual supply (all of which reduce the net load scheduled in the area). This under-scheduling pattern is self-reinforcing to some extent because it increases the need for supplemental commitment, which tends to reduce real-time prices and increases the incentive to under-schedule.

The most effective way to address this problem is to reduce the need for supplemental commitment and out-of-merit dispatch over time by improving the representation of contingency requirements in the market software. The ISO–New England has made strides in reducing supplemental commitment in Connecticut, and this is reflected by increased convergence between day-ahead and real-time scheduling there.

In 2004 in NEMA/Boston, supplemental commitments and out-of-merit dispatch were more significant than in 2003. However, this did not lead to a significant change in day-ahead demand scheduling. The reason is that the majority of supplemental commitments were made in the day-ahead market. This consistency between the day-ahead and real-time models allows better convergence between markets. Thus, to the extent the need for supplemental commitment can be predicted day-ahead, it is beneficial for the operator to take these actions as early as possible.

This section indicates that the overall scheduling patterns are consistent with the economic incentives facing the market participants. A key component of overall scheduling patterns has to do with the quantities of virtual load and virtual supply (collectively “virtual trading”) scheduled in the day-ahead market. These patterns are evaluated in the following subsection.

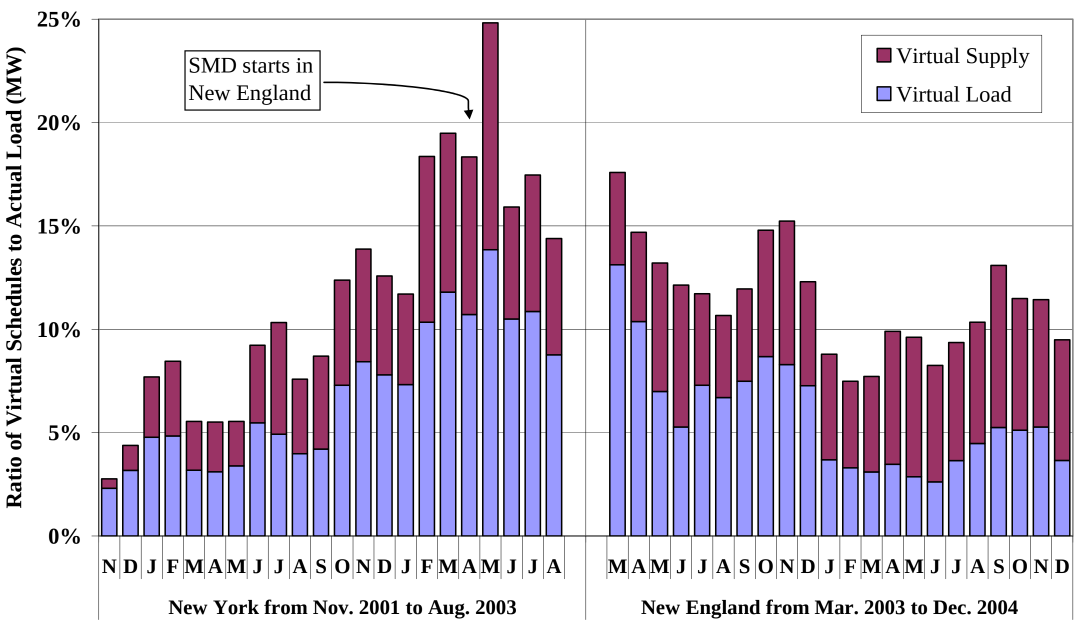
B. Virtual Trading

Virtual trading allows participation in the day-ahead market by entities other than Load-Serving Entities and generators. Virtual trades settle in the real-time market. For example, if the day-ahead prices are lower than a participant expects they will be in the real-time market, the participant can make virtual purchases in the day-ahead market and subsequently sell the purchased energy back into the real-time market. Virtual trading plays an important role in a multi-settlement market by:

- Improving the convergence between the day-ahead and real-time prices;
- Providing additional flexibility for participants to manage their positions and associated risk in the ISO markets; and
- Mitigating market power in the day-ahead market by reducing net day-ahead energy purchases when day-ahead prices would otherwise be artificially inflated.

The following analysis evaluates the trend of scheduled virtual load and virtual supply in New England since the start of SMD. To provide a benchmark for performance, Figure 27 shows the quantities of scheduled virtual transactions as a percent of actual in New England compared with the same ratio in New York during the first 22 months after virtual trading was introduced there.

Figure 27
Average Volume of Virtual Transactions per Hour
New England vs. New York – First 22 Months of Operation



The figure shows that in New England the volumes of virtual purchases and sales as a percent of actual load were highest during the first few months of SMD. The average quantity scheduled in the first month of SMD was 2,600 MW, but it has since fallen and ranged between 1,100 MW and 2,000 MW during 2004. In contrast, the figure shows that virtual trading was initially quite small in New York during the first months—less than 500 MW on average. However, the quantity of virtual transactions has increased significantly over time, ranging from 2,800 MW to 4,000 MW in the last six months shown.

The figure also shows that in March 2003, virtual transaction volumes were similar in New York and New England. However, after the first months, the overall level of virtual trading has declined in New England, while volumes have continued to increase in New York. This is a particularly surprising outcome since New England allows for nodal virtual trading while New York still allows it at only the zonal level.

Virtual trading can play an important role in the day-ahead market by improving price convergence with the real-time market, providing flexibility for participants to hedge commitment and scheduling risks, and mitigating potential market power and gaming opportunities in the day-ahead market. Therefore, it is important to minimize inefficient disincentives for participants to engage in virtual trading. In prior reports, we had identified an issue regarding the allocation of real-time operating reserve credits – namely, that these costs are allocated only to real-time deviations from day-ahead schedules.[13] Such deviations include under-scheduled load that will purchase energy in the real-time market, over-scheduled load that will sell the excess energy in the real-time market, and virtual trades that will settle their position in the real-time market.

Because the virtual load and supply can represent a relatively large share of the deviations, they will bear a corresponding large portion of the real-time uplift costs. This raises concerns to the extent that these costs could serve as a disincentive to engage in virtual trading. To evaluate this issue, Figure 28 shows the average uplift cost allocation per MWh resulting from operating reserve credits in three areas of New England.

13 *Six-Month Review of SMD Electricity Markets in New England*, Potomac Economics, February 2004.

Figure 28
Allocation of Operating Reserve Credits for 1st and 2nd Contingencies
January to December 2004

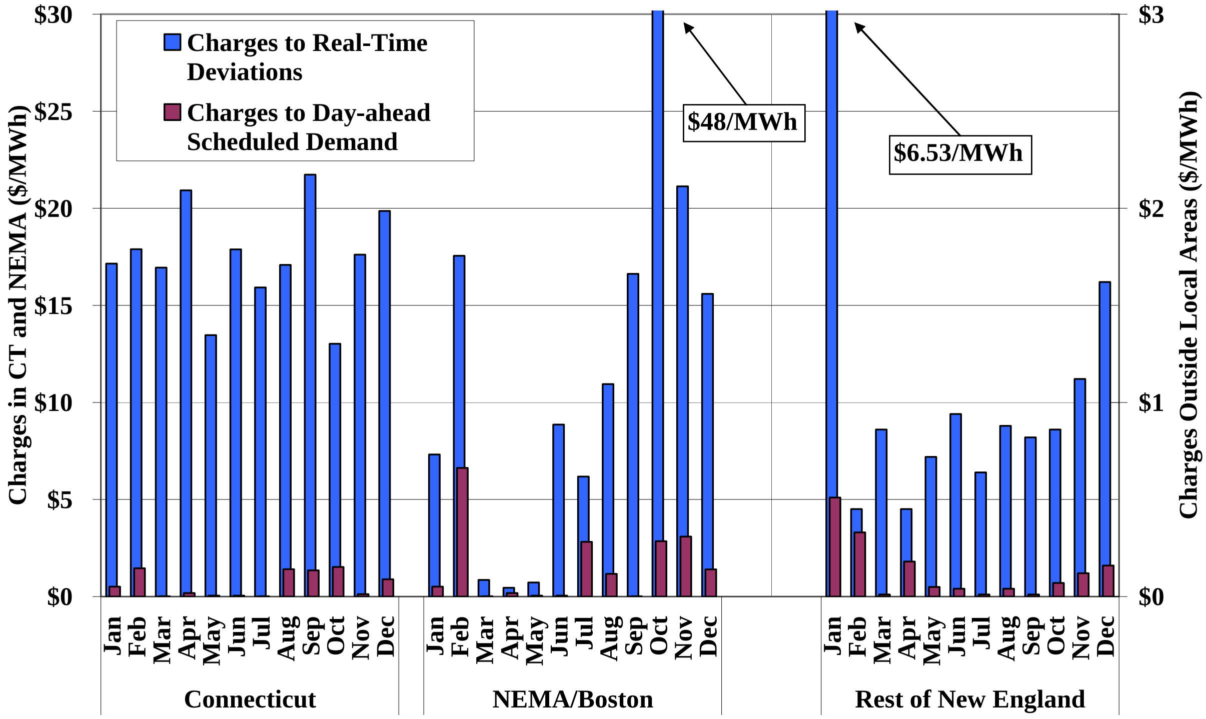


Figure 28 shows that the uplift cost allocation method results in much larger charges to real-time deviations than to physical demand scheduled day-ahead. This imposes significant costs on virtual trades which are charged as real-time deviations. In NEMA and Connecticut, the real-time charges can be quite high due to frequent supplemental commitment of units for local 2nd contingencies. Uplift costs associated with local 2nd contingency units are allocated directly to the zones so that the costs are higher for real-time deviations in these locations. In areas outside NEMA and Connecticut, the charges were due to 1st contingency commitments that averaged $0.45 to $6.53 per MWh for real-time deviations.

In order to arbitrage day-ahead to real-time price differences, virtual traders form expectations regarding the real-time prices at a location the following day. They will try to schedule virtual load when they expect the day-ahead price to be lower than the real-time price and virtual supply when they expect the day-ahead price to be higher. Because these expectations are subject to substantial uncertainty, virtual trades can result in a loss. Virtual traders will only submit offers if they expect profits to exceed what they expect the allocation to be for real-time deviations.

Thus, in December 2004, virtual traders would have to expect to earn at least $1.62 per MWh in order for their trades to be profitable. Real-time deviation charges create substantial disincentives for virtual trading and this is reflected in the relatively low volume of trades shown in Figure 27.

Table 1 in Section II of this report shows the average day-ahead price premium ranged from $0.82 to $2.00 per MWh for different zones in 2004. While risk aversion helps explain the day-ahead premium, it is likely that the difference between charges to day-ahead scheduled physical load and real-time deviations has contributed to the premium as well. According to Figure 28, load-serving entities outside Connecticut and NEMA/Boston paid $0.12 to $1.46 per MWh less on average in uplift charges by scheduling day-ahead. Thus, they would pay less on average by scheduling load at a slightly higher day-ahead price. In Connecticut and NEMA/Boston, charges for operating reserve credits frequently averaged more than $10 per MWh for a month. While this has the potential to create a large day-ahead premium in these areas, the general lack of congestion has prevented the day-ahead premium from varying significantly by zone.

These results suggest that the uplift cost-allocation rules should be altered to eliminate substantial disincentives to engage in virtual trading and price-sensitive demand bidding in the day-ahead market. To address these incentive problems, the ISO changed its Tariff so that charges for RMR operating reserve credits will be allocated to physical load within the zone starting March 1, 2005. This is particularly appropriate given that all loads in these areas are protected by the local-reliability commitments made after the day-ahead market closes. However, the allocation of real-time economic operating reserve credits to real-time deviations continues to be an impediment to virtual trading. Since a large share of this allocation results from supplemental commitment to resolve local 1st contingency limits in local areas, we recommend that the ISO allocate these local reliability costs to physical load in the area as it is doing with RMR costs.

VI. Regulation Market

In this section of the report we evaluate the market for regulation. In particular, we evaluate (a) the practice of self-scheduling regulation after the close of the regulation market, (b) the adequacy of resources capable of supplying regulation in New England, and (c) upcoming enhancements to the regulation market under Phase 1 of the Ancillary Services Market ("ASM").

A. Self-Scheduling by Regulation Providers

Prior to SMD, regulation was cleared in real-time together with energy and other ancillary services. Under SMD, the regulation market is cleared day-ahead after the day-ahead energy market clears. Under this process, owners of regulation-capable resources may submit regulation bids up until 6 PM the day before. An offer to supply regulation consists of a price per megawatt of flexibility and the associated quantity in both the "up" direction" and the "down" direction (required to be equivalent). The offer also must specify the output range within which the unit is capable of providing regulation service.

The regulation optimizing model, known as REGO, clears the regulation market using regulation offers and the anticipated demand for each hour of the next day. The regulation clearing price (the RCP) and the regulation quantities are determined by 10 PM. To arrive at the RCP and regulation quantities, the ISO establishes an economic merit order of supply offers. The economic merit of each unit is determined by the sum of (a) the unit's regulation offer price and (b) an estimate of the opportunity cost that the unit would incur to ensure it stands ready to provide regulation within its regulating range.[14] All regulating units are paid the RCP, but non-self-scheduled units may earn an additional payment if the RCP is not sufficient to cover lost sales in the energy market as a result of being dispatched lower than or higher than the unit's optimal energy quantity. Self-scheduled units, because they are supplying regulation voluntarily, do not incur such opportunity costs and are paid only the RCP.

14 As discussed further below, the opportunity cost can arise when the unit is dispatched at a lower quantity than specified in its energy bid so that it is within the regulation-capable range. If it is dispatched at the lower level, it loses the opportunity to sell more energy at the energy clearing price.

