2005 STATE OF THE MARKET REPORT
NEW YORK ISO

POTOMAC ECONOMICS, LTD.

Independent Market Advisor
to the New York ISO

August 2006

TABLE OF CONTENTS

List of Figures

EXECUTIVE SUMMARY

Our assessment of the New York ISO wholesale electricity markets in 2005 indicates that the markets showed many areas of improved performance, primarily due to the collection of software enhancements implemented on February 1, 2005, dubbed Standard Market Design ("SMD") 2.0. While implementation of SMD 2.0 encountered some difficulties, the NYISO resolved these issues. Energy and ancillary services markets have functioned with no evidence of significant market power or manipulation by market participants. We also update the analysis of a number of continuing issues related to market rules and operations that have been raised in previous reports.

In evaluating the NYISO markets in 2005, we give particularly close attention to aspects of the market that were affected by the implementation of new market software under SMD 2.0. In this report, we also address the following areas:

- Energy Market Prices and Outcomes;
- Day-ahead to Real-time Price Convergence;
- Ancillary Services Markets;
- Market Participant Bid and Offer Patterns;
- Market Operations;
- Capacity Market;
- External Transactions Scheduling; and
- Demand Response Programs

The following subsections provide an overview of the findings of the Report in each of these areas.

A. Market Enhancements Under SMD 2.0

Under SMD 2.0, the NYISO implemented the Real-Time Scheduling ("RTS") system on February 1, 2005. The RTS system uses a common computing platform, algorithms, and network models for both the real-time commitment and real-time dispatch functions, and is comprised of three major components:

- Real-Time Commitment ("RTC") model replaces the Balancing Market Evaluation ("BME") software. RTC co-optimizes energy, reserves and regulation, and commits resources as necessary to meet the demands of the next hour.
  - RTC runs and posts results every 15 minutes (rather than hourly as had BME), and makes commitment decisions that are optimized over a 2 ½ hour period.
  - RTC issues binding commitments to 10-minute and 30-minute gas turbines.
  - It also determines transaction schedules and the dispatch level for off-dispatch resources for the upcoming hour at the top of each hour, although it has the capability to do so on a 15-minute basis.

- Real-Time Dispatch ("RTD") replaces the Security Constrained Dispatch ("SCD") software to dispatch the system.
  - RTD issues a 5-minute basepoint, co-optimizing energy, reserves, and regulation for the forecasted conditions up to 60 minutes ahead.
  - RTD recognizes the transaction schedules, self-committed unit schedules, and units committed by RTC in making dispatch decisions.
  - RTD can also be run in Corrective Action Mode ("RTD-CAM"), which NYISO system operators can run on-demand to address abnormal or unexpected system conditions.

The RTS was a vehicle for introducing a number of market enhancements, including:

- Generators may submit hourly start-up costs, specified as a discrete dollar cost or as a function of elapsed time since the most recent shutdown.

- Generators may submit three-part offers in real-time for purposes of intraday commitments (start-up costs, minimum generation costs, and incremental energy costs);

- A second settlement in real-time was introduced for the reserves and regulation markets, providing better incentives for these services to be provided by the lowest cost suppliers in real time, regardless of the day-ahead schedules;

- Reserve and regulation market-clearing prices set based upon the marginal system cost of providing the service, including the marginal opportunity cost of not providing a different service;

- Demand curves for reserves and regulation services in the Day-Ahead and Real-Time Markets provide efficient prices during shortage conditions when the reserve requirements cannot be satisfied or would be excessively costly to satisfy.

The last element of the RTS system is particularly important because it enables efficient energy and operating reserve prices during periods of shortage. When operating reserves are sacrificed to meet energy demands, the value of the foregone reserves will be reflected in the energy and reserve prices. This change replaces the prior scarcity pricing provisions that had been in use prior to the introduction of RTS.

B. Energy Market Prices and Outcomes

Summary of Outcomes in 2005

Electricity prices were substantially higher in 2005 than in 2004 due to rising input fuel prices and increased demand. Changes in fuel prices are the primary driver of trends in energy prices over extended periods. Even though much of the electricity consumed in New York is generated from hydro, nuclear, and coal-fired generators, natural gas and oil units are usually the marginal generation units that set prices in the market, particularly in peak hours. Therefore, changes in the prices of these fuels directly affect electricity prices.

Natural gas prices were an average of 44 percent higher in 2005 than in 2004. Hurricanes reduced the flow of natural gas from the Gulf Coast region from the end of August through the fall, leading to tight supply and higher prices. Likewise, oil prices rose by 45 percent from 2004 to 2005. During January of both years, natural gas prices spiked briefly due to extreme cold weather and low storage levels.

Load growth was relatively steady across most hours from 2003 to 2005. However, the summer of 2004 was notable because of the lack of days with extreme load levels, while the summer of 2005 experienced higher than normal load levels. In 2005, there were 228 hours when loads exceeded 28 GW compared with just 2 such hours in 2004 and 40 hours in 2003.

Not surprisingly given the number of extreme load days, spikes in hourly real-time prices were more frequent in 2005 (204 hours had prices over $200 per MWh, compared to 7 hours in 2004). There were also more hours with moderately high prices in 2005 due to the increase in fuel prices. There were approximately 6000 hours with prices above $60 per MWh in 2005, compared to fewer than 2000 hours in 2004.

Prices varied at locations throughout the state in 2005 due to transmission congestion and losses. The primary transmission constraints in New York occur at the following four locations:

- The Central-East interface that separates eastern and western New York;
- The transmission paths connecting the Capital region to the Hudson Valley;
- The transmission interfaces into New York City and the load pockets within the City; and

- The interfaces into Long Island.

Congestion on the transmission paths into and within New York City resulted in average day-ahead prices in the City that were $16 per MWh higher than in the eastern upstate region. Transmission losses make up a significant share of the $12 per MWh average difference between East Up-state and West day-ahead prices, while transmission constraints, particularly on the Central-East Interface and in the Hudson Valley, make up the rest of the difference.

Total "all-in" prices, which include the costs of energy, ancillary services, capacity, and other costs, increased substantially for all locations in 2005. Higher energy costs, driven by higher fuel costs and tighter peak demand conditions, led to higher all-in prices. However, there were also reductions in the capacity costs to up-state New York and the uplift from Schedule 1 charges (and rebates) allocated to all loads in the New York Control Area. There were several factors that led to the net reduction in uplift from 2004 to 2005:

- Under SMD 2.0, real-time commitment of gas turbines was more efficient, reducing the uplift needed for "make whole" payments to units that turn out to be uneconomic.
- Higher overall energy prices and improvements to modeling of transmission losses in the day-ahead market led to more surplus revenue from energy and loss residuals.
- Higher overall congestion in 2005 led to corresponding increases in balancing congestion costs, which is the uplift generated when transmission is oversold in the day-ahead market and must be purchased back from the real-time market.

Price Corrections

All real-time energy markets are subject to some level of price corrections to account for metering errors and other data input problems, although the rate of corrections spiked in 2005 due to issues associated with the implementation of new real-time market software under SMD 2.0. Nine significant software issues arose that led to the majority of price corrections, several of which were identified and corrected in February, 2005. The remaining software issues surfaced from March to June and were addressed between June and October. Once these initial software issues were addressed by NYISO, the frequency of price corrections fell to typical levels by November 2005. Temporary spikes in the frequency of price corrections have typically occurred after major software modifications in the past.

Economic Signals Produced in 2005

The economic signals provided by the New York markets can be measured using the net revenue metric. This metric measures the total revenue that a hypothetical new generator would have earned in the New York markets less its variable production costs.[1] In long-run equilibrium, the market should support the entry of new generation by providing average net revenues that are sufficient to finance new entry. This may not be the case in every year since there are random factors that can cause the net revenue to be higher or lower than the equilibrium value (e.g., weather conditions, generator availability, etc.).

Net revenue increased significantly during 2005, primarily due to tighter load conditions. In 2005, there were reserves shortages that resulted in very high energy prices under the shortage pricing provisions, whereas there were no instances of shortage prices in 2003 or 2004. Despite the increased energy prices, net revenue clearly remained below the levels necessary to justify new investment in gas turbines outside New York City and in combined-cycle units in western New York.

Based on market conditions in 2005, there are several locations where it might be profitable to build new capacity. Increased shortage pricing in eastern New York and higher fuel prices raised combined cycle net revenue in the Hudson Valley to levels that might exceed their investment costs. Net revenue for a new gas turbine in 2005 was close to the estimated annual cost (including return on investment) of building a new gas turbine in the City. Likewise, net revenue for a new combined cycle unit ranged from almost $300 to nearly $400 per kW-year, which likely exceeds the net revenue needed to justify investment in New York City. These results are consistent with market conditions in New York City, which was been relatively close to being capacity-deficient in 2005.

Although estimated net revenues grew considerably in 2005 to levels that would likely justify new investment in some areas if the net revenues continued over the long-term, there are other

1 The assumptions for this analysis have been standardized by FERC and the market monitors in the various markets to provide a comparable basis for comparison of the net revenue values from the different markets. However, the net revenue estimates produced using these assumptions are likely to over-estimate the actual net revenues earned by market participants for several reasons, which are discussed later in the report.

factors that affect new investment. The ability to enter into forward contacts is an important factor because it allows the new investor to secure a stable stream of revenues for the project. The regulatory process is also an important factor. Expectations and risk are also important factors. Market participants must anticipate, over the life of the investment, how prices will be affected by the new capacity investment, future load growth, increasing participation in demand response, and the risk associated with changes in the market rules or regulation over the life of the project.

C. Day-Ahead to Real-Time Price Convergence

Energy Price Convergence

The day-ahead market allows participants to make forward purchases and sales of power for delivery in the real-time. This is a valuable financial mechanism that allows participants to hedge their portfolios and manage risk. In a well functioning system with day-ahead and real-time markets, we expect that day-ahead and real-time prices will not systematically diverge from one another. For example, if day-ahead prices were predictably higher than real-time prices, buyers would shift more of their purchases to the real-time. If day-ahead prices were foreseeably lower than real-time prices, buyers would be attracted to the day-ahead market. In each case, sellers would tend to shift in the opposite direction. Price convergence is desirable because it promotes the efficient commitment of generating resources and scheduling of external transactions.

Outside of the summer months, there was a slight day-ahead premium throughout New York, which was consistent with previous years. During the summer, however, average real-time prices were substantially higher than average day-ahead prices. A small number of real-time price spike events account for most of the real-time premium during the summer. These events had a much stronger effect on real-time prices in New York City and Long Island, which exhibited real-time price premiums of 18 percent and 24 percent, than in Western New York and Eastern Upstate New York, which exhibited real-time price premiums of 3 percent and 9 percent. Many of the price spikes were caused by Thunder Storm Alerts ("TSAs"), which require double contingency operation of the ConEd overhead transmission system in real-time but not in the day-ahead market. Substantial real-time price premiums are not likely to persist in the future as participants revise their expectations of real-time prices to include a greater likelihood of shortage prices. Higher expected real-time prices should induce market participants to profit by scheduling additional virtual load in the day-ahead market.

Convergence in New York City Load Pockets

Day-ahead to real-time price convergence varied substantially by load pocket within the City during 2005. During the summer, all of the areas within the City exhibited a substantial real-time price premium, consistent with the rest of New York state. However, the premium was larger in the most import-constrained areas than in the 345 kV system. From September to December, the most constrained pockets continued to exhibit a large *real-time premium*, while the areas in 345 kV system experienced a large *day-ahead premium*. The New York City zone as a whole experienced only a slight day-ahead price premium from September to December due to the averaging of areas with varying patterns of convergence between the day-ahead and real-time. The large systematic differences in the load pockets would provide significant incentives for market participants to arbitrage, but virtual trading is limited to the zonal level.

Ancillary Services Price Convergence

Under SMD 2.0, the New York ISO integrated real-time ancillary services markets with the existing real-time energy market, complementing the day-ahead market that has included markets for energy, reserves, and regulation since 1999. As with the energy market, suppliers who sell ancillary services into the day-ahead market have the obligation to either provide the service from their own resources in real-time or pay the real-time price to buy back the services in the real-time market.

During the summer, real-time prices for ancillary services were significantly higher than other times of year, while day-ahead prices rose only modestly. For instance, during the 4 pm hour, average prices for 10-minute reserves in Eastern New York were $4 per MWh in the day-ahead market and $47 per MWh in the real-time market. The lack of convergence between day-ahead and real-time prices was partly due to the effect of a small number of TSA events. Typically, real-time reserves prices are very close to zero, but under conditions of near-shortage, prices can rapidly rise to hundreds of dollars per MWh. This volatility is difficult for market participants to predict in the day-ahead market, and based on our analysis, the day-ahead market systematically under-valued 10-minute reserves in the east during the summer. In the fall, the price pattern for reserves was reversed as day-ahead prices remained at the levels that prevailed during the summertime, while real-time prices sank below day-ahead prices.

Average real-time prices for ancillary services were generally higher than average day-ahead prices during the first year under SMD 2.0. This is because price spikes related to ancillary services shortages occur more frequently in the real-time than in the day-ahead market. Reserves shortages never occur in the day-ahead market, because sufficient capacity is offered into the day-ahead market. However, unforeseen conditions such as forced outages and short term ramp constraints can occur resulting in real-time reserves shortages. Under-forecasted demand in the day ahead can result in under-commitment that can lead to real-time reserves shortages.

Market participants should respond to the possibility of real-time price spikes by bidding up the clearing price in the day-ahead market to reflect the likelihood that a real-time shortage will occur. However, the current market rules do no allow load serving entities and virtual traders to arbitrage day-ahead to real-time price differences by adjusting their ancillary services purchases or by submitting price-sensitive bids for ancillary services. Only generators have the ability to submit price-sensitive offers to the ancillary services markets in the day-ahead. Persistent real-time price premiums give generators an incentive to raise their day-ahead offer price, which can reduce the efficiency of the day-ahead commitment.

D. Ancillary Services Markets

The prior section evaluates the price convergence in the ancillary services markets. This section describes the changes made in these markets under SMD 2.0 and evaluates the conduct of participants and the efficiency of the ancillary services markets in 2005.

Market Design Changes

The design of the ancillary services markets and their interaction with the energy market changed significantly with the implementation of SMD 2.0. The new design includes the two key features not found in any other wholesale market in 2005. First, reserves and regulation are co-optimized with energy in the real-time spot market auction. Second, the real-time market uses "demand curves" to limit the costs of procuring each ancillary service and to better reflect the value of ancillary services and energy in prices under shortage conditions. We review the performance of the ancillary services markets in 2005, highlighting the effect of the market design changes that occurred under SMD 2.0.

One of the primary benefits of co-optimizing energy dispatch and ancillary services procurement in the real-time market is that the clearing price of each ancillary service is equal to the marginal cost to the system of providing the service. To the system, the marginal cost of purchasing a service is equal to the sum of the marginal generator's (i) availability bid price and (ii) opportunity cost of not providing another product such as energy. Under SMD 2.0, all dispatchable generators offering to sell energy in real-time must also offer to provide reserves with a $0 per MWh availability bid. Thus, the real-time clearing prices for reserves are equal to the opportunity cost of not providing another product. Frequently, it is not necessary to re-dispatch generators in real-time to meet reserves requirements, because excess reserve capacity is available at the tops of on-line units. During these periods, the clearing prices of reserves drop to $0 per MWh because it costs nothing to maintain reserves.

The new market design added ancillary services to the two-settlement system that has existed for energy since 1999. A two-settlement system consists of a spot market (i.e. real-time) and a forward financial market (i.e. day-ahead), whereby day-ahead financial obligations must be reconciled in the real-time market. So, a generator that is paid to sell reserves in the day-ahead market must either (i) physically provide reserve capacity in real-time or (ii) purchase reserves from the real-time market in order to meet the financial obligation. Generators that sell reserves in the day-ahead market and are dispatched by the ISO to provide energy in the real-time market are paid the real-time clearing price for energy but must still buy back reserves in the real-time market. However, since reserves are co-optimized with energy in the real-time market, the profit from selling energy will exceed the replacement price of the reserves. Under SMD 2.0, real-time dispatch instructions are determined in accordance with clearing prices such that generators provide the most profitable service.

Ancillary Services Market Expenses

Ancillary services expenses rose substantially from 2004 to 2005 due to a 54 percent increase in regulation costs and a 62 percent increase in reserves procurement costs. Higher expenses were expected in 2005 as a result of the SMD 2.0 changes for two reasons. First, in a co-optimized market, the prices of ancillary services reflect the opportunity costs of diverting resources from the energy market so that energy price spikes are frequently accompanied with spikes in ancillary services. Without co-optimization, reserves are priced in a manner that does not reflect the economic trade-offs between reserves and energy, raising the risk that a supplier will be harmed by providing reserves. This can discourage suppliers from offering reserves to the market and result in a less efficient dispatch of energy and reserves. Second, higher fuel prices, particularly in the fall of 2005, led to increased opportunity costs for low-cost generators providing ancillary services rather than energy.

Virtually all of the growth in ancillary services market expenses from 2004 to 2005 may be attributed to the fall of 2005. Ancillary services expenses are driven primarily by day-ahead price levels since the full requirement of ancillary services is purchased in the day-ahead market (only deviations are settled in real time). It is not surprising to find that the pattern of the ancillary services expenses was consistent with the pattern of day-ahead prices, being modest during the spring and summer and rising during the fall. The pattern sharply contrasts with real-time prices, which peaked during the summer at levels much higher than day-ahead prices. Although higher fuel prices led to higher opportunity costs for ancillary services providers during the fall of 2005, higher offer patterns for ancillary services in the fall of 2005 may also have reflected the increased volatility of real-time prices during tight conditions.

Ancillary Services Offer Patterns

The quantity of offers to supply 10-minute spinning reserves and 30-minute operating reserves rose substantially from 2004 to 2005, which we attribute to the improved incentives under SMD 2.0. previously, the day-ahead clearing prices were set by the highest-priced accepted offer, so it was possible for the price to be lower than the opportunity cost of not providing energy. Thus, generators risked losing profits in the energy market by providing reserves. There is no such risk under the new design since the reserves clearing price is always greater than or equal to the opportunity cost of generators scheduled for reserves. In other words, generators are always selected to provide whichever is more profitable (assuming they submit energy and ancillary services offers consistent with their marginal costs).

Although participation in reserves markets increased in 2005 relative to the period before SMD 2.0, offer patterns also changed markedly over the course of 2005. Day-ahead offer prices rose during 2005 for several categories of ancillary services. The amount of 10-minute spin offered under $5 per MWh was approximately 1,170 MW during the spring, but this fell to 750 MW by the fall. For other 10-minute reserves, offers under $5 per MWh fell by 710 MW from the spring to the fall. Most of the regulation capacity offered during the spring and summer was priced below $25 per MWh, but regulation offer prices rose significantly in the fall. The higher ancillary services offer prices contributed to a significant rise in day-ahead clearing prices and market expenses from September through December 2005.

The rise in regulation offer prices is attributable to changes in behavior by two market participants in September and October. The New York ISO's market monitoring staff reviewed the rise in offer prices from these market participants and concluded that the behavior does not warrant mitigation under the NYISO Tariff. However, due to limited participation in the market by regulation-capable capacity, the ownership of resources that participate in the market is relatively concentrated. In the short-term, the high concentration may provide incentives for certain market participants to raise their offer prices above marginal cost. In the longer-term, we expect that additional supply would enter the regulation market if prices rise above competitive levels for a sustained period.

The rise in day-ahead 10-minute spinning and non-spinning reserves offer prices is attributable to changes in offers by several market participants. However, this does not raise significant power concerns for several reasons. First, the market participants in question account for only a modest share of the total capability, making it unlikely that they would withhold in an attempt to exercise market power. Second, we would expect that an attempt to exercise market power would have the greatest effect when conditions are tight. Thus, if the rising offers of these suppliers reflect attempts to withhold, it is unclear why offer prices remained elevated throughout the fall. Third, the real-time market helps discipline competitors in the day-ahead market, particularly since these same suppliers must offer $0 per MWh in real-time. Given that real-time prices are not substantially lower than day-ahead reserves prices, there is little reason to think that suppliers have been able to raise day-ahead reserves prices above competitive levels.

Although market power in the markets for 10-minute reserves is not a significant concern, the pattern of escalating offer prices raises concerns that market participants are submitting reserves offers above marginal costs, because such behavior negatively affects market efficiency. The day-ahead market commits and schedules resources for energy and ancillary services in economic merit order, resulting in an efficient commitment when suppliers offer their resources at marginal cost. When suppliers raise their offer prices above marginal costs, other more costly resources may be committed in their place. Thus, it is important for overall market efficiency to address issues that undermine the incentives of suppliers to offer their resources at marginal cost.

The most likely explanation for the pattern of rising offer prices is that market participants are responding competitively to poor convergence between day-ahead and real-time reserves prices. If suppliers predict real-time prices will be higher than day-ahead prices, we can expect they will avoid selling into the day-ahead market and shift more sales to the real-time market. One way to do this is for them to raise their day-ahead offer prices. Our analysis shows that many of the increases in offer prices occurred shortly after real-time reserves shortages that caused significant price spikes. It is likely that these price spikes caused some suppliers to update their expectations of real-time clearing prices and in some cases raise their day-ahead offer prices.

Ancillary Services Market Conclusions

The implementation of SMD 2.0 has lead to two major enhancements to the markets for reserves and regulation. First, the co-optimization of energy and ancillary services in real-time improves market efficiency by allowing the real-time model to consider the costs of ancillary services procurement in the prices of energy, and vice versa. It guarantees that the clearing prices of energy, reserves, and regulation fully reflect the opportunity cost of not providing the other services, eliminating the need for separate lost opportunity cost payments. Second, under SMD 2.0, pricing under shortage conditions is governed by reserve demand curves. The demand curves establish an economic value for reserves that are reflected in energy prices at times when the energy market must bid scarce resources away from the reserve markets. This improves consistency between clearing prices and the operation of the system, and this better reflects the economic value of reliability.

The enhancements made to the ancillary services markets under SMD 2.0 have lead to four notable changes in market outcomes. First, ancillary services market expenses (i.e. total costs paid by LSEs for ancillary services) were closely correlated with day-ahead ancillary services prices which were modest in the spring and summer but rose significantly in the fall. Second, the introduction of SMD 2.0 induced greater participation from operating reserves suppliers. Third, regulation expenses rose during the fall as a result of higher offers from two large suppliers of regulation. The ownership of regulation-capable capacity is relatively concentrated raising concerns that, in the short-term, certain market participants may have the incentive to raise their offer prices above marginal cost.

Fourth, poor convergence between day-ahead and real-time operating reserves prices created substantial opportunity costs of selling reserves day ahead and incentives for suppliers to raise their day-ahead offer prices. Energy and ancillary services were undervalued on average in the day-ahead market reflecting that, on the whole, market participants did not accurately foresee the frequency and magnitude of price spikes during reserves shortages. Eventually, we expect virtual load bidders to increase their day-ahead energy purchases, putting upward pressure on day-ahead prices and bringing them into closer convergence with real-time prices. Additional demand for energy in the day-ahead market should put indirect upward pressure on average day-ahead ancillary services prices and bring them into closer convergence with average real-time prices. Thus, the incentive for reserves suppliers to raise their day-ahead offer prices above marginal cost should diminish with better convergence between day-ahead and real-time clearing prices.

However, if the convergence between day-ahead and real-time operating reserves prices remains poor, suppliers of operating reserves will continue to have an incentive to raise their offer prices. In this case, the NYISO should consider the feasibility of introducing virtual trading of ancillary services in the day-ahead market. This change would promote convergence of ancillary service prices and reduce the incentive for physical suppliers to raise their offer prices for operating reserves above marginal cost. However, the proposal would need to be carefully studied to ensure it will not have unintended consequences on the day-ahead commitment.

E. Analysis of Energy Bids and Offers

In this section of the Report, we analyze the overall patterns of conduct in the New York market, including those that could indicate attempts to exercise market power.

Potential Physical and Economic Withholding

We examined whether there was any correlation of quantities of potential withholding to load levels. The analysis is based in part on the expectation that suppliers in a competitive market should increase bid quantities during higher load periods to sell more power at the higher peak prices. Alternatively, suppliers in markets that are not workably competitive will have the greatest incentive to withhold at peak load levels when the market effect would be largest. Hence, examining how participant conduct changes under different market conditions is an effective means for evaluating the competitive performance of the market.

We first considered potential physical withholding by analyzing generator deratings. A derating occurs when a participant reduces the maximum output available from the plant. This could be for planned outages, long-term forced outages, or short-term forced outages. A derating could be partial (maximum output is reduced, but is greater than zero) or complete (maximum output is zero). We analyze only non-planned outage deratings, eliminating planned outages from our data. The remaining deratings data would then include only long-term and short-term forced deratings.

We focused on the hours with higher demand because, under a hypothesis of market power, we would expect to find that withholding increases as demand increases. We also focused on the locations east of the Central-East interface, which has limited import capability and is more vulnerable to the exercise of market power. We found that no positive relationship existed between deratings and load level in 2005, which would raise concerns that market power was systematically exercised through physical withholding. Focusing only on short-term deratings, we found the same results. In fact, deratings are least frequent when load reaches high levels, which is consistent with workable competition.

We also examined the trend in forced outages in the New York markets to ascertain if generators are responding to economic incentives to increase reliability of their units. The Equivalent Forced Outage Rate ("EFOR") is the portion of time a unit is unavailable due to forced outages, expressed as equivalent hours of full forced outage at its maximum net dependable capability. EFOR declined substantially following the implementation of the NYISO markets. This is consistent with the incentives the deregulated markets provide to maximize availability, particularly during high load conditions. EFOR was relatively high in 2000 due to the outage of an Indian Point nuclear unit. After the Indian Point outage, the EFOR has been consistently close to 4 percent – much lower than the outage rates that prevailed prior to the implementation of the NYISO markets. In 2005, the EFOR was approximately 4.5 percent.

To evaluate economic withholding, we calculated the hourly "output gap". The output gap is the quantity of generation capacity that is economic at the market clearing price, but either is not running due to the owner's offer price or is setting the LBMP with an offer price substantially above competitive levels (excluding capacity scheduled to provide ancillary services). This withholding can be accomplished through high start-up cost offers, high minimum generation offers, and/or high incremental energy offers.

To determine whether an offer is above competitive levels, we use reference values based on the past offers of the participant during competitive periods. We conduct the analysis with thresholds matching the mitigation thresholds ($100/MWh or 300 percent, whichever is lower) and a lower threshold ($50/MWh or 100 percent, whichever is lower).

Like our analysis of deratings, the results would support a hypothesis of withholding if the output gap increases as load increases. We focused our analysis on Eastern New York where market power is most likely. We found that the output gap actually decreases under the highest load conditions, indicating that suppliers responded to high load conditions by making additional efforts to bring resources to the market rather than withholding resources to increase prices. This result is important because prices are most vulnerable to market power under peak load conditions. These results indicate that economic withholding was not a significant concern in 2005.

Market Power Mitigation

Mitigation is applied in the real-time market for units in certain load pockets within New York City using the NYISO's conduct and impact approach. The in-city load pocket conduct and impact thresholds are set using a formula that is based on the number of congested hours experienced over the preceding twelve-month period. This approach permits the in-city conduct and impact thresholds to increase as the frequency of congestion decreases, whether due to additional generation or increases in transmission capability. An in-city offer fails the conduct test if it exceeds the reference level by the threshold or more. In-city offers that fail conduct are tested for price impact by the real-time software. If their price impact exceeds the threshold, they are mitigated.

In the day-ahead, mitigation was most commonly associated with the Dunwoodie-South interface that brings power into New York City from up-state, and the 345/138 kV interface that brings power into the 138 kV portions of New York City. The majority of capacity mitigated in the day-ahead market is associated with the start-up and minimum generation parameters, while relatively little is for incremental energy parameters.

In the real-time, most mitigation occurred for congestion on the load pocket interfaces inside the 138 kV portion of New York City. These load pockets have a limited number of suppliers and experience frequent real-time congestion, making them more susceptible to the exercise market power. The low frequency of real-time mitigation is partly due to the fact that day-ahead mitigated offers are carried into the real-time up to the day-ahead schedule of the unit. Overall, real-time mitigation was less frequent in 2005 relative to the previous year for at least three reasons. First, the impact test implemented in February 2005 under SMD 2.0 is more selective than the proxy impact test that was used previously. Second, two new units came on-line during 2005 at the East River plant, which is located outside the 138 kV area. These likely helped reduce congestion and mitigation related to the Dunwoodie-South interface (i.e. the primary source of imports to New York City). Third, due to a software error, certain units committed and/or dispatched out-of-merit or through the Supplemental Resource Evaluation ("SRE") process were not subjected to mitigation in real-time.

Analysis of Load Bidding

In addition to physical and economic withholding, buyer behavior can strategically influence energy prices. Therefore, evaluating whether load bidding is consistent with workable competition is an important focus of market monitoring. Load can be purchased in four ways:

- Physical Bilateral Schedules – These allow participants to settle transmission charges (i.e., congestion and losses) with the ISO and to settle the energy portion of any underlying contract privately between the parties.
- Day-Ahead Fixed Load – This represents load scheduled in the day-ahead market for receipt at a specific bus regardless of the day-ahead price.
- Price-Capped Load Bidding – This is a price-sensitive load bid into the day-ahead market by a Load Serving Entity ("LSE"). Price-capped load bidding is only allowed at the zonal level while fixed load bidding is allowed at the bus level.
- Net Virtual Purchases – This quantity is equal to the virtual load purchases minus the virtual supply sales. Like price-capped load bidding, virtual purchases are allowed only at the zonal level.

Our analysis indicates New York City and Long Island tend to over-schedule load day ahead, while load is generally under-scheduled day ahead in upstate New York. These results are consistent with the differences between the day-ahead and real-time prices. Generally, real-time prices are lower than day-ahead prices in upstate New York -- the opposite is true in New York City and Long Island. Even during the summer when real-time prices were unexpectedly high throughout New York State, the real-time price premium was much lower in upstate areas than in New York City and Long Island. As discussed in this report, these pricing and scheduling patterns are primarily the result of modeling inconsistencies between the day-ahead and real-time markets. The market will respond to these inconsistencies by adjusting purchases and sales in the day-ahead market. In this case, that arbitrage improves price convergence, but results in over-scheduling within New York City and under-scheduling outside of New York City.

F. Market Operations

Aside from operating the spot markets, a primary role of the ISO is to ensure safe and reliable grid operation. Many of the ISO's operating functions in this regard can have a substantial effect on market outcomes, especially during peak demand conditions. Reliability requires that operators carry out all of these functions, but they should be done in a way that promotes efficient market pricing and behavior. This section evaluates several operating functions and examines how they affect market outcomes.

Transmission Congestion

Congestion can arise in both the day-ahead and real-time markets when transmission capability is not sufficient to accommodate a least-cost dispatch of generation resources. When congestion arises, this will result in higher spot prices at these "constrained locations" than would occur in the absence of congestion.

The NYISO applies congestion charges to day-ahead market transactions by modeling anticipated congestion. These charges are based on the difference between day-ahead spot prices at different locations. Congestion revenues are collected from participants, which include: a) the difference between the total payments by loads and the payments to generators and net imports (excluding losses), and b) the congestion costs collected from physical bilateral schedules. In an LMP system, this revenue will be equal to the marginal value of the transmission capacity times the amount of power flowing across the constrained interface. In the real-time market, only interface flows that are not scheduled in the day-ahead market are assessed balancing congestion charges (or credits).

Congestion charges[2] have grown from $310 million in 2001 to $990 million in 2005 due to (i) the implementation of load pocket modeling in New York City which allowed market-based congestion management, (ii) higher fuel prices which tend to proportionately increase regional price differences associated with congestion, and (iii) the effects of load growth over the period.

Our report includes an analysis that summarizes congestion levels on major interfaces in New York. Congestion is most frequent into the New York City load pockets, while congestion is considerably less frequent on major interfaces in upstate New York, such as the Central-East interface. Our analysis finds that the value of the major upstate transmission interfaces was approximately $250 million, while the value of the most significant downstate interfaces totaled

2 Total congestion charges include day-ahead congestion rents and balancing congestion costs.

$550 million in 2005.[3] While more analysis would be necessary to determine where transmission investment would be most profitable, this analysis suggests that the existing transmission is most valuable in New York City and Long Island.

In a well-functioning market, the level of transmission congestion should generally converge between the TCC market, day-ahead market, and real-time market. Three aspects of convergence are examined in this report. First, day-ahead congestion revenue shortfalls can occur when the revenues collected by the NYISO from congestion in the day-ahead market are less than the payments by the NYISO to the holders of TCCs. Second, balancing congestion revenue shortfalls occur when congestion revenues collected from buyers in the real-time market are not sufficient to cover congestion payments by the NYISO to sellers. Third, the prices paid for TCCs in the auction should be comparable to congestion prices in the day-ahead market that determine payment to TCC holders. The results of our analysis in these areas are as follows:

- Day-Ahead Congestion Revenue Shortfalls: The NYISO experienced substantial day-ahead congestion revenue shortfalls for several years until mid-2004. These shortfalls generally occur when the quantity of TCCs sold exceeds the transmission capability in the day-ahead market. Ideally, the quantity of TCCs would closely match the physical capability of the transmission system. A large share of the shortfall was due to excess TCCs sold into New York City, which were repurchased by the NYISO in July 2004. The NYISO has also implemented provisions to allocate shortfall costs to transmission operators ("TOs") with outages that cause shortfalls, and by allowing TOs to make up to 5 percent of transmission capacity unavailable to the TCC Auctions.

- Balancing Congestion Costs: The NYISO has also experienced balancing congestion revenue shortfalls. If the day-ahead and real-time models are consistent, these costs should be close to zero. Due to modeling inconsistencies, NYISO effectively oversells transmission capability in the day-ahead market and is, therefore, compelled to buy back the over-sold amount, resulting in balancing congestion costs. These modeling inconsistencies have been partly addressed by the use a more detailed network model for New York City load pocket constraints, which was implemented in May 2006.

- TCC Price – Congestion Cost Convergence: In a well-functioning market, the price for the TCC should reflect a reasonable expectation of day-ahead congestion. The auction prices from the auction of 6-month TCCs during the summer capability period for 2005 resulted in a relatively accurate reflection of the value day-ahead congestion.

[3] These values are not consistent with the total congestion costs reported elsewhere in the report because these values are based only on the real-time prices and the analysis did not consider transmission constraints that restrict flows within Long Island or western New York. It also does not quantify the value of flows that the NYISO schedules from western New York through PJM to New York City.

Real-Time Commitment and Scheduling

The NYISO upgraded its real-time commitment model as part of the SMD 2.0 implementation. The Real Time Commitment model ("RTC") is primarily responsible for committing gas turbines and other resources with short start times, such as external transactions. RTC executes every 15 minutes, looking across a two-and-a-half hour time horizon to determine whether it will be economic to start-up or shut down generation. RTC is a significant improvement over its predecessor, the Balancing Market Evaluation model ("BME"), which ran every 60 minutes and evaluated commitment for just one hour.

Convergence between RTC and actual real-time dispatch is a substantial concern because a lack of convergence can result in uneconomic commitment of generation, especially gas turbines, and inefficient scheduling of external transactions. When excess resources are committed or scheduled by RTC, it will generally depress real-time prices and increased uplift costs. Alternatively, when RTC commits insufficient resources, it leads to unnecessary scarcity and price spikes. This section of the report includes several analyses that evaluate the consistency between RTC and actual real-time market outcomes.

The report assesses whether SMD 2.0 improved the efficiency of real-time commitment and scheduling decisions. Decisions to start and to not start gas turbines improved markedly, particularly for gas turbines with longer start times (i.e. 30-minute gas turbines), although opportunities remain for additional improvement. Decisions made after the day-ahead market to adjust imports and exports show modest improvement, although the benefit is relatively minor because the amount of import and export offers submitted by market participants is small compared with the transfer capability of the external interfaces.

More efficient real-time commitment of gas turbines has led to reduced uplift charges for Bid-Production Cost Guarantee ("BPCG") payments to gas turbines that are started by the ISO and turn out to be uneconomic because real-time prices are lower than their offer prices. Our analysis of real-time commitment decisions before and after the implementation of SMD 2.0 leads us to estimate that $22 million was saved due to improved real-time operations under the new market software.

Convergence Between RTC and RTD Prices

Under SMD 2.0, real-time scheduling is accomplished by two models: the Real-Time Dispatch model (“RTD”), which is responsible for matching generation with load and allocating ancillary services on a five-minute basis, and the Real-Time Commitment model (“RTC”), which executes prior to RTD and schedules resources that are not flexible enough to be deployed on a five-minute basis. One way to assess the consistency between RTC and RTD is to examine the degree of convergence between the prices produced by the two models. Convergence between prices in RTC and RTD is important because large price differences indicate inconsistencies that can result in external transactions and off-line gas turbines being scheduled inefficiently, resulting in increased uplift costs and inefficient real-time prices.

Our analysis indicates that convergence between RTC and RTD energy prices tends to be better when load is at low and moderate levels. These are periods when the market supply curve is relatively flat, so small differences between supply scheduled by RTC and RTD do not cause large inconsistencies in prices. Convergence was considerably worse when load was high. During these periods, the market supply curve is steep, so differences between RTC and RTD in the quantities of supply and demand lead to larger pricing inconsistencies. In order to reduce the incidence of inefficient commitment and scheduling, it is important for the NYISO to find ways to remove any sources of bias that may exist from the real-time commitment model.

Pricing Under Shortage Conditions

With the introduction of SMD 2.0 in February 2005, the NYISO enhanced its approach to real-time scheduling and pricing of energy and ancillary services in several ways. First, RTD reduces production costs by efficiently reallocating resources to provide energy and ancillary services every five minutes. The previous model simply reallocated resources to provide energy using a fixed set of ancillary services obligations. Second, by co-optimizing energy and ancillary services, RTD is able to incorporate the costs of maintaining ancillary services into the price of energy. These costs were not reflected in energy prices prior to SMD 2.0. Third, demand curves rationalize the pricing of energy and ancillary services during shortage periods by setting limits on the costs that can be incurred to maintain reserves and/or regulation. Enhancements introduced in RTD provide a more efficient means of setting prices during shortage conditions than the special shortage pricing rules that were used with the previous model.