Until February 20, 2004, participants were allowed to self schedule after the determination of the RCP. These real-time self schedules were queued at the bottom of the economic merit order and would effectively "bump" units towards the top of the merit order. The new marginal regulation unit was thereby less costly than indicated by the RCP. However, the RCP would not change to reflect this and units that provided regulation in real-time were paid the RCP. Because the RCP would not change as a result of self schedules, participants had a strong incentive to self-schedule after the price floor was established at the market close. Only participants that expected to incur significant opportunity costs had the incentive to provide regulation through the market rather than via a self schedule. Other suppliers were able to costlessly induce a higher regulation floor price by withholding resources from the regulation market and then self-scheduling after the market close.

After February 20, 2004, market participants were prohibited from self scheduling after the RCP was determined. Even after the rule change, units not committed to the energy market that offered to self schedule regulation were still able to provide regulation at the RCP if they committed themselves after the regulation auction. While these units had self scheduled regulation prior to the regulation auction, the REGO model was not able to evaluate the offer because the unit had not been committed. At the end of October 2004, the required software change was made to prevent this class of unit from providing regulation in real-time.

As Figure 29 demonstrates, there were obvious improvements to market efficiency after these changes were made because they gave market participants the incentive to submit low-cost offers to the market in order to be selected by the model.

Figure 29
Average Regulation Schedules and Prices
January to December, 2004

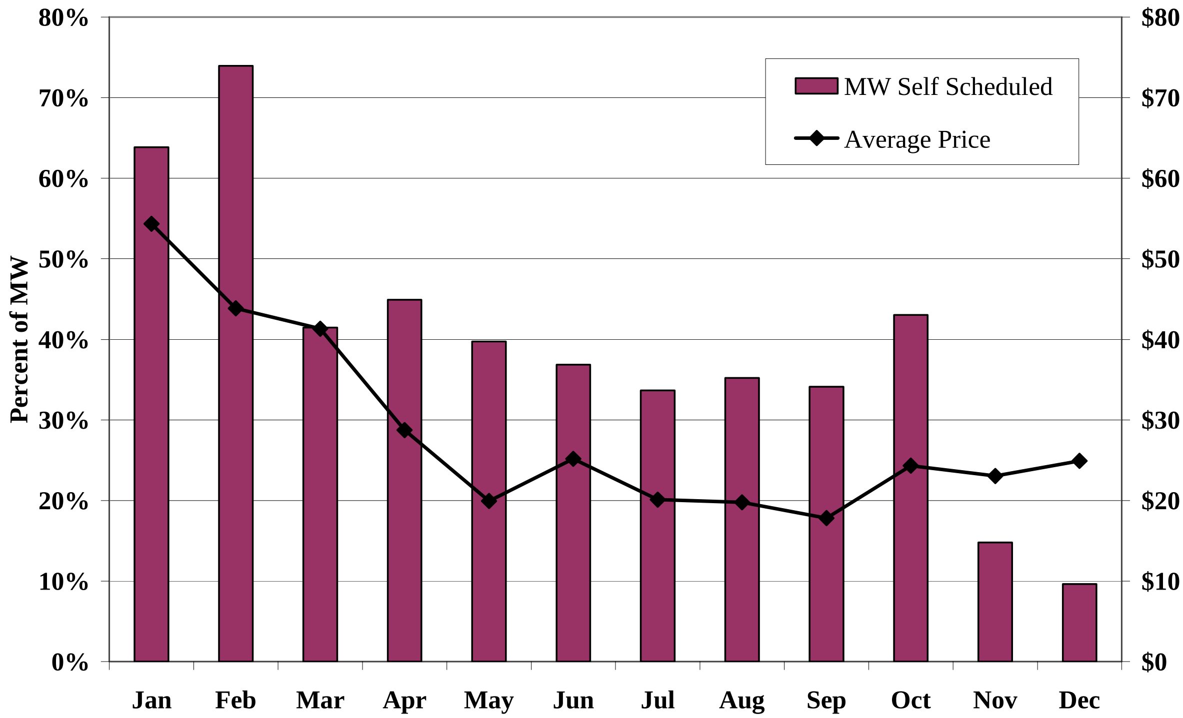


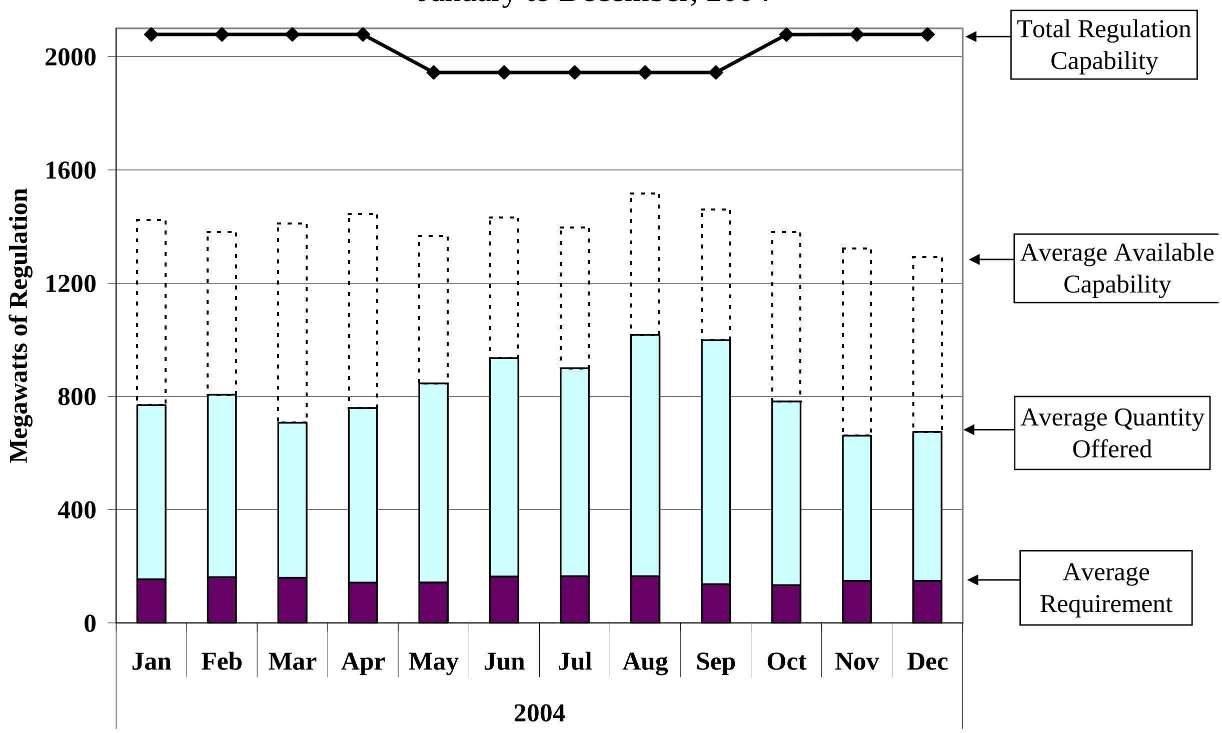
Figure 29 shows the quantity of regulation that is provided via the market versus the quantity that is self-scheduled in each month of 2004. The figure also shows the average RCP during this period. In January and February, more than 60 percent of regulation was provided by self schedules. However in March, once self-scheduling was prohibited after the close of the regulation market, the percentage of regulation provided by self schedules fell substantially. Furthermore, the average RCP dropped precipitously as a result of the rule change from $54 per MW in January to $20 per MW in May. For the remainder of 2004, the RCP fluctuated mostly in the $20 to $25 per MW range.

After the February rule change, there was an immediate drop in self scheduling, while the drop in price was gradual. We attribute the gradual reduction in price to a slow recognition by the participants of the incentives provided by the improved regulation market rules. Since the REGO model bases the merit order in the auction partly on an estimate of the unit's out-of-merit costs and/or opportunity costs, the rule change induced market participants to submit offers that minimize this estimate. This is explained in more detail, in the following sub-section.

B. Regulation Market Participation

Competition should be robust in New England's regulation market because in most hours the amount of regulation capability in New England far exceeds the amount required by the ISO. Figure 30 shows monthly averages of the total quantity of regulation-capable capacity, the available regulation-capable capacity, the regulation-capable capacity offered into the market, and the amount of regulation procured by the ISO.

Figure 30
Regulation Capability, Offers, and Requirements
January to December, 2004



On average, 30 percent of regulation-capable capacity is effectively unavailable to the market. Regulation-capable capacity can be unavailable in a given hour for at least two reasons: (a) the capacity is on a resource that has not been committed prior to the regulation auction, or (b) the capacity is held on a portion of a resource that was self-scheduled for energy.

Naturally, more regulation capacity tends to be available during the high-load portion of the day because more units have been committed and are on-line. This is partly mitigated by the fact that energy self schedules tend to increase during high-load hours and, therefore, the output ranges that are self-scheduled for energy are not available for regulation service. During the summer, an average of 65 percent of available regulation capacity was offered into the market (including self-schedules). This is better than during mild weather periods when the portion of available capability that was offered ranged as low as 50 percent.

During 2004, an average of five times more regulation was offered into the market than was actually procured by the ISO. This limits concerns about market power in the regulation market because demand can easily be supplied without the largest regulation supplier. However, supply may be 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.

Competition between the regulation market and energy market for capacity is reflected in the opportunity cost calculated from the unit's regulation offer. As noted above, the regulation market software, REGO, calculates opportunity costs by estimating the lost revenues or out-of-merit dispatch costs arising from the need to change a unit's energy dispatch level in order for it to provide regulation. Currently, REGO can only accept the entire quantity of regulation offered from a unit. Hence, the adjusted energy schedule, *set point*, must be located so that the resource can regulate up or down (on a symmetric basis) without going outside its operating range. This prevents REGO from selecting the lowest-cost regulation resources that would minimize the cost of satisfying the ISO's regulation requirements.

In the same way that the energy markets optimize the energy dispatch levels of all units, the regulation market should optimize the quantities of regulation taken from each resource. Even though the current regulation market does not optimize the quantity of regulation accepted from each unit, market participants have an incentive to optimize it themselves. They do this by offering a quantity that they predict will be competitive when the REGO model evaluates opportunity costs. To the extent market participants do not perfectly forecast conditions in the regulation market, they may make errors by offering too much or too little. Offering too much may result in not being selected in the auction. Conversely, if the market participant offers a quantity that is smaller than the optimal amount, the participant loses the opportunity to sell more into the market. Both types of forecast errors will lead to higher production costs and higher prices for regulation.

C. Conclusions

While participation in the market for regulation service far exceeds demand, the market outcomes were much more competitive after the ISO eliminated a market flaw in February 2004 that had significantly undermined competitive incentives. In October 2005, the ISO plans to implement Phase 1 of ASM which includes significant enhancements to the regulation market, including:

- Using a more comprehensive set of criteria for estimating opportunity costs;
- Running the regulation market at the beginning of each hour rather than the night before; and
- Incorporating mileage (i.e. actual fluctuation in output by the regulating unit) in the compensation for regulating units.

It would be beneficial to optimize the quantity of regulation accepted in the auction, but initially, it will not be operationally feasible to do so. Thus, we also recommend that the ISO examine the likely costs and benefits as well as impacts on real-time operations of making enhancements to the new regulation model that would allow it to optimize the quantity of regulation taken from each resource.

VII. Competitive Assessment

This section evaluates the competitive performance of the New England wholesale markets in 2004. This type of assessment is particularly important now that LMP markets are in operation. We identify geographic areas and market conditions that are most vulnerable to the exercise of market power. We use a methodology for measuring and analyzing potential withholding that was developed in prior assessments of the competitive performance in the New England markets.[15] In this section we address four main areas:

- 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.

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. Physically withholding occurs when the output of a resource is not offered to the market when it is economic, and is accomplished practically by "derating" a generating unit (i.e., reducing the unit's high operating limit).

While many suppliers can cause prices to go up by withholding, not every supplier can actually 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 only

[15] See "2002 Competitive Assessment of the Energy Market in New England" and "2001 Competitive Assessment of the Energy Market in New England"

pays off when the price impact overwhelms the opportunity cost of lost sales for the supplier. If a supplier is very large, it can withhold a substantial quantity but still sell enough to profit.

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 effectively reduce their incentive to raise price in the spot market. 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 that when demand levels are high, a given quantity of withholding has a larger price impact than when demand levels are lower. 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.

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 this is established, it is possible to assess conditions where one or more large suppliers could profitably raise price. This sub-section of the report examines structural aspects of supply and demand in the relevant market in order to focus the behavioral analysis 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 satisfy the demand in real time. This limits inter-temporal substitution between spot and forward 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) on-line and quick start capacity available for deployment in the real-time spot market, and (ii) off-line non-quick start 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 real-time spot market. The timeframe in which the market is most vulnerable to the exercise of market power is the real-time spot market when only on-line and quick-start capacity is available for deployment. Hence, we define the relevant product as energy produced in real time.