We examined the consistency of real-time prices with the availability of 10-minute reserves in Eastern New York. The two key findings are as follows:

- A large majority (89 percent) of shortage pricing intervals associated with the Eastern 10-minute reserves requirement occurred during authentic periods of physical shortage.
- However, only 50 percent of the intervals with physical shortages were accompanied by shortage pricing. During the remaining intervals (those that did not exhibit shortage prices), the average price of reserves was $113 per MWh.

Differences between real-time prices and the physical availability of reserves are attributable to Hybrid Pricing, which has been a key element of the real-time market software since 2002. Hybrid Pricing was specially designed to address the problems posed by gas turbines in a nodal pricing market. While gas turbines can be started quickly, they are relatively inflexible (i.e., tend to have a narrow dispatch range), which causes them to frequently be ineligible to set the energy price absent the hybrid pricing provisions. This creates challenges for pricing energy efficiently when the gas turbines are the marginal source of supply. This issue is particularly acute in New York City where gas turbines account for 34 percent of dispatchable capacity, and in the 138kV load pocket where gas turbines account for 50 percent of dispatchable capacity.

Hybrid Pricing is a methodology that treats gas turbines as inflexible resources for the purpose of determining physical dispatch instructions and as flexible for pricing purposes. While this improves the efficiency of energy prices when gas turbines are deployed in-merit, it results in certain inconsistencies between the physical dispatch and the pricing dispatch. Ideally, these inconsistencies should be limited such that: (i) prices reflect scarcity during physical shortage conditions, and (ii) high prices are only set when the system is physically short of either energy or ancillary services.

Hybrid Pricing Under Shortage Conditions

The primary difference between the pricing and physical dispatch passes under Hybrid Pricing is that output from each on-line gas turbine is treated as flexible. However, there are at least two other elements that may account for the low clearing prices that sometimes accompany reserves shortages. The first element pertains to how ramp rate constraints are formulated in each pass of RTD. The physical dispatch pass constrains the instructed output level of each resource according to its ramp rate offer relative to its actual output level, whereas the pricing pass constrains the output level of each resource according to its ramp rate offer relative to its output level in the previous RTD interval's pricing pass. Although Hybrid Pricing was designed this way in order to facilitate treating gas turbines as flexible in the pricing pass, large inconsistencies can arise when a steam unit does not respond immediately to its physical dispatch instructions.

The second element that contributes to inconsistencies between the physical dispatch pass and the pricing pass of RTD has to do with the treatment of gas turbines that have reduced capability due to high ambient temperatures. Since the output capability of gas turbines is inversely related to ambient air temperatures, gas turbine capability tends to be lowest on hot summer days, especially during the afternoon. Given the difficulty of predicting weather, this creates uncertainty about how much capacity is available to RTD. To ensure that RTD matches load with the correct amount of generation, the physical dispatch pass of RTD uses a reduced upper operating limit for gas turbines that appear to be limited by the ambient temperature for several intervals after starting up. The pricing pass of RTD uses the upper operating limit submitted by the market participants, which is typically higher than the ambient temperature limit. The treatment of ambient temperature restrictions on output generally leads to a greater availability of supply in the pricing pass than in the physical dispatch pass.

Significant differences can arise between the physical dispatch pass and the pricing pass of RTD, resulting in a substantial number of intervals when there was a shortage of Eastern 10-minute reserves while the Eastern 10-minute reserves price was low. We examined the major sources of deviations between the pricing pass and the physical dispatch pass of RTD, focusing on the afternoon hours of ten days with most frequent shortage conditions. Our analysis showed:

- During these periods, deviations associated with resources not following dispatch instructions and the hybrid pricing of gas turbines were at a minimum.
- The most significant contributor to deviations during these periods was the inconsistent output limitations of gas turbines due to ambient temperature restrictions. The net effect was that the pricing pass has an average of 150 MW of additional resources available to it during afternoons of the days with most frequent shortages.

The inconsistent output limits for gas turbines due to ambient temperature restrictions largely explains why there were a large number of intervals with physical reserves shortages that did not exhibit shortage pricing.

Some differences between the pricing pass and the physical dispatch pass of RTD are necessary to implement the Hybrid Pricing regime. However, unnecessary differences will generally lead to inaccurate prices and increased uplift. Improving the consistency of the pricing pass and the physical dispatch pass of RTD will improve the efficiency of New York's energy and ancillary services pricing (particularly during shortages) and reduce uplift.

To address these issues, we have two recommendations. In the short-term, we recommend that the NYISO reduce of the upper operating limit of gas turbines according to ambient temperature restrictions in the pricing pass just as it does in the physical dispatch pass. This change, which the NYISO implemented on June 1, 2006, is expected to substantially address the problems with pricing efficiently during reserves shortages. We will evaluate the effect of this change after the summer of 2006. In the long-term, we recommend that the NYISO change the ramp rate constraint in the pricing pass of RTD for generators that do not respond to dispatch instructions.

Transmission Shortages and Congestion Price Spikes

Real-time transmission price spikes occur when the re-dispatch costs necessary to resolve a transmission constraint reach extremely high levels. Spikes in the shadow prices of transmission constraints typically occur for brief periods when there is not sufficient ramp capability within a transmission-constrained area. This may result in large amounts of re-dispatch that provide little reliability benefit when only remote generators are available to be re-dispatched. In such cases, relieving the transmission constraint can cause shortages of operating reserves. Therefore, the actions taken to maintain reliability by resolving a transmission constraint may actually undermine reliability.

Similar to ancillary services demand curves, transmission demand curves could be used to prevent costly re-dispatch when there is little reliability benefit. Use of transmission demand curves could result in a more reliable prioritization of transmission constraints and reserves requirements when it is not possible to meet all requirements and constraints simultaneously.

Under the current operation of the system, the lack of transmission demand curves results in transmission constraints always having priority over reserves requirements. Use of transmission demand curves would also result in more efficient price signals during periods of shortage. Therefore, we recommend that the NYISO continue to evaluate the effect on reliability of using transmission demand curves.

Out-of-Merit Commitment and Dispatch

Uplift costs are incurred when market revenues are not sufficient to satisfy all market obligations while ensuring that all suppliers recover sufficient revenue to cover their full as-offered costs. Generators that do not receive sufficient revenue are given Bid Production Cost Guarantee ("BPCG") payments (i.e. make whole payments). We report uplift costs for four categories of BPCG payments, which include payments for local reliability and non-local reliability operation in both the day-ahead market and the real-time market. Our findings regarding these classes of uplift costs are summarized as follows:

- Day-Ahead Local Reliability Uplift: This is paid to generators committed in the local reliability pass of SCUC, which commits generators out-of-merit in New York City to protect against second contingencies. This increased approximately 35 percent from 2004 to 2005, primarily due to more frequent commitments and higher fuel prices.

- Real-Time Local Reliability Uplift: Most of this is paid to generators committed by local transmission operators in the SRE process. This increased approximately 40 percent from 2004 to 2005, primarily due to higher fuel prices.

- Other Day-Ahead Uplift: This is paid to units committed economically that do not receive sufficient revenue to cover start-up costs and minimum generation costs over their entire run time. Commitments in the local reliability pass of SCUC have a downward effect on clearing prices and increase the costs of this category.

- Other Real-Time Uplift: This goes primarily to gas turbines committed economically by RTC and RTD that do not receive sufficient revenue to cover start-up costs and other running costs over their entire run time. This amount was greatly reduced by the enhancements to the real-time commitment process made under SMD 2.0.

Local reliability uplift is charged to specific transmission owners or load serving entities while non-local reliability uplift is allocated to all loads in the New York Control Area.

Out-of-merit dispatch and commitment have significant market effects. Primarily, they inefficiently reduce prices in both the day-ahead market and real-time market. When this occurs in a constrained area, it will inefficiently dampen the apparent congestion into the area. OOM commitments also may increase uplift payments as units committed economically will be less likely to recover their full bid production costs in the spot market.

Out-of-merit commitment by the local reliability pass of the day-ahead market model increased in 2005, because the resources needed in New York City were committed less frequently on an economic basis. In the long-run, it would be superior to include local reliability constraints into the initial economic commitment pass of SCUC. In the short-run, we recommend that the ISO consider the feasibility and benefits of allowing operators to pre-commit units needed for NOx compliance and other predictable conditions.

Both of these recommendations would require the NYISO to work with participants to revise the cost allocation methodology for uplift associated with the local reliability requirements. Currently, the uplift costs for BPCG payments to units committed for local reliability are allocated locally, while BPCG payments to other units are allocated throughout NYCA. If the recommendations were implemented, a methodology would need to be developed to identify units that were committed as a result of the local reliability requirements.

G. Capacity Market

The capacity market is intended to ensure that sufficient capacity is available to meet New York's electricity demands reliably. This market provides economic signals that supplement the signals provided by the NYISO's energy and operating reserve markets. Load Serving Entities ("LSEs") can meet their capacity obligations by self-scheduling, bilateral purchasing, or through one of the NYISO's forward procurement auctions. Any remaining obligations are settled against the NYISO's monthly spot auction where clearing prices are determined by a capacity demand curve. Currently, the capacity auctions have three distinct locations within New York state: New York City, Long Island, and Rest-of-State. The clearing prices in New York City and Long Island are generally much higher than those in the Rest-of-State.

We evaluate the performance of the capacity market from May 2004 through April 2006. This includes four six-month capability periods from the Summer 2004 capability period through the Winter 2005-06 capability period. Over this period, clearing prices in the Rest-of-State area have ranged from just above $0 per kW-month to $3 per kW-month according to modest variations in imports and exports as well as the timing of retirements and new investments. Clearing prices in New York City have been consistently close to $7 per kW-month during the winter capability periods and near $12 per kW-month during the summer capability periods, reflecting that most generators have higher capabilities during the winter due to lower ambient temperatures.

Prior to January 2006, virtually all of the capacity in New York City was sold. However, after the addition of approximately 500 MW of new capacity in January 2006, there was virtually no increase in the amount of scheduled capacity, and correspondingly, no reduction in clearing prices from the In-City suppliers' price cap. After examining the data on capacity and energy outcomes, we found that the unsold capacity still participated in the energy market. These results raise concerns about potential withholding from the capacity market. The lack of additional sales after the installation of new capacity in January 2006 had a substantial effect on clearing prices in the New York City UCAP market. However, it is also the case that the lack of additional sales in New York City caused Rest-of-State prices to be higher since New York City capacity also contributes toward meeting the state-wide capacity requirement.

The capacity market in New York City is highly concentrated and these results are consistent with one or more suppliers having incentives to withhold from the market. Most of the capacity in New York City is owned by several Divested Generation Owners ("DGOs") that purchased the capacity from ConEd when it was required to divest itself of most of its generation in 1998. Regulators foresaw the potential for market power and imposed market power mitigation measures at that time. These mitigation measures primarily consisted of caps on the revenue that DGOs could earn from the capacity market. To the extent that individual DGOs exercise market power in the capacity market, it is limited by the revenue caps imposed since the initial divestiture.

H. External Transactions

The performance of the wholesale electricity markets depends not only on the efficient use of internal resources, but also the efficient use of transmission interfaces between New York and other areas. Absent transmission constraints, trading should occur between neighboring markets to cause prices to converge.

Based on our analysis in this report, the real-time markets continue to not be arbitraged efficiently by participants. The dispersion in prices during unconstrained hours is considerable. In a significant number of hours for each interface, power is scheduled from the high-priced market to the lower-priced market. These results are similar to results presented in prior years. However, in 2005, there were more frequent price spikes associated with shortage conditions so that efficient utilization of the interfaces would have brought correspondingly greater benefits. Frequently during peak demand conditions, a small amount of additional imports can substantially reduce the magnitude of price spikes.

Several factors prevent real-time prices from being fully arbitraged between New York and adjacent regions.

- Market participants do not operate with perfect foresight of future market conditions at the time that transaction bids must be submitted. Without explicit coordination between the markets by the ISOs, complete arbitrage will not be possible.

- Second, differences in scheduling procedures and timing in the markets serve as barriers to full arbitrage.

- Third, although export fees have been eliminated, there are remaining transaction costs associated with scheduling imports and exports that diminish the returns from arbitrage. Risks associated with curtailment and congestion reduce participants' incentives to engage in external transactions at small price differences.

- A significant portion of imports and exports reflect long-term bilateral agreements (rather than arbitrage of hourly prices) which tend to be insensitive to real-time prices and contribute to the price divergence.

We continue to encourage New York and New England to develop and implement new scheduling procedures, such as "intrahour transaction scheduling." Intrahour transaction scheduling is a process that would allow the physical interchange to be adjusted within an hour when prices diverge at the interface between the two markets. These adjustments would ensure that the interchange levels are efficient, eliminating the price distortions and other inefficiencies caused by poor market arbitrage. We performed simulations to estimate the potential benefits of perfect arbitrage of the main interface with New England. If flows across the interface had been optimal in 2004 and 2005, it could have resulted in production cost savings of $34 million and reduced costs to New York and New England consumers by approximately $200 million.

I. Demand Response

The New York ISO has some of the most effective demand response programs in the country. There are currently three demand response programs in New York State:

- Special Case Resources ("SCR") – These are loads that must curtail within two hours. SCR participants are paid the higher of a strike price that they bid (up to $500/MWh) or the real-time clearing price.

- Emergency Demand Response Program (EDRP) – Loads that curtail on two hours notice on a voluntary basis. EDRP resources are paid the higher of $500/MWh or the real-time clearing price.

- Day-Ahead Demand Response Program ("DADRP") – This program schedules physical demand reductions for the following day, allowing resources with curtailable load to offer into the day ahead market as any supply resource.

The EDRP and SCR programs are among the most effective of their kind in achieving actual load reductions during peak conditions. The total registered quantity of more than 2000 MW is much larger than comparable programs in other ISOs. In the summer of 2005, the quantity of SCR/ICAP subscribers that sold capacity were 279 MW in NYC, 111 MW in Long Island, and 667 MW in upstate New York. These demand response resources have had a substantial effect on prices in the capacity market. They also help increase the competitiveness of the capacity market in New York City where ownership of generation is relatively concentrated.

In 2005, there was one afternoon, July 27th, when the NYISO called upon SCR and EDRP resources in five zones to curtail. The NYISO estimates these resources provided 819 MWh of response during four hours when the average real-time price exceeded $600 per MWh. Nearly 90 percent of the response came from resources located in New York City and Long Island. The day-ahead demand response program has provided considerably less valuable demand reduction than the EDRP and SCR programs. Although approximately 2000 MWh of day-ahead demand response was scheduled in 2005, this was dispersed across many hours with moderate prices.

I. Energy Market Prices and Outcomes

The New York ISO operates a multi-settlement wholesale market system consisting of financially-binding day-ahead markets and real-time markets for energy, operating reserves, and regulation (i.e. automatic generator control). Through these markets, the NYISO commits generating resources, dispatches generation, procures ancillary services, schedules external transactions, and sets market-clearing prices based on supply offers and demand bids. On February 1, 2005, the NYISO upgraded its market software and systems under the Standard Market Design 2.0 project ("SMD 2.0"). Under SMD 2.0, the NYISO became the first wholesale market operator to:

- Co-optimize energy dispatch and ancillary services allocation every five minutes in the real-time market;
- Use demand curves for real-time ancillary services procurement under shortage conditions; and
- Re-evaluate the economic efficiency of real-time commitment and scheduling decisions every 15 minutes rather than hourly.

This report assesses the overall performance of the New York electricity markets in 2005, focusing on the effects of the enhancements made under SMD 2.0.

This section of the report provides an overview of market results in 2005 and evaluates the performance of these markets. This evaluation includes an assessment of the long-term economic signals provided by the New York markets that govern new investment and retirement decisions.

A. Summary of 2005 Outcomes

We begin in this sub-section by summarizing the 2005 energy price trends, load levels, and trends in individual components of the market expenses.

1. Energy Prices

Electricity prices were substantially higher in 2005 than in 2004 due to rising input fuel prices and increased demand. Changes in fuel prices are the primary driver of trends in electricity prices over extended periods. Even though much of the electricity used by New York consumers is generated from hydro, nuclear, and coal-fired generators, natural gas and oil units are usually the marginal generation units which set prices in the market, especially during peak hours. Therefore, changes in the prices of these fuels will directly affect market prices. Figure 1 shows average natural gas prices and electricity prices on a monthly basis during 2004 and 2005.

Figure 1: Electricity and Natural Gas Prices
2004 – 2005

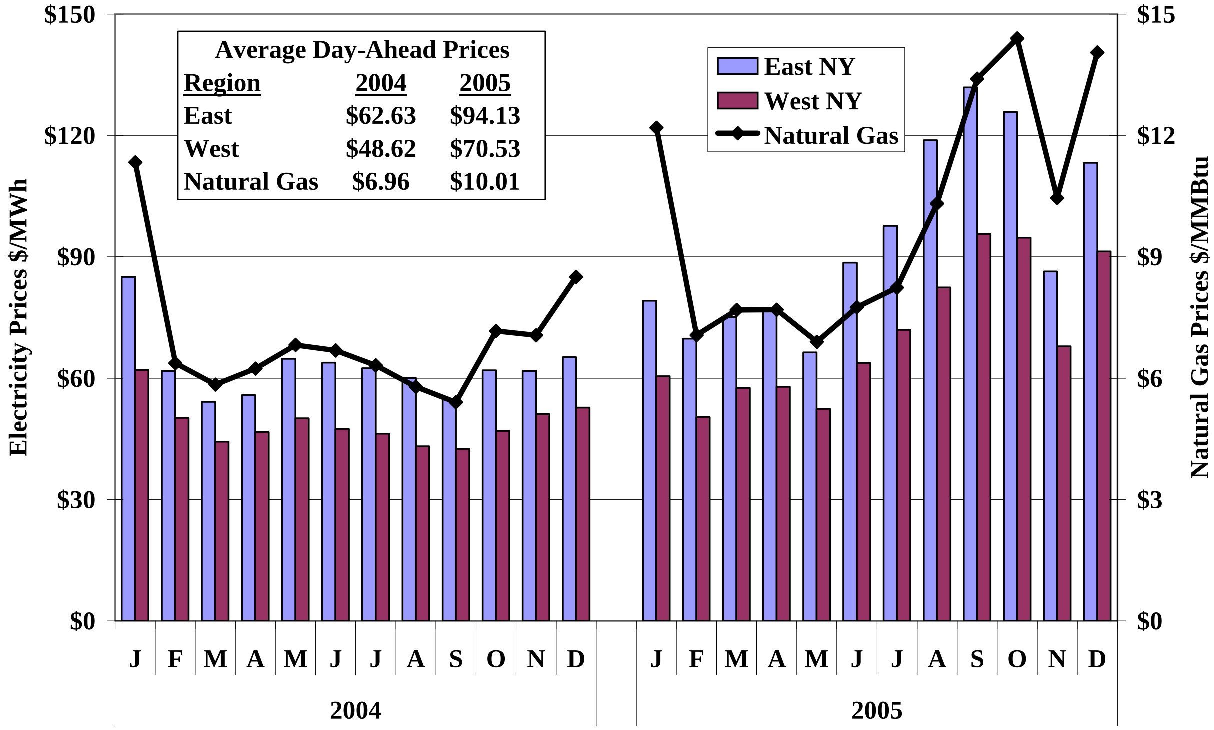


Figure 1 demonstrates that monthly average electricity prices are closely correlated with natural gas prices. Rising fuel prices led to corresponding changes in electricity prices in 2005. Relative to 2004, natural gas prices were an average of 44 percent higher in 2005. Hurricanes reduced the flow of natural gas from the Gulf Coast region from the end of August through the fall, leading to tight supply and higher prices. Likewise, oil prices rose by 45 percent from 2004 to 2005. During January of both years, natural gas prices spiked briefly due to extreme cold weather and low storage levels. We expect a strong correlation between electricity prices with oil and gas prices since fuel costs account for the majority of most generators' variable production costs, and oil and gas units are on the margin in most hours.

Prices were significantly higher in East New York due to transmission congestion and, to a lesser extent, losses. In 2005, average prices in East New York were almost $24/MWh higher (33 percent) than in West New York. Prices also varied significantly within these two regions, particularly the East. Hot weather contributed to a significant rise in electricity prices during the summer of 2005, while the summer of 2004 was comparatively mild. Figure 1 shows that prices began to rise significantly in the summer of 2005 ahead of the escalation in natural gas prices, however, this is largely hidden by the large variations in fuel prices.

To clearly identify changes in electricity prices that are not driven by changes in natural gas prices, the following figure shows day-ahead electricity prices, adjusted for natural gas price fluctuations. The following figure shows the monthly average marginal heat rate that would be implied if natural gas were always on the margin. The *Implied Marginal Heat Rate* equals the *Day-Ahead Electricity Price* divided by the *Natural Gas Price*. The Implied Marginal Heat Rate is shown for East and West New York on a monthly average basis for 2004 and 2005 in Figure 2.

Figure 2: Average Implied Marginal Heat Rate
Based on Day-Ahead Electricity and Natural Gas Prices
2004 – 2005

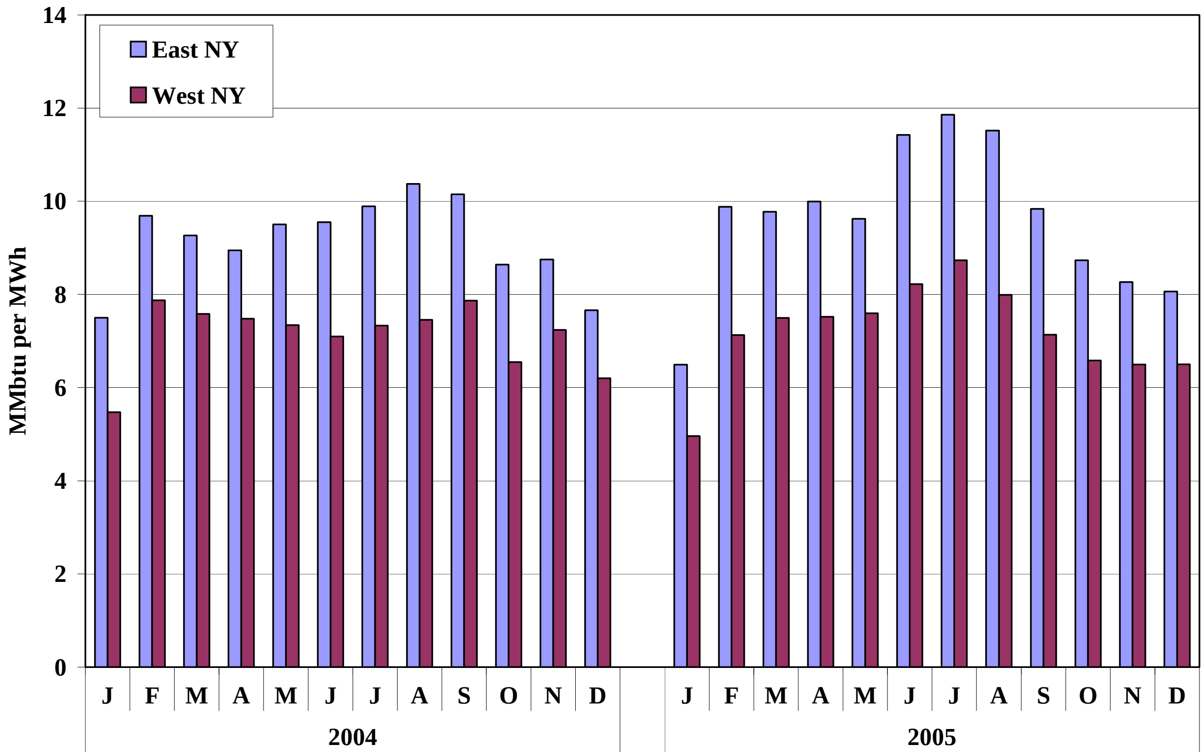
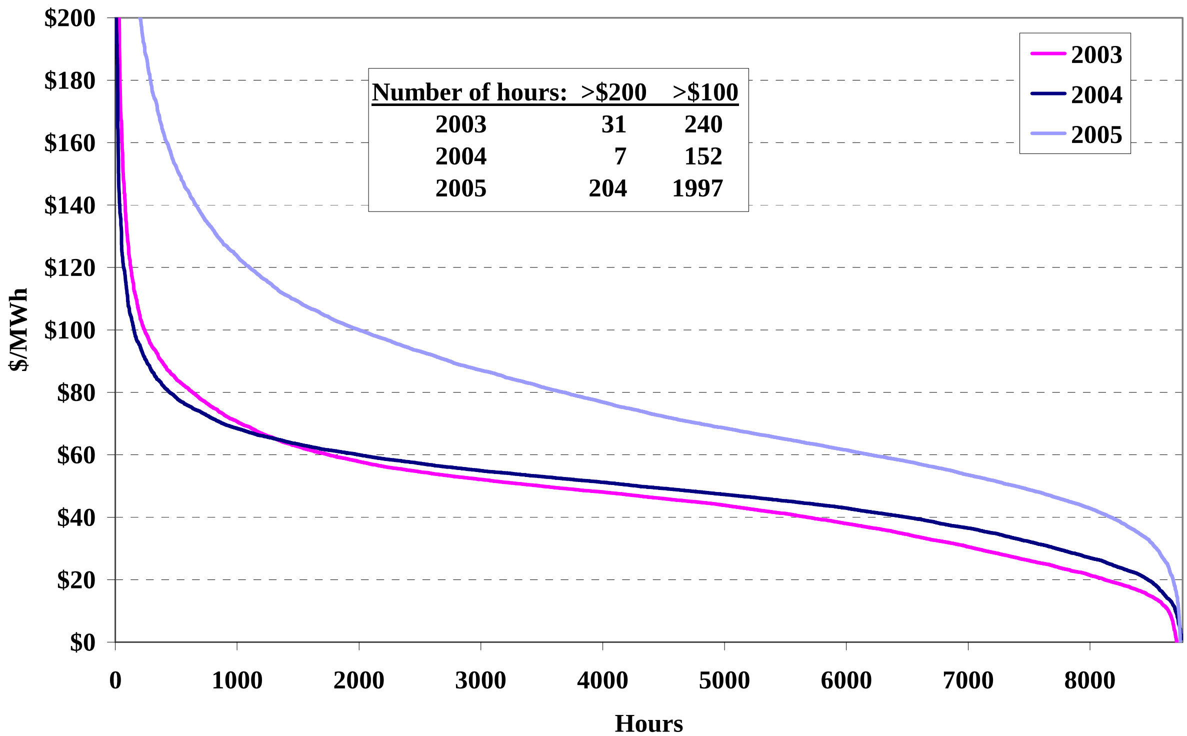


Figure 2 shows with more clarity that higher summer demand levels in 2005 had a notable effect on prices. During the summer months, implied heat rates in East New York were 15 to 20 percent higher in 2005 than in 2004. In the non-summer months, implied heat rates in East New

York were generally comparable between 2004 and 2005. Implied heat rates dropped significantly during months with extreme natural gas prices. This generally indicates that natural gas was frequently not on the margin during these periods (when prices are set by units burning lower-cost fuels, implied heat rates based on natural gas prices will fall).

The highest electricity demand levels occur during the hot summer months and typically result in relatively high electricity prices, particularly during periods of supply shortages. During periods of shortage, prices can rise to more than 10 times the average price levels. Hence, a small number of price spikes can have a significant effect on overall price levels. Thus, the higher demand experienced in 2005 affected overall prices through a relatively small number of hours. Conversely, the rise in electricity prices from higher fuel costs is felt in the majority of hours. The following two figures show price duration curves that characterize the impact of these factors on electricity prices from 2003 to 2005.

Figure 3: Price Duration Curve
Average Real-Time Price
2003 – 2005

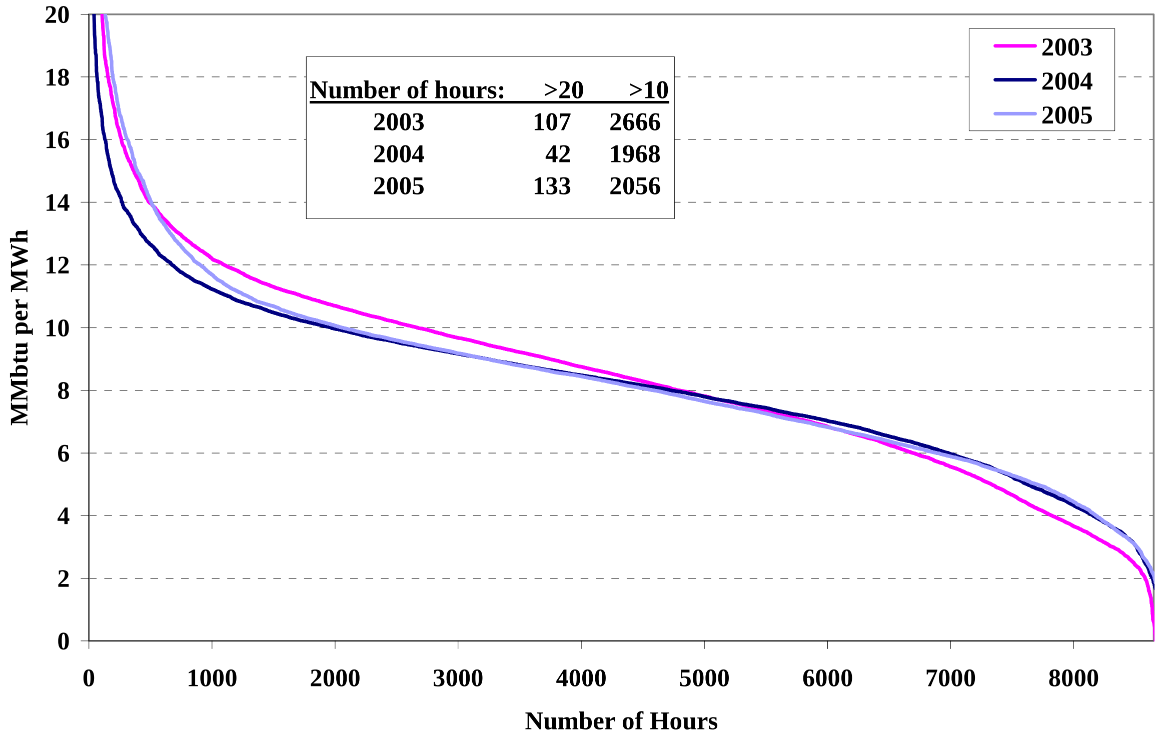


Note: Hours during the August 2003 blackout period are excluded.

Figure 3 presents a price duration curve, which shows the number of hours on the x-axis in which the market settled at or above the price level shown on the y-axis. The price shown in the figure is a state-wide average of the real-time prices. In 2005, there were 1997 hours with prices above $100, compared to 152 such hours in 2004. This rise in prices across a large number of hours reflects the effects of higher input fuel prices. In 2005, there were 204 hours with prices above $200, compared to 7 such hours in 2004. This difference is partly due to higher fuel prices, but primarily driven by the higher peak demand levels experienced in a relatively small number of hours during 2005.

To isolate factors other than fuel price changes, the following figure shows the implied heat rate duration curve for 2003 to 2005. This shows the number of hours on the x-axis in which the market settled at or above a given implied heat rate level which is shown on the y-axis. In this case, the implied heat rate is the state-wide average real-time price divided by the natural gas price.

Figure 4: Implied Heat Rate Duration Curves
Based on State-wide Average Real-Time Price and Natural Gas Price
2003 – 2005



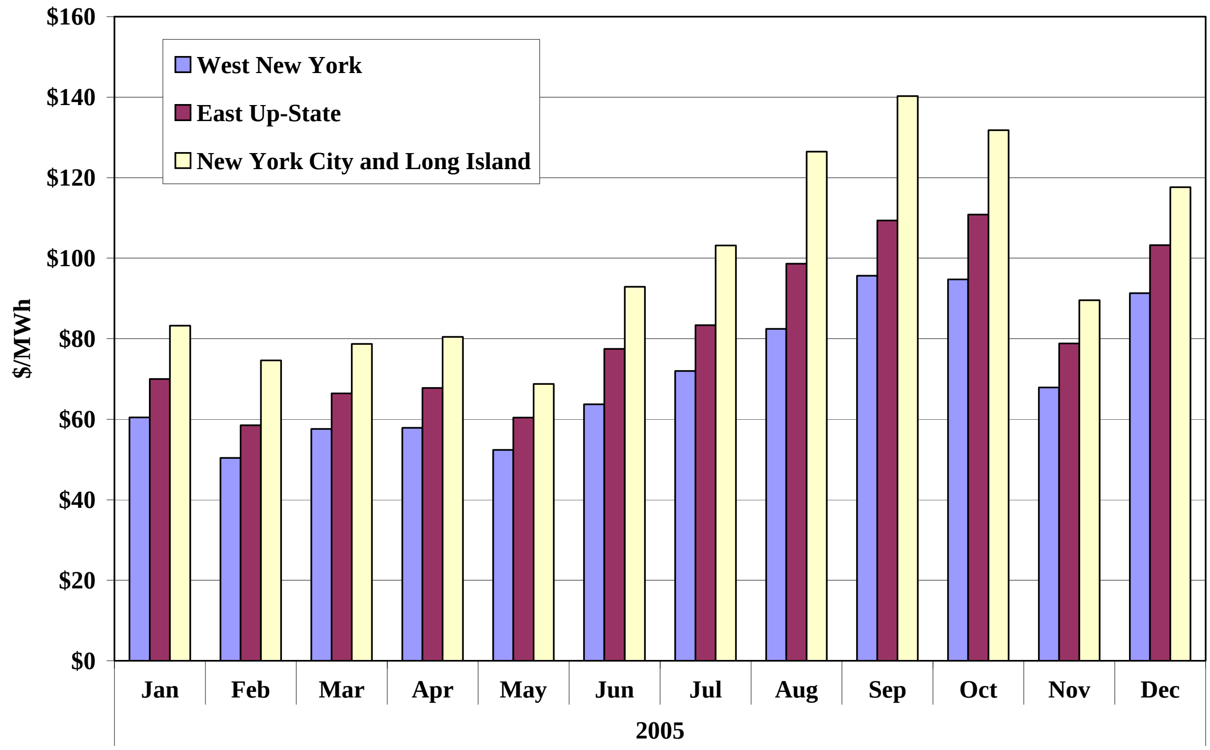
Number of hours:	>20	>10
2003	107	2666
2004	42	1968
2005	133	2056

Note: Hours during the August 2003 blackout period are excluded.

In 2005, higher demand led to 133 hours where implied heat rates exceeded 20 MMbtu/MWh, compared to only 42 such hours in 2004. Outside of peak conditions, implied heat rates were comparable between 2004 and 2005. The substantial increase in high-priced hours (i.e. hours with implied heat rates > 20 MMbtu per MWh) is due, primarily, to the hotter weather that contributed to higher peak load in many hours. Furthermore, the shortage pricing provisions lead to approximately 20 hours of shortage prices in 2005 during reserve shortages, while shortage pricing did not occur in 2004.

Figure 5 presents the monthly day-ahead energy prices in three regions in the State for 2005. Prices are lowest in Western New York, which exports significant amounts of power to Eastern New York. The prices are highest in New York City and Long Island which import a large portion of their power. Most of the power that flows from Western New York to New York City and Long Island passes through the Eastern upstate portion of the New York system. These west-to-east flows result in transmission losses and congestion that cause the pricing patterns shown in the figure.

Figure 5: Day-Ahead Energy Prices by Region 2005



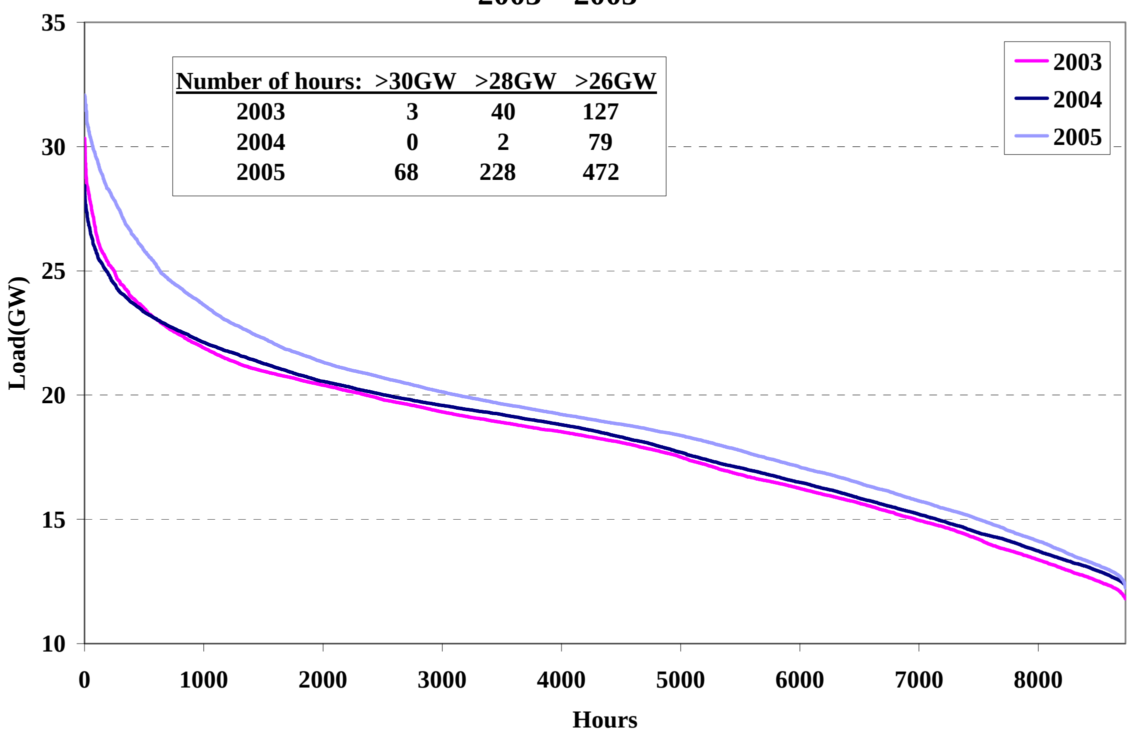
Transmission losses make up a significant share of the $12 per MWh average difference between East and West prices, while transmission constraints, particularly on the Central-East Interface and in the Hudson Valley, make up the rest of the difference. There are transmission constraints into New York City and Long Island from upstate New York, as well as local load pockets within the City which resulted in price differences between the City and the eastern upstate region averaging $16 per MWh.