The second dimension of the market that must be defined is the geographic area in which suppliers compete to sell the relevant product, referred to as the relevant geographic market. In electricity markets, the relevant geographic market is generally defined by the transmission network constraints. When a transmission constraint is binding, there are limits on the extent to which power can flow between regions. In these situations, a supplier within the geographic area faces competition from fewer suppliers. There are a small number of geographic areas in New England that are generally recognized as being 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. These geographic markets are:

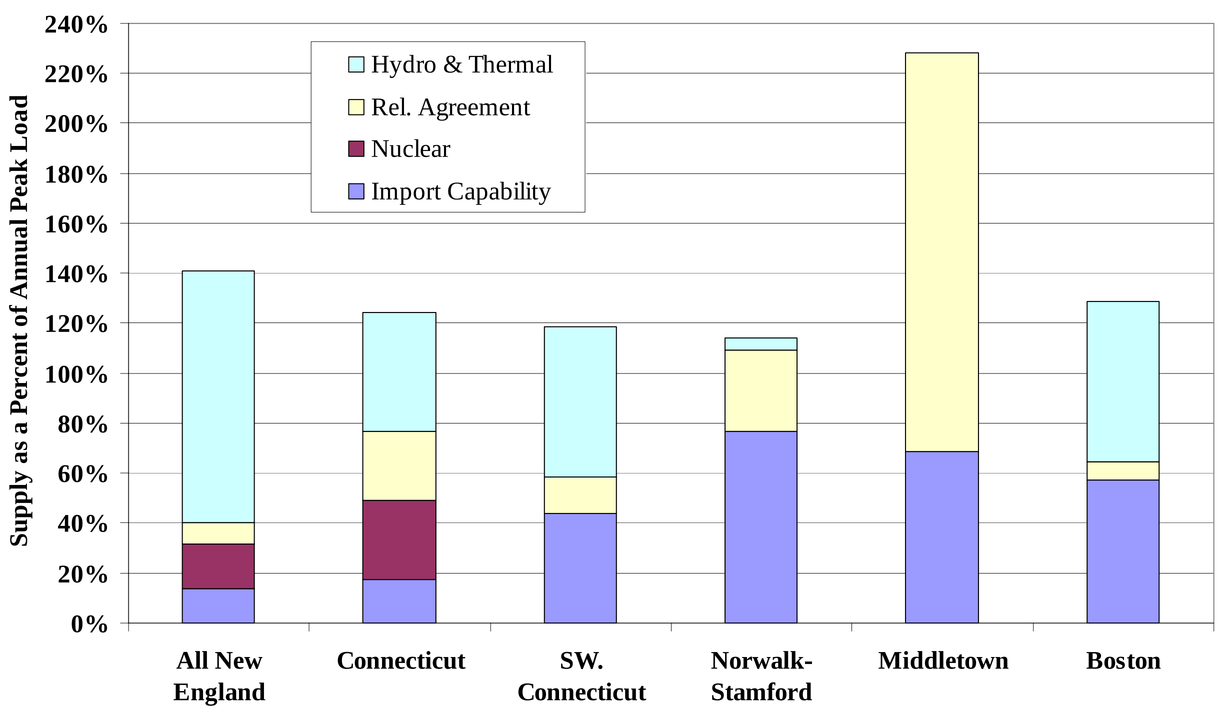
- All of New England;
- Connecticut;

- The southwest portion of Connecticut;
- Norwalk-Stamford which is contained in southwest Connecticut;
- The Middletown portion of Connecticut; and
- The areas of Boston and northeast Massachusetts.

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 transfer capability of the transmission grid. Figure 31 shows several categories of supply relative to the load in each of the six regions of interest.

Figure 31
Supply Resources versus Summer Peak Load in Each Region
2004



For each region under summer peak load conditions, Figure 31 shows import capability and three categories of installed summer capability: (i) nuclear units, (ii) units with reliability agreements, and (iii) all other generators. These resources are shown as a percentage of demand, although a substantial quantity of additional capacity is also necessary for reserves in New England. The figure shows that while the New England control area can import no more than 15 percent of its load, the five load pockets can serve larger shares of their load with imports into the constrained area. In particular, Norwalk-Stamford, Middletown, and Boston can rely on imports to serve more than 50 percent of their load under peak conditions. Alternatively, imports can supply only 42 percent of the load in Southwest Connecticut and less than 20 percent of the load in all of Connecticut.

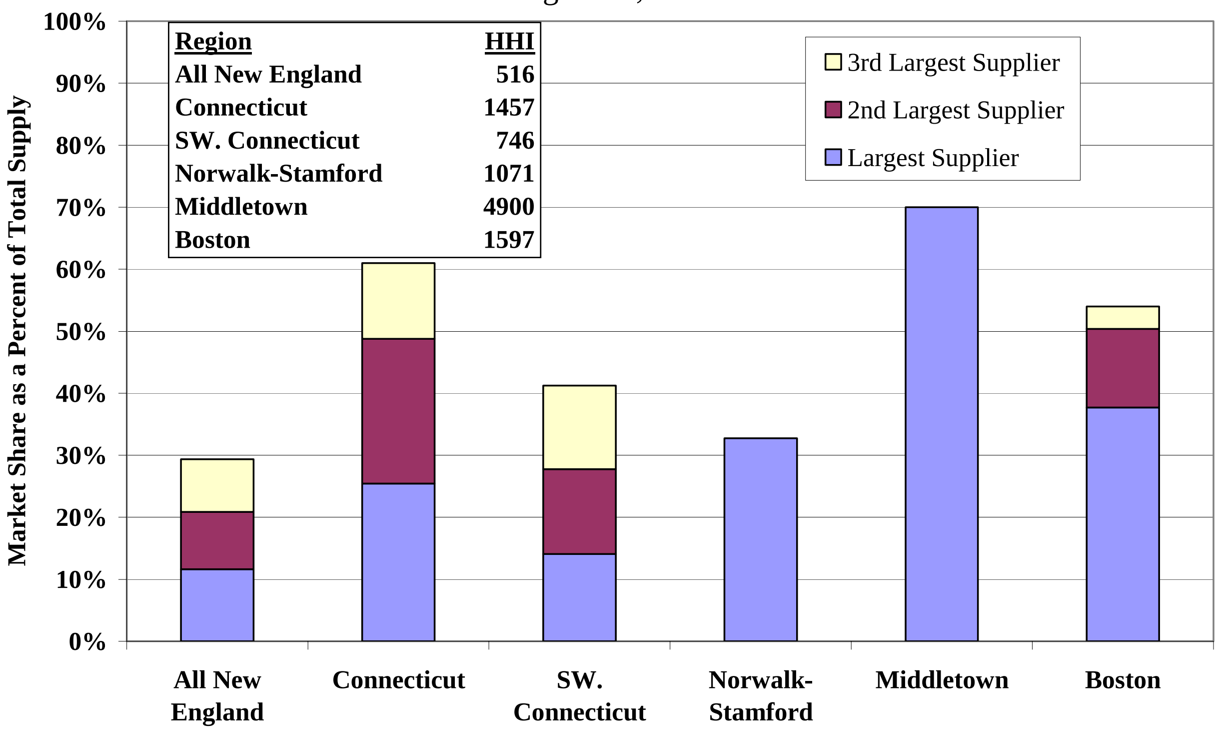
The figure also shows the margin between the total available supply, including both imports and native resources, and the peak load. Areas with lower margins may be more susceptible to withholding than other areas. For example, the total supply able to serve Norwalk-Stamford exceeds the annual peak load by only 14 percent. Thus, even a small reduction in supply or import capability to Norwalk-Stamford can cause a shortage under peak conditions.

Native generation is shown separately for nuclear capacity and capacity under reliability agreements because these resources are likely to pose fewer market power concerns. In order to exercise market power successfully in an electricity market, it is important to be able to withhold capacity only at times when it will be profitable because the lost revenue on withheld units can be very costly. Nuclear generators cannot be dispatched up and down in a way that would allow the owner of the unit to profitably withhold. Units with reliability agreements would not find it profitable because of the large fixed cost payments they receive under their contracts. Under these contracts, they agree to offer their units at short-run marginal costs which make it unlikely that they could be used to economically withhold.

While it is possible for a market participant to physically withhold from a unit that is under a reliability agreement, these units are subject to unusual scrutiny. If withholding by the owner of a reliability agreement were to be detected, the unit's reliability agreement status would be in jeopardy. This provides a substantial disincentive to withhold a unit with a reliability agreement. The areas within Connecticut rely heavily on nuclear and units under reliability agreements, particularly Norwalk-Stamford and Middletown. This significantly reduces the fraction of capacity that could be used to exercise market power.

The previous figure shows that the capacity margins can be as low as 14 percent in some areas, and 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 32 shows the market shares of the largest three suppliers coinciding with the annual peak load hour on August 30, 2004. The remainder of supply to each region comes from smaller suppliers as well as 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 and so treat the sum of all imports as having zero market share. This assumption will tend to understate the true level of concentration.

Figure 32
Installed Capacity Market Shares for Three Largest Suppliers
August 30, 2004



The figure indicates a substantial variation in market structure across regions. The largest suppliers have market shares ranging from 12 percent in all New England and 14 percent in Southwest Connecticut to 70 percent in Middletown. Likewise, there is variation in the number of suppliers that have significant market share. For instance, Norwalk-Stamford and Middletown have only one native supplier while the top three suppliers in Southwest Connecticut have virtually the same market share. While Norwalk-Stamford and Middletown have only a single supplier, the benefit from exercising market power is substantially mitigated by the fact that nearly all of the capacity in those areas is under reliability agreements.

Of the regions with more than one supplier, Boston is the area with the most significant single supplier, although Connecticut's largest two suppliers both have close to 25 percent market share. Based on market shares, all of New England and Southwest Connecticut appear to be of less concern since in each area the largest supplier accounts for 14 percent or less of total resources. The HHI figures suggest that only Middletown is highly concentrated, which raises potential market power concerns.[16] The HHI for Norwalk-Stamford is very low, which is counter-intuitive since the only supplier possesses 37 percent of the resources and the capacity margin is less than 20 percent. Of the four areas not dominated by units under reliability agreement, Connecticut and Boston have the highest HHI statistics with 1457 and 1597, respectively. While HHI statistics can be instructive, in the next sub-section we introduce a pivotal supplier analysis which is more appropriate for evaluating market power in electricity markets.

3. Pivotal Supplier Analysis

While HHI statistics can provide reliable competitive inferences for many types of products, this is not generally the case in spot electricity markets.[17],[18] The HHI's usefulness is limited by the

16 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.

17 It is true that the DOJ and FTC evaluate the *change* in HHI as part of its merger analysis. However, this is only a preliminary analysis that would typically be followed by a more rigorous simulation of the likely price effects

fact that it reflects only the supply-side, ignoring demand-side factors that affect the competitiveness of the market. The most important demand-side factor is the level of demand. Since electricity cannot be stored economically, production must match demand on a real-time basis. When demand rises, an increasing quantity of generating capacity is utilized to satisfy the demand, leaving less capacity 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 energy market prices to arbitrarily 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 they 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.[19] This section of the report assesses which suppliers were pivotal in the real-time energy market 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 be deemed to have market power. For a supplier to have the ability to substantially

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.

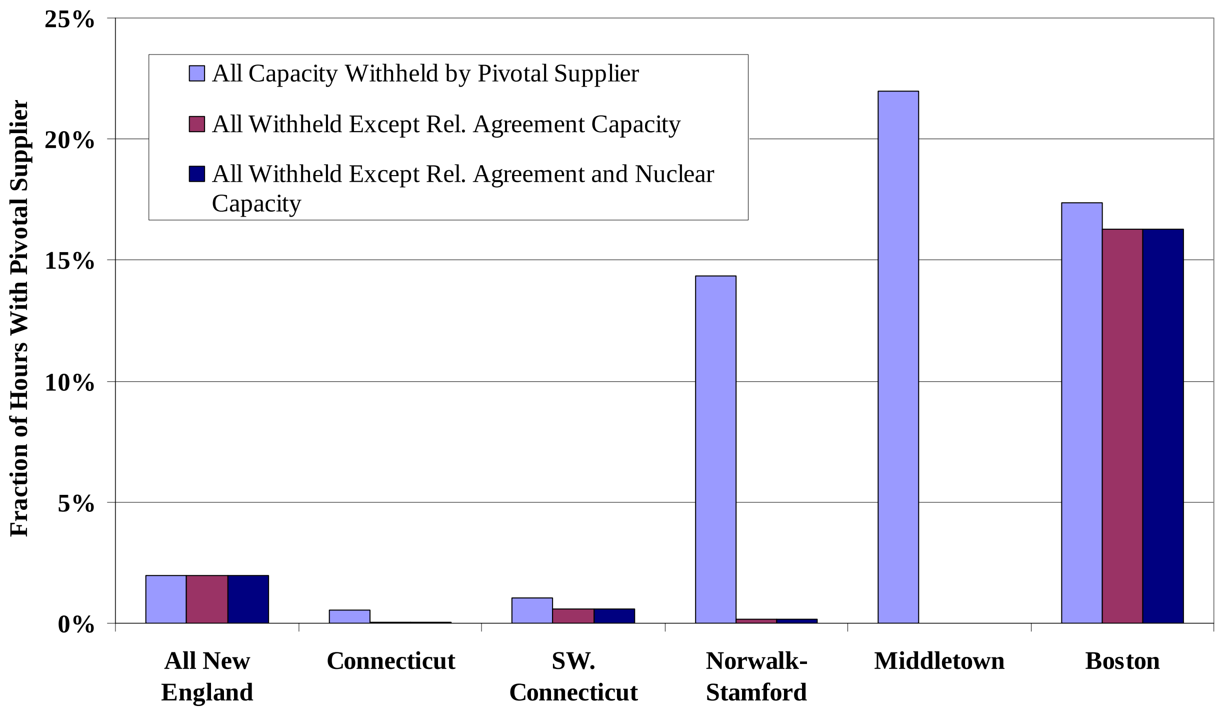
18 For example, see Severin Borenstein, James B. Bushnell, and Christopher R. Knittel, "Market Power in Electricity Markets: Beyond Concentration Measures," *Energy Journal* 20(4), 1999, pp. 65-88.

19 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.

raise the balancing 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 assess which areas have the most frequent conditions where market power might be a concern, Figure 33 shows the portion of hours where at least one supplier was pivotal in each region during 2004. The figure also shows the impact of assuming that nuclear units and units under reliability agreements are never withheld by a large supplier.

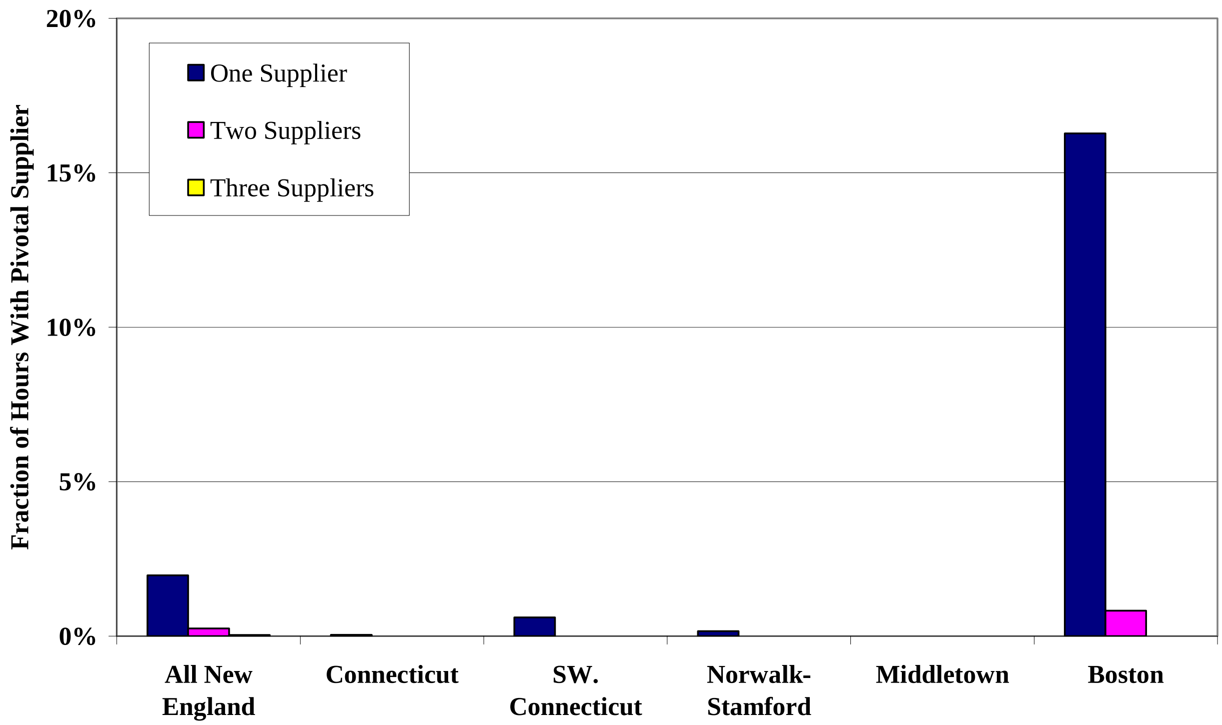
Figure 33
Frequency of One or More Pivotal Suppliers Based on Type of Withheld Capacity
All Hours – 2004



The figure indicates that if all capacity is potentially withheld, Norwalk-Stamford, Middletown, and Boston will have at least one pivotal supplier in a large portion of hours. However, All New England, Connecticut, and Southwest Connecticut have a pivotal supplier in no more than 2 percent of all hours. The frequency of a pivotal supplier decreases substantially for Connecticut and the three load pockets inside of it due to the large amount of capacity under reliability agreement there. Moreover, Connecticut also has a significant block of nuclear capacity. After accounting for units under reliability agreement and nuclear units, the area that has the most hours with a pivotal supplier is Southwest Connecticut, but there is a pivotal supplier there in only 0.6 percent of hours.

Since there are no nuclear units and only one unit under reliability agreement in Boston, the largest supplier there is still pivotal in 16 percent of hours. Likewise, the largest suppliers in All of New England did not own any nuclear or RMR capacity, so that a supplier is pivotal there in 2 percent of hours. The pivotal supplier summary indicates very significant market power potential in Boston while all of New England is likely to be a concern under limited market conditions. The market shares in Figure 32 indicate that there are areas with several dominant suppliers, suggesting that during certain periods, several suppliers might be pivotal at one time. Figure 34 shows the number of pivotal suppliers during hours where one or more supplier is pivotal in each region.

Figure 34
Frequency of One or More Pivotal Suppliers
No Withholding of Nuclear and Rel. Agreement Units -- All Hours in 2004



The frequency of one or more pivotal suppliers is the same as the previous figure (Figure 33). But the current this figure also shows the frequency of two and three suppliers being pivotal in a single hour. It is very uncommon for more than one supplier to be pivotal at the same time. And there were virtually no periods with three or more pivotal suppliers.

Since the relevant market includes capacity able to serve demand in the real-time market, it excludes non-quick-start capacity that is off-line. Thus, there will be some variation in the market shares on a daily basis due to differences in the unit commitments. There was virtually no variation in the order of suppliers from largest to smallest during 2004. The only exception was in a case where a large amount of capacity was transferred from one entity to another. Therefore, each area had a single supplier that was most capable of exercising market power. Accordingly, the next sub-section will compare the behavior of the largest single supplier with that of other suppliers under various market conditions.

As stated 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 withholding is most likely to be profitable, Figure 35 shows the fraction of hours where a supplier is pivotal according to various load levels. The left most bar in each load range shows the fraction of hours with a pivotal supplier in Boston. The next bar represents All New England, the next is Southwest Connecticut, and the last bar represents Norwalk-Stamford. These are displayed according to the frequency with which a pivotal supplier emerges with Boston on the left having the highest frequency and the areas with lesser frequency to the right. Middletown and Connecticut are not shown because their frequencies of pivotal suppliers were less than 0.1 percent of hours during 2004.

Figure 35
Frequency of One or More Pivotal Suppliers by Load Level
No Withholding of Nuclear and Rel. Agreement Units – All Hours in 2004

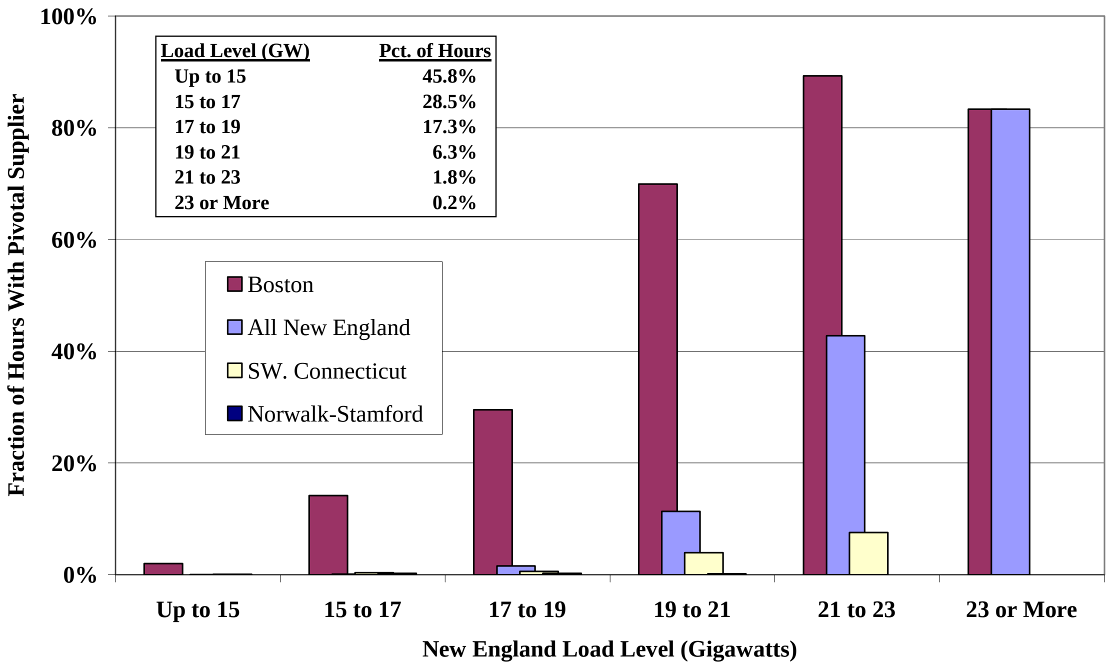


Figure 35 indicates that the largest supplier in Boston is pivotal in the majority of hours when load is greater than 19 GW in New England. Load was above 19 GW in 8.3 percent of the hours during 2004. In all of New England, the largest supplier was pivotal in 83 percent of the hours when load exceeded 23 GW, although this includes only 0.2 percent of the hours. Due to the pattern of commitment under various load conditions, the largest supplier in Southwest Connecticut was not pivotal during the highest load conditions, but was pivotal when aggregate load in New England was between 19 GW and 23 GW. During these hours, a supplier was pivotal in Southwest Connecticut less than 10 percent of the time. Similarly, Norwalk-Stamford's supplier was pivotal during a very small share of relatively low load hours.

Based on the pivotal supplier analysis in this sub-section, market power is most likely to be a concern in Boston when New England load rises above 19 GW, while All of New England is a concern when load is above 21 GW and becomes an increasing concern when load is above 23 GW. 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 our analysis. This is necessary to determine the quantity of output that is potentially economically withheld. The ISO's Internal Market Monitoring Unit calculates generator cost reference levels pursuant to Attachment A of Section III of the ISO's Tariff. The Internal Market Monitoring Unit has provided us with cost reference levels, which can be used as a competitive benchmark for our analysis of economic withholding.[20]

20 In the case of one unit, variable cost estimates were used instead of the reference level because the reference level substantially understated the unit's marginal costs.

1. Measuring Economic Withholding

We measure economic withholding by estimating an output gap for units that fail conduct tests on their start-up, no-load, and incremental energy offer parameters.[21] 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 look at 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 stage we examine whether the unit would have been economic *for commitment* on that day if it had offered its true 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 dispatch. Finally, we determine the economic level of incremental output in hours when the unit was economic to run. 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

21 For incremental energy offers, a unit fails if the offer exceeds (i) the reference level plus $25/MWh or (ii) 150 percent of the reference level. For an offer at the minimum generation level, a unit fails if the offer exceeds 150 percent of the reference level. For the start-up parameter, a unit fails if the offer exceeds 150 percent of the reference level.

on day-ahead market outcomes for non-quick start units, and for quick start units this assessment is based on real-time market outcomes.

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. Hence the output gap formula used for this report 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, units that are dispatched down due to transmission constraints or subject to ramp limitations are excluded from the output gap.

We make a further key adjustment to the output gap to better reflect potential economic withholding. Portions of resources that are offered above marginal costs due to a forward reserve market obligation are not included in the output gap.

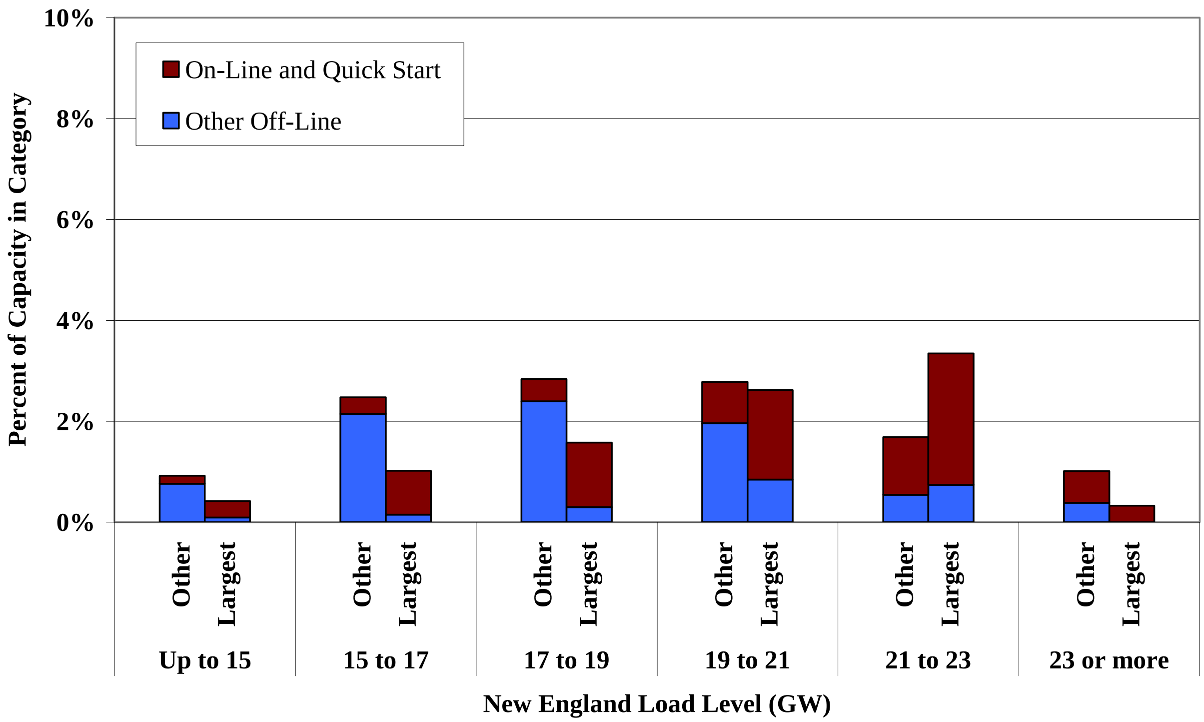
In this section we evaluate the output gap results relative to various market conditions and participant characteristics. The objective is to determine whether the output gap increases when those factors prevail that can create the ability and incentive for a pivotal supplier to exercise market power. 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 sub-section, 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 for the following six geographic markets:

- All of New England;
- Connecticut;
- The southwest portion of Connecticut;
- Norwalk-Stamford which is contained in southwest Connecticut;
- The Middletown portion of Connecticut; and
- The areas of Boston and north-east Massachusetts.