The price differences between regions were driven up proportionately by fuel cost increases during the latter part of the year. However, the largest percentage differences between regions occurred during the summer months. Load is highest during the summer months, which tends to increase the transmission flows between the West and New York City. This leads to more frequent congestion and a higher proportion of transmission losses.

2. Energy Demand

One factor that has a major effect on changes in energy prices in the New York markets is the duration and timing of electricity demand. The hours with the highest loads are important because a disproportionately large share of the market costs to consumers and revenues to generators occur in these hours. Figure 6 shows load duration curves for 2003 to 2005.

Figure 6: New York State Load Duration Curves 2003 – 2005



Number of hours:	>30GW	>28GW	>26GW
2003	3	40	127
2004	0	2	79
2005	68	228	472

Note: The blackout hours during August 2003 were excluded from this figure.

The figure suggests that load growth was relatively steady in most hours from 2003 to 2005, however the peak load levels in 2005 were substantially higher than in 2004.

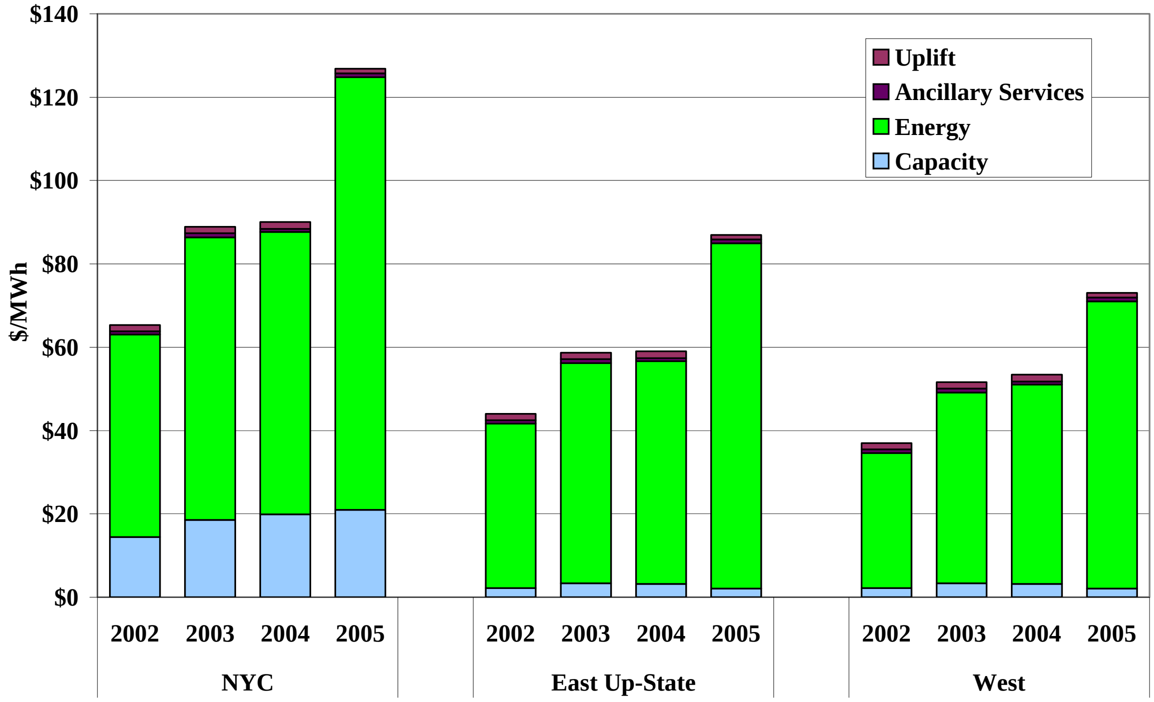
the increased frequency of tight market conditions in 2005 was generally caused by higher peak load levels. These curves show the number of hours in which the load is greater than the level indicated on the vertical axis. The absence of severe price spikes during 2004 was primarily due to mild summer demand, while in 2005, there were far more hours with extreme demand levels. In 2005, there were 68 hours when actual loads exceeded 30 GW, and no such hours in 2004. In 2005, there were 228 hours when actual loads exceeded 28 GW, and just 2 of these hours in 2004.

3. Total Market Expenses

Next we examine the all-in price for electricity which includes the costs of energy, uplift associated with Bid-Production Cost Guarantee (“BPCG”) payments, capacity, ancillary services, congestion, and losses. The all-in price is calculated for various locations within New York because both capacity and energy prices vary substantially by location. The energy prices used for this metric are real-time energy prices. The capacity component is calculated by multiplying the average capacity price (based on a weighted average of the six-month strip, monthly, and spot auctions) by the load obligations in each area, and dividing by total energy consumption. For the purposes of this metric, costs other than energy and capacity are distributed evenly for all locations. Figure 7 presents the average annual all-in price of electricity for the past four years.

Figure 7 shows that the all-in price rose considerably in 2003 and again in 2005 in the three regions. In 2003, the increase was primarily caused by higher energy prices in 2003, which rose 36 percent due to higher fuel prices. The capacity component also rose in 2003 due primarily to rising forecasted peak load resulting in a higher capacity obligations, and additional purchases under the demand curve which was instituted during 2003. In 2005, the higher energy price component is primarily due to higher fuel prices and more frequent price spikes during peak demand periods.

Figure 7: Average All-In Price
2002 – 2005



In addition to the expenses for energy, capacity, and ancillary services, there are several other categories of costs (and surpluses) that must be charged (or rebated) to consumers. The following figure summarizes the main sources of costs and surpluses that are allocated across all loads in New York State on a monthly basis. This excludes charges that are assessed only to certain locales such as the BPCG payments associated with local reliability. The following costs and surpluses are shown in Figure 8:

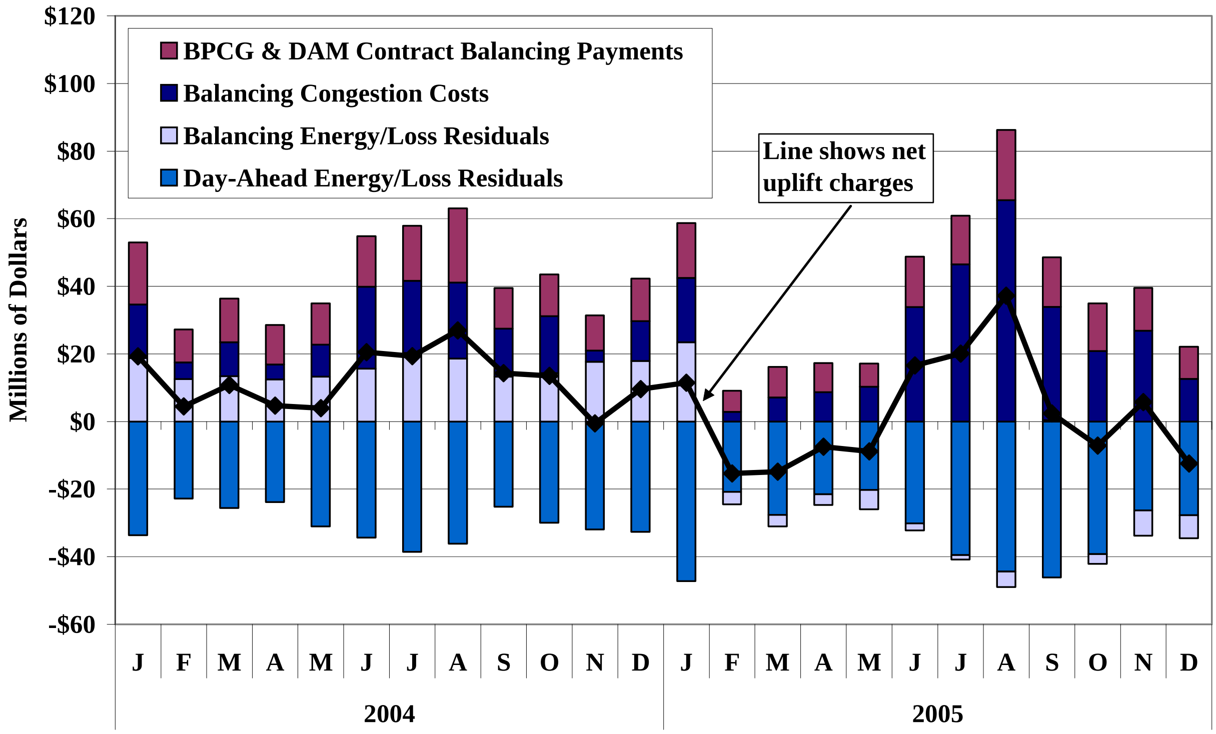
- *BPCG and DAM Contract Balancing Payments* – BPCG payments are paid to generators that operate according to ISO instructions but the resulting market revenues are not sufficient for them to recoup their as-bid costs. DAM contract balancing payments are made to generators that sell into the day-ahead market and, as a result of operating according to ISO instructions in real-time, are forced to buy back their day-ahead sale at a loss.

- *Balancing Congestion Costs* – When the ISO allows day-ahead market participants to schedule more flows across a particular interface than are feasible in real-time, any resulting real-time congestion across the interface leads to a revenue shortfall. Conversely, if the real-time flows exceed the day-ahead scheduled flows, any resulting congestion across the interface leads to a revenue surplus.

- *Energy and Loss Residuals* – In any market with locational marginal pricing ("LMP"), the marginal cost of losses exceeds the average cost of losses. As a result, the total payments for energy and losses by purchasers must generally exceed the total payments to sellers by approximately the difference between marginal cost and average cost of losses. However, these may be larger or smaller based on the scheduling pattern in the day-ahead market.

The charges and rebates spread across all NYCA load are shown on a monthly basis for 2004 and 2005 in Figure 8.

Figure 8: Uplift Charged/Rebated to All Load in NYCA 2004 – 2005



Note: The portion of Balancing Congestion Costs due to Thunder Storm Alerts is charged to the local TO.

Figure 8 shows a significant decline in net uplift charges after the implementation of SMD 2.0 for at least two reasons. First, BPCG payments were reduced due to more efficient commitment of gas turbines in real-time. When the real-time commitment models anticipates the need for gas turbines, but they are uneconomic in real-time, they may require large "make whole" payments from the NYISO. Second, balancing residual charges were largely eliminated due to improvements in the consistency of day-ahead and real-time loss modeling.

After the implementation of SMD 2.0, balancing congestion charges dropped for several months. However, tighter operating conditions in the summer led to increased uplift from balancing congestion charges. Real-time congestion price differentials were higher in 2005, contributing to an increase in balancing congestion costs. Likewise, Thunder Storm Alerts ("TSAs") reduced transmission capability on many days during the summer, leading to real-time pricing events and significantly more balancing congestion. A portion of these costs are assessed to the local Transmission Owner ("TO"). Several other factors contribute to these fluctuations in monthly uplift charges and rebates including (i) increased reliance during the summer on gas turbines, which receive a large share of the BPCG payments, and (ii) higher fuel prices that can increase a resources' out-of-merit costs and, hence, BPCG payments.

Under SMD 2.0, the ISO substantially improved the efficiency of starting-up and shutting-off gas turbines in real-time. This has led to large reductions in uplift charges for BPCG payments to units that are started by the ISO but turn out to be uneconomic based on real-time prices. We performed an analysis of decisions made by the real-time commitment software before and after the implementation of SMD 2.0 to estimate the reduction in uplift costs in 2005 attributable to the new commitment software. These estimates are summarized in Figure 9 for three types of gas turbine generators ("GTs") : quick-start GTs, older 30-minute GTs, and recently installed 30-minute GTs.

Figure 9: Estimated Reduction in Uplift from Uneconomic Gas Turbines February to December, 2005

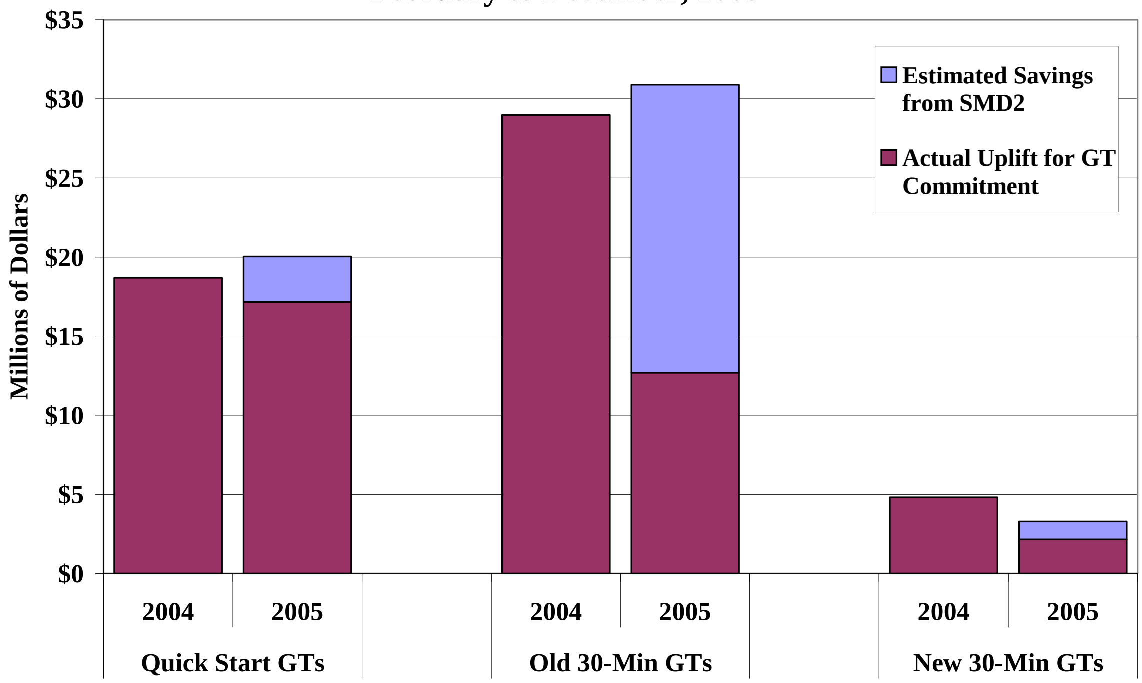
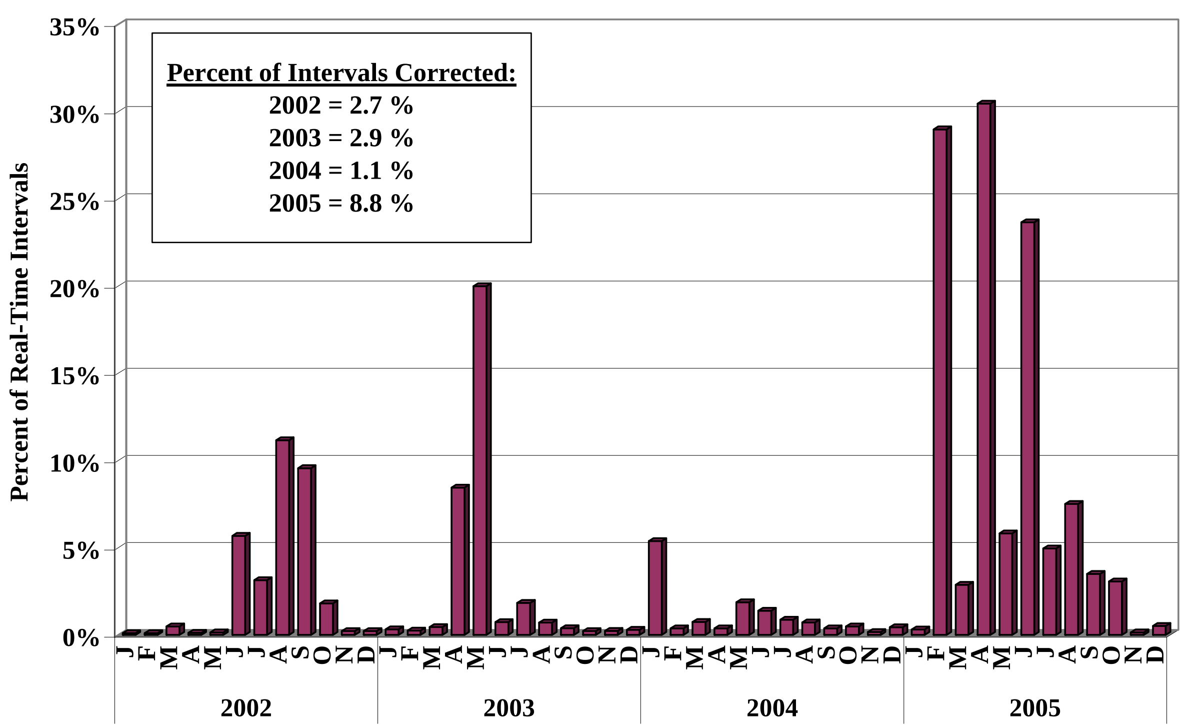


Figure 9 shows that BPCG payments for uneconomic gas turbines committed for non-local reliability reasons decreased from $52.5 million in 2004 to $32.0 million in 2005. We estimate that the implementation of SMD 2.0 reduced uplift by $22 million in 2005, with most of the reduction coming from more efficient commitment of older 30-minute gas turbines. The methodology used to estimate the uplift reduction is described in greater detail in Section V.B of this report.

B. Price Corrections

All real-time energy markets are subject to some level of price corrections to account for metering errors and other data input problems. Accurate prices are critical not only for the obvious need to settle market transactions fairly, but also for sending reliable real-time price signals to market participants. Price corrections are required when flaws in the market software or flaws in operating procedures cause prices to be posted erroneously. It is important to resolve these errors as quickly as possible to maximize price certainty. Figure 10 summarizes the frequency of price corrections in the real-time energy market from 2002 to 2005.

Figure 10: Percentage of Real-Time Prices Corrected



The rate of corrections spiked in 2005 due to issues associated with the implementation of new real-time market software under SMD 2.0. Nine significant software issues arose that led to the majority of price corrections, several of which were identified and corrected in February, 2005. The remaining software issues surfaced from March to June and were addressed between June and October. Once these initial software issues were addressed by NYISO, the frequency of price corrections fell to more typical levels by November 2005.

Temporary spikes in the frequency of price corrections have typically occurred after major software modifications in the past. The frequency of price corrections was relatively high in 2000 after the market opened, but then decreased steadily until the summer of 2002. The frequency of price corrections increased substantially in June, 2002 as a result of changes to the modeling of New York City load pockets. Once the modeling issues related to load pockets were addressed, the level of corrections returned to the low frequency that was experienced prior to the summer of 2002.

C. Net Revenues Analysis

Revenues from the energy, ancillary services, and capacity markets provide the key signals for investment in new generation and retirement of existing generation. The decision to build or retire a generation unit will depend on the expected net revenues that unit will receive in the market from sales of energy, ancillary services, and capacity. Net revenue is defined as the total revenue that a generator would earn in the New York markets less its variable production costs.

If there is not sufficient net revenue in the short-run from these markets to justify entry of a new generator, then one or more of the following conditions may be present:

- New capacity is not needed because there is sufficient generation already available;
- Load conditions are below expectations due to mild weather or reduced demand, leading to lower energy prices than expected; and/or
- Market rules are causing revenues to be reduced inefficiently.

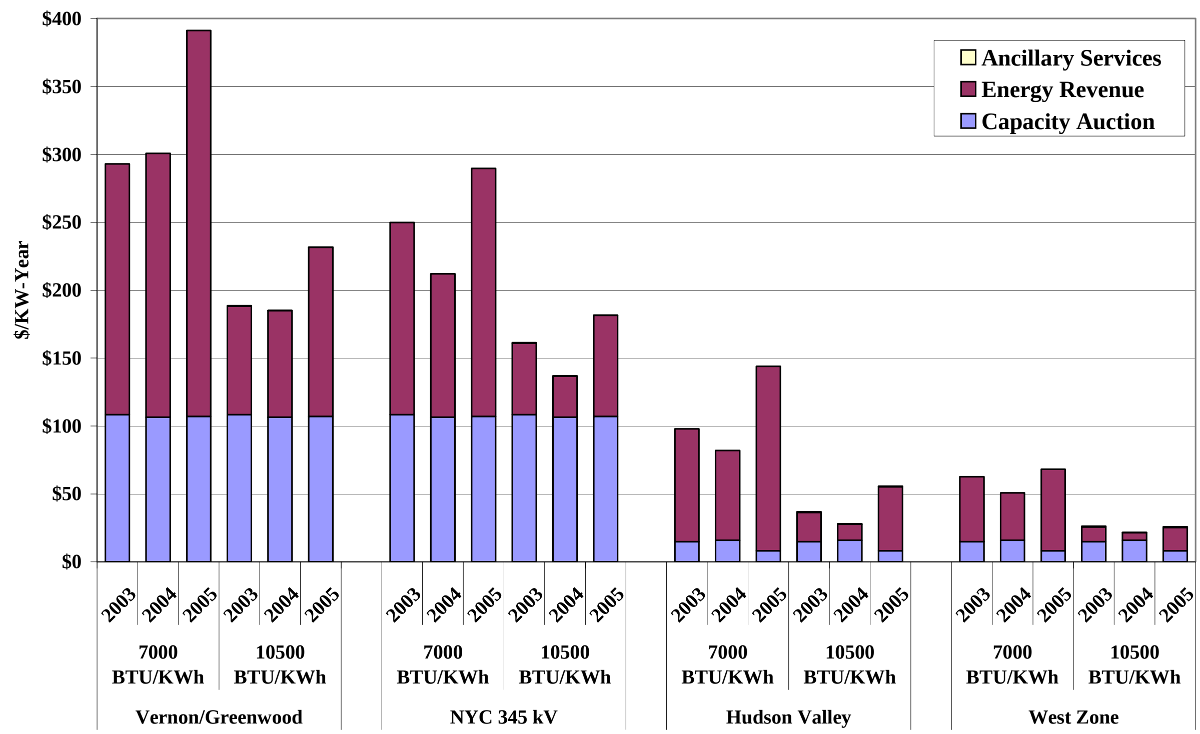
Likewise, if prices provide excessive revenues in the short-run, it might imply the opposite of one of these factors. If a revenue shortfall persists for an extended period, without an excess of capacity, this is a strong signal that markets need modifications.

In this section we analyze the net revenues that would have been received by various types of generators at four different locations: the 345kV portion of New York City, the Vernon/Greenwood load pocket in New York City, the West Zone, and the Hudson Valley Zone. We calculated the net revenue the markets would have provided to two different types of units at these locations for the last three years. The two types of units are:

- Gas combined-cycle: assumes a heat rate of 7 MMBTU per MWh and variable O&M expenses of $3 per MWh, and
- New gas turbine: assumes a heat rate of 10.5 MMBTU per MWh and variable O&M expenses of $1 per MWh.

For both unit types, the analysis assumes a forced outage rate of 5 percent. In this analysis, net revenue is equal to the average hourly real-time price minus variable production expenses in hours when the price is greater than the variable production expenses. The assumptions for this analysis have been standardized by FERC and the market monitors in the various markets to provide a comparable basis for comparison of the net revenue values from the different markets. Figure 11 shows net revenue estimates for four locations in New York from 2003 to 2005.

Figure 11: Estimated Net Revenue in the Day-Ahead Market
2003 – 2005



The figure indicates that net revenue increased significantly during 2005, primarily due to tighter load conditions. In 2005, there were reserves shortages that resulted in very high energy prices under the shortage pricing provisions, whereas there were no instances of shortage prices in 2003 or 2004. Capacity market revenues remained relatively constant over the three years. Despite the increased energy prices, net revenue clearly remained below the levels necessary to justify new investment in gas turbines outside New York City and in combined-cycle units in western New York.

Based on market conditions in 2005, there are several locations where it might be profitable to build new capacity. Increased shortage pricing in eastern New York and higher fuel prices raised combined cycle net revenue in the Hudson Valley to almost $150 per kW-year, which exceeds most estimates of the their investment costs.[4] A new gas turbine would earn net revenue of approximately $180 per kW-year outside of the load pockets in New York City and more than $200 per kW-year inside the load pockets. These levels are close to the estimated annual cost (including return on investment) of building a new gas turbine in the City.[5] Likewise, net revenue for a new combined cycle unit ranged from almost $300 to almost $400 per kW-year, which likely exceeds the net revenue needed to justify investment in New York City. These results are consistent with market conditions in New York City, which has been relatively close to being capacity-deficient.

Although estimated net revenue grew considerably in 2005 to levels that would likely justify new investment in some areas, there are other factors that help to determine incentives for new investment. Market participants must anticipate how prices will be affected by the new capacity investment, future load growth, and increasing participation in demand response. Market participants must project these factors over the life of the investment, which involves risks that may be hedged through bilateral contracts. Over the long-term in a well-functioning market, we expect net revenue estimates and other signals for investment to be consistent with the adequacy of supply.

---

4 Energy Information Administration Annual Energy Outlook 2005 includes an estimate that does not take into account the extra costs of building in a densely populated area such as New York City.

5 Levitan & Associates, Inc., *Independent Study to Establish Parameters of the ICAP Demand Curves for the New York Independent System Operator (2004).*

However, the net revenue estimates produced using the standardized net revenue assumptions described above may differ from the actual net revenues earned by market participants for several reasons. First, the assumptions over-simplify certain aspects of the operations of combined cycles and gas turbines. Gas turbines may have significant start-up costs, while combined cycle generators have significant start-up costs, lengthy start-up lead times, and minimum run time requirements that exceed one hour. Because net revenue is calculated on an hourly basis ignoring, commitment considerations, the net revenues shown in this section tend to be overstated.

Second, the variable fuel expenses in this analysis are based on day-ahead natural gas price indices, although some generators may incur higher costs to obtain natural gas. Gas turbines frequently purchase natural gas in the intraday market, which generally trades at a slight premium to the day-ahead price used in this analysis. Therefore, the net revenue for gas turbines is likely to be slightly higher than an actual new unit would realize. Combined cycle units may also incur additional fuel charges when the amount of fuel they burn in real-time differs from the amount of fuel they nominated day-ahead.

Third, the net revenue estimates are based on real-time clearing prices, which may not be representative for generators that typically sell into the day-ahead market. Combined cycle generators sell most of their output in the day-ahead market because it allows them to ensure that they will recoup their commitment costs. Gas turbines have less substantial commitment costs and generally sell more of their output into the real-time market, although many gas turbines sell significant amounts in the day-ahead market. Since day-ahead prices were lower than real-time prices in 2005 on average, the net revenues tend to be over-stated for combined cycle units.

II. Convergence of Day-Ahead and Real-Time Prices

The day-ahead market allows participants to make forward purchases and sales of power for delivery in the real-time. This is a valuable financial mechanism that allows participants to hedge their portfolios and manage risk. Loads can insure against volatility in the real-time market by purchasing in the day-ahead market. Suppliers can avoid the risk of starting-up their generator on an unprofitable day since the day-ahead auction market will only accept their offer when they will profit from being committed. However, suppliers that sell in the day-ahead market are exposed to some risk, because they are committed to deliver physical quantities in the real-time market and an outage could force them to purchase replacement energy from the spot market during a price spike.

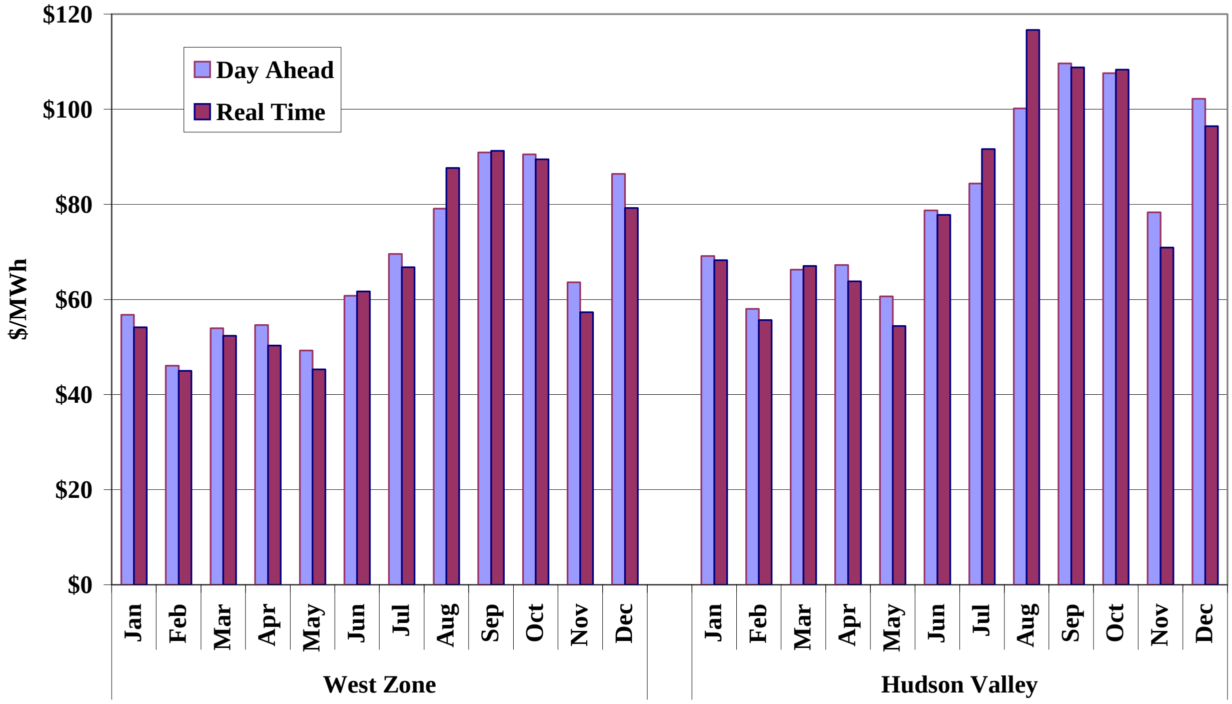
In a well functioning system with day-ahead and real-time markets, we expect that day-ahead and real-time prices will not systematically diverge from one another. If day-ahead prices were predictably *higher* than real-time prices, buyers would shift more of their purchases to the real-time. Alternatively, if day-ahead prices were foreseeably *lower* than real-time prices, buyers would increase their purchases in the day-ahead. Sellers would show the opposite tendencies. These incentives have historically led the average day-ahead prices to be relatively consistent with the average real-time prices in New York and in multi-settlement markets in other regions, although it has been common for day-ahead prices to carry a slight premium over real-time prices. Price convergence is desirable because it promotes the efficient commitment of generating resources and scheduling of external transactions.

In this section, we evaluate three aspects of convergence in prices between day-ahead and real-time markets: (i) energy prices at the zone level, (ii) energy prices in load pockets within the New York City zone, and (iii) ancillary services prices.

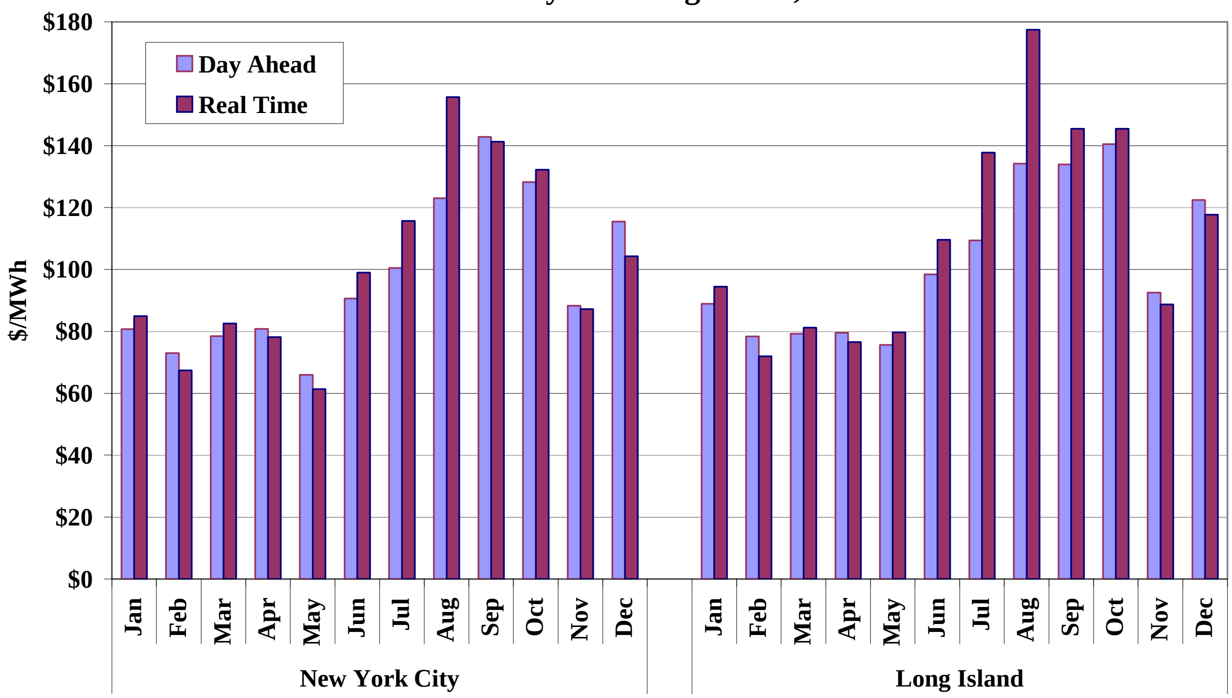
A. Energy Price Convergence

Figure 12 shows monthly comparisons of the average day-ahead and real-time energy prices in the West zone, Hudson Valley, New York City zone, and Long Island during 2005.

Figure 12: Day-Ahead and Real-Time Energy Price Convergence
West Zone and Hudson Valley, 2005



New York City and Long Island, 2005



Outside of the summer months, all four regions exhibited a slight day-ahead premium, consistent with previous years. However, during the summer (particularly in August), real-time prices were substantially higher than day-ahead prices. In the West Zone and the Hudson Valley, the real-time price premiums averaged 3 percent and 9 percent during the summer. The real-time premiums were considerably larger in New York City and Long Island during the summer, averaging 18 percent and 24 percent, respectively.

A close review of the underlying data indicates that a small number of real-time price spike events account for most of the real-time premium during the summer. The average premium would be just 5 percent for the summer in New York City if the eight afternoons with real-time premiums exceeding $175/MWh were excluded. Many of the price spikes were caused by Thunder Storm Alerts ("TSAs"), which require double contingency operation of the ConEd overhead transmission system in real-time but not in the day-ahead market. TSA operation was a major reason why real-time premiums exceeded $175/MWh on five of the eight afternoons.

TSAs require double contingency operation of the ConEd overhead transmission system in real-time. This effectively reduces the amount of power that can flow from Up-state New York through the Hudson Valley to New York City and Long Island. When TSAs coincide with high demand levels, price spikes can occur in the areas that have reduced import capability. For instance, TSA operation caused the Leeds-to-Pleasant Valley transmission line to be constrained with shadow prices exceeding $1000 per MWh in 159 intervals. This is a major path for power flowing from the Capital region toward New York City. Furthermore, 51 percent of the 259 price spike intervals resulting from Eastern 10-minute reserve shortages occurred when a TSA was in effect. Day-ahead prices were far below real-time prices on these afternoons, indicating that market participants did not anticipate the resulting shortages. Otherwise, market participants would have scheduled additional virtual load in the day-ahead market in order to arbitrage day-ahead and real-time prices.

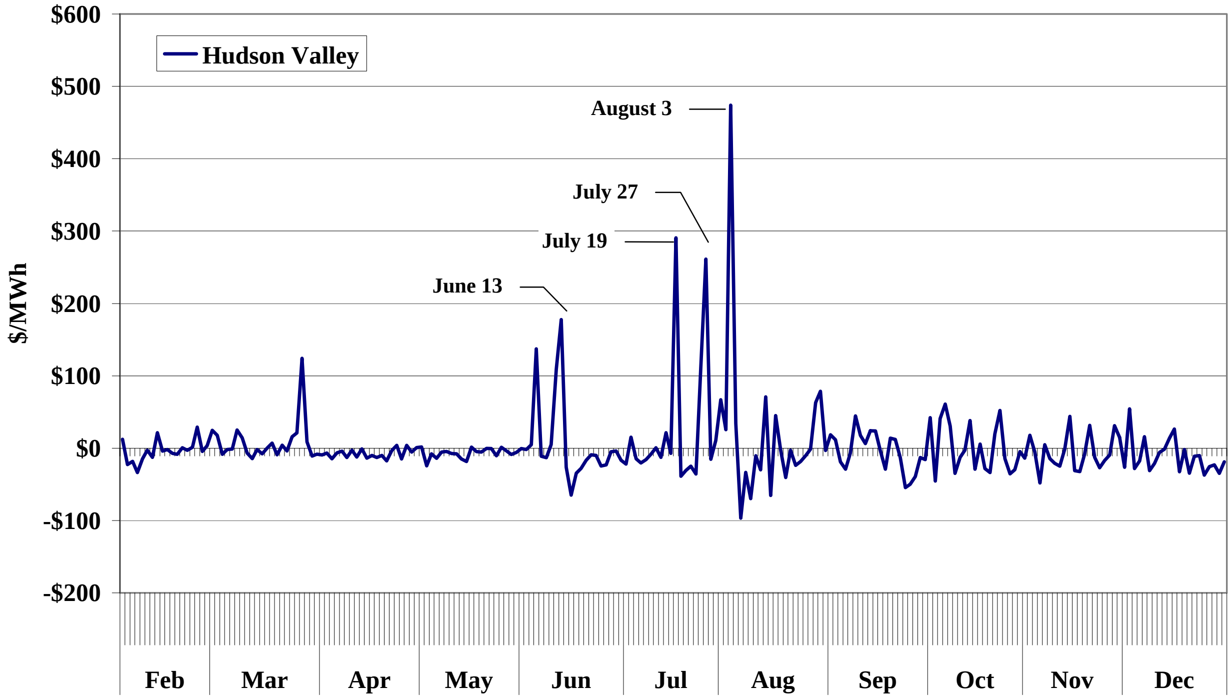
Substantial real-time price premiums are not likely to persist in the future as participants revise their expectations of real-time prices to include a greater likelihood of shortage prices. Higher expected real-time prices induce market participants to profit by scheduling additional virtual load. Additional virtual load will raise prices in the day-ahead market, bringing them into better convergence with real-time prices.