2. Output Gap in NEMA/Boston

Figure 36 shows output gap results for the NEMA/Boston area for various load levels. Based on the pivotal supplier analysis in the previous sub-section, the dominant supplier can expect to be pivotal in most hours when load exceeds 19 GW. Output gap statistics are shown for the dominant supplier compared with all other suppliers in the area.

Figure 36
Average Output Gap by Load Level and Type of Supplier
NEMA/Boston – 2004

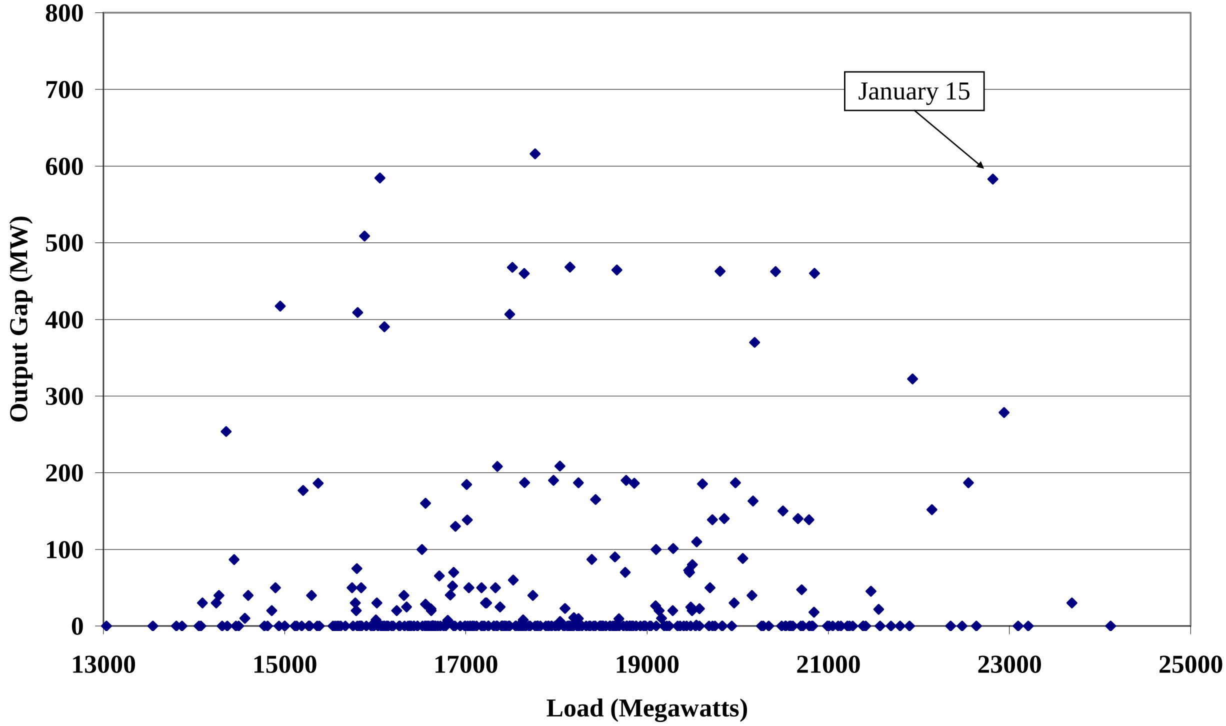


The figure shows that the output gap for the largest supplier and other suppliers in Boston are very low, which do not indicate economic withholding concerns. In neither case does the output gap rise substantially as load rises. However, the output gap related only to online and quick-start resources is larger for the largest supplier than the others and does tend to increase as load increases with the exception of moving to the highest load level. As outlined in the discussion of the relevant market in subsection B above, the more narrow definition of the relevant market consists of capacity that is on-line in real-time or off-line capacity that is able start quickly. Capacity in the other output gap category is outside this narrower market definition, but is included in the wider set of resources that compete in the day-ahead market.

The average output gap quantity of the largest supplier in Boston is approximately 100 MW when load is between 19 GW and 23 GW. In 98 percent of the more than 1,300 hours when the largest supplier was pivotal during 2004, it would have needed to withhold more than 1,000 MW of on-line and quick start resources to significantly increase prices in Boston. Thus, the average output gap in Boston would not have been sufficient to have a substantial market impact.

To assess whether the output gap was large in even a small number of hours, Figure 37 shows the output gap associated with on-line and quick start units of the largest supplier on an hourly basis for the peak load hour of each day during 2004.

Figure 37
Hourly Output Gap of Largest Supplier in NEMA/Boston by Load Level
On-Line and Quick Start Units
Daily Peak Load Hour – 2004



The largest supplier in NEMA/Boston had an output gap of more than 500 MW on just one day when load exceeded 19 GW. This occurred on January 15th when there was considerable uncertainty in the natural gas market. Otherwise, there were three days when the output gap was 470 MW while load ranged between 20 GW and 21 GW. The majority of days showing an output gap in excess of 300 MW occurred during moderate demand periods when load ranged between 15 GW and 19 GW.

The pivotal supplier analysis from the previous sub-section indicates that the probability of the largest supplier being pivotal was highly correlated with load, and that the largest supplier was usually pivotal when load rose above 19 GW. However, the figure shows that the output gap occurred during a wide range of load conditions. If the output gap were indicative of economic withholding, we would expect to see it concentrated at higher load levels. However, the fact that the pattern of higher output gap levels occurred over a wide range of load levels is more consistent with pay-as-bid incentives. Based on the results in this subsection, particularly the quantities and patterns of output gap identified for the largest supplier, we do not find a competitive concern regarding the conduct of the supplier. In addition, real-time prices in Boston in 2004 did not reflect significant congestion.

However, it is important to recognize that the output gap measure does not address economic withholding related to supplemental commitments needed to meet local reliability requirements. Section IV of this report describes that large amounts of capacity are typically committed supplementally for local reliability reasons in the Boston area. Because most generators committed for local reliability are paid their offer prices rather than the LMP, they face significantly different offer incentives (i.e., “pay-as-bid” incentives). Even in perfectly competitive pay-as-bid markets, firms with no market power will rationally raise their offer above marginal costs since they do not receive a market-clearing price. Although offers by competitive suppliers in pay-as-bid markets will rise above marginal costs, one cannot conclude that all increases in offer prices by suppliers that face pay-as-bid incentives are justified.

To the extent that suppliers hold resources needed to meet local reliability requirements and do not face competition to meet those requirements, they may have local market power that can be exercised by inflating the guarantee payments the ISO must make to utilize the resources. Although they face pay-as-bid incentives, if such suppliers increase their offer prices by more than they would if they faced competition from other suppliers, then one may conclude that they are exercising local market power.

Based on our review of the commitment patterns, offers, and uplift payments made in the Boston area, we conclude that in late 2004 a significant exercise of local market power began that has continued into 2005. This conduct has not been effectively mitigated under the existing mitigation measures for economic withholding due primarily to the inflated reference prices used for the resources in question.[22] However, the ISO filed for a change in the reference price

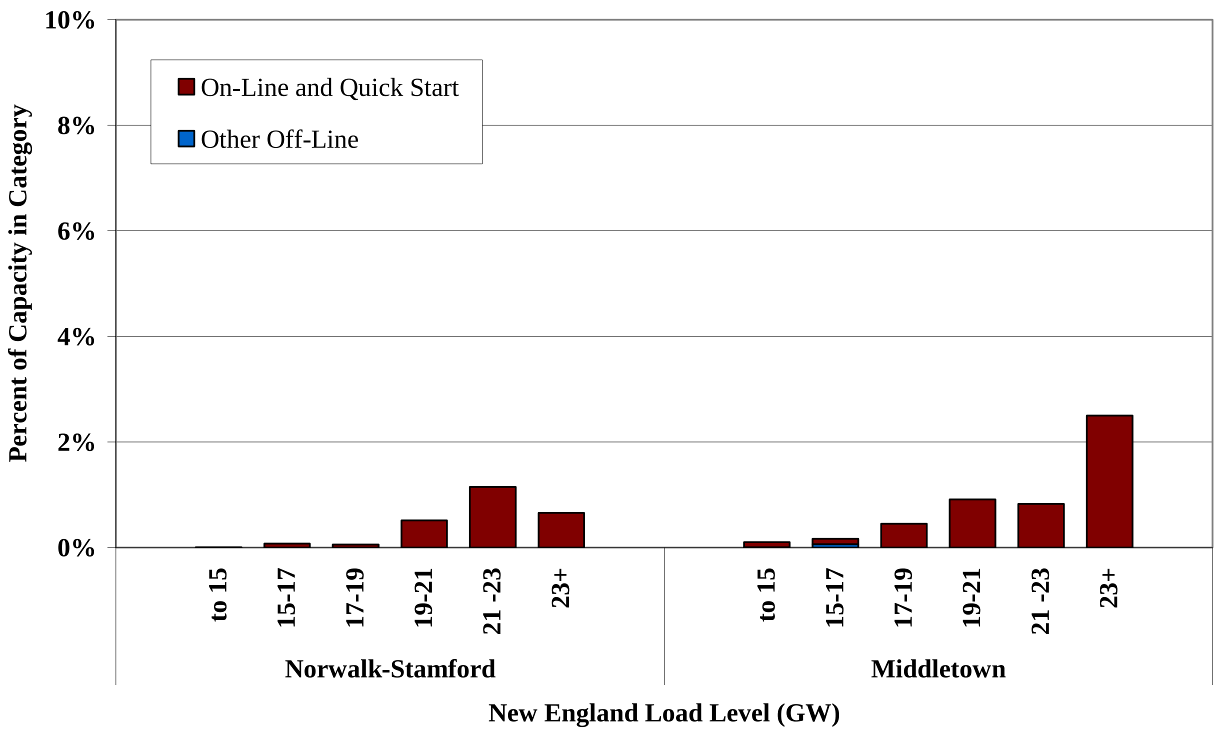
[22] Even with reference levels that perfectly reflect a unit’s marginal costs, the mitigation measures utilize conduct thresholds that can allow a supplier uniquely positioned to satisfy local reliability requirements to extract significant economic rent from the market.

calculation provisions that should substantially improve the effectiveness of the mitigation measures in addressing this conduct.[23] In addition, our recommendation to remove the flexibility for units in constrained areas to self-schedule after the RAA process will further mitigate the exercise of this form of local market power.

3. Output Gap in Connecticut

There are four areas within Connecticut that are examined in this assessment. Figure 38 and Figure 39 summarize output gap results for these areas by load level. Based on the pivotal supplier analysis in the previous sub-section, it is unlikely that any supplier in Connecticut and/or Middletown would profit from withholding. However, the pivotal supplier analysis did indicate a potential for market power in Southwest Connecticut and Norwalk-Stamford.

Figure 38
Average Output Gap by Load Level
Norwalk-Stamford and Middletown – 2004



[23] The Commission accepted the proposed reference price change in an order issued on May 6, 2005 in Docket No. ER05-767. *ISO New England Inc. and New England Power Pool*, 111 FERC ¶ 61,184 (2005).

Figure 38 shows the average output gap quantities for Norwalk-Stamford and Middletown during 2004. In both areas, there is only a single supplier, which would raise serious concerns under ordinary circumstances. However, virtually all of the capacity is under reliability agreements, which significantly mitigates the incentive for this supplier to strategically withhold. The output gap quantities for this supplier generally average less than 1 percent of the supplier's total capacity in these areas. Although the output gap rises above 2 percent during hours where load is greater than 23 GW, there is no significant evidence of strategic economic withholding in Norwalk-Stamford and/or Middletown.

Figure 39 summarizes output gap results for the Connecticut and Southwest Connecticut areas by load level.

Figure 39
Average Output Gap by Load Level and Type of Supplier
Connecticut and Southwest Connecticut – 2004

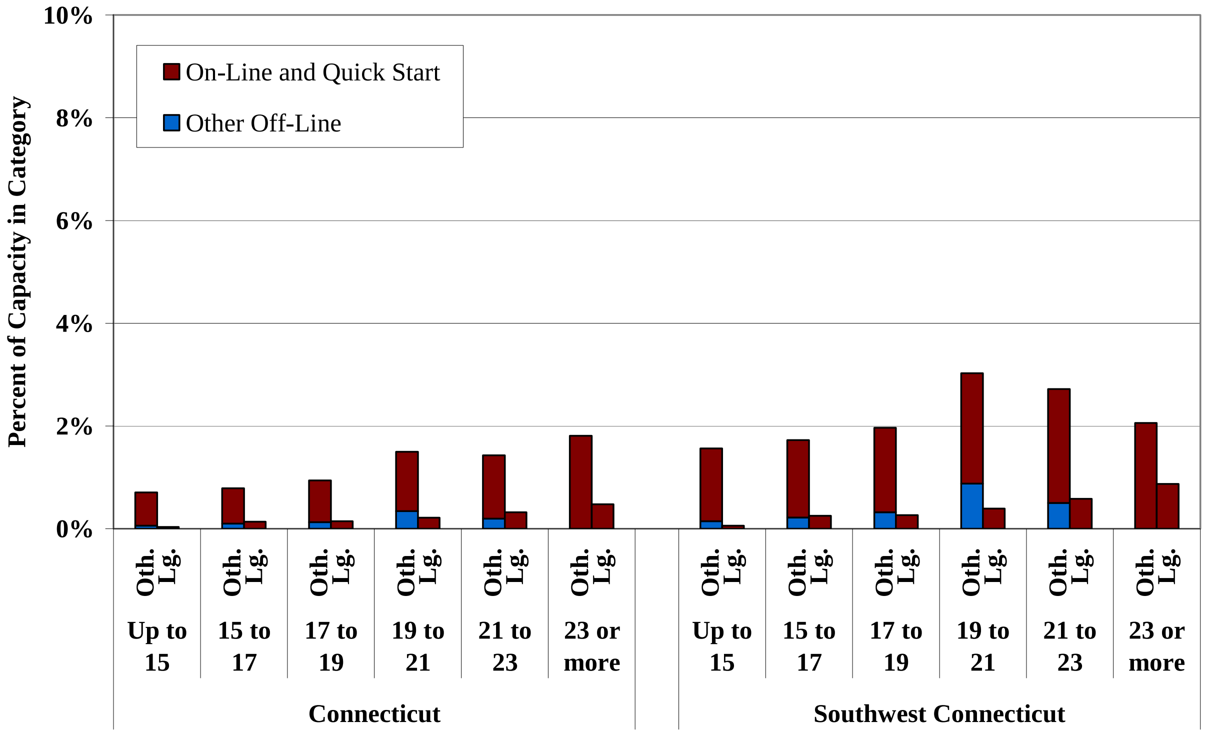
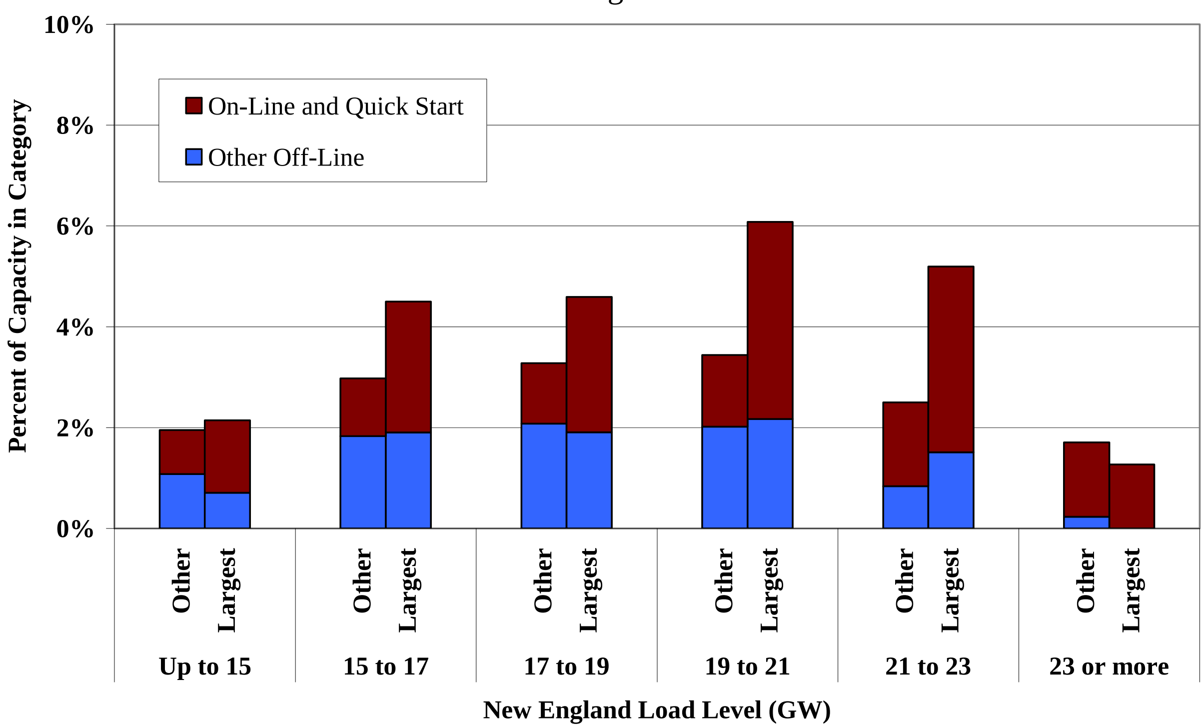


Figure 39 displays average output gap quantities for Connecticut and Southwest Connecticut. These are shown for the largest supplier of non-reliability agreement capacity and non-nuclear capacity in each area compared with other suppliers. These areas are grouped together because the largest supplier of non-reliability agreement capacity and non-nuclear capacity is the same entity in both areas. The average output gap for the largest supplier was less than 1 percent under all load conditions, while other suppliers showed somewhat larger quantities of output gap. The import-constrained areas within Connecticut do not show evidence of strategic withholding behavior on the part of the largest suppliers there. Due to tight supply conditions and concentrated ownership, the ISO should continue to monitor the competitive conditions within Connecticut on a periodic basis. This will be particularly important after the introduction of locational capacity markets when the reliability agreements are expected to terminate.

4. Output Gap in All New England

Figure 40 summarizes output gap results for all of New England by load level, comparing the supplier that is sometimes pivotal with all other suppliers.

Figure 40
Average Output Gap by Load Level and Type of Supplier
All of New England – 2004

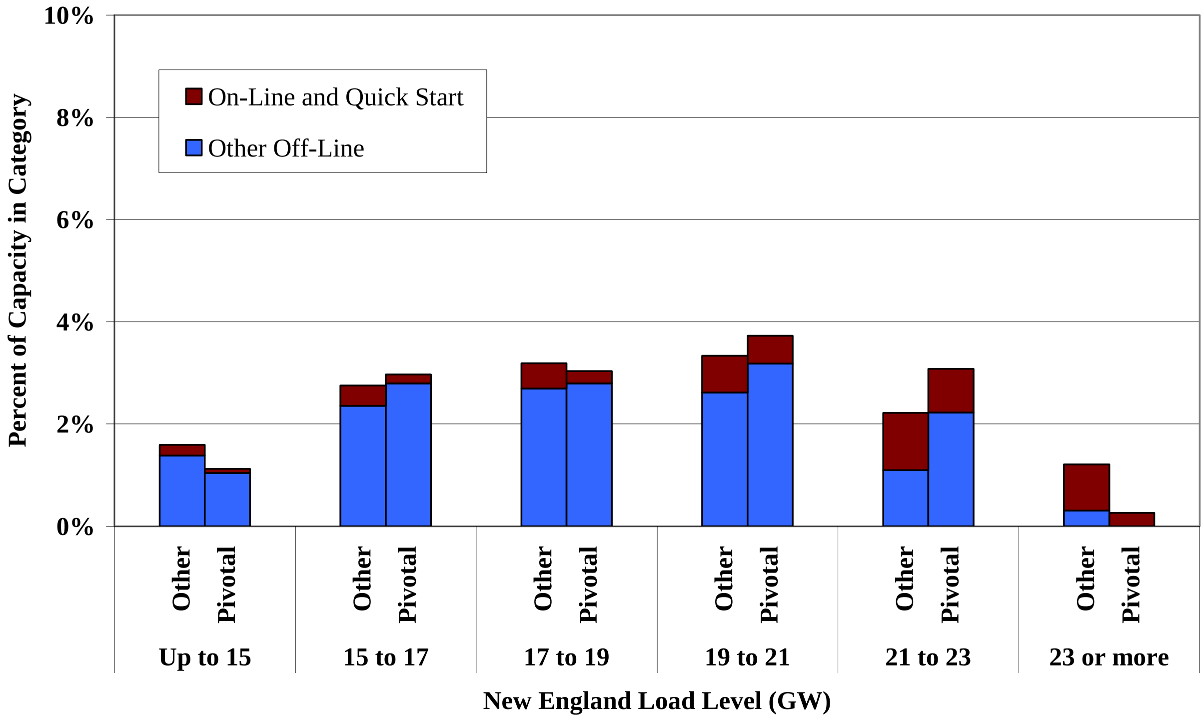


The figure shows that the largest supplier's output gap is largest when load ranges from 19 GW to 23 GW, but declines significantly when load rises above 23 GW. Based on the pivotal supplier analysis above, the largest supplier in New England was pivotal during most of the hours when load was above 23 GW and approximately 42 percent of the hours where load was between 21 GW and 23 GW. The output gap of other suppliers is substantially lower than for the largest supplier in all but the highest load categories. Furthermore, a substantial share of the largest supplier's output gap is from on-line and quick start units. While the output gap quantities shown above include a relatively small share of the total capacity in the market, Figure 40 does raise some competitive concerns.

During 2004, the largest supplier in New England owned a substantial amount of flexible hydro capacity. Typically, hydro generators incur little or no variable production costs but have energy limitations that restrict their output over a certain amount of time. The owner of a hydro resource will raise its offer price in order to run during the most profitable periods. Thus, the true marginal costs of a hydro generator fluctuate based on water storage levels, water conditions, environmental restrictions, and expectations of prices during subsequent hours.

It is difficult to quantify the opportunity costs faced by a hydro supplier, so there is a tendency for hydro resources to exhibit larger output gap quantities than other technology types. Since 25 percent of the largest supplier's portfolio is from flexible hydro generators, a large portion of its output gap is likely due to the difficulty of estimating their marginal costs. To characterize the effect of hydro units on output gap results, Figure 41 shows the output gap for just fossil generators.

Figure 41
Average Output Gap by Load Level and Type of Supplier
Fossil Units in All of New England – 2004

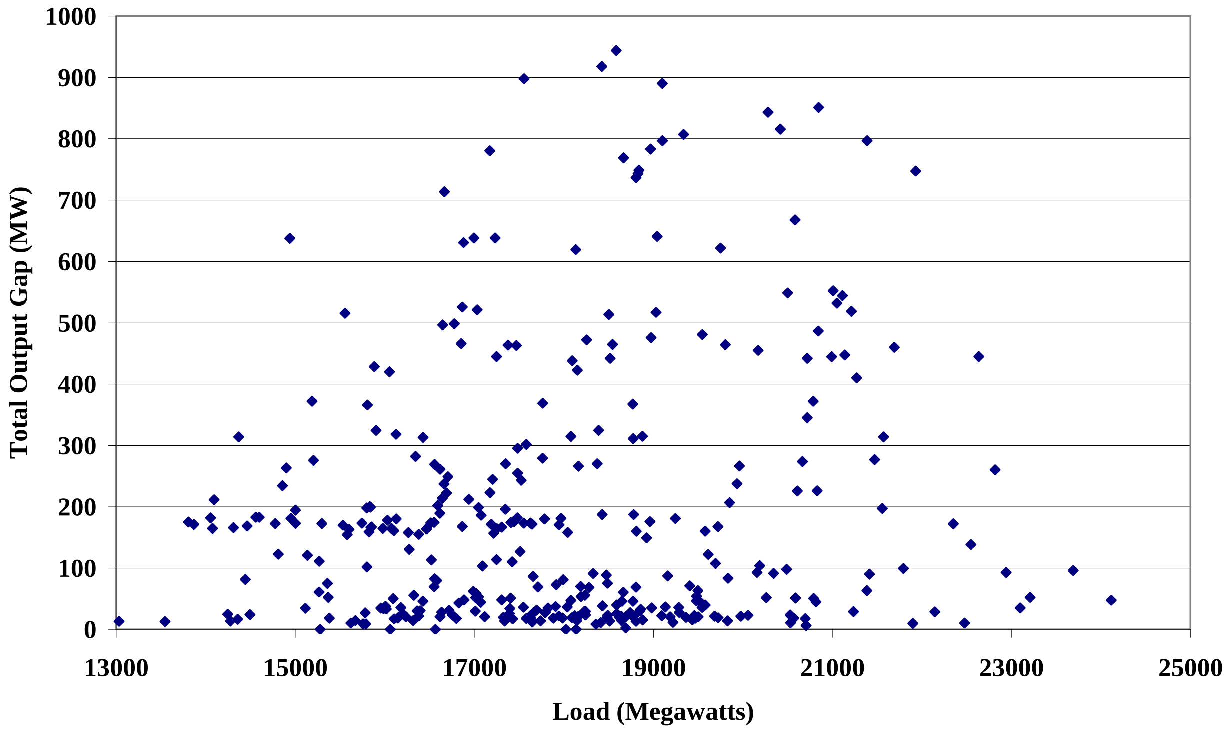


The analysis still shows that the large supplier's output gap is largest when load ranges from 19 GW to 21 GW, and it declines significantly when load rises above 23 GW. However, the output gap quantities associated with on-line and quick start units for the large supplier shrink from approximately 4 percent of its portfolio (see Figure 40) to less than 1 percent of its fossil portfolio. This suggests that the higher output gap quantities for the large supplier's on-line and quick-start resources might be due to the difficulty measuring the marginal costs for the hydro resources in its portfolio. While this significantly mitigates concerns that the largest supplier might have been withholding at high load levels, the following analysis examines whether the output gap was large enough to raise significant concerns in a subset of hours.

Figure 41 shows an increase in the output gap of the largest supplier at high load levels, but the output gap quantity is relatively small on average. Under the current market rules, shortage pricing is invoked when New England does not have sufficient market-wide reserves to cover its largest contingency plus half of its second largest contingency (approximately 1700 MW).