The factors that dictate real-time prices on a particular day are inherently random and difficult to predict, leading day-ahead and real-time prices to differ significantly from one another on individual days. On the following page, Figure 13 shows the variation in these differences on a daily basis in the Hudson Valley and New York City zone during afternoon hours from February to December, 2005. The figure shows that while day-ahead and real-time premiums occurred with similar frequency, there were more afternoons during which real time prices were substantially higher than day ahead prices, than afternoons with day ahead prices much higher than real time prices. The average real-time premium exceeded $100/MWh on nine afternoons in the Hudson Valley, and on 13 afternoons in New York City. The figure demonstrates that a small number of peak pricing events in the real-time market are primarily responsible for the poor overall convergence.

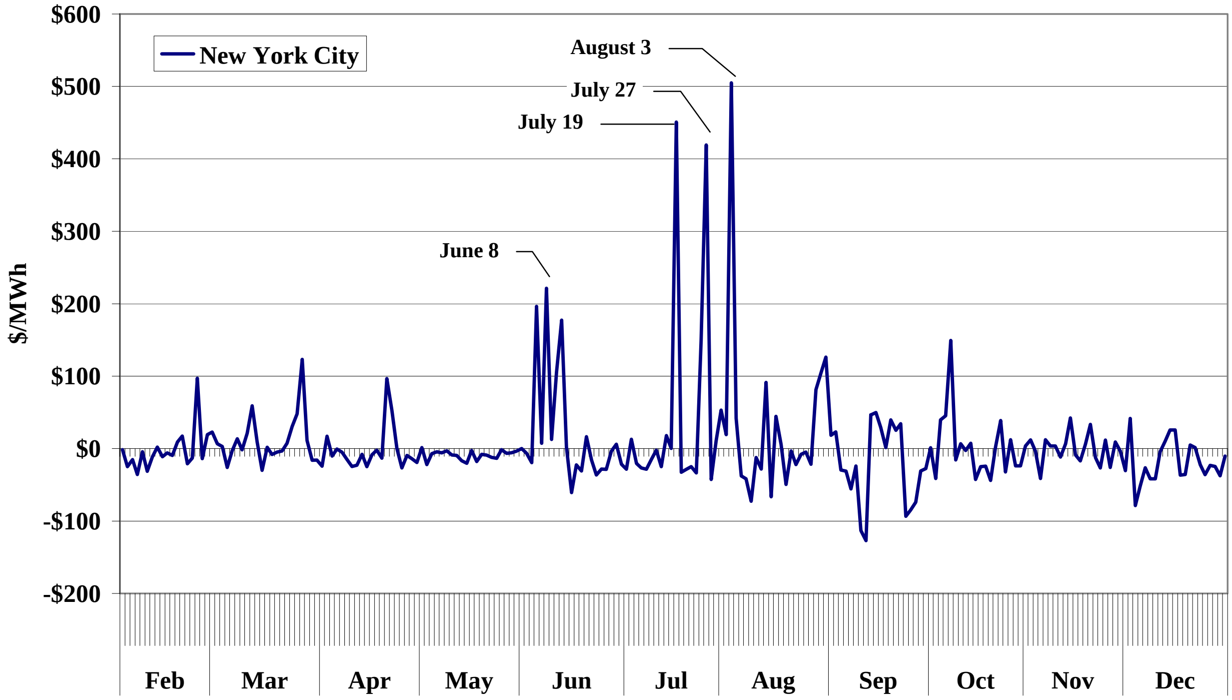
Differences between day-ahead and real-time prices increased relative to the spring months. For instance, in the Hudson Valley, the average daily difference shown in the following figure was $11 per MWh during the spring and $24 per MWh during the fall. This increase may be attributed to two factors. First, increased fuel prices resulted in higher overall prices and proportionately larger day-ahead to real-time price differences. Secondly, fuel prices were more volatile during the fall and inherently harder to predict, leading to greater differences between day-ahead expectations and real-time outcomes. The average daily change in the natural gas price was 5.5 percent during the fall, and just 2.8 percent during the spring. Over time, market participants will gain experience that should improve their ability to forecast real-time condition, and lead to better convergence between day-ahead and real-time prices.

Figure 13: Average Daily Real-Time Energy Price Premium 1 p.m. to 7 p.m., Weekdays – February to December, 2005

Hudson Valley



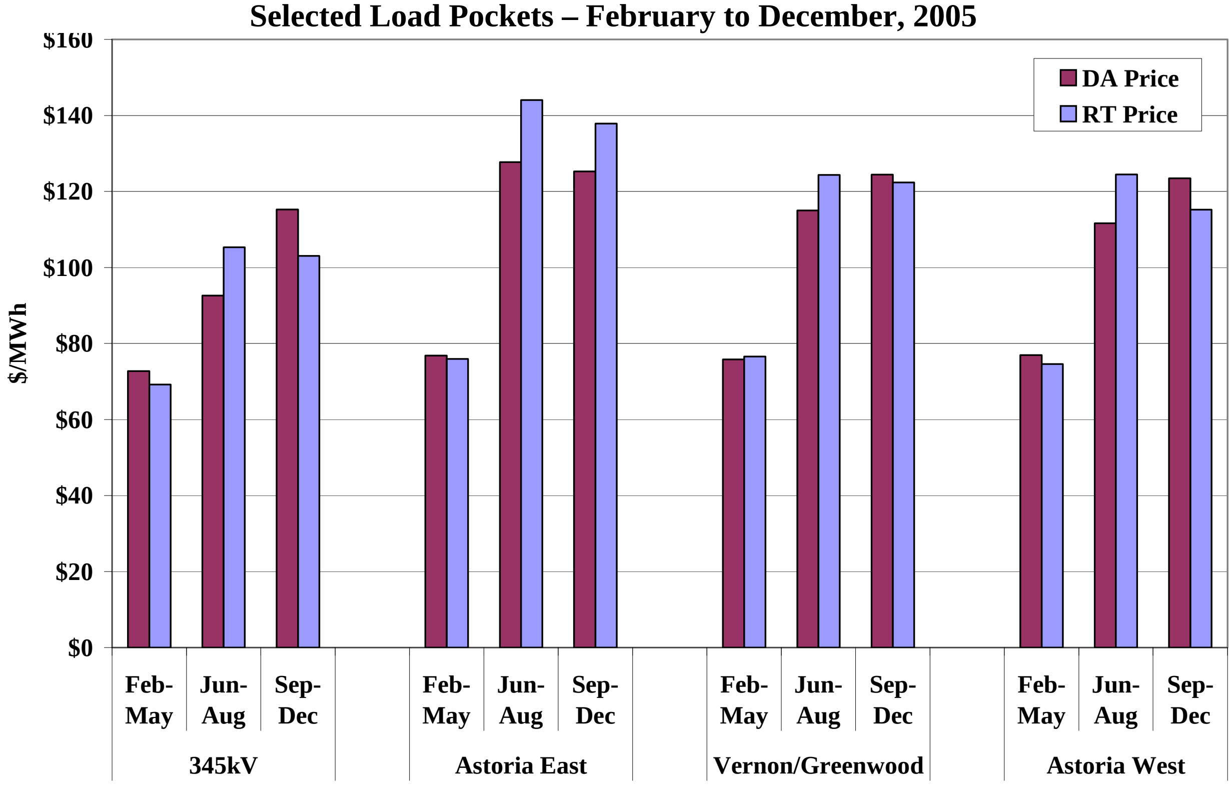
New York City



B. Price Convergence in New York City Load Pockets

Transmission congestion can be significant within New York City leading to a wide variation in prices throughout the zone. The New York City zone price is a load-weighted average price based on the locational prices in each of the load pockets in the City. Therefore some locations may experience significant divergence in day-ahead and real-time prices that are off-set by divergences in the opposite direction at other locations. In general, virtual trading and price-sensitive load bidding help improve convergence by facilitating arbitrage of day-ahead and real-time prices. However, the NYISO is currently unable to allow market participants to submit virtual trades and price sensitive load bids at the load pocket level, only at the zonal level, making it possible for price convergence to be good at the zonal level, but not in individual load pockets. This sub-section examines price statistics for various New York City load pockets to assess the extent to which day-ahead and real-time prices converge there. Figure 14 shows average day-ahead and real-time prices for four areas which account for more than 80 percent of the load within New York City.

Figure 14: Day-Ahead and Real-Time Energy Prices in New York City Selected Load Pockets – February to December, 2005



The 345kV system is usually the lowest-priced area in the City, while the Astoria East load pocket is typically a high-priced area. Indeed, during the summer months, the average real-time price was $105 per MWh in the 345kV system and $142 per MWh in Astoria East. The Vernon/Greenwood and Astoria West/Queensbridge load pockets are also shown in Figure 14.

Figure 14 shows that during the summer, all four regions exhibited a substantial real-time price premium, consistent with the rest of New York state. Between September and December, Astoria East exhibited a $13 per MWh *real-time premium*, while the area outside the 345 kV system experienced a $12 per MWh *day-ahead premium*. Over the same period, the New York City zone experienced a $2.50 per MWh day-ahead price premium. The large systematic differences in the load pockets would provide significant incentives for market participants to arbitrage, but virtual trading is limited to the zonal level. This is primary reason why price convergence has been far better at the zonal level than at the load pocket level.

Inconsistencies between the day-ahead and real-time market models may also contribute to the lack of consistency between day-ahead and real-time prices at the load pocket level. Through early 2006, a simplified representation of the intra-New York City constraints has been used in the real time market, while a more detailed representation of the New York City system has been used in the day-ahead market. The more detailed representation generally allows more efficient use of the transmission network, resulting in greater transfer capability in the day-ahead market. This difference has likely contributed to the divergence between the day-ahead and real-time prices within New York City. In June, 2006, the NYISO implemented the modeling of individual transmission lines and contingencies in New York City (rather than simplified interfaces) in the real-time market. In addition to allowing better use of the transmission system, it should also help improve convergence in the load pockets.

C. Ancillary Services Price Convergence

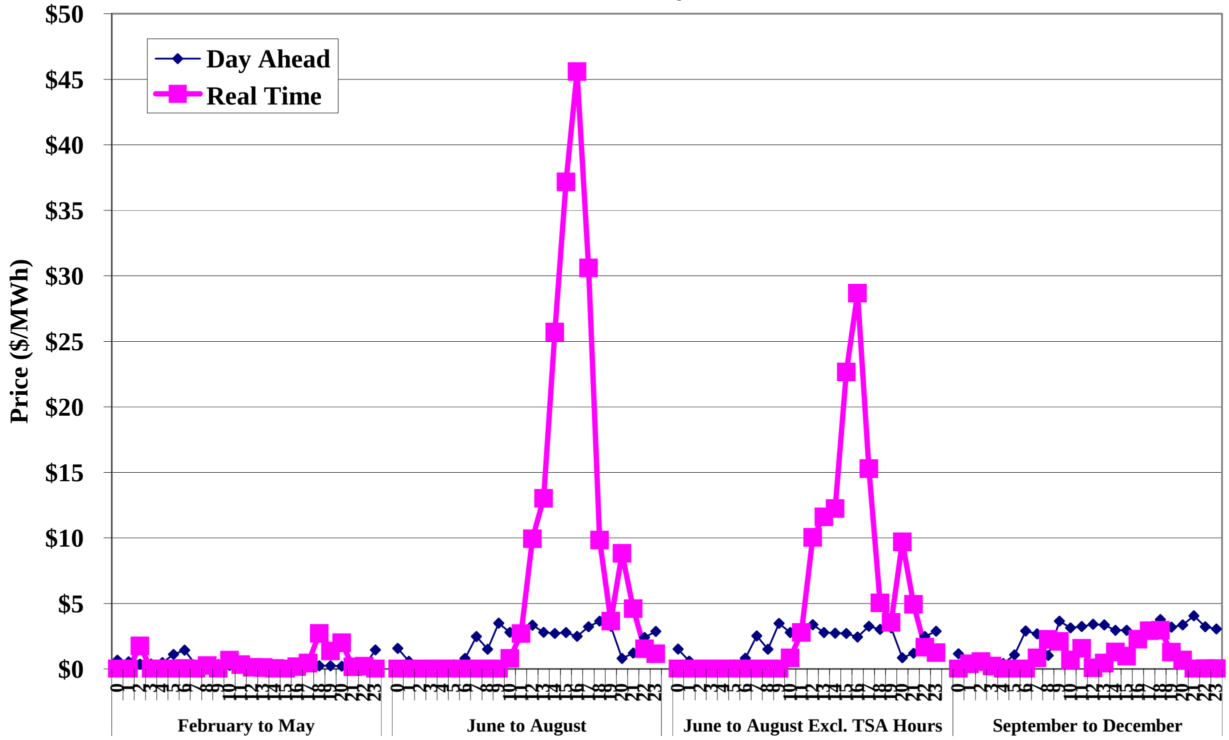
Under SMD 2.0, the New York ISO integrated real-time ancillary services markets with the existing real-time energy market, complementing the day-ahead market which has included markets for energy, reserves, and regulation since 1999. The energy and ancillary services markets place demand on the same supply resources, so prices for energy and ancillary services are highly correlated, and scarcity in the energy market is generally accompanied by a scarcity of ancillary services. This sub-section examines ancillary services price statistics to assess how well day-ahead and real-time prices converge.

As with the energy market, suppliers who sell ancillary services into the day-ahead market have the obligation to either provide the service from their own resources in real-time or pay the real-time price in order to provide the service from the pool. When systematic differences arise between day-ahead and real-time prices, ancillary services suppliers are the only entities that arbitrage them by shifting sales to the higher priced market. Unlike the market for energy, load serving entities cannot submit price sensitive bids to purchase ancillary services in the day-ahead market. The same amount of ancillary services must be purchased in the day-ahead and real-time markets in a non-price sensitive manner, thereby preventing load serving entities from arbitraging day-ahead to real-time ancillary services price differences. Moreover, virtual trading of ancillary services products is not currently allowed, so market participants without physical resources cannot arbitrage day-ahead to real-time price differences.

The following chart shows day-ahead and real-time 10-minute reserves prices in eastern New York by hour of the day for several periods during 2005. The eastern 10-minute reserves price is important because the ISO requires 1,000 MW of 10-minute reserves east of the Central-East interface.[6] This particular requirement is typically the most costly for the ISO to satisfy due to the relative scarcity of capacity in eastern New York. The summer months are shown with and without TSA periods, demonstrating that TSAs had a substantial impact in 2005.

6 The NYISO actually requires 1,200 MW of 10-minute reserves in eastern New York, however, 200 MW of this requirement is generally satisfied through a reserves sharing agreement with New England.

**Figure 15: Day-Ahead and Real-Time 10-Minute Reserves Prices
Eastern New York – February to December, 2005**

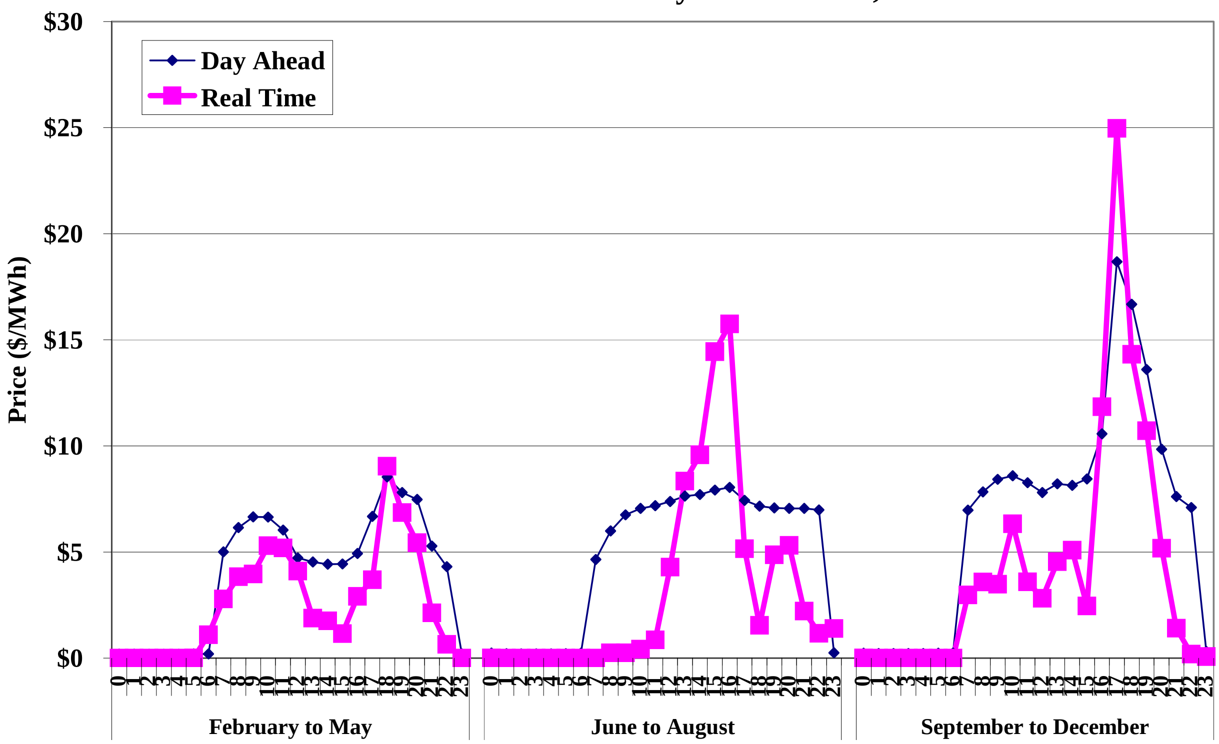


During the summer, real-time prices for 10-minute reserves were significantly higher than other times of year, while day-ahead prices rose only modestly. During daytime hours, average day-ahead prices ranged from $3 to $5 per MWh, while average real-time prices ranged from $0 to $47 per MWh. The figure shows that the lack of convergence between day-ahead and real-time prices was substantially affected by a small number of TSA events. Typically, real-time reserves prices are very close to zero, but under conditions of near-shortage, prices can rapidly rise to hundreds of dollars per MWh. The $47 per MWh price in the peak afternoon is an average across many days when the price was very low and a small number of peak pricing events. This volatility is difficult for market participants to predict in the day-ahead market, and based on the figure above, the day-ahead market systematically under-valued 10-minute reserves in the east during the summer.

In the fall, the price pattern was reversed as day-ahead prices remained at the levels that prevailed during the summertime, while real-time prices sank below day-ahead prices. Real-time prices fell due to the infrequency of real-time peak pricing events. Since reserves suppliers must submit $0 per MWh availability offers in the real-time market, the real-time price is based on the opportunity costs of generators whose energy production is backed-down in order to provide reserves. In the majority of hours, excess reserves are available on on-line generators and off-line quick start resources, leading the real-time price of reserves to be close to $0. Reserves suppliers are able to submit availability offers to the day-ahead market that exceed $0, and we find that a change in day-ahead offer patterns explains the rise in day-ahead prices. This analysis is discussed in Section III.

The following figure shows day-ahead and real-time 10-minute synchronous reserves prices in western New York by hour of the day for several periods during 2005. This price depends primarily on the state-wide 10-minute synchronous reserves requirement of 600 MW. We show western rather than eastern New York because the eastern New York price is heavily influenced by other reserves requirements.

Figure 16: Day-Ahead and Real-Time 10-Minute Spinning Reserves Prices Western New York – February to December, 2005



The relationship between day-ahead and real-time prices varied greatly by time of day and the season. In the Spring, day-ahead prices were generally slightly higher than real-time prices. During the Summer, day-ahead prices substantially exceeded real-time prices during the morning and evening hours while real-time prices were much higher on average in two of the afternoon hours. In the fall, day-ahead and real-time prices rose relative to the summer, which is counter-intuitive given that energy demand is lower during the fall.

During the fall, 10-minute spinning reserves prices were likely pushed up by higher overall fuel prices which generally lead to higher opportunity costs for on-line generators. Reserves offer patterns play no role in real-time reserves prices since generators must submit $0 availability offers. However, day-ahead 10-minute spinning reserves prices are based on the availability offers of individual generators as well as the opportunity costs of providing reserves rather than energy. There has been a rise in offer prices which has likely had an upward effect on clearing prices. This pattern is examined in greater detail in Section III.

Figure 17 shows day-ahead and real-time regulation prices in New York by hour of the day for several periods during 2005. This price depends primarily on the state-wide regulation requirement which typically varies between 150 MW to 275 MW according to the time of day. More regulating capability is needed in the morning when the system is ramping up and in the evening when the system is ramping down.

Figure 17: Day-Ahead and Real-Time Regulation Prices
State-wide Price – February to December, 2005

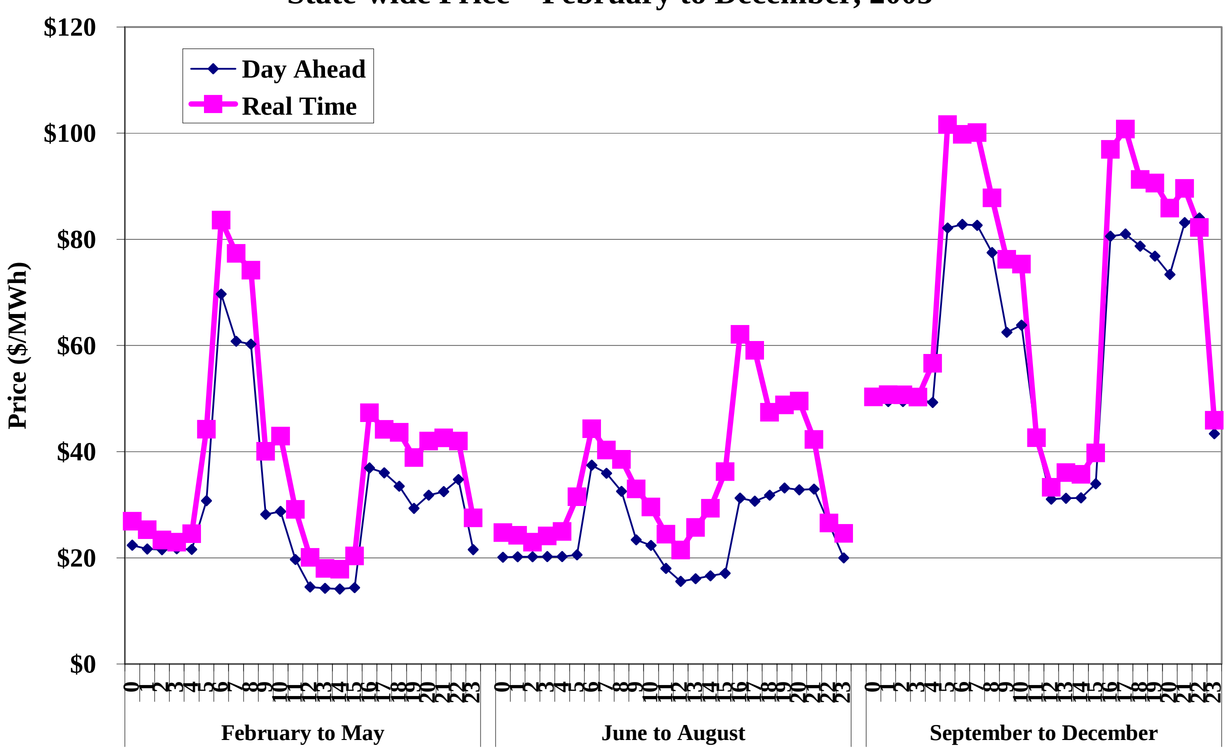


Figure 17 shows that regulation clearing prices fluctuate significantly according to the time of day. During the morning and evening ramping hours, the regulation capacity requirement rises. This leads to higher prices in these hours while prices are relatively low overnight and in the afternoon. Day-ahead and real-time regulation prices are highly correlated to one another across the day, with real-time prices being consistently higher than day-ahead prices. This reflects that real-time spikes in the regulation price occur with greater frequency than spikes in the day-ahead.

Figure 17 also shows a marked increase in regulation prices from the summer to the fall at all times of day for at least two reasons. First, there was a significant rise in regulation offer prices during September which contributed to higher clearing prices. This pattern is discussed in greater detail in the following section. Secondly, higher fuel prices during the fall led inexpensive energy producers to have higher opportunity costs when providing regulation.

Overall, the analyses in this sub-section indicate that average real-time prices for ancillary services were generally higher than average day-ahead prices during the first year under SMD 2.0. This is because price spikes related to reserves shortages occur more frequently in the real-time than in the day-ahead market. Reserves shortages never occur in the day-ahead market, because sufficient capacity is offered into the day-ahead market. However, unforeseen conditions such as forced outages and short term ramp constraints can occur resulting in real-time reserves shortages. Under-forecasted demand in the day ahead can result in under-commitment and real-time reserves shortages.

Market participants should respond to the possibility of real-time price spikes by bidding up the clearing price in the day-ahead market according to the likelihood that a real-time shortage will occur. However, the current market rules do no allow load serving entities and virtual traders to arbitrage day-ahead to real-time price differences by adjusting their ancillary services purchases or by submitting price-sensitive bids. Only generators have the ability to submit price-sensitive offers to the ancillary services markets in the day-ahead. Persistent real-time price premiums give generators an incentive to raise their day-ahead offers price, which can reduce the efficiency of the day-ahead commitment. The following section examines ancillary services offer patterns to determine whether their behavior is influenced by the persistent real-time price premiums.

III. ANCILLARY SERVICES MARKETS

A. Background

The NYISO operates ancillary services markets in conjunction with the day-ahead and real-time energy markets. Under the SMD 2.0 market enhancements, the NYISO became the first wholesale market operator to co-optimize energy and ancillary services every five minutes in the real-time market, and to use demand curves for real-time ancillary services procurement under shortage conditions. These initiatives substantially improved the efficiency of energy and ancillary service dispatch levels and prices in real time, particularly during shortage conditions. This section reviews the performance of the ancillary services markets in 2005, highlighting the effect of the market design changes.

1. Operating Requirements

New York procures three types of operating reserves: ten-minute spinning reserves, ten-minute total reserves (can be spinning or non-synchronous reserves), and 30-minute reserves. Ten-minute spinning reserves are held on generating units that are on-line and can provide additional output within 10 minutes. Ten-minute total reserves can be supplied by ten-minute spinning resources or ten-minute non-spinning resources, which are typically gas turbines that are not on-line but can be turned on and be producing within 10 minutes. 30-minute reserves may be supplied by any unit that can be ramped up in 30-minutes or that can be on-line and be producing within 30 minutes.

New York also purchases regulation services, necessary for the continuous balancing of resources (generation and NY Control Area interchange) with load and to assist in maintaining scheduled interconnection frequency at 60 Hz. This service is accomplished by committing on-line generators whose output is raised or lowered (predominately through the use of Automatic Generation Control) as necessary to follow moment-by-moment changes in load.

In each hour, the New York ISO purchases approximately 1800 MW of operating reserves. Of this 1800 MW, at least 1200 MW must be ten-minute reserves (at least 600 MW must be spinning reserves and the balance may be either spinning or non-spinning). Consequently, the NYISO may purchase up to 600 MW of 30-minute reserves. There is no limit on how much spinning reserves is purchased – all 1200 MW of total ten-minute reserves (indeed, all 1800 MW of the total operating reserves) could be spinning reserves.

The procurement of reserves is also subject to locational requirements to ensure that they will be fully available to respond to possible system contingencies. Because the Central-East Interface is often constrained, maintaining reliability requires that a substantial portion of the reserves be procured in Eastern New York. Likewise, constraints on the interface between Long Island and the rest of New York has resulted in a requirement that specified amounts of operating reserves must be purchased from within Long Island.

For total ten-minute reserves (spinning and non-spinning), 1000 MW must be purchased east of the Central-East constraint, including at least 300 MW of 10 minute spinning reserves. Prior to 2002, the eastern requirement was 1200 MW. However, the requirement was lowered to 1000 MW after the NYISO and ISO-NE entered into a reserve-sharing agreement. The locational reserve requirements for Long Island oblige the NYISO to designate at least 60 MW of ten-minute spinning, 120 MW of total ten-minute, and 540 MW of total reserves (ten-minute and 30-minute) on Long Island.

Regulation capability can be purchased from anywhere within the New York Control Area. The NYISO purchased 275 MW of regulation during high-ramp hours and 150 MW during low-ramp hours in 2005. The amount of regulating capability a generating resource may sell is equal to the amount of output it can produce within 5 minutes (ramp rate per minute times 5). In addition, to qualify as a regulating unit, the unit must be able to receive and respond to a continual dispatch signal and have the ability to ramp at a rate of 1 percent of the unit's total capability per hour.

2. Ancillary Services Market Design

The design of the ancillary services markets and their interaction with the energy market changed significantly with the implementation of SMD 2.0. The new design includes the two key features not found in any other wholesale market in 2005. First, reserves and regulation are co-optimized with energy in the real-time spot market auction. Every five minutes, the model re-evaluates the most efficient allocation of energy and ancillary services based on supplier offers and real-time operating conditions. Clearing prices determined in each auction reflect the value of energy and ancillary services based on the scarcity of supply given the level of demand and transmission constraints. Prior to SMD 2.0, reserves and regulation allocations were determined on an hourly basis. Since system conditions can change quickly and unexpectedly, this sometimes resulted in reserves being held inefficiently or in areas that could not provide reliability benefit.

Second, the real-time market uses “demand curves” to limit the costs of procuring each ancillary service and to better reflect the value of ancillary services and energy in prices under shortage conditions. Without demand curves, the model would incur unlimited costs in order to maintain each megawatt of reserves or regulation. In cases of shortage when sufficient reserves did not exist, a model without demand curves would have to sacrifice reserves and this would not be reflected in the clearing prices for energy or ancillary services. In shortage conditions under SMD 2.0, the real-time model reduces reserves and regulation purchases when the procurement costs rise to extreme levels. The demand curves provide the model with a rational basis for prioritizing high-value reserves over lower-value reserves under shortage conditions and setting clearing prices that reflect this hierarchy of value appropriately.

One of the primary benefits of co-optimizing energy dispatch and ancillary services procurement in the real-time market is that the clearing price of each ancillary service is equal to the marginal cost to the system of providing the service. To the system, the marginal cost of purchasing a service is equal to the sum of the marginal generator’s (i) availability bid price and (ii) opportunity cost of not providing another product such as energy. Under SMD 2.0, all dispatchable generators offering to sell energy in real-time must also offer to provide reserves with a $0 per MWh availability bid. Thus, the real-time clearing prices for reserves are equal to the opportunity cost of not providing another product. Frequently, it is not necessary to re-dispatch generators in real-time to meet reserves requirements because excess reserve capacity is available at the tops of on-line units. During these periods, the clearing prices of reserves drop to $0 per MWh because it costs nothing to maintain reserves.

The new market design added ancillary services to the two-settlement system that has existed for energy since 1999. A two-settlement system consists of a spot market (i.e. real-time) and a forward financial market (i.e. day-ahead), whereby day-ahead financial obligations must be reconciled in the real-time market. So, a generator that is paid to sell reserves in the day-ahead market must either (i) physically provide reserve capacity in real-time or (ii) purchase reserves from the real-time market in order to meet the financial obligation. Generators that sell reserves in the day-ahead market and are dispatched by the ISO to provide energy in the real-time market are paid the real-time clearing price for energy but must still buy back reserves in the real-time market. However, since reserves are co-optimized with energy in the real-time market, the profit from selling energy will exceed the replacement price of the reserves.[7] Under SMD 2.0, real-time dispatch instructions are determined in accordance with clearing prices such that generators provide the most profitable service.

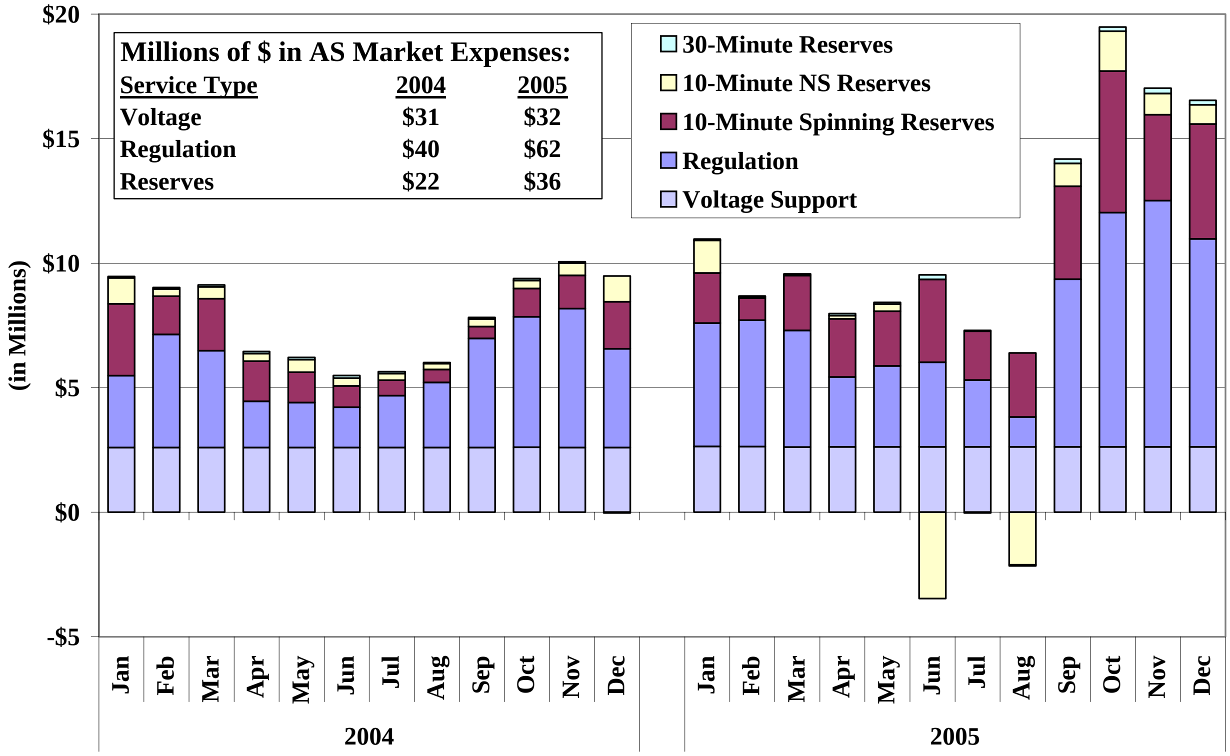
B. Ancillary Services Expenses

Since there are markets for operating reserves and regulation, expenditures for these services flucuate in response to market conditions. The NYISO also procures voltage support as an ancillary service, but voltage support is procured through contract agreements with generators. The nature of these agreements makes voltage support expenditures consistent throughout the year. Since there are markets for reserves and regulation, the expenditures for these services fluctuate based on market conditions.

Ancillary services expenses rose substantially from 2004 to 2005 due to a 54 percent increase in regulation costs and a 62 percent increase in reserves costs. Higher expenses were expected in 2005 as a result of the SMD 2.0 changes for two reasons. First, in a co-optimized market, the prices of ancillary services reflect the opportunity costs of diverting resources from the energy market so that energy price spikes are frequently accompanied with spikes in ancillary services. Second, higher fuel prices, particularly in the fall of 2005, led to increased opportunity costs for low-cost generators providing ancillary services rather than energy. Figure 18 shows expenses for regulation, voltage support, and operating reserves on a monthly basis during 2004 and 2005.

7 Suppose a generator is at a location where the price of energy is \$150 per MWh and the price of spinning reserves is \$10 per MWh, and it offers to supply energy for \$100 per MWh. The real-time model would dispatch the generator for energy since it would earn \$50 per MWh based on its offer and this is greater than the value of spinning reserves. This determination is made independent of whether the generator sold energy or reserves in the day-ahead market. If the generator sold spinning reserves in the day-ahead market, it would be paid \$150 per MWh for energy in real-time, but it would still have to purchase back its spinning reserves obligation at \$10 per MWh.

Figure 18: Ancillary Services Costs 2004 – 2005



In addition to the increase in expenses described above, Figure 18 shows another market outcome of note. First, expenditures for 10-minute non-spinning reserves were negative in June and August, 2005, meaning that there was a surplus rather than a net expense. The phenomenon results from periods when the day-ahead price is much lower than the real-time price but a larger quantity of the product was scheduled day-ahead than in real-time. This can be illustrated with the following example.

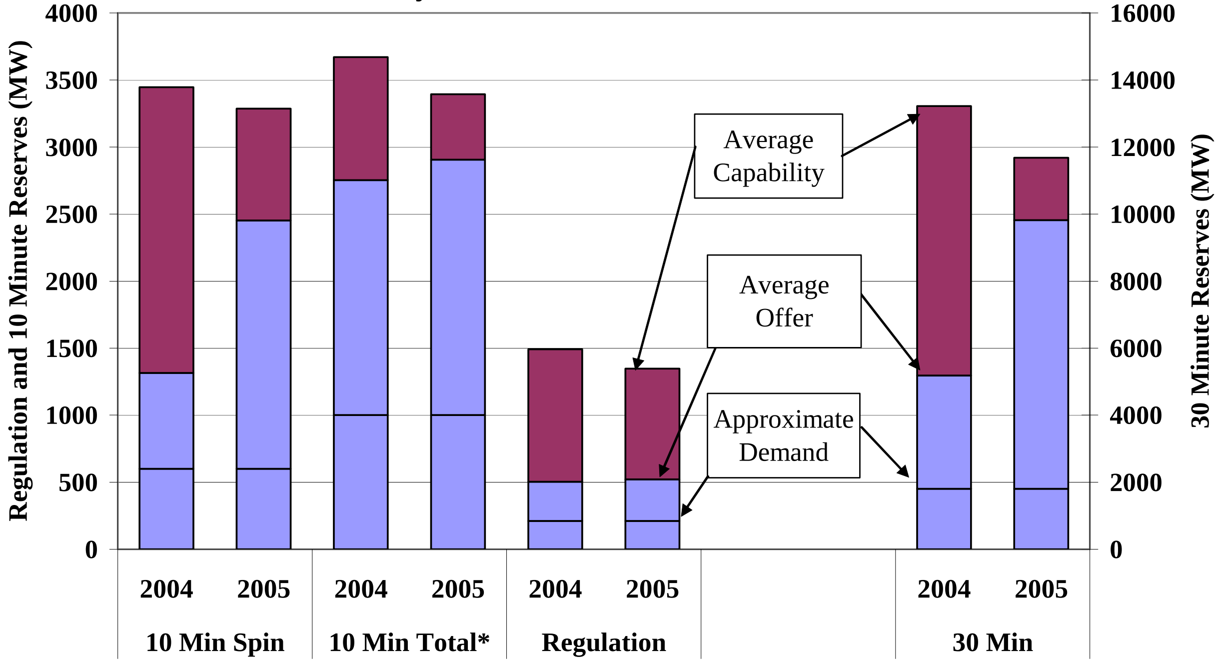
Assume that a gas turbine sells 20 MW of reserves for $5 per MWh in the day-ahead market, but is dispatched to provide energy in real-time when the price of the same reserves product is $500 per MWh. In this case, the ISO would have paid $100 (= 20 MW * $5 per MWh) to the generator for reserves in the day-ahead, and collected $10,000 (= 20 MW * $500 per MWh) back from the generator for reserves in the real-time, generating a net surplus of $9,900. Ordinarily, the ISO would have to purchase 20 MW from another generator in real-time to maintain sufficient reserves, however, the ISO might not be able to during periods of reserves shortage. In some cases, the ISO would maintain sufficient reserves by purchasing a higher value service from another generator, so that the surplus in one type of reserves would be offset by an expense for the higher value service.

Figure 18 shows dramatic growth in expenses from the summer to the fall for every ancillary service product except voltage support. Indeed, virtually all of the growth in expenses from 2004 to 2005 may be attributed to the fall of 2005. This pattern is consistent with that of day-ahead prices that were modest during the spring and summer and rose during the fall. This pattern is in sharp contrast with real-time prices that peaked during the summer at levels much higher than day-ahead prices. Since the full requirement of ancillary services is purchased in the day-ahead market, ancillary services expenses are driven primarily by day-ahead rather than real-time price levels. Although higher fuel prices led to higher opportunity costs for ancillary services providers during the fall of 2005, the primary reason for the rise in expenses was an increase in day-ahead offer prices. These offer patterns are summarized in further detail in the following sub-section.

C. Offer Patterns

Our findings in previous analyses in New York have indicated that a substantial portion of the capability of certain services was not offered in the day-ahead ancillary services markets, particularly for 30-minute reserves and regulation. This section examines ancillary services offer patterns to determine how participation has changed since the introduction of SMD 2.0. Figure 19 summarizes the average levels of capacity, offers to supply, and demand for three reserves products and regulation service in the day-ahead market. Because of the nature of the locational requirements, ten-minute reserves are shown only for the region east of the Central-East Interface.

Figure 19: Summary of Ancillary Services Capacity and Offers Day-Ahead Market, 2004 – 2005



* Eastern side of the Central-East Interface only

The quantity of offers to supply 10-minute spinning reserves and 30-minute operating reserves rose substantially from 2004 to 2005 due to the improved incentives under SMD 2.0. Under the old design, the day-ahead clearing prices were set by the highest-priced accepted offer, so it was possible for the price to be lower than the opportunity cost of not providing energy. Thus, generators risked losing profits in the energy market by providing reserves. There is no such risk under the new design since the reserves clearing price is always greater than or equal to the opportunity cost of generators scheduled for reserves. In other words, generators are always selected to provide whichever is more profitable (assuming they submit energy and ancillary services offers consistent with marginal cost).

Participation in the eastern 10-minute non-spinning reserves market did not change significantly, because most generators capable of providing the service were already required to do so. Regulation participation is still relatively low and did not change significantly after the implementation of SMD 2.0. Some generators must incur fixed costs to enable their facilities to provide regulation and may rationally choose not to make the necessary investment if they do not anticipate significant profits by participating.

Generally, participation was enhanced in 2005 relative to the period before SMD 2.0 by the redesign of the ancillary services markets. However, offer patterns have also changed markedly over the course of 2005. The following figure summarizes day-ahead offers to supply three ancillary services market requirements during the spring, summer, and fall. Offer quantities are shown in categories according to offer price level.

Figure 20: Summary of Ancillary Services Offers in the Day-ahead Market February to December, 2005

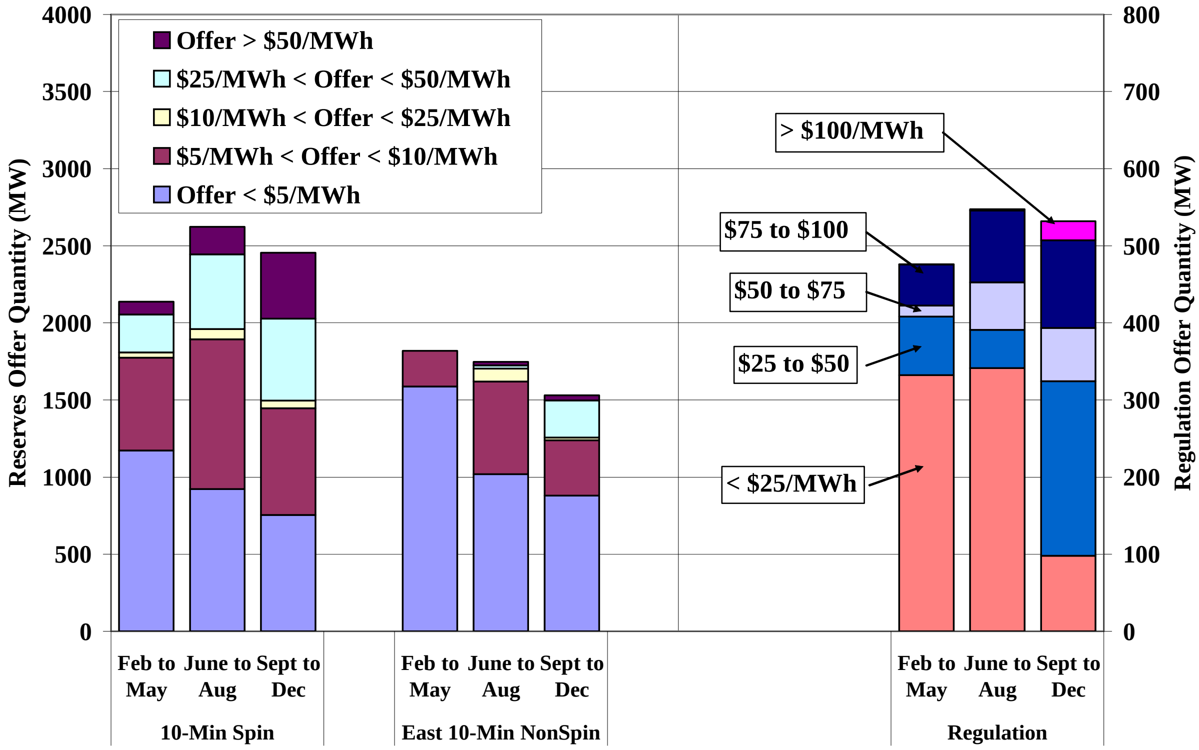


Figure 20 indicates that day-ahead offer prices rose during 2005 for several categories of ancillary services. The amount of 10-minute spin offered under $5 per MWh was approximately 1,170 MW during the spring, but this fell to 750 MW by the fall. From the spring to the fall, there was a 710 MW reduction in 10-minute reserves offers priced below $5 per MWh. Most of the regulation capacity offered during the spring and summer was priced below $25 per MWh, but the offer prices rose significantly in the fall. The higher ancillary services offer prices contributed to a significant rise in day-ahead clearing prices and market expenses from September through December, 2005.

The rise in regulation offer prices is attributable to changes in behavior by two market participants in September and October. The New York ISO's market monitoring staff reviewed the rise in offer prices from these market participants and concluded that the behavior does not warrant mitigation under the NYISO Tariff. However, due to limited participation in the market by regulation-capable capacity, the ownership of resources that participate in the market is relatively concentrated. In the short-term, the high concentration may provide incentives for certain market participants to raise their offer prices above marginal cost. In the longer-term, we expect that additional supply would enter the regulation market if prices rose above competitive levels for a sustained period.

The rise in day-ahead 10-minute spinning and non-spinning reserves offer prices, which was less dramatic than for regulation, is attributable to changes in offers by several market participants. This does not raise significant competitive concerns, since the shares of reserves capability held by these market participants are not sufficiently large to confer market power under normal market conditions. To help explain the reasons for rising 10-minute spinning and non-spinning offer prices, the following two figures summarize the offers of several suppliers from February to December 2005.

Figure 21 summarizes the day-ahead 10-minute non-spinning reserves offers of two suppliers with capacity in eastern New York, while Figure 22 summarizes the day-ahead 10-minute spinning reserves offers of two other suppliers with capacity throughout the state. The offers shown in Figure 21 and Figure 22 are characteristic of reserves offer patterns under SMD 2.0.

Figure 21: Eastern 10-Minute Reserves Offer Prices and Quantities Selected Suppliers in Afternoon Hours Day-ahead Market -- February to December 2005

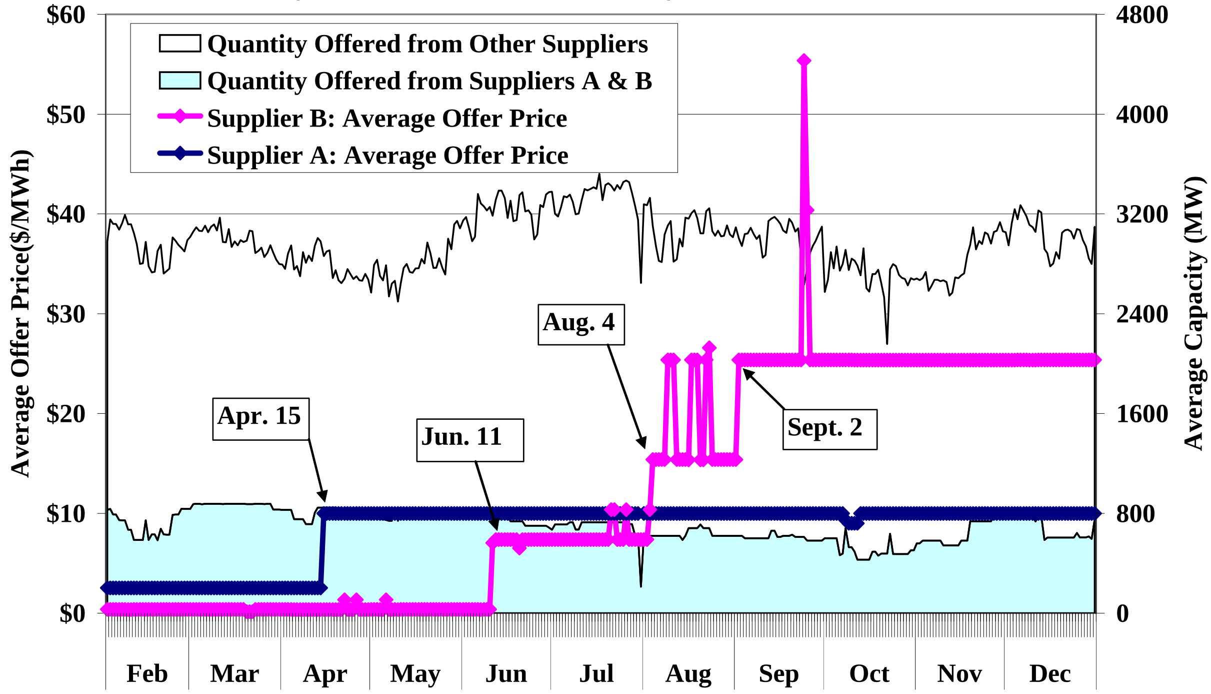
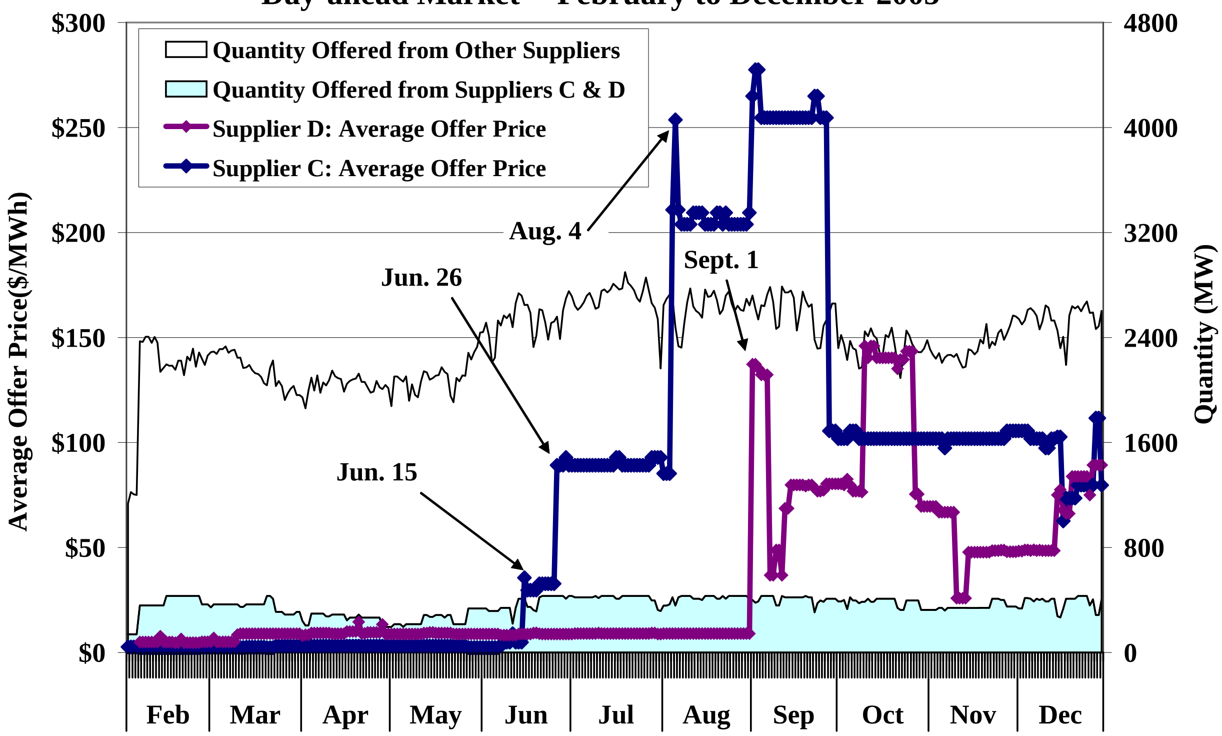


Figure 22: 10-Minute Spinning Reserves Offer Prices and Quantities Selected Suppliers in Afternoon Hours Day-ahead Market -- February to December 2005



Both of the figures above show large quantities offered at less than $5 per MWh for several months after the start of SMD 2.0 that rose substantially in subsequent months. During the summer months, large offer price increases are focused on a small number of days, primarily in mid June, on August 4, and in early September. Most of the days with significant increases followed shortly after days with significant real-time price spike events where real-time reserves prices were far greater than day-ahead reserves prices. The two suppliers of spinning reserves shown in the east adjusted their offers at very high levels a number of times during the fall.

The offer patterns shown in Figure 21 and Figure 22 do not raise significant concerns that market participants are increasing their offer prices to exercise market power. The figures show that these four suppliers account for only a modest share of the total capability. Moreover, we would expect that an attempt to exercise market power would have the greatest effect when conditions are tight. Thus, if the rising offers of these suppliers reflect attempts to withhold, it is unclear why offer prices remained elevated throughout the fall. Furthermore, the real-time market helps discipline competitors in the day-ahead market, particularly since these same suppliers must offer $0 per MWh in real-time. Given that real-time prices are not substantially lower than day-ahead reserves prices, there is little reason to think that suppliers have been able to raise day-ahead reserves prices above competitive levels.

The most likely explanation for the pattern of rising offer prices is that market participants are responding competitively to poor convergence between day-ahead and real-time reserves prices. As shown in Figure 15 and Figure 16 in the previous section, average real-time prices were systematically higher than average day-ahead prices during the summer of 2005. If suppliers predict real-time prices will be higher than day-ahead prices, we can expect they will avoid selling into the day-ahead market and shift more sales to the real-time market. One way to do this is for them to raise their day-ahead offer prices. The figures above show that many of the increases in offer prices occurred shortly after real-time reserves shortages that caused significant price spikes. It is likely that these price spikes caused some suppliers to update their expectations of real-time clearing prices and in some cases raise their day-ahead offer prices.

Although market power in the markets for 10-minute reserves is not a significant concern, the pattern of escalating offer prices raises concerns that market participants are submitting reserves offers above marginal costs, because such behavior negatively effects market efficiency. The day-ahead market commits and schedules resources for energy and ancillary services in economic merit order, resulting in an efficient commitment when suppliers offer their resources at marginal cost. When suppliers raise their offer prices above marginal costs, other more costly resources may be committed in their place. Thus, it is important for overall market efficiency to address issues that undermine the incentives of suppliers to offer their resources at marginal cost.

D. Conclusions

The implementation of SMD 2.0 has lead to two major enhancements to the markets for reserves and regulation. First, the co-optimization of energy and ancillary services in real-time improves market efficiency by allowing the real-time model to consider the costs of ancillary services procurement in the prices of energy, and vice versa. It guarantees that the clearing prices of energy, reserves, and regulation fully reflect the opportunity cost of not providing the other services. These improved price signals eliminate the need for separate lost opportunity cost payments which were necessary prior to SMD 2.0.

Second, pricing during shortage conditions under SMD 2.0 is governed by reserve demand curves. The demand curves establish an economic value for reserves that are reflected in energy prices at times when the energy market must bid scarce resources away from the reserve markets. Reserves prices are based on the shadow prices of reserves constraints, but reserves purchases are reduced when necessary to prevent the shadow prices from exceeding the prices set forth by the demand curves.[8] The demand curve values have been set at levels that are consistent with the actions normally taken by the NYISO operators in reserve shortage conditions. This practice improves consistency between clearing prices and the operation of the system, and better reflects the economic value of reliability.

The enhancements made to the ancillary services markets under SMD 2.0 have lead to four notable changes in market outcomes. First, ancillary services market expenses (i.e. total costs paid by LSEs for ancillary services) were closely correlated with day-ahead ancillary services

---

8 The total value of a reserve in a location is the sum of the reserve demand curve values for each reserve requirement constraint that the reserve contributes to relieving. In other words, because reserves should generally be substituted to maintain the highest quality reserve, the total value of a specific reserve type generally includes the sum of the demand curve values of the lower quality reserves.

prices, which were modest in the spring and summer but rose significantly in the fall. Loads were relatively unscathed by the high prices for 10-minute reserves experienced in the real-time market since average day-ahead prices were considerably lower.

Second, the introduction of SMD 2.0 induced greater participation from operating reserves suppliers. Relative to the preceding year, the average quantity of day-ahead offers for 10-minute spinning reserves rose by more than 80 percent, while the quantity of 30-minute operating reserves offers rose by more than 90 percent. Third, regulation expenses rose during the fall as a result of higher offers from two large suppliers of regulation. The ownership of regulation-capable capacity is relatively concentrated raising concerns that, in the short-term, certain market participants may have incentives to raise their offer prices above marginal cost.

Fourth, poor convergence between day-ahead and real-time operating reserves prices gave suppliers incentives to raise their offer prices for operating reserves. Energy and ancillary services were undervalued on average in the day-ahead market reflecting that, on the whole, market participants did not accurately foresee the frequency and magnitude of price spikes during reserves shortages. Eventually, we expect virtual load bidders to increase their day-ahead energy purchases putting upward pressure on day-ahead prices bringing them into closer convergence with real-time prices. Additional demand for energy in the day-ahead market should put indirect upward pressure on average day-ahead ancillary services prices bringing them into closer convergence with average real-time prices. Thus, the incentive for reserves suppliers to raise their day-ahead offer prices above marginal cost should diminish with better convergence between day-ahead and real-time clearing prices.

However, if the convergence between day-ahead and real-time operating reserves prices remains poor, suppliers of operating reserves will continue to have an incentive to raise their offer prices. In this case, the NYISO should consider the feasibility of introducing virtual trading of ancillary services in the day-ahead market. This change would promote convergence of ancillary service prices and reduce the incentive for physical suppliers to raise their offer prices for operating reserves above marginal cost. The proposal would need to be carefully studied to ensure it will not have unintended consequences on the day-ahead commitment.

IV. ANALYSIS OF ENERGY BIDS AND OFFERS

In this section, we examine bidding patterns to evaluate whether market participant conduct is consistent with efficient and effective competition. On the supply side, the analysis seeks to identify potential attempts to withhold generating resources as part of a strategy to increase prices. On the demand side, we evaluate load-bidding behavior to determine whether load bidding has been conducted in a manner consistent with competitive expectations. We also analyze virtual trading.

A. Analysis of Supply Offers

Wholesale electricity production is attributable primarily to base-load and intermediate-load generating resources. Relatively high-cost resources are used to meet peak loads and comprise a very small portion of the total supply. The marginal cost of base-load and intermediate-load resources do not vary substantially relative to the marginal cost of resources used at peak times. This causes the market supply curve to be relatively flat at low and moderate output levels and steeply sloped at high output levels. Therefore, as demand increases from low load levels, (as an almost vertical demand curve shifts along the supply curve) prices remain relatively stable until demand approaches peak levels, where prices can increase quickly as the more costly units are required to meet load. The shape of the market supply curve has critical implications for evaluating market power.

Suppliers exert market power in electricity markets by withholding resources and increasing the market clearing price. This can be accomplished through physical withholding or economic withholding. Physically withholding occurs when a resource is derated or not offered into the market when it is economic to do so. Economic withholding occurs when a supplier raises the offer price of a resource to reduce its output below competitive levels or to otherwise raise the market price. Demand must be high enough that withholding a resource has the potential to significantly affect market price. When the market clears along the flat portion of the supply curve, prices will be relatively insensitive to withholding.

An analysis of withholding must distinguish between strategic withholding aimed at exercising market power and competitive conduct that could appear to be strategic withholding.

Measurement errors and other factors can erroneously identify competitive conduct as market power. For example, a forced outage of a generating unit may be either legitimate or a strategic attempt to raise prices by physically withholding the unit.

To distinguish between strategic and competitive conduct, we evaluate potential withholding in light of the market conditions and participant characteristics that would tend to create the ability and incentive to exercise market power. Under competitive conditions, suppliers maximize profits by increasing their offer quantities during the highest load periods to sell more power at the higher peak prices. Alternatively, a supplier that possesses market power will find withholding to be profitable during periods when the market supply curve becomes steep (i.e., at high-demand periods). Therefore, examining the relationship between the measures of potential withholding and demand levels will allow us to test whether the conduct in the market is consistent with workable competition.

1. Deratings and Physical Withholding

We first consider potential physical withholding by analyzing generator deratings. A derating occurs when a participant reduces the maximum output available from the plant. This could be for planned outages, long-term forced outages, or short-term forced outages. A derating could be partial (maximum output is reduced, but is greater than zero) or complete (maximum output is zero). We analyze only the summer months to effectively eliminate planned outages from our data. By eliminating planned outages, we implicitly assume that planned outages are legitimate and are not aimed at exercising market power.[9] The remaining deratings data would then include only long-term and short-term deratings. We first analyze both long-term and short-term deratings together. In our second analysis, we focus on short-term deratings because short-term deratings are more likely to reflect attempts to physically withhold since it is more costly to withhold via long-term deratings or outages.

---

9 Planned outages are usually scheduled far in advance, and are almost always scheduled for a period during the year when demand is historically at low levels, in New York, typically spring and autumn months. Since weather forecasters are currently incapable of predicting unusual weather events, like record setting heat waves in May, the fact that a planned outage results in higher prices in those circumstances is not evidence of the exercise of market power. Thus, only outages which occur during periods when the supplier can anticipate a benefit from withholding are relevant to the market power analysis.

We focused on the afternoon hours which have higher demand because, under a hypothesis of market power, we would expect to find that withholding increases as demand increases. We also limited ourselves to the locations east of the Central-East interface, as this area, which includes two-thirds of the State's load, has limited import capability, and is more vulnerable to the exercise of market power.

Figure 23 and Figure 24 show our analysis for all deratings and for short-term deratings, respectively.

Figure 23: Relationship of Deratings to Actual Load
Day-Ahead Market - East New York
Weekdays, Noon to 6 PM, Summer 2005

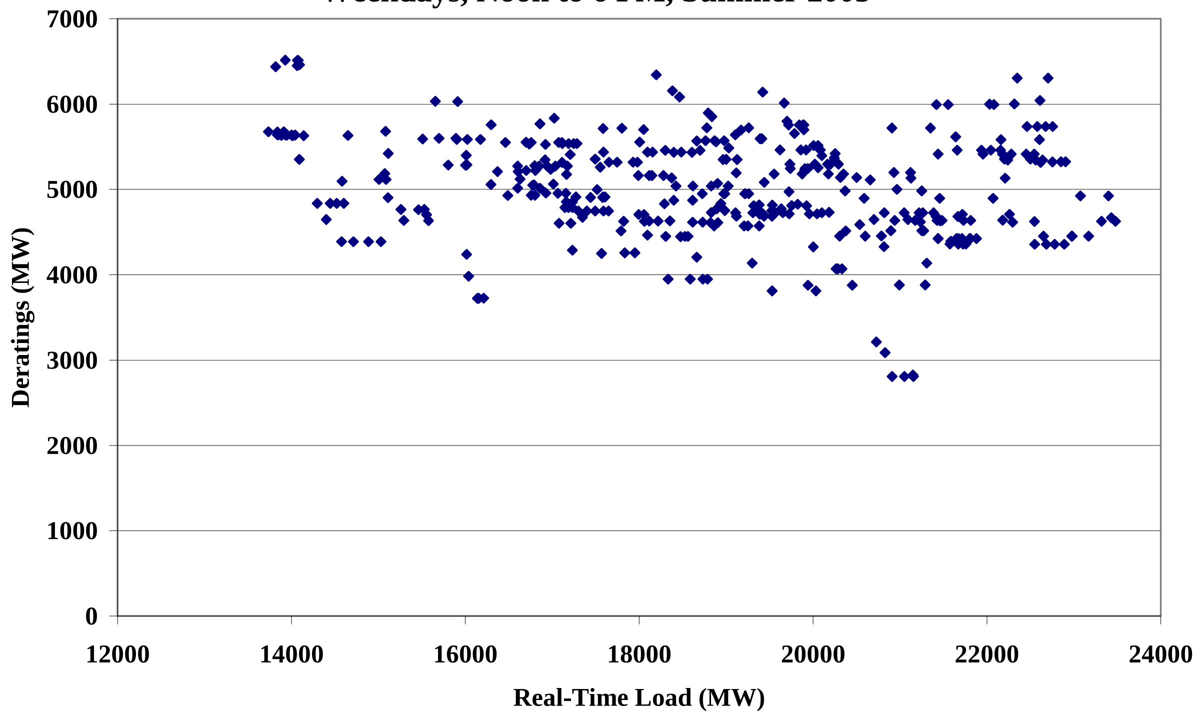


Figure 24: Relationship of Short-Term Deratings to Actual Load
Day-Ahead Market – East New York
Weekdays, Noon to 6 PM, Summer 2005

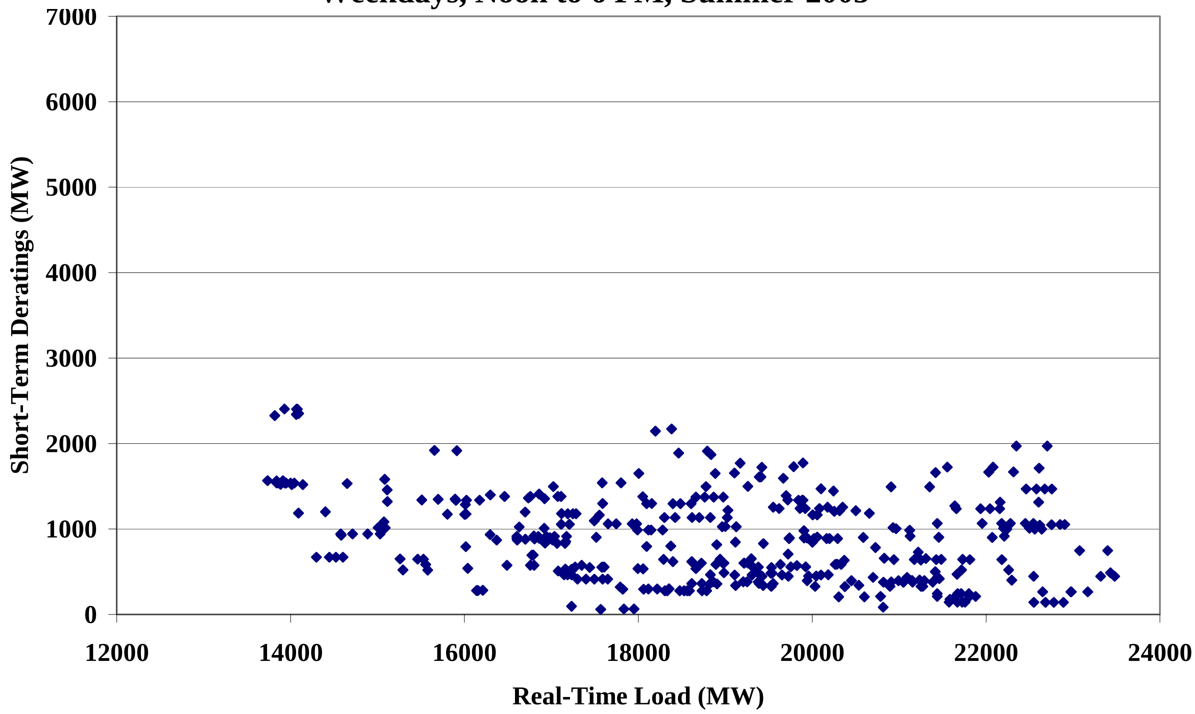
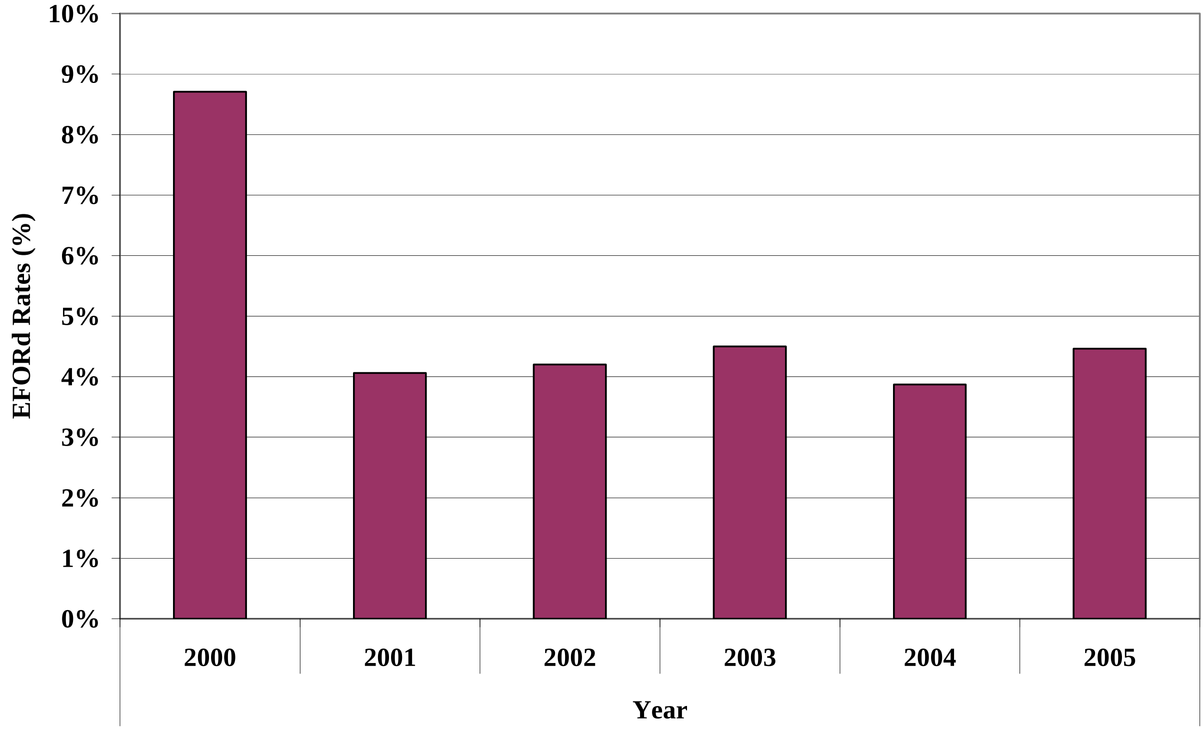


Figure 23 indicates that the total quantity of deratings is generally smaller when demand reaches very high levels. Focusing on short-term deratings in Figure 24, we also found less deratings during high demand periods. This indicates good competitive performance since the incentive to physically withhold resources would generally increase under high demand conditions for participants with market power. Furthermore, while forced outages are generally random, the total forced outages would be expected to rise slightly under peak demand conditions when the ISO must call on units that operate infrequently. Therefore, we find that the overall pattern of outages and deratings was consistent with workable competition during the summer of 2005.

2. Forced Outages

We examined the trend in forced outages in the New York markets to ascertain if generators are responding to economic incentives to increase availability of their units. Figure 25 shows the Equivalent Forced Outage Rate ("EFOR"), which is used as a measure of forced outages. The EFOR is the portion of time a unit is unavailable due to forced outages, expressed as equivalent hours of full forced outage at its maximum net dependable capability.

Figure 25: Equivalent Demand Forced Outage Rates 2000 – 2005



EFOR declined substantially following the implementation of the NYISO markets. This is consistent with the incentives the deregulated markets provide to maximize availability, particularly during high load conditions. EFOR was relatively high in 2000 due to the outage of an Indian Point nuclear unit. After the Indian Point outage, the EFOR has been consistently close to 4 percent – much lower than the outage rates that prevailed prior to the implementation of the NYISO markets. In 2005, the EFOR was approximately 4.5 percent.

3. Output Gap and Economic Withholding

To evaluate economic withholding, we calculated the hourly "output gap". The output gap is the quantity of generation capacity that is economic at the market clearing price, but is not running due to the owner's offer price or is setting the LBMP with an offer price substantially above competitive levels (excluding capacity scheduled to provide ancillary services). This withholding can be accomplished through high start-up cost offers, high minimum generation offers, and/or high incremental energy offers.

To determine whether an offer is above competitive levels, we use reference values based on the past offers of the participant during competitive periods. A supplier will normally offer at levels near marginal cost, because during periods when market power is unlikely to be exercised, excessive offers will cause the unit not to be dispatched and cost the owner lost profits. We allow considerable tolerance in our threshold. An offer parameter is indicated as above competitive levels if it exceeds the reference values by a given threshold. We conduct the analysis with thresholds matching the conduct threshold used by the state-wide automated mitigation procedure ($100/MWh or 300 percent, whichever is lower) and a lower threshold ($50/MWh or 100 percent, whichever is lower).

Like our analysis of deratings, we examine the relationship of the output gap to the market demand level. We focus our analysis on Eastern New York where market power is most likely. Figure 26 shows the output gap using the state-wide mitigation thresholds of $100/MWh or 300 percent. To assess whether there have been significant attempts to withhold by offering just below the state-wide mitigation threshold, Figure 27 shows the output gap results using a lower threshold of $50/MWh or 100 percent.