During 2004, there were no occasions when market-wide reserves dropped below 3200 MW, which implies that the largest supplier would need to withhold at least 1500 MW to drive the market into shortage pricing conditions. The average output gap for the largest supplier was under 250 MW in hours with 21 GW to 23 GW of load, which is unlikely to have a significant impact on prices. To assess whether the output gap was large in even a small number of hours, Figure 42 shows the total output gap from all on-line and off-line resources for the largest supplier on an hourly basis for the peak load hour of each day during 2004.

Figure 42
Hourly Output Gap for Largest Supplier in New England by Load Level
Daily Peak Load Hours – 2004



The figure shows the highest output gap of 944 MW occurred during a period when the peak load for the day was less than 19 GW. The pivotal supplier analysis in the previous sub-section suggests that the largest supplier could only be pivotal in hours when the load exceeded 21 GW, and only during a majority of hours when load exceeded 23 GW. The figure further indicates that the highest output gap was 800 MW when load exceeded 21 GW, and approximately 100 MW when load exceeded 23 GW. These results for individual hours are not large enough to raise significant competitive concerns. Overall, the analyses in this sub-section provide little evidence of systematic economic withholding by the largest supplier in New England.

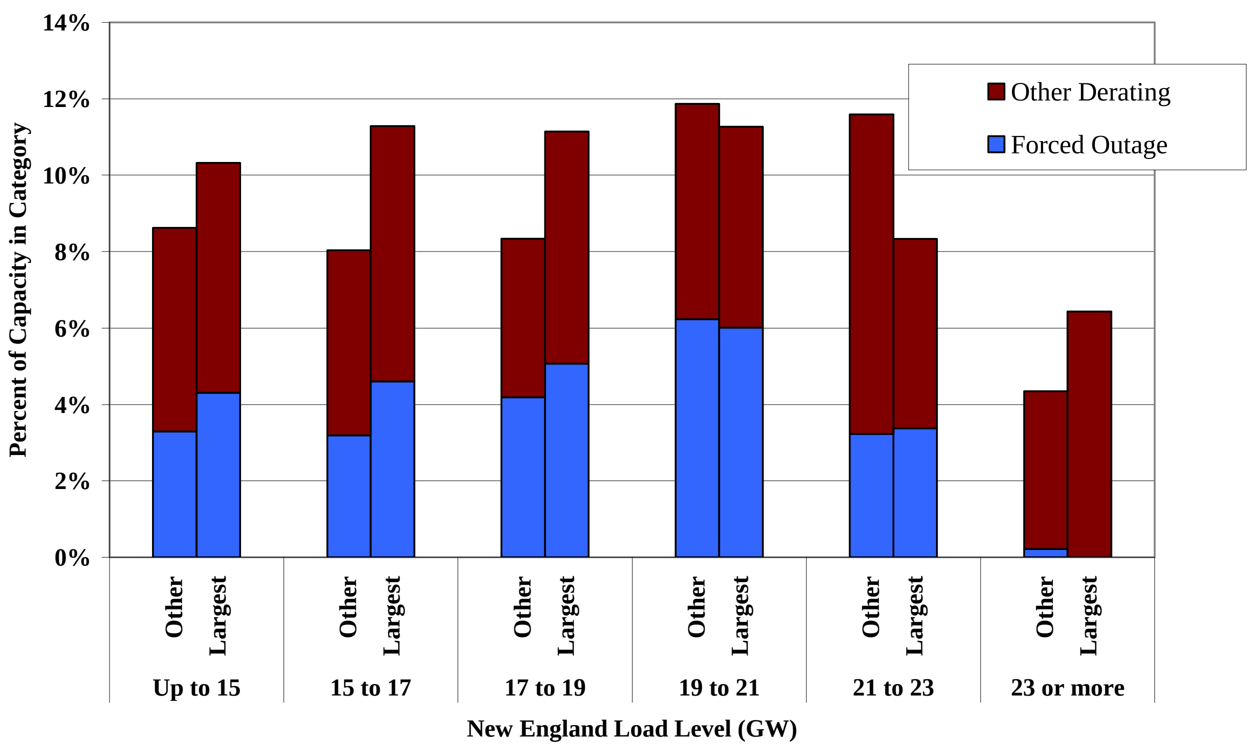
D. Physical Withholding

This sub-section of the report examines forced outages and other non-planned deratings to assess whether they have occurred in a manner that is consistent with the exercise of market power. For our analysis, we use each of the six areas examined in the pivotal supplier analysis above.

1. Potential Physical Withholding in NEMA/Boston

Figure 43 shows forced outages and other deratings in the NEMA/Boston area for various load levels. Outage and derating statistics are shown for the dominant supplier compared with all other suppliers in the area. Based on the pivotal supplier analysis in the previous sub-section, the dominant supplier can be expected to be pivotal in most hours when load exceeds 19 GW.

Figure 43
Forced Outages and Other Deratings by Load Level and Type of Supplier
NEMA/Boston - 2004



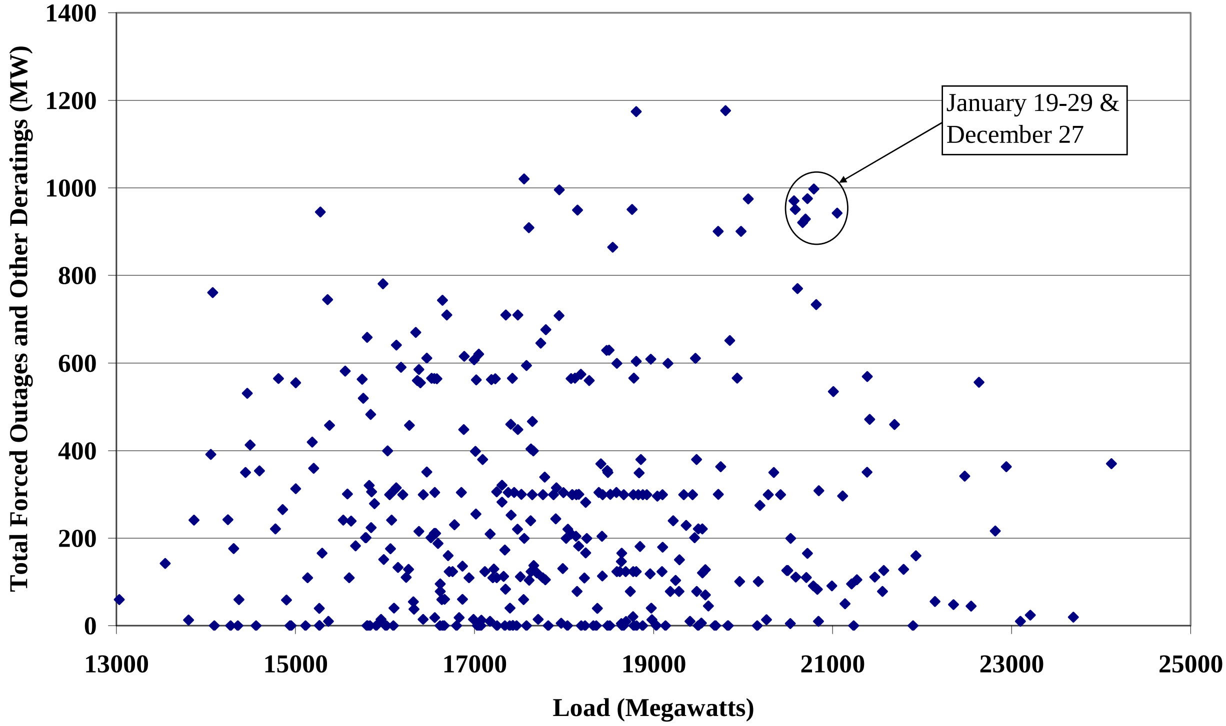
The "other deratings" shown in the figure include reductions in the hourly capability of a unit from its maximum seasonal capability that are not logged as forced outages or planned outages. These deratings are frequently the result of ambient temperatures or other factors that affect the maximum capability of a unit.

The figure shows the largest supplier's physical deratings as a percentage of its portfolio, which generally range between 10 percent and 12 percent. However, deratings and outages decline to 8 percent when load is between 21 GW and 23 GW, and fall to 6 percent when load rises above 23 GW. The average physical deratings of other suppliers is lower when load is below 19 GW, rises above the largest supplier's share when load is between 19 GW and 23 GW, and falls dramatically at the highest load level. Forced outages account for nearly half of total deratings under most circumstances, except when load rises above 23 GW, when forced outages accounted for almost none of the deratings.

Overall, Figure 43 suggests that the pattern of deratings and outages is consistent with a competitive market for at least two reasons. First, the large supplier shows levels of outages and deratings that are not significantly higher than for other suppliers. Second, both the large supplier and other suppliers show a general decline in the level of outages and deratings as load increases 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 market participants have tried to keep capacity available during periods of high load when it is most valuable, but also when the market is more susceptible to the exercise of market power.

To make a further inquiry, our next analysis examines whether large amounts of capacity could have been physically withheld during a small number of hours. According to the pivotal supplier analysis in the previous section, the largest supplier was pivotal in more than 1300 hours during 2004. However, in 98 percent of these hours, it would have needed to withhold more than 1,000 MW of on-line and/or off-line quick start capacity to drive the market into a shortage. Thus, the average quantity of physical deratings and forced outages in Boston would not have been sufficient to have a substantial market impact. To assess whether the physical withholding could have been large in even a small number of hours, Figure 44 shows the forced outages and other non-planned deratings for the large supplier's portfolio on an hourly basis for the peak load hour of each day during 2004.

Figure 44
Hourly Non-Planned Outages and Deratings for Largest Supplier in NEMA/Boston
Daily Peak Load Hours – 2004



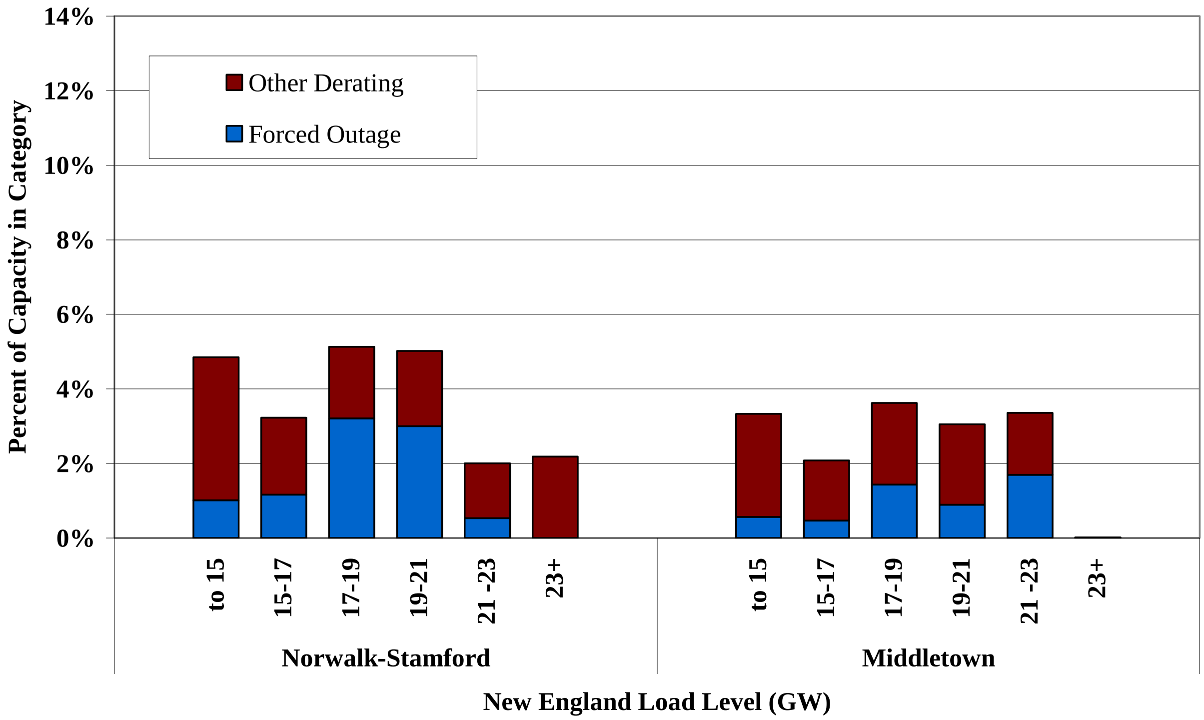
The figure shows that the highest derating of nearly 1,200 MW occurred when load was just below 20 GW. Also, there were seven days where deratings ranged from 900 MW to 1,000 MW when load was close to 21 GW. These are highlighted by a circle in the figure. At higher load levels, the amount of deratings decreases considerably.

The seven days that raise the most significant competitive concerns occurred during extreme cold weather rather than during the summer. A large majority of the physical deratings on these days were due to one or two base load units being forced out. These forced outages as well as reasons for them were logged with the ISO. While it is outside the scope of this report to assess the validity of a particular forced outage, the hourly deratings shown in the figure do not raise significant concerns.

2. Potential Physical Withholding in Connecticut

There are four import-constrained areas within Connecticut that are under review in this report. Figure 45 and Figure 46 summarize forced outages and other deratings in these areas by load level. The pivotal supplier analysis in Part B of this section indicates there is no supplier in Connecticut and/or Middletown that would profit from withholding. However, it suggested a slight potential for market power in Southwest Connecticut and Norwalk-Stamford. The following figure shows forced outages and other deratings for the only supplier in Norwalk-Stamford and Middletown.