Figure 26: Relationship of Output Gap at Mitigation Threshold to Actual Load
Real Time Market – East New York
Weekdays, Noon to 6 PM

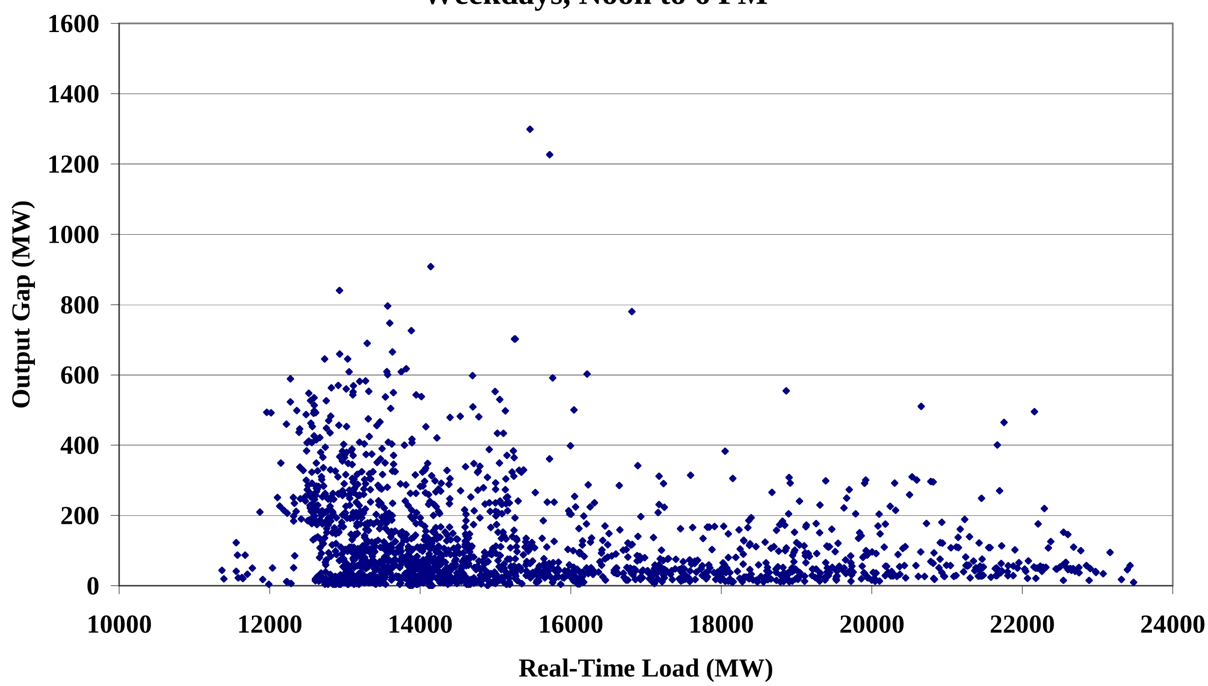
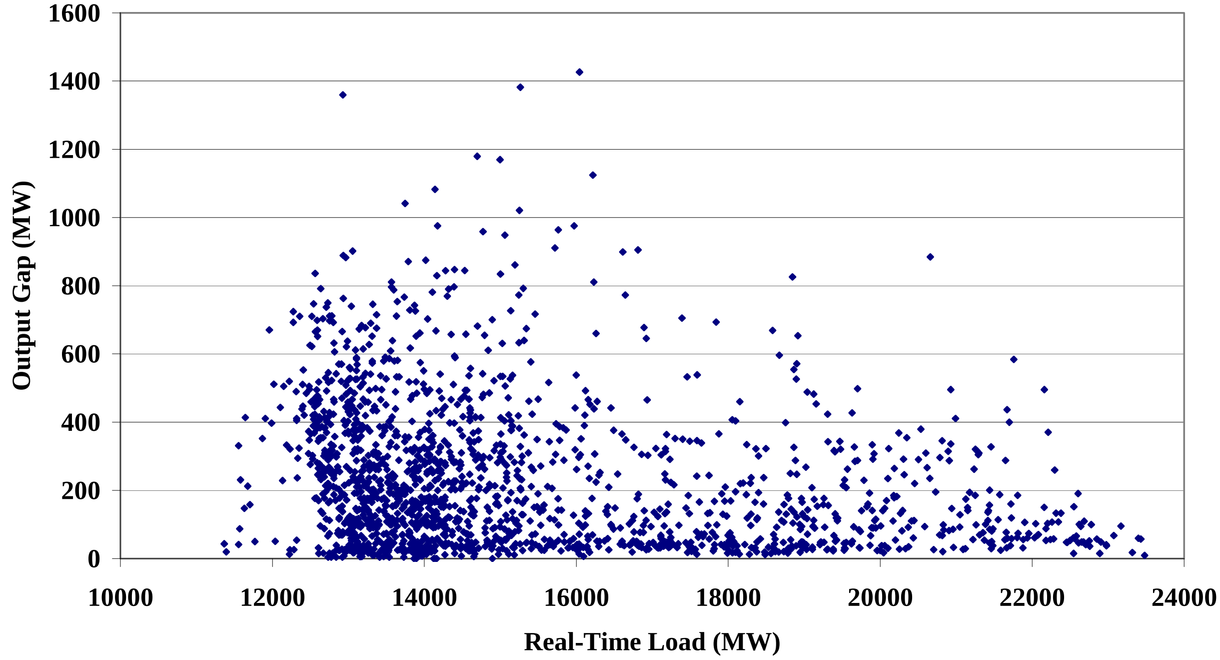


Figure 27: Relationship of Output Gap at Low Threshold to Actual Load
Real Time Market – East New York
Weekdays, Noon to 6 PM



These figures both show that the output gap decreases to relatively low levels under the highest load conditions. Figure 26 shows that the output gap measured at the high threshold was greater than or equal to 300 MW during just eleven hours when Eastern New York load exceeded 18 GW during 2005. More than 70 percent of the output gap in these hours was associated with generation in up-state areas which are generally more competitive than New York City and Long Island. Some of the output gap shown in the two figures could be the result of actual withholding, although the quantities are relatively small compared to the total load in Eastern New York and are negatively correlated with actual load. Thus, the results of the output gap analysis indicate that the market generally functioned competitively under peak demand conditions and do not raise significant concerns about economic withholding during 2005.

4. Market Power Mitigation

The NYISO applies a conduct-impact test that can result in mitigation of participant bid parameters (i.e., energy offers, start-up and no-load offers, and physical parameters). The conduct test first determines whether bid parameters exceed pre-defined conduct thresholds. If at least one of the participant's bid parameters exceeds a conduct threshold, the bid parameter may be mitigated if the conduct results in sufficient effect on the energy price. While the NYISO tariff allows for mitigation to be invoked manually according to pre-defined criteria, this rarely occurs. Instead, the day-ahead and real-time market software is automated to perform most mitigation according to pre-defined conduct and impact thresholds.

Mitigation is applied in the real-time market for units in certain load pockets within New York City using the NYISO's conduct and impact approach. The in-city load pocket conduct and impact thresholds are set using a formula that is based on the number of congested hours experienced over the preceding twelve-month period.[10] This approach permits the in-city conduct and impact thresholds to increase as the frequency of congestion decreases, whether due to additional generation or increases in transmission capability. An in-city bid fails the conduct test if it exceeds the reference level by the threshold or more. In-city bids that fail conduct are tested for price impact by the real-time software, and if their price impact exceeds the threshold, they are mitigated.[11]

Prior to May 1, 2004, the day-ahead software used the conduct and impact test framework only for determining whether to mitigate outside New York City. Inside New York City, the day-ahead software would mitigate all units to their reference level (based on variable production expenses) whenever it detected at least a small amount of congestion between Indian Point and New York City. These mitigation procedures were referred to as the Consolidated Edison or "Con Ed" mitigation procedures as they were developed when Con Ed divested its generation. Under the Con Ed procedures, mitigation occurred nearly every day.

The Con Ed procedures were replaced on May 1, 2004 by the conduct and impact mitigation framework which was already being applied to the rest of the state. This framework significantly reduced the frequency of mitigation by making it more focused on potential market power in the NYC load pockets. This prevents mitigation from occurring when it is not necessary to address market power and allows high prices to occur during legitimate periods of shortage.

---

10 $$\text{Threshold} = \frac{2\% * \text{Avg. Price} * 8760}{\text{Constrained Hours}}$$

11 Prior to SMD 2.0, the real-time software did not have the capability to perform a price impact test, so it used a proxy impact test. The proxy impact test imposed mitigation whenever the cost of congestion into New York City load pockets exceeded the threshold. Thus, the real-time software imposes mitigation more selectively under SMD 2.0.

Figure 28 summarizes the extent of real-time mitigation from February to December 2005 according to load pocket. Conduct and impact testing is performed according to the threshold that applies to each load pocket. For each load pocket shown in the figure below, the line indicates the percent of hours when mitigation was imposed on one or more units. The bars indicate the average amount of capacity mitigated in hours when mitigation occurred. Mitigated quantities are shown separately for the flexible output ranges of units (i.e. Energy) and the non-flexible portions (i.e. Mingen/Start-Up).[12]

Figure 28: Frequency of Real-Time Mitigation in the NYC Load Pockets February to December 2005

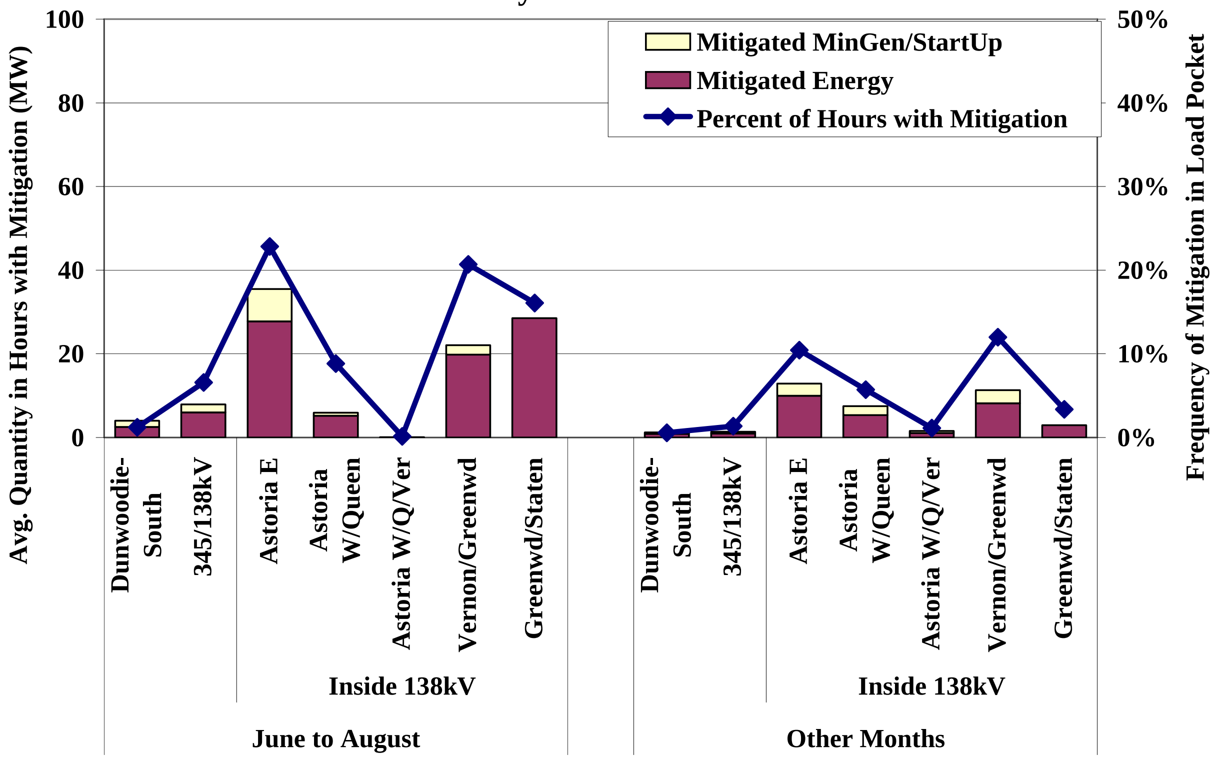


Figure 28 shows that most real-time mitigation occurred within the 138 kV portion of New York City. These load pockets have a limited number of suppliers and experience frequent real-time congestion making them more susceptible to the exercise market power. More mitigation took place during the summer due to higher loads and more frequent congestion. The majority of real-time mitigation was associated with incremental energy bid parameters rather than Mingen bid parameters. This is because a large share of real-time mitigation was of gas turbines, which

12 For modeling purposes, gas turbines are treated as flexible from zero to full output, although they generally run at full output when on-line.

do not submit Mingen offers. The low frequency of real-time mitigation is partly due to the fact that day-ahead mitigated offers are carried into the real-time up to the unit's day-ahead schedule.

Outside of the load pockets in the 138 kV portion of New York City, real-time mitigation was less frequent in 2005. There are three factors that may explain the reduced mitigation. First, the impact test implemented in February 2005 under SMD 2.0 is more selective than the proxy impact test that was used previously. Second, two new units come on-line during 2005 at the East River plant which is located outside the 138 kV area. These likely helped reduce congestion and mitigation related to the Dunwoodie-South interface (i.e. the primary source of imports to New York City). Third, certain units committed and/or dispatched out-of-merit or through the SRE process were not subjected to mitigation in real-time due to a software error.

Figure 29 summarizes the extent of day-ahead mitigation from February to December 2005 according to load pocket. For each load pocket shown in the figure below, the line indicates the percent of hours when mitigation was imposed on one or more units, while the bars indicate the average amount of capacity mitigated in hours when mitigation occurred.

Figure 29: Frequency of Day-ahead Mitigation in the NYC Load Pockets February to December, 2005

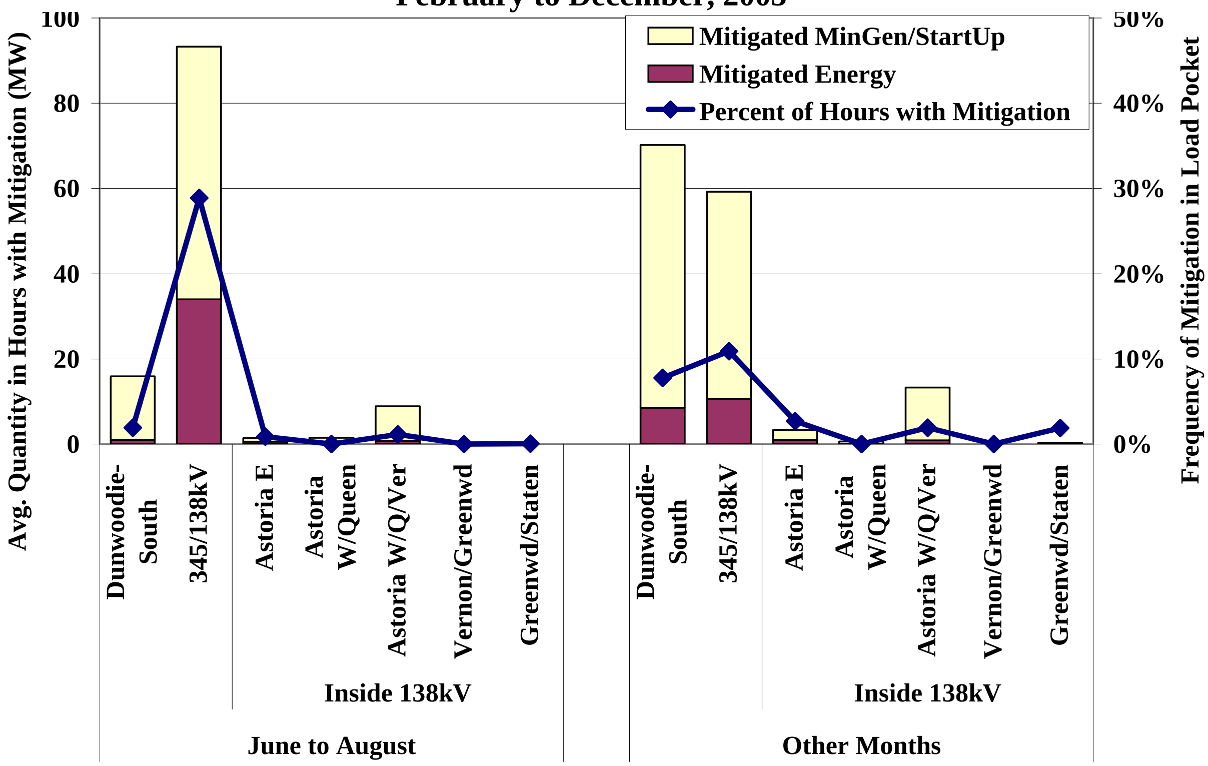


Figure 29 shows that the pattern of day-ahead mitigation was substantially different from the pattern of real-time mitigation shown in previous figure. In the day-ahead, mitigation was most commonly associated with the Dunwoodie-South interface which brings power into New York City from up-state and the 345/138 kV interface which brings power into the 138 kV portions of New York City. The bids of units mitigated in the day-ahead market are carried forward into the real-time market.

Figure 29 indicates that the majority of capacity mitigated in the day-ahead market is associated with the start-up and Mingen parameters, while relatively little is for incremental energy parameters. This relationship shows that units with significant minimum run times are sometimes mitigated for price impact in a relatively small number of hours. For instance, a unit with a 12 hour minimum run time might raise its Mingen bid parameter above the conduct threshold. However, if this conduct would cause the unit to not be committed resulting in a price impact above the applicable threshold for one hour, the unit's Mingen parameter would be mitigated for the duration of its minimum run time, while its incremental energy parameter would be mitigated only in the hour with impact.

B. Analysis of Load Bidding and Virtual Trading

In addition to physical and economic withholding, buyer behavior can strategically influence energy prices. Therefore, evaluating whether load bidding is consistent with workable competition is an important focus of market monitoring. Load can be purchased in one of the following four ways:

- *Physical Bilateral Contracts*. These are schedules that the NYISO provides to participants that allow them to settle transmission charges (i.e., congestion and losses) with the ISO and to settle on the commodity sale privately with their counterparties. It does not represent the entirety of the bilateral contracting in New York, however, because participants have the option of constructing identical arrangements by other means that would settle through the NYISO. In particular, participants may sign a "contract-for-differences" ("CFD") with a counterparty to make a bilateral purchase. Financial bilateral contracts such as CFDs are settled privately and generally would show up as day-ahead fixed load.

When the CFD is combined with a TCC, the participant can create a fully-hedged forward energy purchase. Therefore, the trends in the quantity of physical bilateral contracts scheduled with the NYISO do not indicate the full extent of forward contracting.

Day-Ahead Fixed Load. This represents load scheduled in the day-ahead market for receipt at a specific bus regardless of the day-ahead price. It is the equivalent of a load bid with an infinite bid price, which is difficult to rationalize from an economic perspective.

Price-Capped Load Bids. This represents load bid into the day-ahead market with a bid price indicating the maximum amount the Load-Serving Entity ("LSE") is willing to pay. For example, an LSE may make a price-capped bid for 500 MW at $60 per MWh. If the day-ahead market at its location clears above $60, the energy would not be purchased in the day-ahead market. If the load is actually realized in real-time, it would be served with energy purchased in the real-time market. This is a more rational form of load-bidding than the non-price sensitive fixed load schedules. However, price-capped load bidding is only allowed at the zonal level while fixed load bidding is allowed at the bus level.

Virtual Load Bids. These are bids to purchase load in the day-ahead market with a bid price indicating the maximum amount the bidder is willing to pay. Virtual load purchased in the day-ahead market is automatically sold back to the real-time market. So, the virtual load purchaser earns the quantity of the purchase in megawatt-hours multiplied by the real-time price minus the day-ahead price. Like price-capped load bidding, this is currently allowed at the zonal level but not the bus level.

Virtual Supply Offers. These are offers to sell energy in the day-ahead market with an offer price indicating the minimum amount the market participant is willing to accept. Virtual supply sold in the day-ahead market is automatically purchased back from the real-time market. So, the virtual seller earns the quantity of the sale in megawatt-hours multiplied by the day-ahead price minus the real-time price. Like virtual load, this is currently allowed at the zonal level but not the bus level.

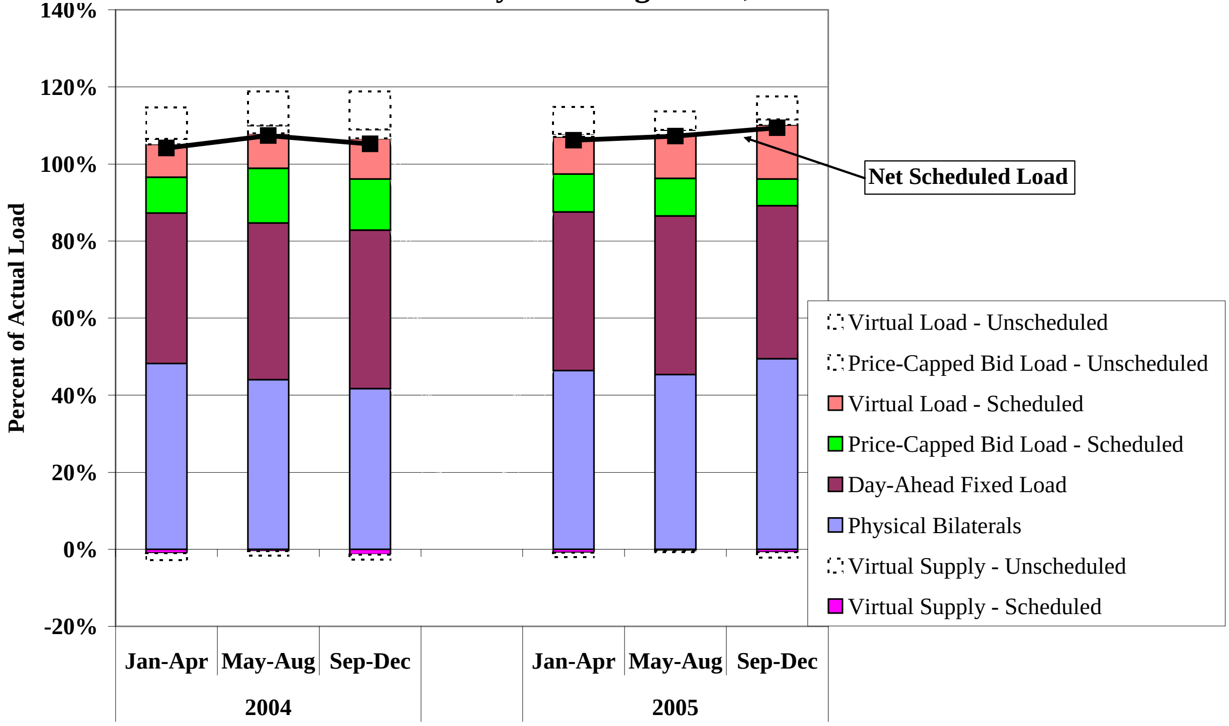
1. Day-Ahead Scheduling

Many generating units have lengthy start-up times and substantial commitment costs. Their owners must decide whether to commit well in advance of real-time before they can be certain that the unit will be economic. The day-ahead market provides these suppliers with a way of deciding to commit only when it is economic to do so. These suppliers are willing to sell into the day-ahead market if day-ahead prices are generally consistent with real-time prices. Thus, efficient unit commitment relies on consistency between the day-ahead market and the real-time market. The analyses in this subsection examine the consistency of day-ahead load scheduling patterns and actual load which provide an indication of the over efficiency of the day-ahead market.

We expect that day-ahead load schedules will be generally consistent with actual load in a well-functioning market. Under-scheduling load will tend to result in day-ahead prices lower than real-time with insufficient commitment for real-time needs. Over-scheduling tends to bid up day-ahead prices above the level of real-time prices. Thus, natural market incentives give market participants strong incentives to schedule amounts of load that are consistent with real-time load. The following figures show day-ahead load schedules and offers as a fraction of real-time load during 2004 and 2005 at various locations in New York. Virtual load and load schedules of LSEs have the same effect on day-ahead prices and resource commitment, so virtual load is treated the same for this analysis. Conversely, virtual supply has the same effect on day-ahead prices and resource commitment as a reduction in physical load, so we treat it as a negative load for the purposes of this analysis.

Figure 30 shows a comparison of day-ahead load scheduling to actual load in New York City and Long Island on a seasonal basis in 2004 and 2005. For each period, it shows scheduled and unscheduled quantities of physical load and virtual load. Scheduled and unscheduled virtual supply is shown by bars in the negative direction because they have the same effect on day-ahead commitment and pricing as a reduction in scheduled load. Net scheduled load, indicated by the line, is the sum of physical and virtual load minus virtual supply.

Figure 30: Composition of Day-Ahead Load Schedules New York City and Long Island, 2004 - 2005



Load is generally over-scheduled by 5 to 10 percent in New York City and Long Island relative to actual load. While this pattern might be expected to raise day-ahead prices above real-time prices, day-ahead prices were generally lower than real-time prices in both years. Thus, the pattern of over-scheduling is induced by depressed day-ahead prices and actually contributes to better price convergence by bidding up day-ahead prices. The over-scheduling implies a higher level of imports to constrained areas in the day-ahead market than in real time. As discussed in Section V of this report, these pricing and scheduling patterns are partly the result of modeling inconsistencies between the day-ahead and real-time markets.

The next two figures compare day-ahead load scheduling to actual load in areas outside New York City and Long Island on a seasonal basis in 2004 and 2005. This comparison is shown for East Up-State New York in Figure 31 and Western New York in Figure 32.

Figure 31: Composition of Day-Ahead Load Schedules East Up-State New York, 2004 - 2005

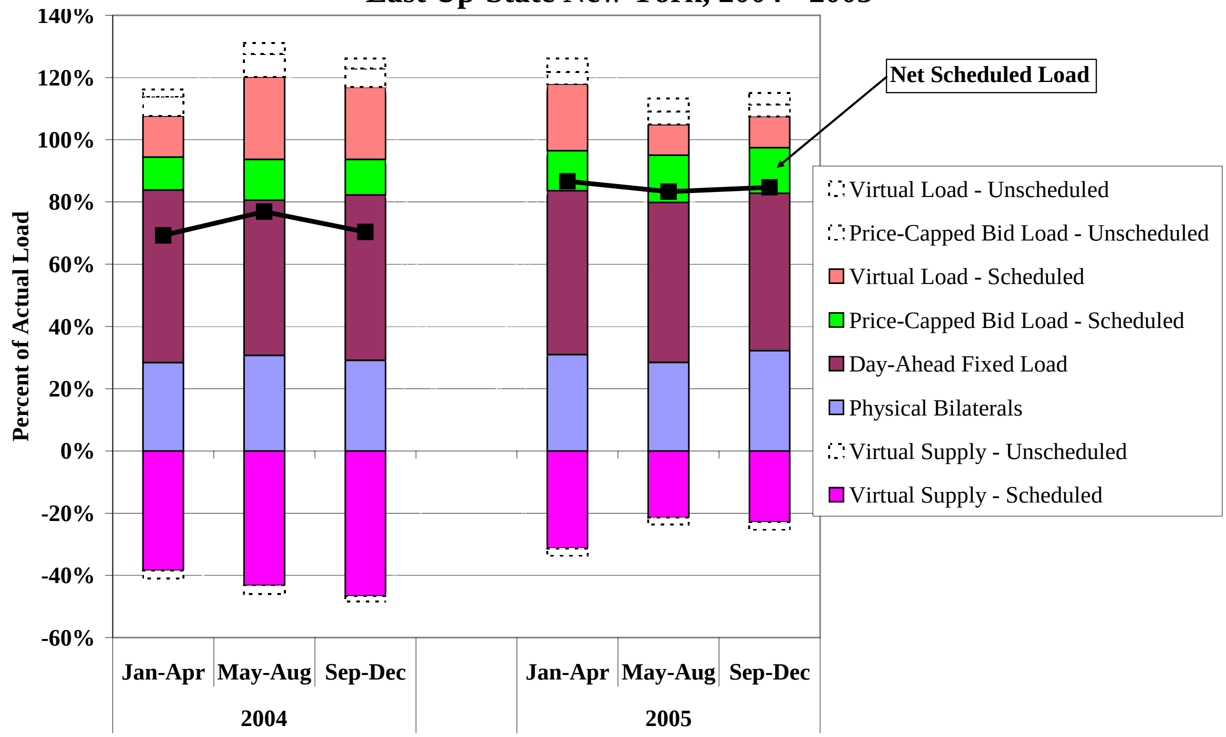


Figure 32: Composition of Day-Ahead Load Schedules West New York, 2004 - 2005

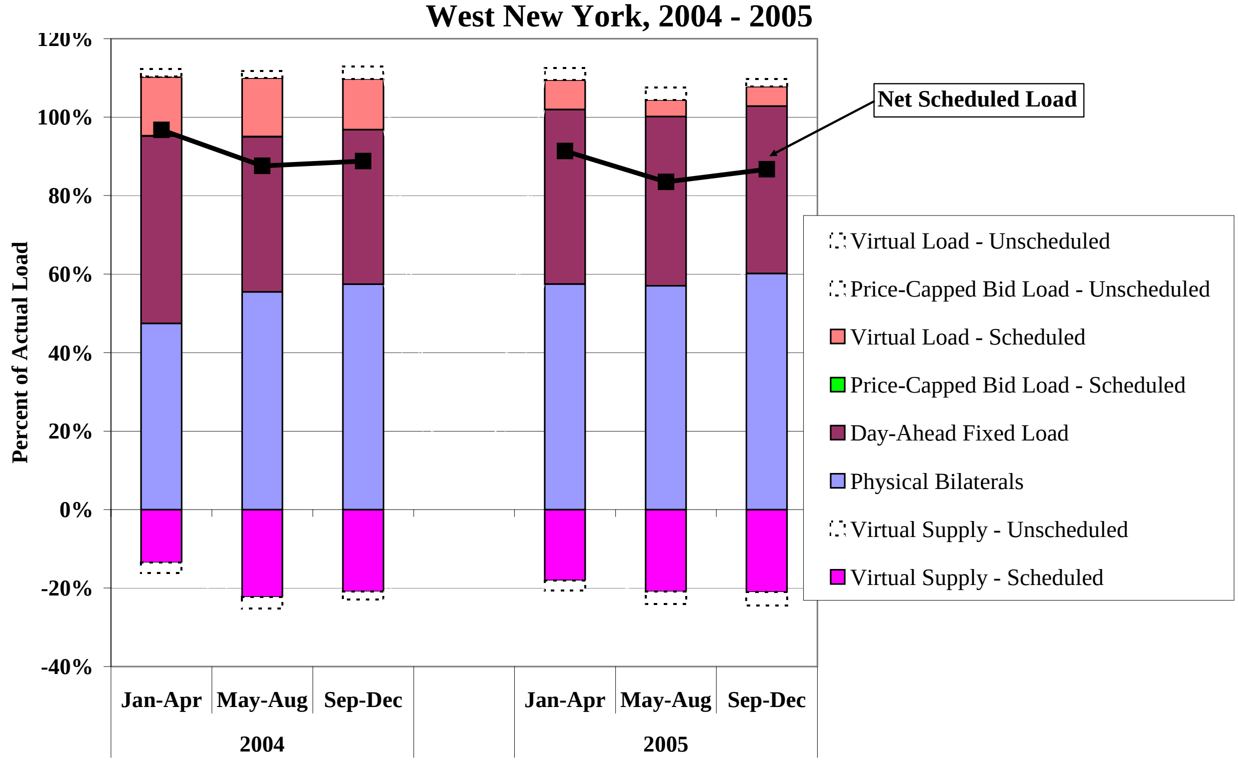


Figure 31 and Figure 32 show a pattern of load scheduling in up-state areas, which contrasts sharply with the pattern in Figure 30 for New York City and Long Island. Although the sum of physical and virtual load exceeded actual load in up-state New York on average, large amounts of virtual supply led to net under-scheduling of load. While this might be expected to lead to depressed day-ahead prices, it is actually a response to the persistent day-ahead price premium observed in Figure 12 in Section II. Thus, the lack of scheduling convergence in up-state New York caused by virtual trading activity has improved price convergence. Explanations for these pricing and scheduling patterns are discussed further in Section V of this report.

2. Virtual Trading

Virtual trading was introduced in November 2001 to allow participation in the day-ahead market by entities other than LSEs and generators. The motivation was to improve arbitrage between the day-ahead and real-time markets as well as allowing flexibility for all participants in managing risk. Virtual energy sales or purchases in the day-ahead market settle in the real-time market, allowing participants to arbitrage price differences between the day-ahead and real-time markets. For example, a participant can make virtual purchases in the day-ahead market if the participant expects prices to be higher in the real-time market, and then sell the purchased energy back into the real-time market. The result of this intertemporal arbitrage would be to raise the day-ahead price slightly and decrease the real-time price slightly to improve convergence.

We analyzed the quantities of virtual load and supply that have been offered and scheduled on a monthly basis during the past two years. These patterns are shown for New York City and Long Island in Figure 33 and up-state New York in Figure 34.

Figure 33: Hourly Virtual Load and Supply New York City and Long Island 2004 – 2005

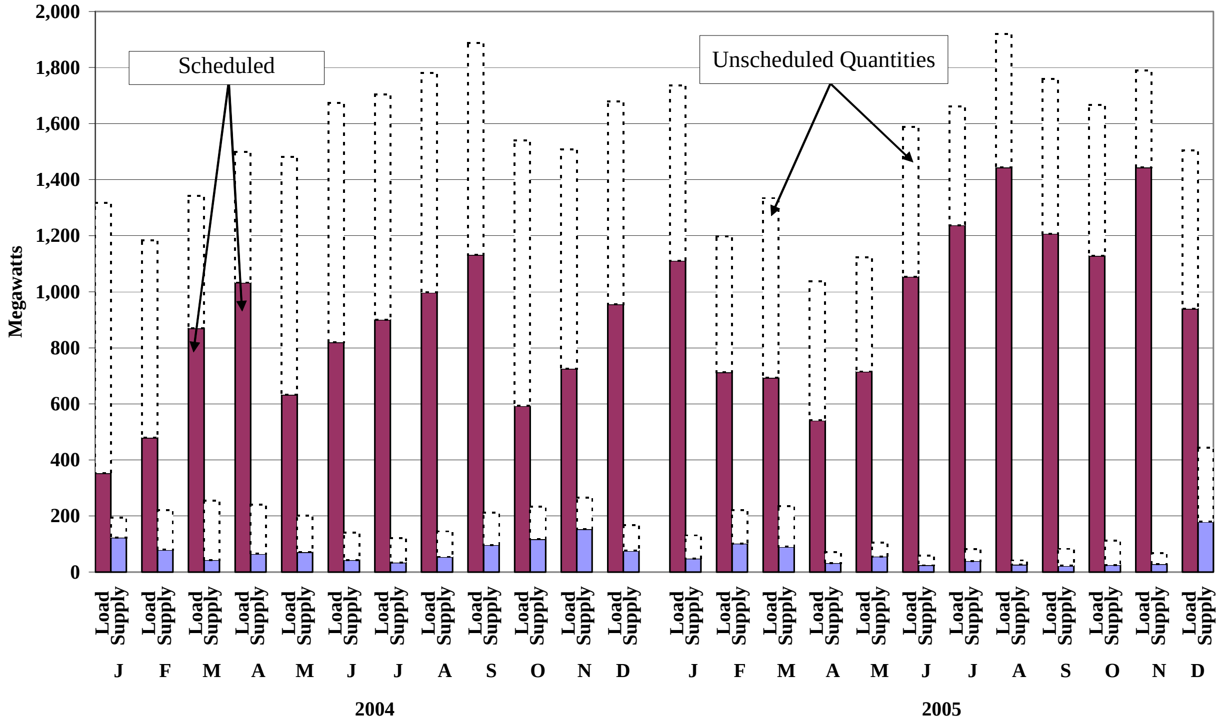
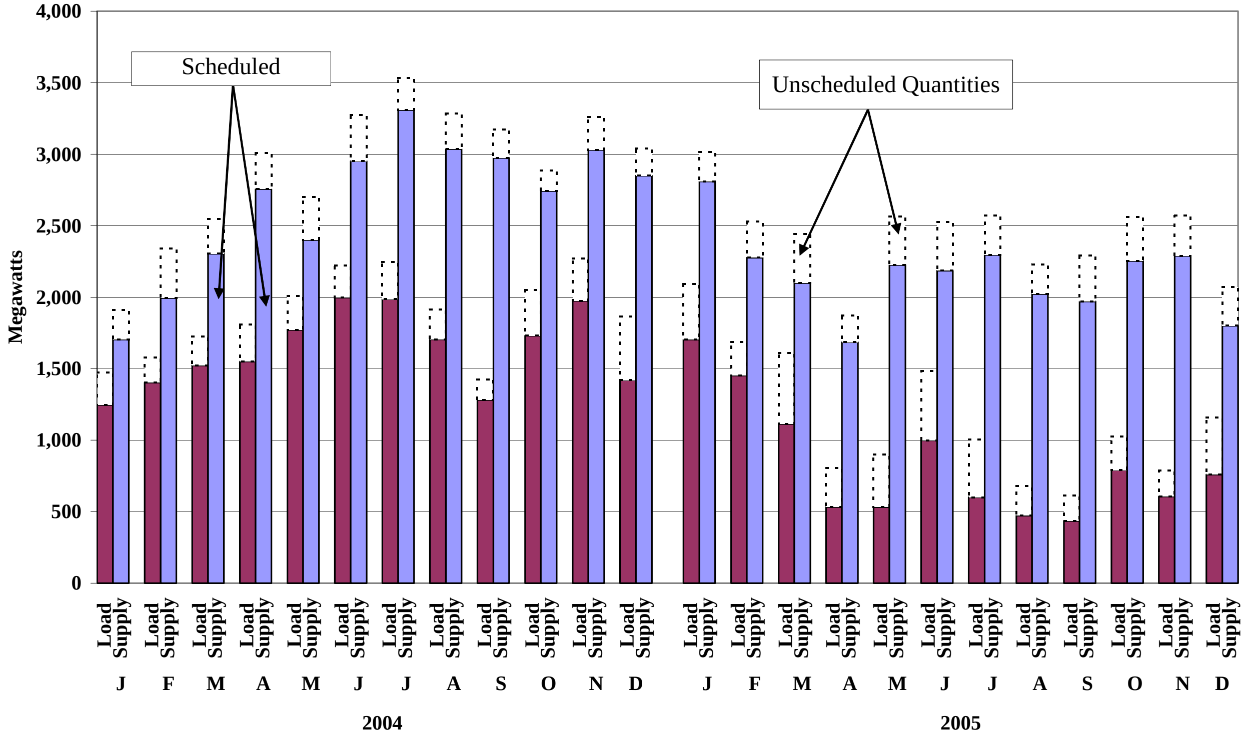


Figure 34: Hourly Virtual Load and Supply in Up-state New York 2004 – 2005



In New York City and Long Island, virtual load schedules rose during the summer of 2005 and remained at unusually high levels during the fall months. This was most likely in response to the trend that emerged of real-time prices being higher on average than day-ahead prices. Virtual trading was generally lower in Up-State New York in 2005 compared with 2004, although this was primarily due to a reduction in virtual supply. Thus, the difference between virtual supply and virtual load scheduling actually widened in 2005. The figures show that a substantial share of the virtual bids and offers in New York City and Long Island were not scheduled, while a much smaller portion of the virtual bids and offers in up-state New York were not scheduled.

The net virtual purchases in New York City and Long Island and net virtual sales in up-state New York contribute to the pattern of over-scheduling down-state and under-scheduling up-state which is discussed in the prior section. These virtual trading patterns have contributed to better convergence between the day-ahead and real-time prices. These scheduling patterns imply that the day-ahead market routinely schedules greater flows from net exporting up-state areas to net importing down-state areas than actually occurs in real-time. This pattern is consistent transmission modeling issues discussed in the next section.

V. MARKET OPERATIONS

Aside from operating the spot markets, a primary role of the ISO's operations is to ensure safe and reliable grid operation. Many of the ISO's operating functions in this regard can have a substantial effect on market outcomes, especially during peak demand conditions. Operating functions that affect clearing prices and other market outcomes include:

- Modeling a security-constrained transmission system in the day-ahead and real-time markets;
- Real-time commitment and dispatch of gas turbines;
- Operating the transmission system during periods of contingency reserves shortages;
- Redispatch of generation to resolve transmission constraints; and
- Committing supplemental resources not selected by the day-ahead market to maintain reserves in local areas;

Reliability requires that operators carry out these functions, but they should be done in a way that promotes efficient market pricing and behavior. This section evaluates operating functions and examines how they affect market outcomes.

A. Transmission Congestion

Congestion can arise in both the day-ahead and real-time markets when transmission capability is not sufficient to accommodate a least-cost dispatch of generation resources. When congestion arises, both the day-ahead and real-time market software establish spot prices based on the cost of meeting load at each location, which reflects the fact that higher-cost generation may be required at locations where transmission constraints prevent the free flow of available resources. This will result in higher spot prices at these "constrained locations" than would occur in the absence of congestion. Furthermore, transmission losses greatly affect the cost of serving load in each area, and are also reflected in locational spot prices.

The day-ahead market is a forward market, facilitating financial transactions among participants that are binding in real-time. The NYISO applies congestion charges to these transactions, which are both bilateral transactions and spot transactions, by modeling anticipated congestion. Bilateral transactions are charged based on the difference between day-ahead spot prices at the two locations (the price at the sink less the price at the source). Buyers and sellers pay congestion charges implicitly equal to the difference in prices between the locations where power is injected and withdrawn from the transmission network.

Congestion charges may be hedged in the day-ahead market by owning TCCs, which entitle the holder of the TCC to payments corresponding to the congestion charge between two locations. A TCC consists of a directional pair of points (locations or zones) and a MW value. For example, if a participant holds 150 MW of TCC rights from point A to zone B, this participant is entitled to 150 times the congestion price at zone B less the congestion price at location A. Excepting losses, a participant can perfectly hedge its bilateral contract if it owns a TCC between the same two points over which it has scheduled the bilateral contract.

In the real-time market, participants with day-ahead contracts do not pay real-time congestion charges. Only transactions not scheduled in the day-ahead market are assessed real-time congestion charges. As in the day-ahead market, charges for bilateral transactions are based on the difference between the locational prices at the two locations of the bilateral contract. For real-time spot market transactions, the congestion charge is paid by the purchaser through the congestion component of the LMP. There are no TCCs for real-time congestion because the real-time spot market is a balancing market where congestion charges should be zero on average.

This subsection addresses several aspects of transmission congestion management and locational pricing. It examines the following three areas:

- Congestion Across Major Transmission Interfaces: This analysis summarizes growth in the frequency and value of congestion across major interfaces during the past four years.
- Congestion Revenue Shortfalls: Congestion revenues collected in the day-ahead and real-time market by the NYISO are sometimes not sufficient to cover congestion payments. We examine the size of these shortfalls, discuss factors that increase shortfalls, and highlight steps taken by the NYISO to reduce shortfalls.
- Price Convergence Between TCCs and the Day-ahead Market: We review the consistency of prices paid for TCCs and congestion prices in the day-ahead market that determine payments to TCC holders.

1. Congestion Across Major Transmission Interfaces

Supply resources in New York City and Long Island generally have higher costs than in up-state New York. The physical capability of the transmission system limits the amount of power that can be transferred from lower cost resources to load pockets in New York City and Long Island, making the economic value of major transmission interfaces considerable. Thus, is important that the transmission planning process and incentives for transmission investment lead to efficient new investment. The analyses in this sub-section summarize the frequency and value of congestion on several key interfaces in New York.

Figure 35: Frequency of Real-Time Congestion on Major Interfaces 2002 - 2005

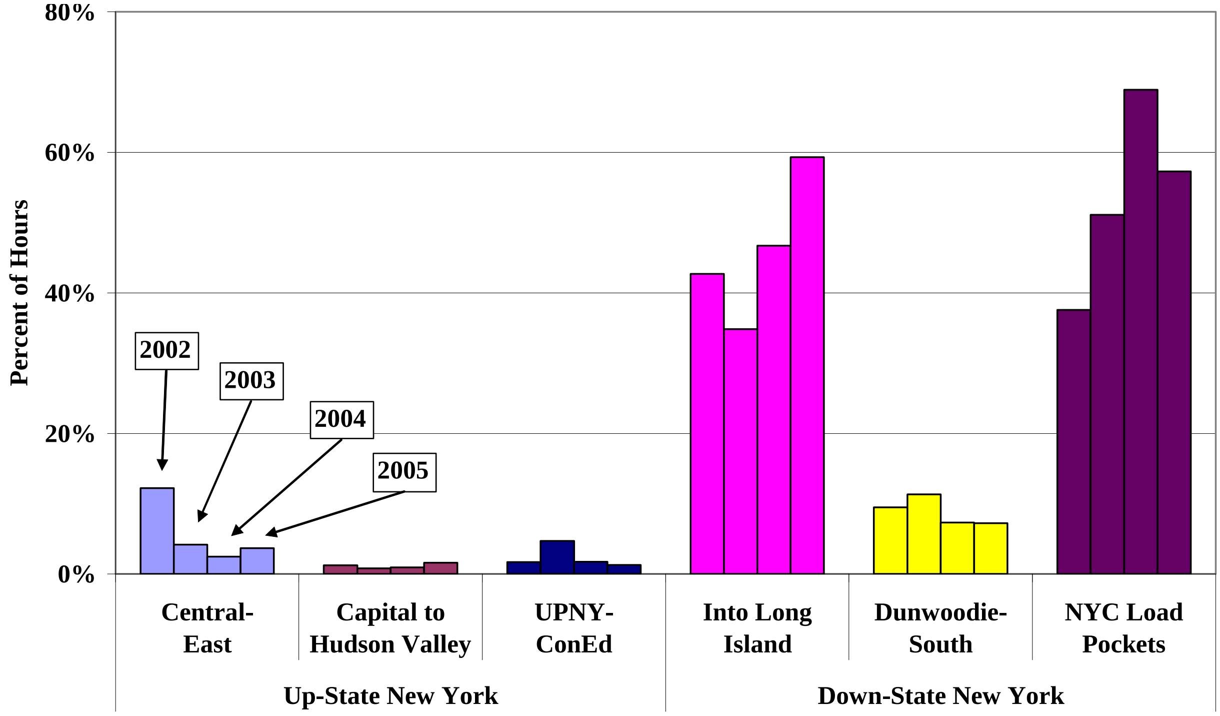


Figure 35 shows the frequency of congestion on select interfaces in up-state and down-state New York. From up-state New York, the figure includes constraints that (i) are part of the Central-East Interface, (ii) limit southward flows from the Capital region through the Hudson Valley, and (iii) make up the interface between up-state New York and the Con Ed transmission area. From down-state New York, the figure includes (i) transmission constraints from up-state New York into Long Island, (ii) the Dunwoodie-South constraint that limits flows from upstate New York into New York City, and (iii) the group of constraints that limit flows within New York City. This analysis excludes constraints within Western New York and also within the Long Island zone.

The results of Figure 35 show the preponderance of congestion occurs into and within down-state areas. Congestion into New York City load pockets increased for several years but decreased from 2004 to 2005. Congestion in Long Island became substantially more frequent in 2005. While congestion became more frequent across the three pathways shown for Up-State New York above, congestion was still far less frequent in up-state areas than in down-state areas. At least four factors influence the trends in congestion shown above. First, load pocket modeling was introduced to New York City in June 2002. The NYC Load Pockets were constrained during more than 60 percent of the intervals in 2002 after load pocket modeling was introduced (which is comparable to the results after 2002).

Second, increasingly frequent thunderstorm alerts ("TSAs") resulted in more congestion on the up-state interfaces shown above. Since the blackout of August 2003, TSAs have required double contingency operation of the ConEd overhead transmission system in real-time. This effectively reduces the amount of power that can flow from Up-state New York through the Hudson Valley to New York City and Long Island resulting in more frequent congestion.

Third, there have been significant transmission outages that have affected congestion patterns. The Central-East interface experienced large outages during the spring of 2002 which contributed to the frequency of congestion in that year. There was also significant congestion on the Dunwoodie-South interface early in 2003 as a result of maintenance work, resulting in lower import levels to New York City. Frequently the need for more New York City generation to resolve the Dunwoodie-South constraint was met with generation in load pockets, which reduced the frequency of congestion within New York City.

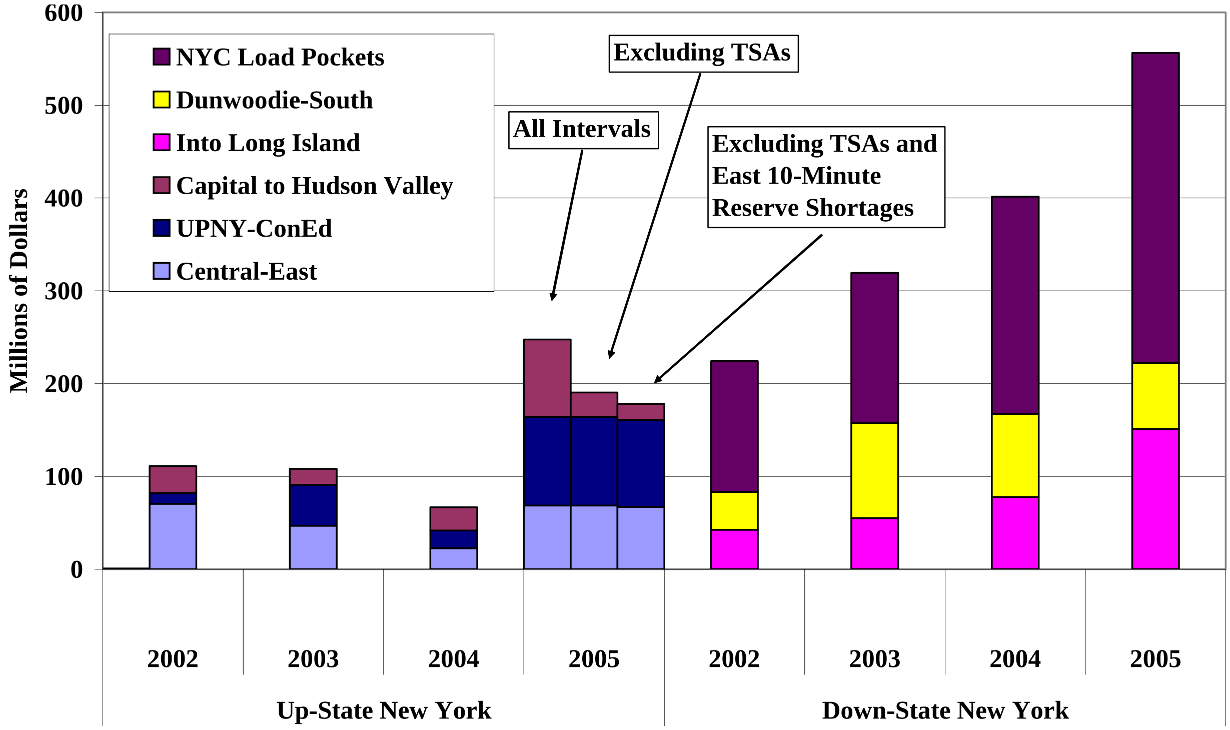
Fourth, generating capacity additions and other changes in supply have influenced congestion patterns. The Athens plant in the Capital region began operation during 2004 along with a substantial amount of new generation in New England in 2003 and 2004, which have together reduced the flows over the Central-East interface. In addition, imports from Hydro-Quebec, have decreased substantially since 2002, reducing the loadings on the Central-East interface.

Figure 36 measures the approximate value of congestion in real-time for the interfaces shown in the previous figure. For this analysis, the value of congestion is measured as the shadow price[13]

[13] The shadow price of a transmission constraint represents the marginal value to the system of one megawatt of transfer capability. However, when the real-time dispatch model is unable to solve due to insufficient resources, real-time location-based marginal prices may not be consistent with constraint shadow prices.

of the interface in the real-time market multiplied by the flow. To illustrate their importance, the following figure also shows the value of congestion across up-state interfaces excluding TSA events and eastern 10-minute reserves shortages.

Figure 36: Value of Real-Time Congestion on Major Interfaces 2002 - 2005



In 2005, the value of the up-state transmission interfaces was approximately $250 million, while the value of the down-state interfaces totaled $550 million.[14] The value of congestion was considerably higher in 2005 than in previous years for at least two reasons. First, the rise in congestion costs was consistent with higher overall prices for electricity, which were driven by the dramatic rise in oil and natural gas prices. Second, there were several hundred real-time intervals with acute congestion on the Leeds-to-Pleasant Valley transmission line resulting from TSA operation. These intervals account for approximately $60 million of the congestion costs incurred in the Capital to Hudson Valley corridor. This indicates that the double-contingency

In such cases, this analysis estimates the value of congestion from location-based marginal prices rather than constraint shadow prices.

14 These totals do not equal actual congestion costs paid by market participants because it values congestion based only on the real-time market results which can differ from day-ahead results.

criteria used during TSAs added significantly to congestion costs based on a small number of hours.

2. Congestion Revenue Shortfalls

This sub-section evaluates the congestion revenue shortfalls that arise from differences between the real-time market, the day-ahead market, and the Transmission Congestion Contract ("TCC") market. In this section, we examine the two sources of congestion revenue shortfalls:

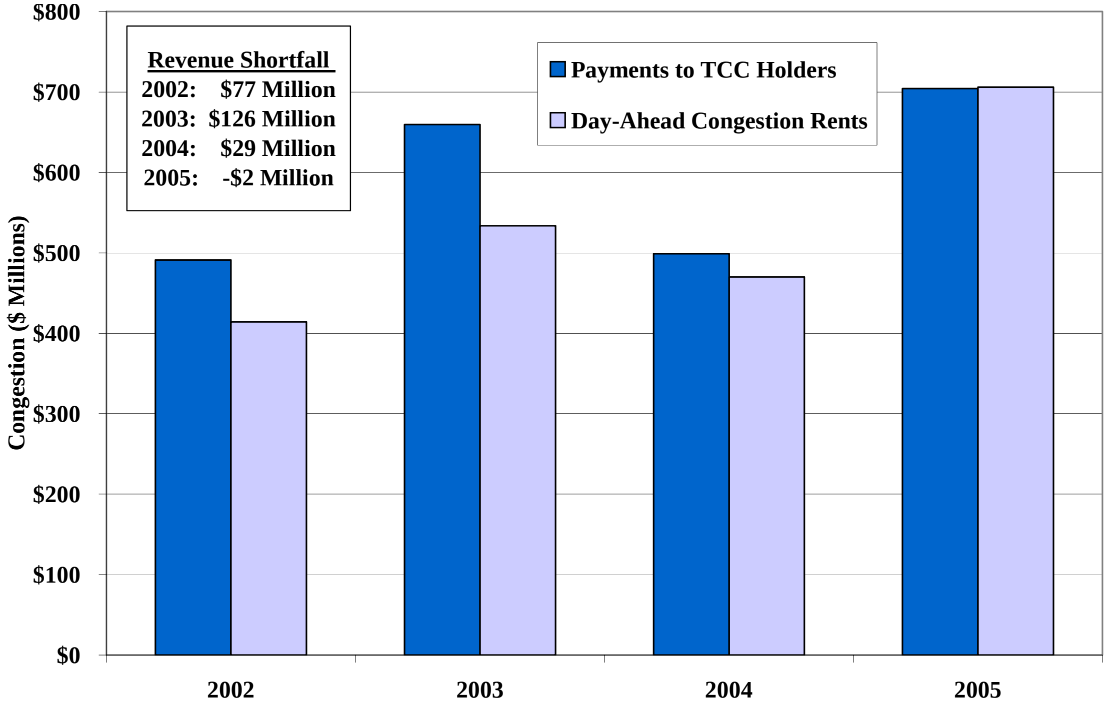
- Day-ahead Congestion Revenue Shortfalls: Revenues collected by the NYISO from congestion in the day-ahead market compared with payments by the NYISO to the holders of TCCs. These arise when the quantity of TCCs sold for a particular path exceeds the transfer capability of the path as modeled in the day-ahead market during periods of congestion.
- Balancing Congestion Revenue Shortfalls: Congestion revenues collected from buyers in the real-time market are not sufficient to cover congestion payments by the NYISO to sellers. These arise when the flow modeled in the day-ahead across a particular line or interface exceeds the actual transfer capability during periods of real-time congestion.

Figure 37 evaluates the significance of day-ahead congestion revenue shortfalls, while Figure 38, Figure 39, and Figure 40 examine the magnitude of balancing congestion revenue shortfalls.

The NYISO conducts auctions to sell the TCCs to market participants. In order to determine the maximum quantity of TCCs that can be sold in a TCC Auction, the transmission system must be modeled to ensure that the TCCs are simultaneously feasible. The NYISO uses a power flow model that includes an assumed configuration of the transmission system. The simultaneous feasibility condition requires that the TCCs awarded be feasible in a contingency constrained economic dispatch of the NYISO transmission system. If this condition is satisfied, the congestion rents collected should be sufficient to fully fund awarded TCCs.

If transmission outages occur that were not modeled in the TCC auction, then the congestion rents collected may be insufficient to meet the TCC obligations. To fully fund TCCs under these conditions, the congestion rent shortfall is charged to transmission owners and passed through to final customers through the transmission owners' service charge. To the extent that these charges are "socialized," they do not provide efficient incentives to minimize the congestion effects of transmission outages. To evaluate the significance of day-ahead congestion revenue shortfall amounts over the past four years, Figure 37 shows day-ahead congestion costs and TCC payments.

Figure 37: Day-Ahead Congestion Costs and TCC Payments 2002-2005



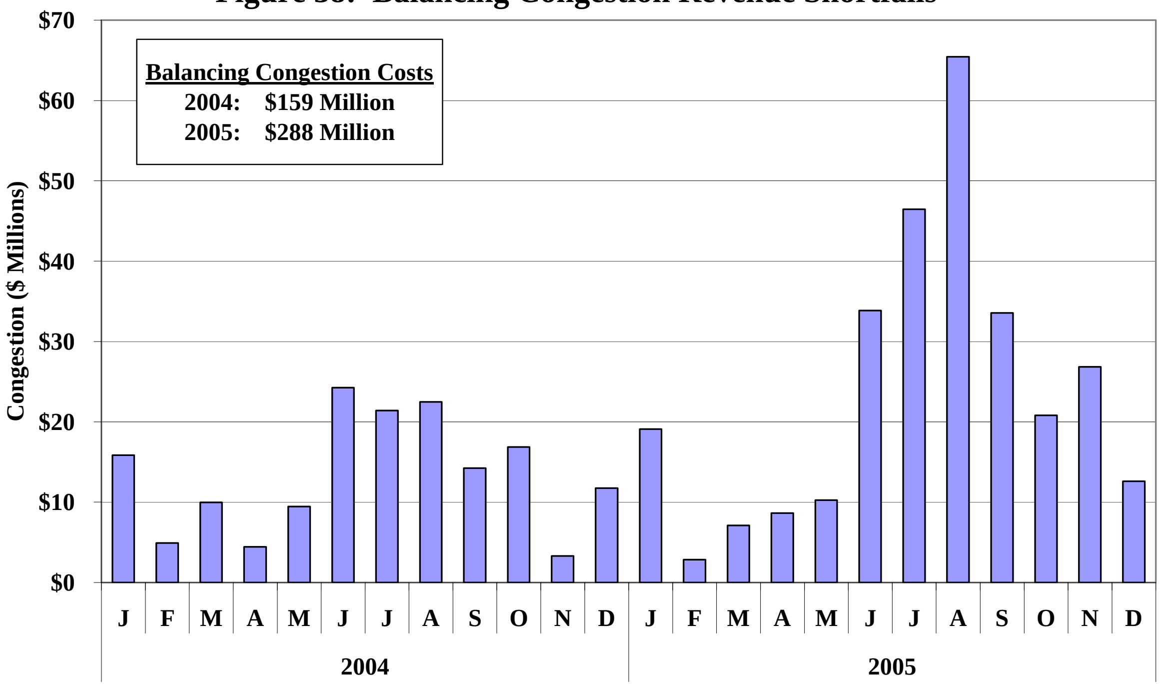
The figure shows that congestion revenues were substantially lower than payments to TCC holders until 2004. This occurred because the transmission capability assumed in the TCC auction generally exceeded the capability modeled in the day-ahead market. The pattern of consistent congestion revenue shortfalls decreased significantly in 2004 when the NYISO took several actions. First, a large share of the shortfall was due to excess TCCs mistakenly sold from up-state New York to New York City. The excess TCCs were repurchased in July 2004.

Second, on December 15, 2003, the FERC approved NYISO's proposal to employ cost-causation principles in assigning responsibility for TCC revenue shortfalls and surpluses to transmission owners ("TOs"). The NYISO now assesses shortfall costs resulting from maintenance to individual transmission owners. This encourages TOs to schedule outages in a manner that minimizes their market effect. Third, the NYISO implemented two mechanisms that allow TOs to retain up to 5 percent of transmission capacity in the form of six-month TCCs not made available in TCC Auctions. The first mechanism permits TOs that hold Existing Transmission

Capacity for Native Load ("ETCNL")[15] to reserve a limited amount of this capacity. The second mechanism, allows all TOs to reserve a limited portion of the residual transmission capacity[16] between contiguous pairs of load zones. Congestion payments for the reserved TCCs will help to offset the TOs' share of a Congestion Rent Shortfall. The FERC approved these measures, subject to minor changes, effective February 2, 2004. Figure 37 shows that together, these provisions reduced the congestion revenue shortfalls in 2004 and eliminated them in 2005.

The next analysis summarizes the additional congestion revenue shortfalls incurred in the real-time market (balancing congestion costs). Balancing congestion costs arise when the flow modeled in the day-ahead across a particular line or interface exceeds the actual transfer capability during periods of real-time congestion. When this occurs, the ISO must purchase additional generation in the constrained area and sell back energy in the unconstrained area (i.e., purchase counter-flow to offset the day-ahead schedule). The cost of this re-dispatch is collected from loads through uplift charges. We examined the congestion revenue shortfall incurred in the balancing market on a monthly basis during 2004 and 2005 in the following figure.

Figure 38: Balancing Congestion Revenue Shortfalls



15 TOs were allocated ETCNLs to facilitate the transition to locational marginal pricing.

16 Once ETCNLs and grandfathered transmission rights are accounted for, the NYISO sells any remaining transmission capacity as Residual TCCs.

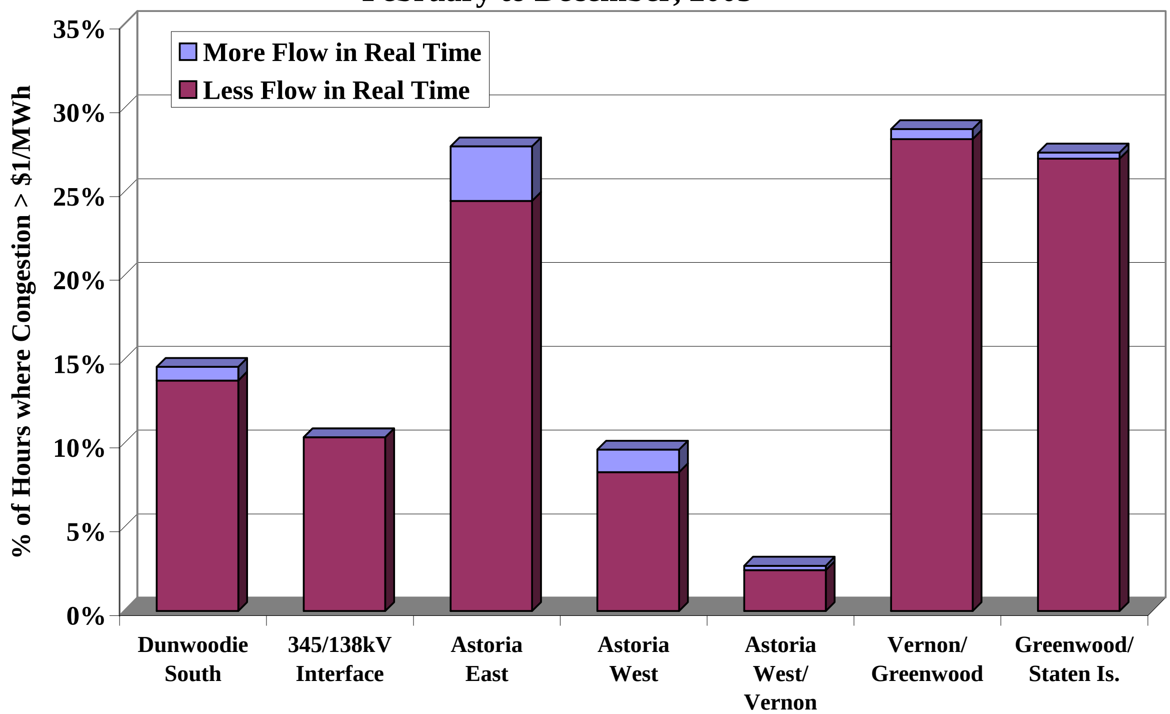
Prior to SMD 2.0 in February 2005, the day-ahead market model did not fully incorporate the effect of losses on transmission usage. Although this was fixed under SMD 2.0 and tends to reduce balancing congestion costs, several other factors have contributed to the rise in balancing congestion. First, total congestion costs increased in 2005, contributing to the increase in balancing congestion costs. Second, current reliability rules require the NYISO to reduce real-time flows across certain key interfaces during TSA events. This forces the NYISO to purchase counterflows to make up the difference between the day-ahead scheduled flows and the actual real-time flows. In 2005, this resulted in approximately $20 million in balancing congestion costs, which were charged to Consolidated Edison whose customers benefit most directly from the additional reliability.

Third, there are differences between day-ahead and real-time transmission modeling resulting in higher effective interface capability in the day-ahead market. The day-ahead market model uses a detailed network of line constraints and a contingency analysis to determine the feasible flows across the network. In 2005, the real-time model used a more simplified network of interface constraints to determine actual flows. Using a more detailed representation of the transmission system allows the day-ahead model to more fully utilize transfer capability. In May 2006, the NYISO began using a more detailed network model to manage some New York City congestion. This change should allow greater transmission usage and help decrease balancing congestion revenue shortfalls.

There may be additional factors that lead actual real-time flow limits to be lower than day-ahead scheduled flows. Although some factors may not be known, it is possible to quantify the systematic reductions in flows from day-ahead schedules to real-time operations. The following two analyses summarize these differences for seven key interfaces during hours with real-time congestion. This analysis only looks at hours with real-time congestion, because the counterflows that generate balancing congestion revenue shortfalls are costless during unconstrained periods.

Figure 39 shows how frequently actual real-time flows were above and how frequently they were below day-ahead scheduled flows in hours with real-time congestion. Periods with actual real-time flows that were lower than day-ahead scheduled flows generated balancing congestion revenue shortfalls, while periods with higher actual real-time flows generated balancing congestion revenue surpluses.

Figure 39: Interface Flows During Hours with Real Time Congestion February to December, 2005



The seven interfaces examined in this analysis accounted for nearly all of the congestion in New York City and a substantial share of the congestion in the state. The Astoria East/Corona/ Jamaica, Vernon/Greenwood, and Greenwood/Staten Island interfaces experienced congestion most frequently, in more than 25 percent of hours. Although the frequency of congestion varied by interface, the figure above indicates that day-ahead scheduled flows exceeded actual flows in the vast majority of hours with real-time congestion for every interface. In these hours, the resulting balancing congestion revenue shortfall depended on the real-time congestion price across the interface and the quantity (in megawatts) of counterflows that had to be purchased by the NYISO.

The following analysis summarizes the average quantity of counterflows that were purchased for each interface during hours with real-time congestion. In particular, Figure 40 shows the average reduction from the day-ahead scheduled flow to the actual real-time flow in hours with real-time congestion for the seven interfaces examined in the previous analysis.

Figure 40: Interface Flows Reductions After the Day-Ahead Market During Hours with Real Time Congestion February to December, 2005

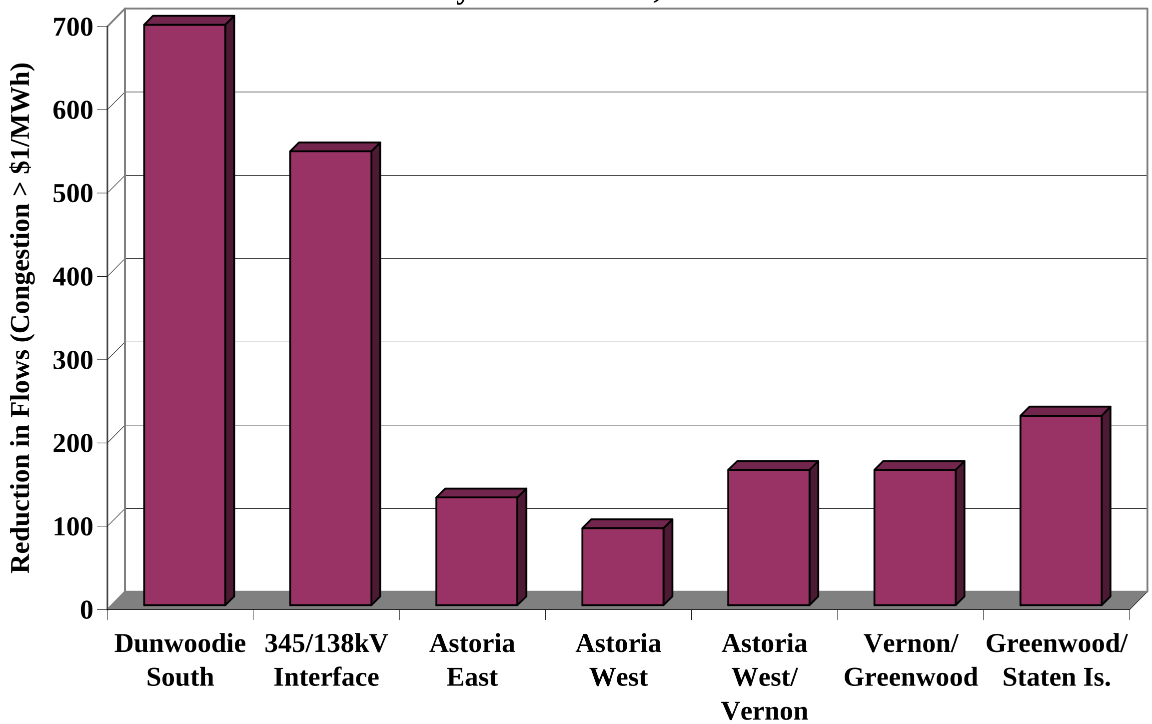


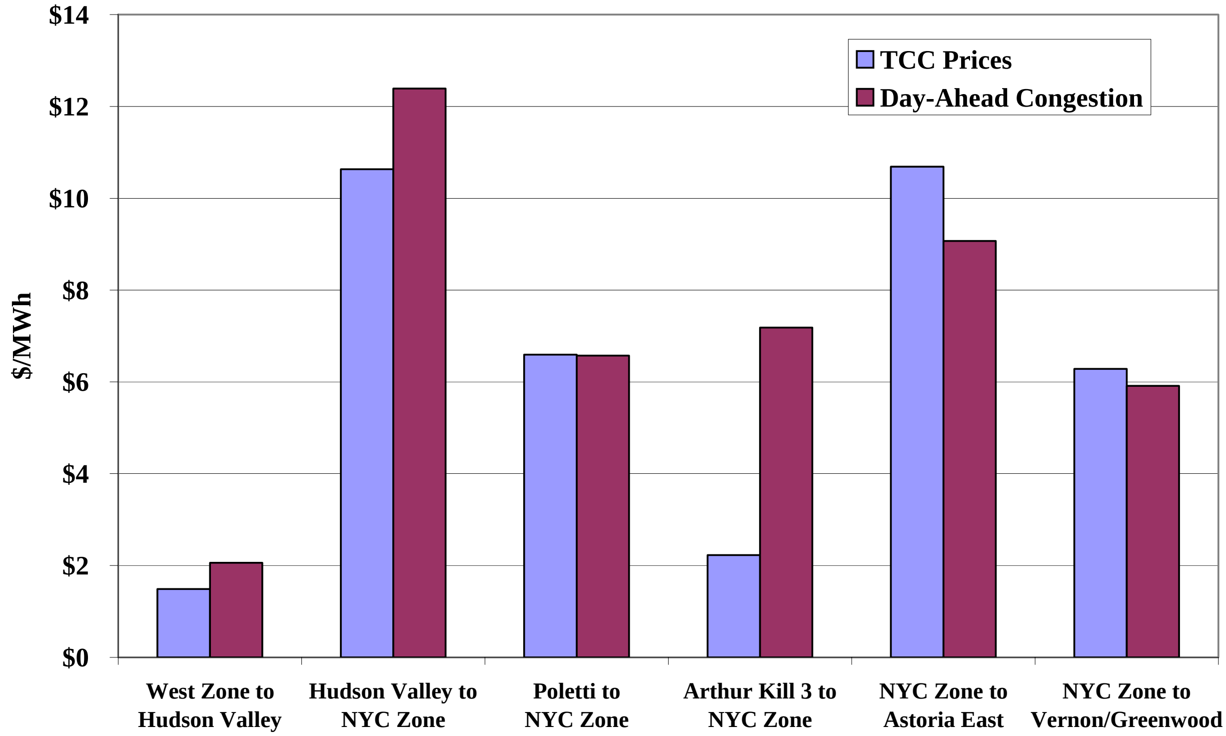
Figure 40 shows that systematically more flows were scheduled in the day-ahead market than actually flowed in the real-time. The figure indicates that the larger load pocket interfaces (i.e. Dunwoodie-South and 345/138 kV constraints) experience larger reductions in capability after the day-ahead market. These systematic reductions from the day-ahead to the real-time contribute to balancing congestion revenue shortfalls.

3. Price Convergence Between TCCs and the Day-ahead Market

Our final analysis in this area is designed to evaluate whether the TCC prices that have emerged from the NYISO's markets converge with the outcomes in the day-ahead energy market. TCCs provide an entitlement to the holder for the day-ahead congestion between two points. Hence, in a well-functioning market, the price for the TCC should reflect a reasonable expectation of the day-ahead congestion.

To evaluate this, Figure 41 compares the auction prices from the auction of 6-month TCCs during the 2005 summer capability period to the day-ahead congestion that actually occurred.

Figure 41: TCC Prices and Day-Ahead Congestion
May to October 2005



For five of the six paths shown in Figure 41, the analysis shows that the TCC auctions predicted the value of the day-ahead congestion relatively accurately. However, the Arthur Kill 3 to New York City zone path was grossly undervalued in the TCC market. Market participants likely anticipated that the Staten Island load pocket, where Arthur Kill 3 resides, would be import-constrained with significant frequency. In the day-ahead market, the Staten Island load pocket was rarely import constrained causing a larger price differential between Arthur Kill 3 and the New York City zone than anticipated in the TCC market.

It is important to recognize that perfect convergence cannot be expected in any one year. When the TCC auctions occur, a number of factors affecting congestion are not known (including forced outages of generators and transmission, fuel prices, weather, etc.). Alternatively, TCC prices should reflect a reasonable expectation of day ahead congestion,

B. Real-Time Commitment and Scheduling

The NYISO upgraded its real-time commitment model as part of the SMD 2.0 implementation. The Real Time Commitment model ("RTC") is primarily responsible for committing gas turbines and other resources with short start times such as external transactions. RTC executes every 15 minutes, looking across a two-and-a-half hour time horizon to determine whether it will be economic to start-up or shut down generation. RTC is a significant improvement over its predecessor, the Balancing Market Evaluation model ("BME") that ran every 60 minutes and evaluated commitment for just one hour.

Convergence between RTC and the Real-Time Dispatch model ("RTD") is important because a lack of convergence can result in uneconomic commitment of generation, especially gas turbines, and inefficient scheduling of external transactions. When excess resources are committed or scheduled by RTC, the results are depressed real-time prices and increased uplift costs. Alternatively, when RTC commits insufficient resources, it leads to unnecessary scarcity and price spikes. This section of the report includes several analyses that evaluate the consistency between RTC and actual real-time market outcomes.

1. Efficiency of Real-Time Commitment and Scheduling

The following two analyses show how SMD 2.0 has improved the efficiency of real-time commitment and scheduling decisions. The first analysis, shown in Figure 42, evaluates decisions to start and to not start gas turbines in 2004 and 2005. This shows substantial improvement, particularly for gas turbines with longer start times (i.e. 30-minute gas turbines). The second analysis, shown in Figure 43, evaluates decisions after the day-ahead market to schedule and to not schedule imports and exports. Although this shows modest improvement, the amount of import and export offers submitted by market participants is small compared with the transfer capability of the external interfaces.

The following figure measures the efficiency of gas turbine commitment by comparing the offer price (energy plus start-up) to the real-time LBMP over the unit's commitment period. When these decisions are efficient, the offer price components of committed turbines are usually lower than the real-time LBMP while the offer price components of off-line turbines are generally higher than the real-time LBMP. However, when a gas turbine that is committed efficiently is close to the margin, it is possible for the offer price components to be greater than the LBMP. Thus, the following figure will tend to understate the fraction of decisions that were economic.

The left panel of Figure 42 shows the volume of gas turbines that were started between June and December in 2004 and in 2005. Based on the comparison of the sum of offer price components and the real-time LBMP over the initial commitment period, these are broken into several categories: (a) offer < LBMP (these commitments were clearly economic), (b) offer > LBMP by up to 25 percent, (c) offer > LBMP by 25 to 50 percent, and (d) offer > LBMP by more than 50 percent. Some of the gas turbines in the latter three categories (i.e. with offers greater than the LBMP) were also economic, because gas turbines that are started efficiently may sometimes not recover their start-up costs. The right panel of Figure 42 shows the quantity of gas turbines that were not started but most likely would been economic if they had been committed. These are off-line gas turbines with energy and start-up offers that were lower than the LBMP for the minimum commitment period of one hour.