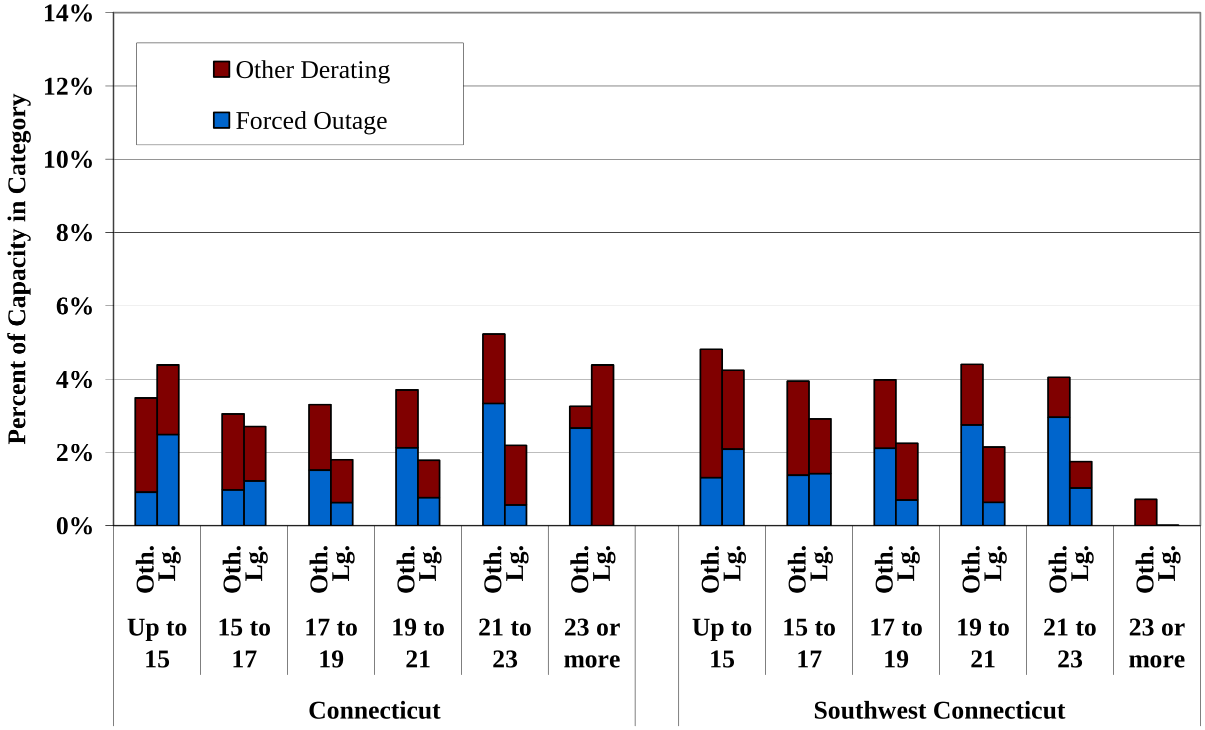
Figure 45
Forced Outages and Other Deratings by Load Level and Type of Supplier
Norwalk-Stamford and Middletown – 2004



Ordinarily, the existence of only a single supplier in market would raise serious market power concerns, but nearly all of the capacity in these two locations is under reliability agreements, which substantially mitigates market power incentives. Figure 45 shows that the physical derating quantities are quite small on average for this supplier, they generally decrease as load rises, and do not suggest evidence of strategic physical withholding in these load pockets.

Figure 46 summarizes physical deratings results for the Connecticut and Southwest Connecticut areas by load level.

Figure 46
Forced Outages and Other Deratings by Load Level and Type of Supplier
Connecticut and Southwest Connecticut – 2004

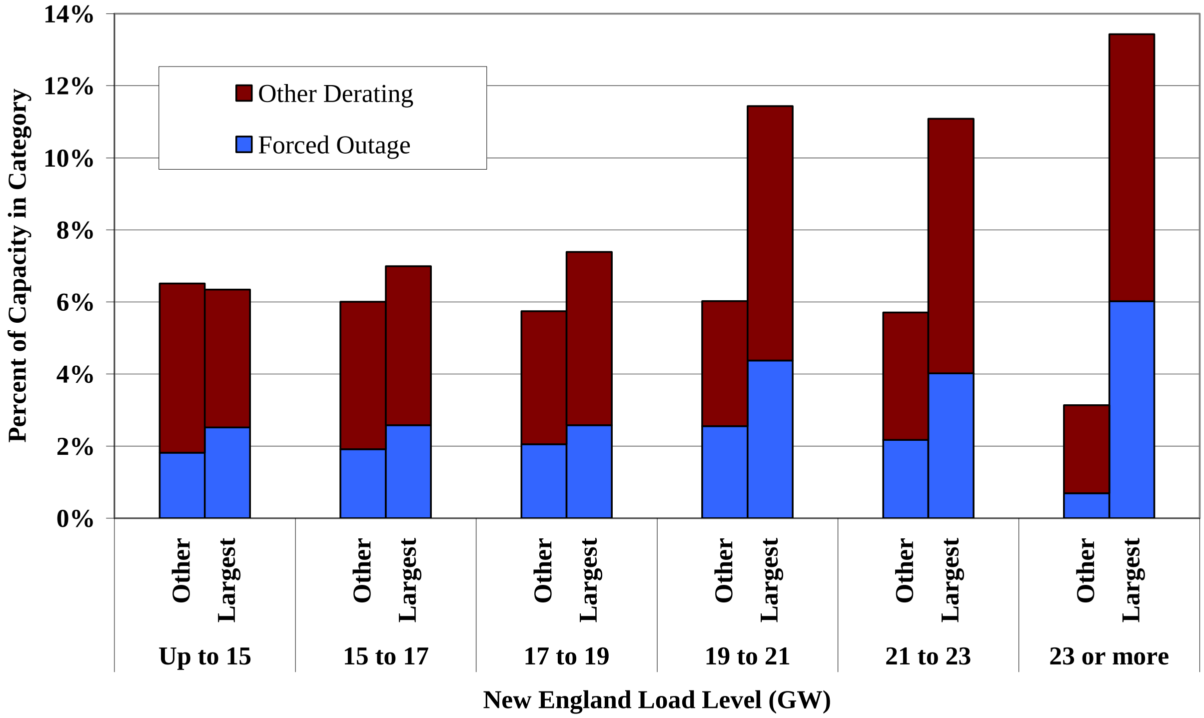


The physical deratings were relatively low in Connecticut and Southwest Connecticut, accounting for an average of 2 percent to 4 percent of capacity when load is below 23 GW. When load is greater than 23 GW, physical deratings within Southwest Connecticut virtually fall to zero, although the other non-planned deratings for the large supplier in Connecticut go up above 4 percent. The same supplier is most pivotal in Connecticut and Southwest Connecticut, although the analysis in the previous section indicates that the large supplier was pivotal in Connecticut in less than 0.1 percent of hours. Generally, the quantities shown in the figure above are relatively small and do not provide evidence of systematic physical withholding.

3. Potential Physical Withholding in All New England

Figure 47 summarizes physical withholding analysis for all of New England by load level, comparing the largest supplier that is sometimes pivotal with all other suppliers.

Figure 47
Forced Outages and Other Deratings by Load Level and Type of Supplier
All New England – 2004



The figure shows that the largest supplier's forced outages and other non-planned deratings grow from 6 percent to 8 percent of its capacity when load is below 19 GW to more than 10 percent when load is greater than 19 GW. On the other hand, the potential physically withheld capacity for other suppliers consistently averages 6 percent when load is below 23 GW and just 3 percent at the highest load levels. This trend raises concerns about possible physical withholding.

The large forced outages at high load levels for the largest supplier were due to the outage of a large base load unit on August 30, 2004. This forced outage was logged with the ISO, and it is beyond the scope of this report to assess whether the outage was valid. However, we did review the real-time market outcomes during the period of the outage and found that prices at the generator's location did not exceed $100 per MWh for the day. Given that load exceeded 24 GW and this was the highest load day of the year, we find little evidence that this forced outage constituted an exercise of market power.

The primary source of other deratings under high load conditions was hydro resources operated by the largest supplier. Under some circumstances, operators of hydro resources may manage energy limitations by adjusting the maximum output level in real-time. Since generators must submit their real-time offers prior to 6 PM on the evening before real-time, they may not have sufficient flexibility to manage their energy limitations efficiently. One way to manage the output of the unit over the day when prices are uncertain is to revise the offer quantities in real-time by adjusting the minimum and maximum output levels in real time.

The overall pattern of forced outages and other deratings does not provide conclusive evidence of physical withholding, but does raise some concerns. As described in the previous sub-section, shortage pricing is invoked when New England does not have sufficient market-wide reserves to cover its largest contingency plus half of its second largest contingency, approximately 1700 MW. There were no occasions when market-wide reserves dropped below 3200 MW during 2004. Therefore, the largest supplier would need to withhold at least 1500 MW to drive the market into shortage pricing conditions.

Figure 48 examines whether the largest supplier could have withheld such a large quantity on any particular day. Thus, this figure combines all potential economic and physical withholding from the output gap, forced outages, and other non-planned deratings.

Figure 48
Hourly Output Gap, Forced Outages, and Other Non-Planned Deratings for the Largest Supplier in New England
Daily Peak Load Hours – 2004

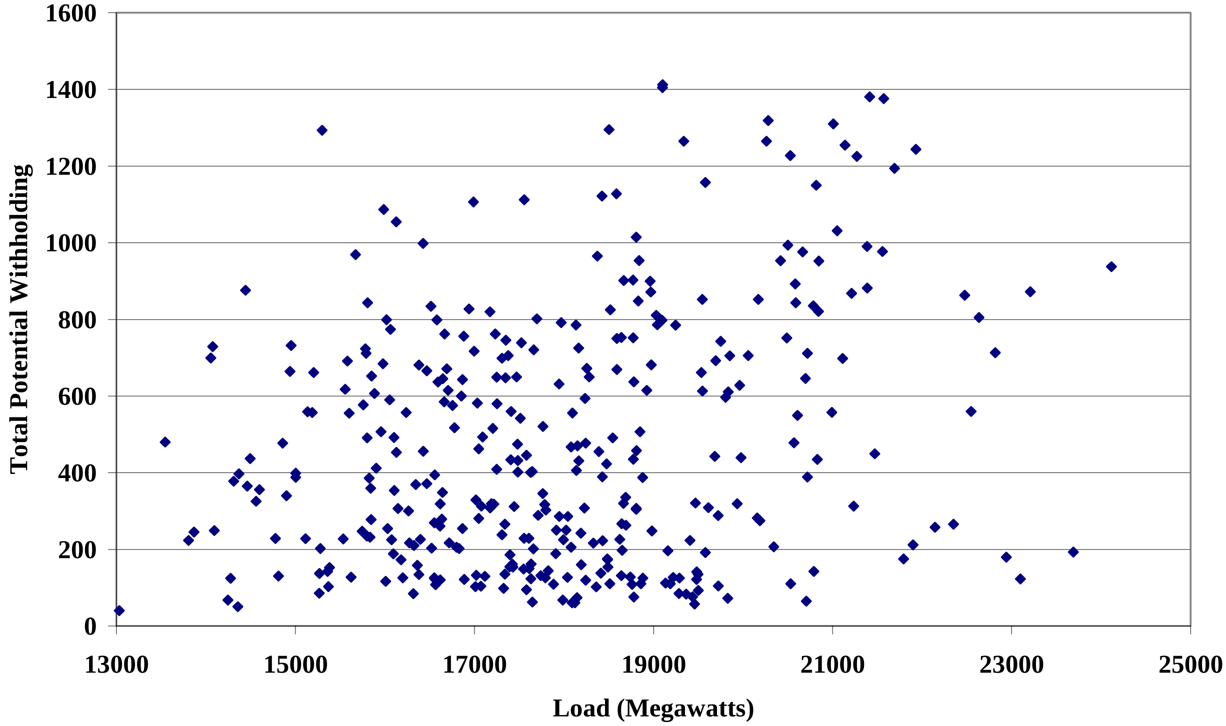


Figure 48 shows four hours when the total potential withholding from the largest supplier was approximately 1,400 MW. Two occurred when load was close to 19 GW while the other two occurred when load was above 21 GW. When load was above 21,500 MW, the total potential withholding was always below 1,000 MW. Because the pivotal supplier analysis suggested that the largest supplier would need to withhold more than 1,500 MW under peak demand conditions to substantially affect prices, we do not find significant evidence of attempts to exercise market throughout New England.

Based on the analyses of potential economic and physical withholding in this section, we find little evidence of significant withholding that might indicate market power abuses in any of the six regions during 2004. However, peak demand conditions were very mild in 2004 relative to historical standards. Furthermore, the pivotal supplier analysis suggests that market power is limited by the large amount of capacity under reliability agreements. Nonetheless, the ISO should continue to monitor for potential economic and physical withholding, particularly in constrained areas after the reliability agreements expire.

Finally, although we found little evidence that participants had engaged in economic or physical withholding to raise energy prices in New England, the report identifies a local market power issue that arose in late 2004 related to supplemental commitments made by the ISO to satisfy local reliability requirements in the NEMA/Boston area. We have consulted with the ISO's Internal Market Monitoring Unit regarding this issue. The ISO has already filed and FERC has approved a change to the reference level calculation provisions under its market power mitigation framework to more effectively mitigate this conduct. In addition, we have proposed a change in the self-scheduling rules for constrained areas that would further limit this conduct.