Figure 42: Efficiency of Gas Turbine Commitment
Comparison of SMD and SMD 2.0
June to December, 2004 & 2005

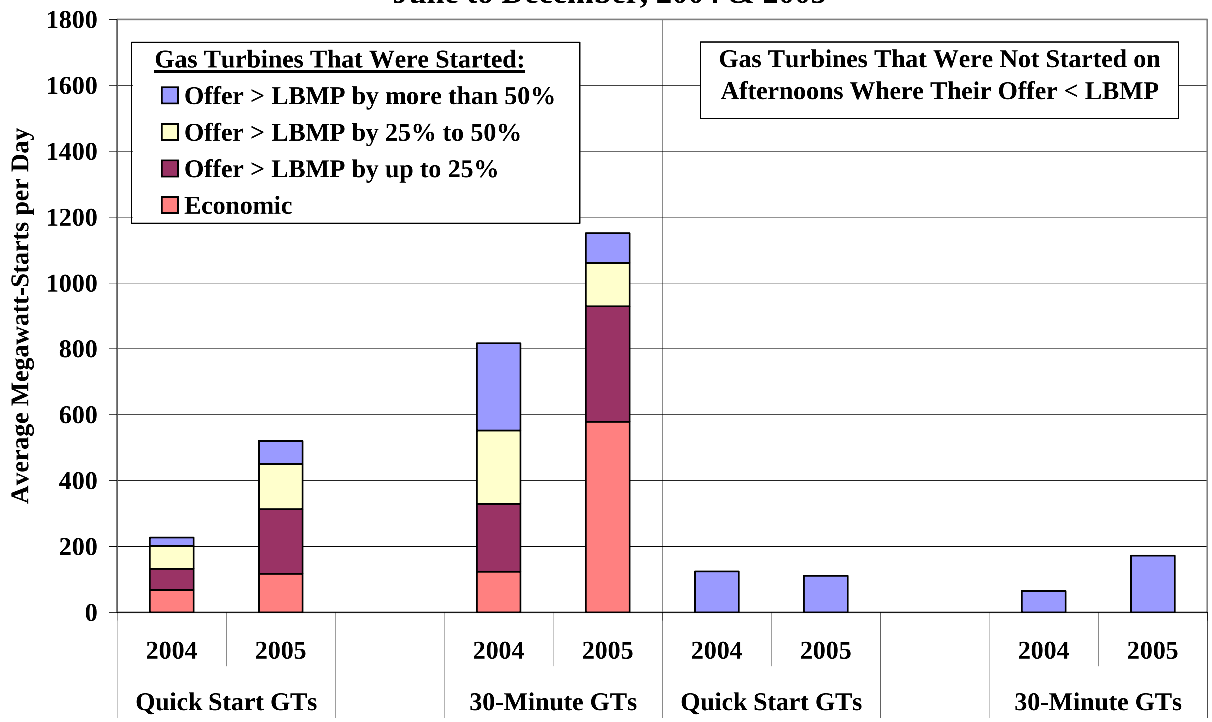


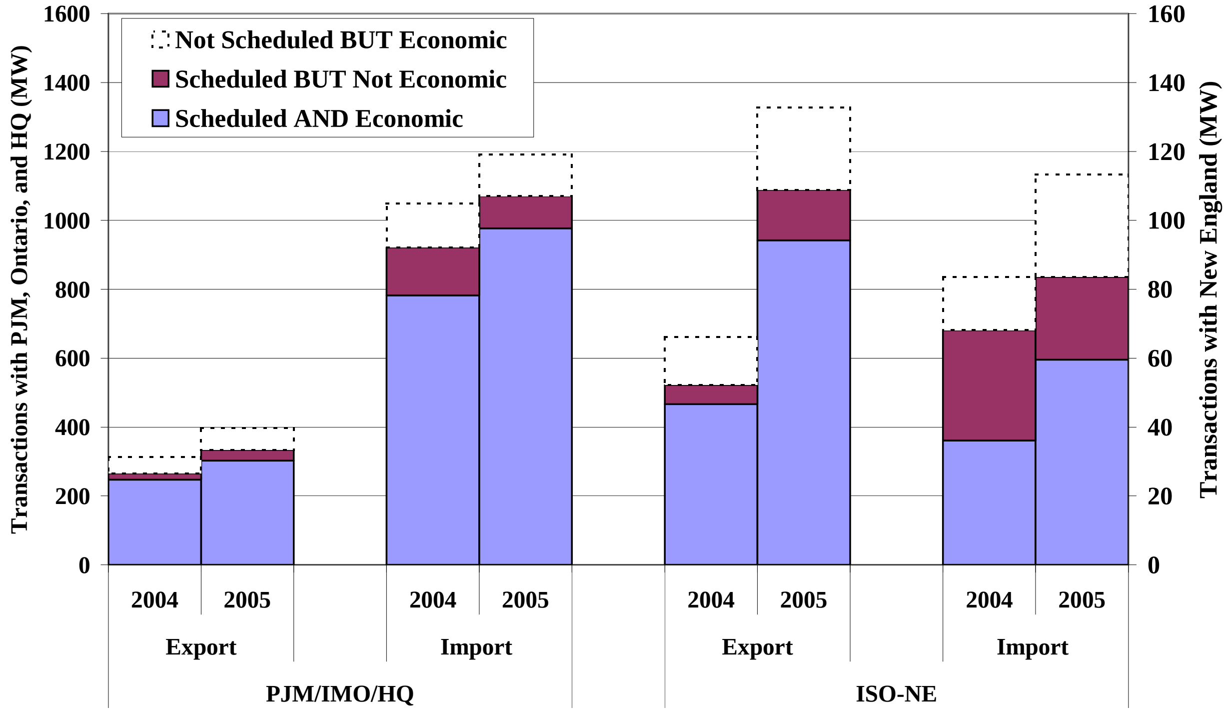
Figure 42 indicates that gas turbine commitment has been far more efficient under SMD 2.0 than in the previous year. The share of gas turbine starts that were clearly economic grew from 18 percent in 2004 to 42 percent in 2005. To the extent that gas turbines were started when their offers were greater than the LBMP, the average margin between the offer and the LBMP was much smaller in 2005. In 2004, the fraction of gas turbine starts occurred when the offer exceeded the LBMP by at least 25 percent exceeded 55 percent compared to only 26 percent in 2005. The right panel of Figure 42 indicates that the average quantity of off-line gas turbines that would most likely have been economic if they had been started grew by 50 percent from 2004 to 2005. This is not surprising given that the higher loads and more frequent congestion lead to a similar increase in the number of gas turbines that were started.

The most substantial area of improvement shown in Figure 42 was in the efficiency of decisions to start 30-minute gas turbines. Prior to SMD 2.0, these units were usually committed by the BME, which ran 75 minutes before the beginning of the hour that it was evaluating. The BME committed resources from the top of one hour to the top of the next hour and did not have the capability to start-up or shutdown a unit midway through the hour. Under SMD 2.0, RTC makes the decision to start these units 45 minutes before the time they are expected to reach full output. RTC repeats this evaluation every 15 minutes, while BME did so only once per hour.

Figure 42 shows less improvement in the efficiency of commitment decisions for quick-start gas turbines. Ordinarily, RTC makes the decision to start these units 30 minutes before the time they expected to reach full output. However, in August 2005, the RTD model was enhanced with the capability of starting quick-start gas turbines when they are expected to be economic for at least one hour. Since the RTD evaluation takes place much closer to real-time, it should be more efficient than RTC for committing resources that have short lead times. This is software enhancement is expected to further improve the efficiency of quick-start gas turbine commitment.

Figure 43 measures the efficiency of external transaction scheduling by comparing the import and export offer prices to the real-time LBMP at the border. The importance of efficient external transaction scheduling in New York is diminished by the fact that most offers to import or export are submitted to the real-time market in a non-price sensitive manner. The following analysis evaluates offers submitted in a price sensitive manner, those with offer prices between $0 and $900 per MWh. Three categories of price sensitive offers to import and export are shown in the figure below: (a) offers that are scheduled by RTC and economic at the real-time price, (b) offers that are scheduled by RTC but not economic at the real-time price, and (c) offers that are not scheduled by RTC but would have been economic at the real-time price. The first category includes offers that were scheduled efficiently, while the second and third categories indicate instances when RTC did not make an efficient decision.

Figure 43: Efficiency of External Transaction Scheduling Evaluation of Price Sensitive Offers* 2004 & 2005



* Includes real-time offers to import or export that are priced between $0 and $900/MWh.

Figure 43 shows that the volume of price sensitive offers that were scheduled grew significantly from 2004 to 2005, particularly for exports to New England, which grew from an average of 66 MW in 2004 to 133 MW in 2005. The fraction of price sensitive offers that were both scheduled and economic stayed relatively consistent from 2004 to 2005, except for imports from New England. The fraction of scheduled imports from New England that were economic rose from 53 percent in 2004 to 71 percent in 2005. Overall, the figure above shows modest improvement in the efficiency of external transaction scheduling. However, the benefit to the market is limited because market participants submit a small amount of import and export offers price sensitively compared with the total transfer capability of the interfaces.

2. Uplift Costs Under SMD 2.0

Under SMD 2.0, the ISO substantially improved the economic efficiency of gas turbine commitment in real-time. This has led to reduced uplift charges for Bid-Production Cost Guarantee ("BPCG") payments to gas turbines that are started by the ISO and turn out to be uneconomic because real-time prices are lower than their offer prices. We performed an analysis of real-time commitment decisions before and after the implementation of SMD 2.0 to estimate how much of the uplift cost reduction was attributable to the new commitment software. We estimate $22 million was saved due to improved real-time operations under the new market software. The findings of this analysis are summarized in the following four figures.

To estimate the uplift savings from more efficient gas turbine commitment under SMD 2.0, we quantify the improved efficiency while controlling for higher natural gas prices, high demand levels, and more frequent price spikes. In the first analysis, we examine how frequently gas turbines ran when their offer (including average start-up cost) was greater than the LBMP, and we find that a larger fraction of production was economic under SMD 2.0. In the second analysis, we show that the vast majority of uplift derives from occasions when the gas turbine's offer is far above the LBMP (i.e. LBMP is less than 60 percent of the gas turbine's offer). In the third analysis, we show how gas turbines that run for a mix of economic and uneconomic hours on a particular day receive substantially lower BPCG payments per unit of uneconomic production. The importance of these factors to the estimated savings is discussed below in greater detail.

The following figure summarizes the efficiency of gas turbine production by comparing the offer prices (energy plus average start-up) of on-line gas turbines to the average hourly real-time LBMP. Each megawatt-hour is classified according to the ratio of the real-time LBMP to the offer price. For instance, if the real-time LBMP is $100 per MWh and an on-line gas turbine has a $150 per MWh offer price, its output will appear in the "60% to 70%" category in the figure below. The analysis allows us to assess the change in the efficiency of gas turbine production between 2004 and 2005.

Figure 44: Efficiency of Production by Gas Turbines Based on Comparison of LBMPs and Offer Prices February to December, 2004 & 2005

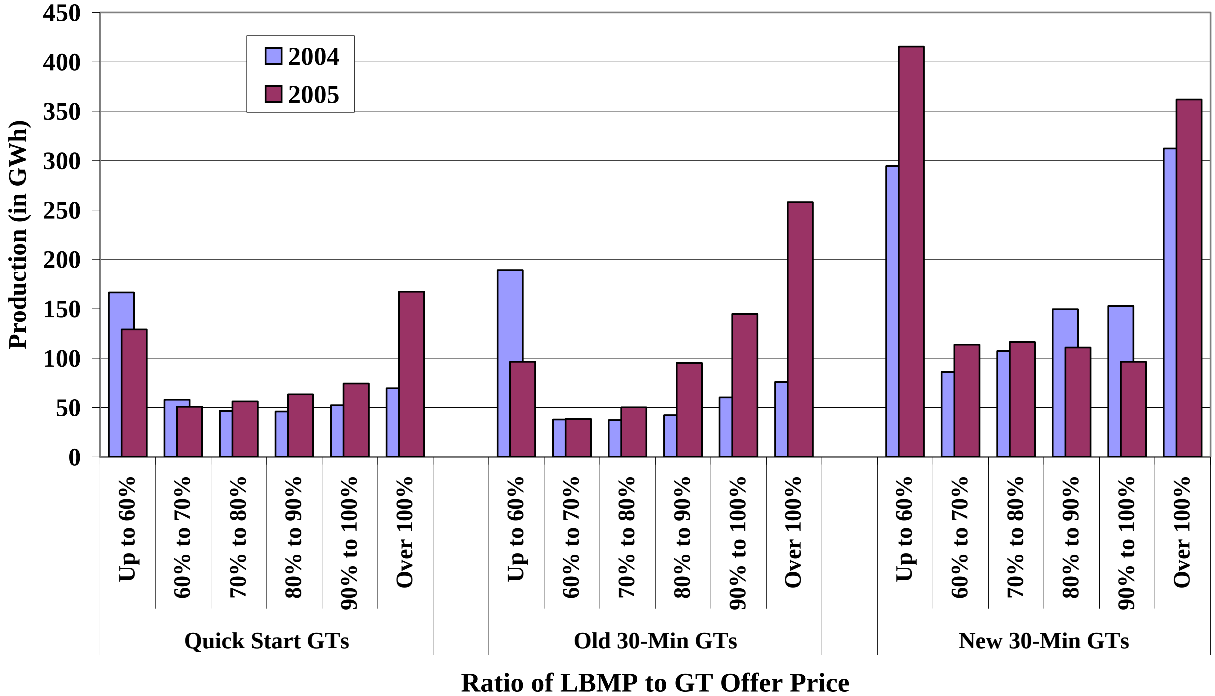
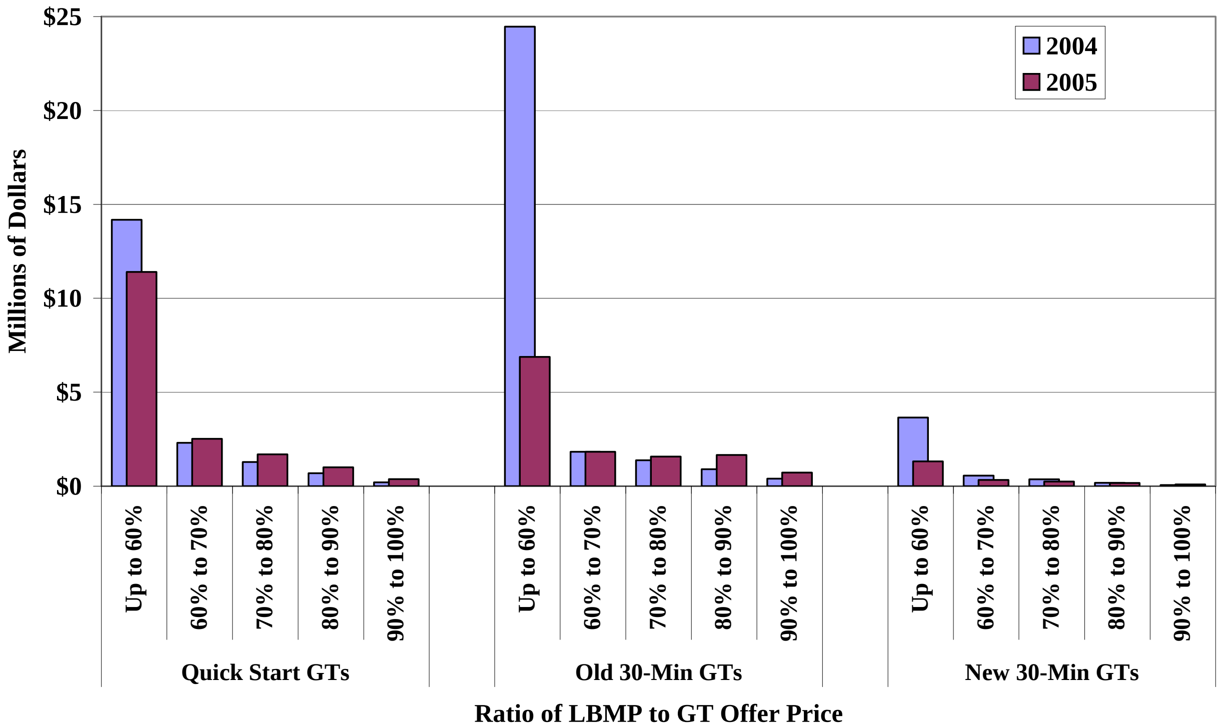


Figure 44 shows that production by quick start gas turbines and older 30-minute gas turbines was clearly more efficient in 2005. For instance, older 30-minute gas turbines produced 240 percent more output in hours that were clearly economic (i.e. when the LBMP was greater than their offer price) and 50 percent less output in hours that were most clearly uneconomic (i.e. when the LBMP was less than 60 percent of their offer price). Figure 44 does not indicate any efficiency improvement for newer 30-minute gas turbines (i.e. ones installed since 2001), which are characterized by lower running costs than older gas turbines. While these account for just 14 percent of the gas turbine capacity in New York, they account for 52 percent of the total output from gas turbines due to their relatively low costs. Due to higher demand during the summer months, electricity production by each type of gas turbine was higher in 2005 than in 2004.

BPCG payments are calculated on a daily basis for gas turbines that are started by ISO but do not receive sufficient revenue to recoup their as-bid costs. Since the calculation is performed daily, a generator does not receive a BPCG payment if its losses in one hour are exceeded by its gains in another hour. However, to assess factors contributing to uplift costs, we have allocated the uplift costs of BPCG payments to the hourly level for each unit.[17] The following figure summarizes these uplift costs according to the efficiency of gas turbine production for three categories of gas turbines in 2004 and 2005. Like the previous figure, efficiency is based on the ratio of the LBMP and the gas turbine's offer price in each hour.

Figure 45: Summary of BPCG Payments to Uneconomic GTs Based on Comparison of LBMPs and Offer Prices February to December, 2004 & 2005



The majority of uplift for BPCG payments comes from hours when gas turbines are producing while the LBMP is less than 60 percent of the gas turbine's offer price. In 2004, 82 percent of the uplift may be attributed to these hours, although they accounted for just 33 percent of the output. In 2005, 63 percent of the uplift may be attributed to these hours, although they accounted for just 26 percent of the output. In 2005, there was a dramatic reduction in BPCG payments associated with hours where the LBMP was less than 60 percent of the gas turbine's offer price, particularly for older 30-minute gas turbines.

17 For generators that are uneconomic in each hour of a particular day, this is simply the difference in each hour between (i) the real-time energy revenue and (ii) the as-bid cost of the generator with the start-up cost amortized over the run time. For generators that are economic for at least one hour of a particular day, the BPCG payments in uneconomic hours are offset by the gains from economic hours. These gains are allocated evenly across megawatt-hours produced in uneconomic hours.

In order to estimate the effect of the new market software on uplift costs for inefficient commitment, it is necessary to control for exogenous factors. Electricity demand and natural gas prices were significantly higher in 2005 compared with the previous year. To isolate these factors, we examined statistics on the average uplift cost per megawatt of uneconomic production under various circumstances. The following figure shows BPCG payments per megawatt-hour of uneconomic production according to the ratio of the LBMP to the gas turbine's offer price. Average uplift costs from BPCG payments are shown separately for hours that occurred on days when at least one other hour was economic. The uplift costs are lower for hours that occurred on days when at least one other hour was economic, because BPCG payments are calculated on a daily basis and gains from high-priced hours go to defray losses from low-priced hours.

Figure 46: Summary of BPCG Payments to Uneconomic GTs
BPCG Payments per Megawatt-Hour of Uneconomic Production
February to December, 2004 & 2005

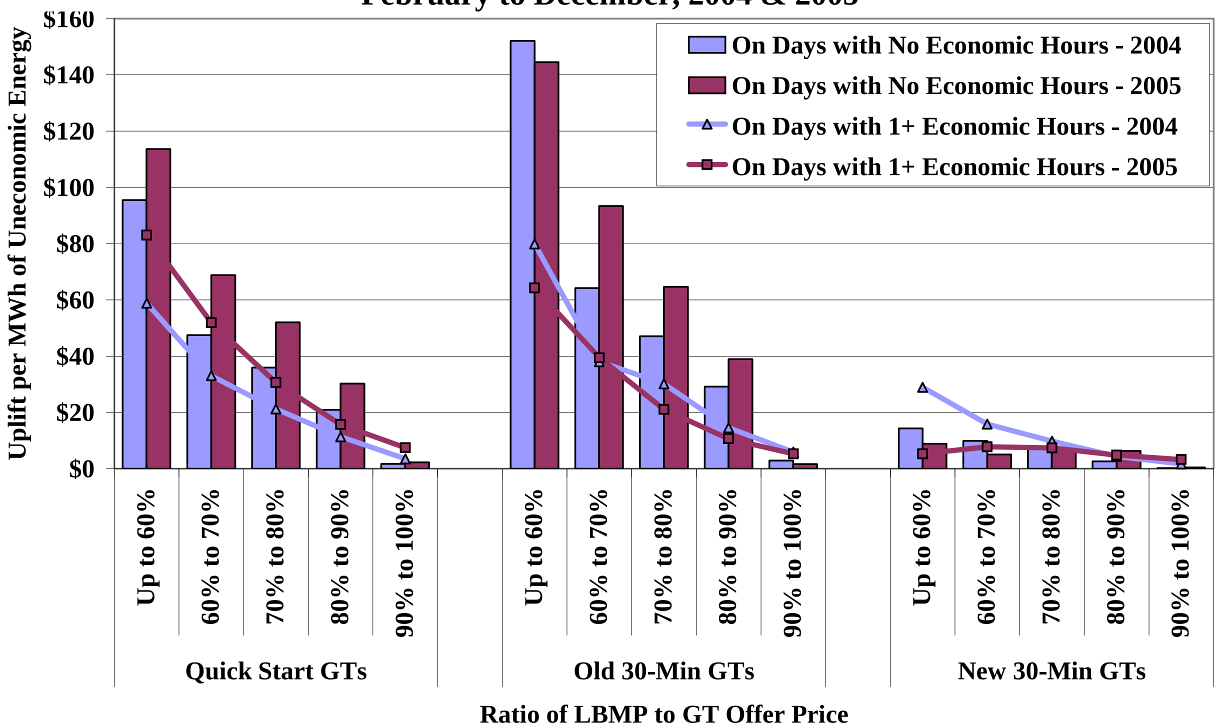


Figure 46 highlights at least two general conclusions regarding uplift costs for uneconomic commitment. First, uplift costs for BPCG payments were significantly lower per megawatt-hour on days when the gas turbine was economic for at least one hour. For instance, in hours when the LBMP was less than 60 percent of the offer price during 2005, quick start gas turbines received an average of $112 per MWh on days where the unit was not economic in any hour and $83 per MWh on days where the unit was economic for at least one hour. Second, due to higher fuel prices in 2005, BPCG payments were generally higher per megawatt-hour of uneconomic production than in 2004. For instance, in hours when the LBMP was less than 60 percent of the offer price on days where the unit was not economic in any hour, quick start gas turbines received an average of $112 per MWh in 2005 compared with $95 per MWh in 2004. Although BPCG payments in 2005 were higher per unit of uneconomic production, overall uplift costs were lower due to a reduction in the volume of uneconomic production.

However, Figure 46 shows several noteworthy exceptions to these general patterns. First, for older 30-minute gas turbines on days where the unit ran economically for at least one hour, the average BPCG payment per unit of production was lower in 2005 than in 2004. The lower BPCG in 2005 can be attributed to greater real-time energy price volatility and more frequent high-priced hours in 2005, which caused these gas turbines to earn substantially more net revenue during economic hours. This helped offset losses in uneconomic hours. Second, BPCG payments per unit of production of the older 30-minute gas turbines were substantially higher in 2004 than in 2005 in the hours when the LBMP was less than 60 percent of the offer price on days where the unit was not economic in any hour. This is because the average ratio of the LBMP to the offer price was lower in 2004 than in 2005, particularly within this limited category of uneconomic hours. Third, the pattern of BPCG payments to newer 30-minute gas turbines differs from those of the other two categories shown in Figure 46. This is primarily because these units sold substantial amounts of energy in the day-ahead market and, therefore, did not receive real-time BPCG payments to the extent that they were uneconomic while fulfilling day-ahead obligations.

Additionally, two general conclusions may be drawn from the analyses in this section of the report regarding the efficiency of gas turbine commitment under SMD 2.0 compared with the prior period. First, the frequency of uneconomic commitment and production decreased, particularly for older 30-minute gas turbines, which resulted in lower BPCG payments. Second, BPCG payments were higher in 2005 per unit of inefficient production, largely due to higher fuel prices. We estimated the uplift savings from more efficient gas turbine commitment under SMD 2.0 from February to December, 2005. This was done by estimating the BPCG payments that would have occurred using the old market software based on the following criteria: (i) the rate of commitment efficiency in 2004 compared to 2005, (ii) in 2005, LBMPs were generally higher relative to gas turbine offer prices, which helped push uplift down, (iii) holding other factors constant, higher fuel prices in 2005 led to higher payments per unit for a particular degree of inefficient commitment. Figure 47 shows uplift costs for BPCG payments in 2004 and 2005 as well as the estimated savings from SMD 2.0.

Figure 47: Estimated Reduction in Uplift from Uneconomic Gas Turbines February to December, 2005

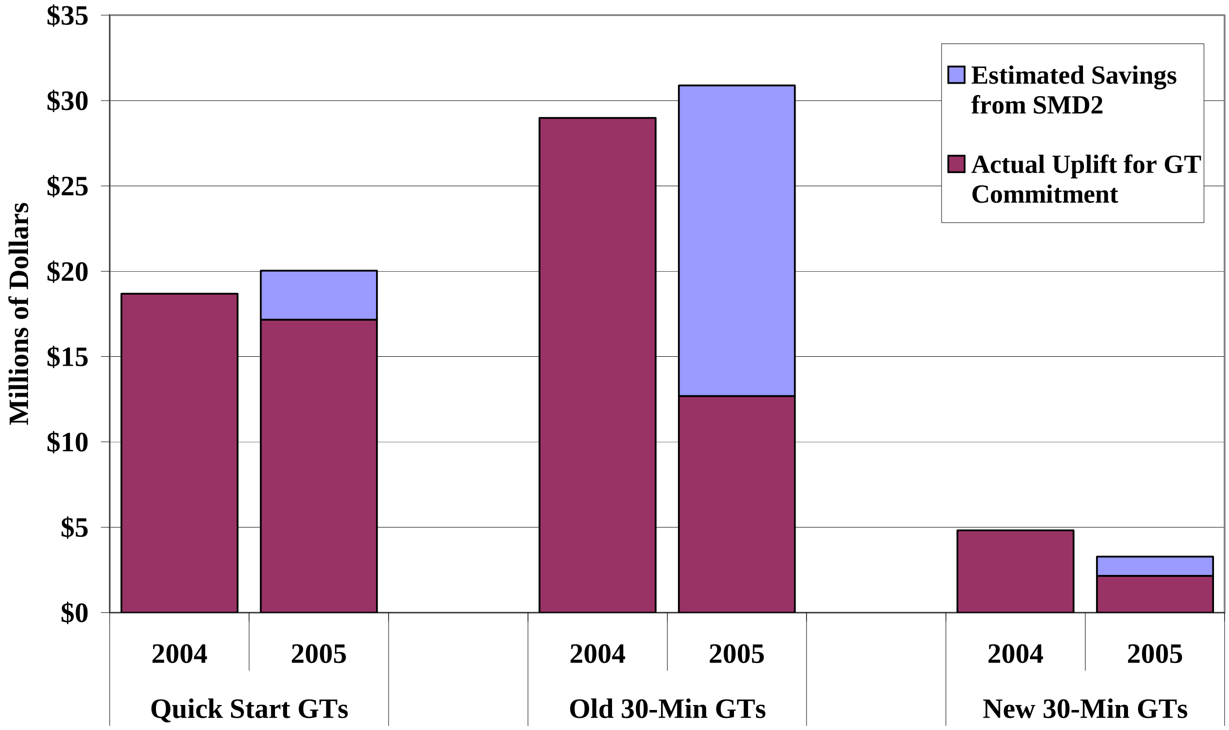


Figure 47 shows that BPCG payments to uneconomic gas turbines committed for non-local reliability reasons decreased from $52.5 million in 2004 to $32.0 million in 2005. We estimate that the implementation of SMD 2.0 reduced uplift by $22 million in 2005, with most of the reduction coming from more efficient commitment of older 30-minute gas turbines.

3. Convergence Between RTC and RTD Prices

Under SMD 2.0, real-time scheduling is accomplished by two models: the Real-Time Dispatch model (“RTD”), which is responsible for matching generation with load and allocating ancillary services on a five-minute basis, and the Real-Time Commitment model (“RTC”), which executes prior to RTD and schedules resources that are not flexible enough to be deployed on a five-minute basis. External transactions are scheduled prior to each hour for the duration of one hour and must be coordinated with the neighboring control area. Off-line gas turbines may take several intervals to start-up and reach full output. RTC executes every 15 minutes committing resources and producing advisory schedules across a two-and-a-half hour window. Consistent with RTD, RTC performs an economic evaluation that commits and schedules the least expensive resources available to meet forecasted demand and ancillary services requirements. This evaluation produces advisory clearing prices for each 15 minute interval across the scheduling horizon. Lack of convergence between RTC and RTD prices can be a substantial concern because large price differences point to inconsistencies that can result in external transactions and off-line gas turbines being scheduled inefficiently; resulting in increased uplift costs and inefficient real-time prices.

With the introduction of SMD 2.0, RTC replaced the Balancing Market Evaluation model (“BME”), which previously scheduled external transactions and committed gas turbines. RTC, which executes every 15 minutes optimizing across a two-and-a-half hour time horizon, is a significant improvement over the BME, which executed hourly producing one set of schedules for one hour. This should lead to closer convergence with the real-time market. One aspect that did not change with the upgrade from the BME to the RTC is that both models cooptimize energy and ancillary services. RTD was designed to cooptimize energy and ancillary services, while its predecessor model optimized energy only, assuming fixed ancillary services schedules. This enhancement should also contribute to better convergence between RTC and the real-time market.

To measure consistency between RTC and RTD, Figure 48 shows a comparison of RTC and RTD energy prices in Eastern New York during 2005. Each RTC execution optimizes across ten 15-minute intervals, and therefore, produces ten sets of prices. The following figure evaluates energy prices from the first of the ten periods, the one closest to the time RTC executes. The evaluation is also limited to the four RTC intervals that occur during the peak demand hour of each day, because efficient commitment is most critical in the peak hour.

Figure 48: Convergence Between RTC and RTD Energy Prices
Eastern New York – Daily Peak Load Hours
February to December, 2005

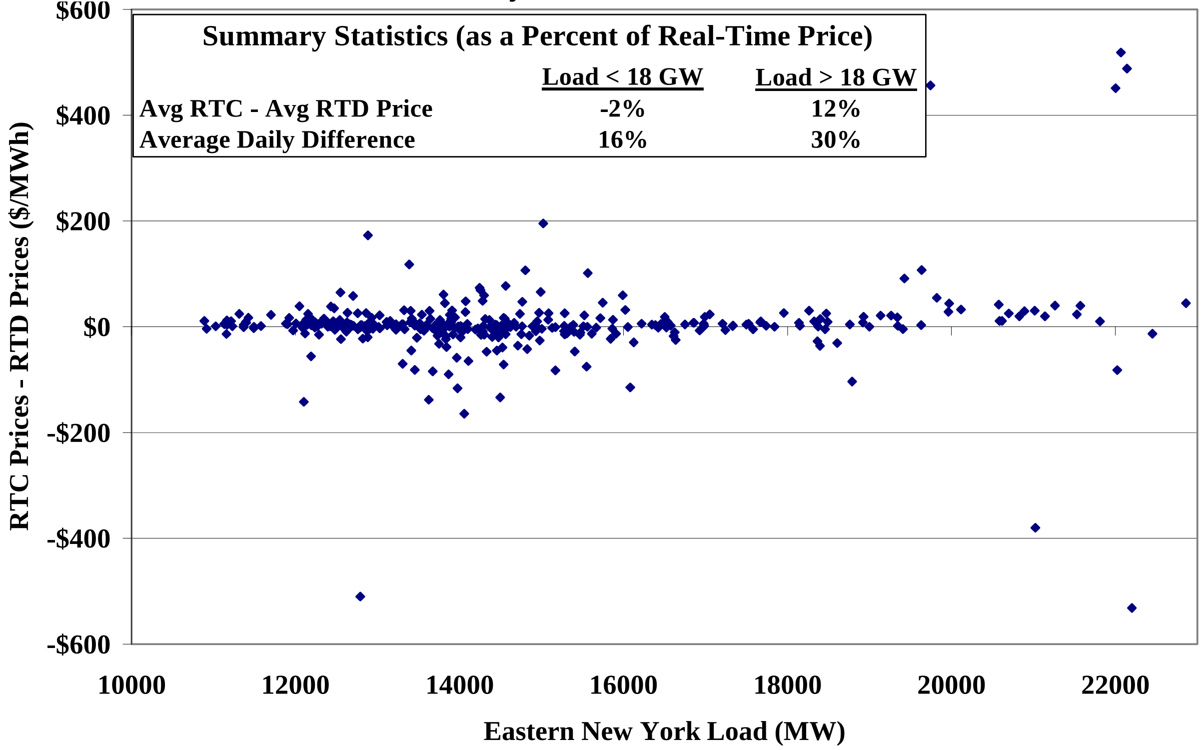


Figure 48 shows the dispersion of price differences on the y-axis compared with the actual load level on the x-axis. Points above $0 per MWh correspond to hours when the RTC price was higher while points below $0 per MWh denote hours when the RTD price was higher. The RTC prices that are shown in the figure above are for the first of ten periods that are evaluated by each execution of RTC. While the other periods are also important to the commitment and scheduling of resources by RTC, the figure provides a general measure of convergence between RTC and RTD.

The figure above clearly shows that convergence between RTC and RTD energy prices tends to be better when load in Eastern NY is less than 18 GW. At low to moderate load levels, the market supply curve is relatively flat, so small differences between supply scheduled by RTC and RTD do not cause large inconsistencies in prices. Convergence was considerably worse when load exceeded 18 GW. Under these high load conditions, the RTC prices were systematically higher than RTD prices by 12 percent. The market supply curve is steep at high load levels, so differences between RTC and RTD in the quantities of supply and demand lead to larger pricing inconsistencies. During high demand periods, the commitment and scheduling decisions of the RTC are particularly important for both preventing unnecessary real-time shortages and not over-committing uneconomic generation that can lead to substantial uplift costs. We plan to investigate factors that lead to inconsistencies between RTC and RTD in a subsequent report.

C. Market Operations under Shortage Conditions

With the introduction of SMD 2.0 in February 2005, the NYISO enhanced its approach to real-time scheduling and pricing of energy and ancillary services. RTD co-optimizes the procurement of energy and ancillary services on a 5-minute basis. In comparison to the Security Constrained Dispatch ("SCD") that was previously used to produce the real-time market outcomes, RTD is a significant improvement in the following ways. First, RTD reduces production costs by efficiently reallocating resources to provide energy and ancillary services every five minutes, whereas SCD simply dispatched resources to provide energy using a fixed set of ancillary service schedules that were produced hourly by BME. Second, RTD is able to incorporate the costs of maintaining ancillary services into the price of energy by co-optimizing energy and ancillary services. These costs were not reflected in energy prices prior to SMD 2.0. Third, demand curves rationalize the pricing of energy and ancillary services during shortage periods by establishing an economic value for the reserves and regulation. The enhancements introduced in RTD provide a more efficient means of setting prices during shortage conditions than the shortage pricing rules that were used with SCD.

This section evaluates the consistency between reserves prices in the real-time market and the actual physical scarcity of reserves after the implementation of SMD 2.0. In particular, this section examines the real-time prices and physical holdings of Eastern 10-minute reserves. The NYISO maintains a requirement of 1,000 MW of 10-minute reserves inside Eastern New York with a demand curve set at $500 per MWh. Thus, the real-time market software will incur a cost of up to $500 per MWh to maintain the required level of reserves. The Eastern 10-minute reserves constraint was chosen for this analysis because it exhibited the highest market value of any reserves requirement during 2005.

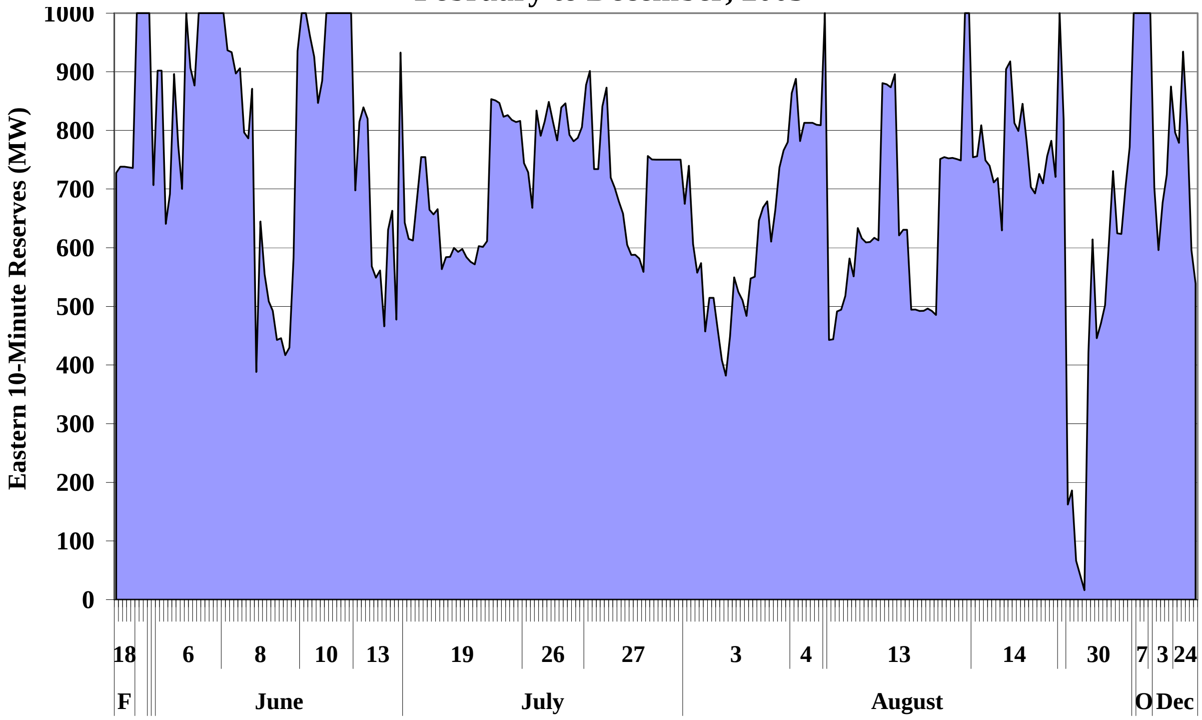
1. Real-Time Prices under Shortage Conditions

Under SMD 2.0, co-optimization of energy and reserves has been integrated with the existing Hybrid Pricing approach, which has been a key element of the real-time market software since 2002. Hybrid Pricing was specially designed to address the problems posed by gas turbines in a marginal cost pricing market. While gas turbines can be started quickly, they are relatively inflexible in the variable operating range. This creates challenges for pricing energy efficiently when the gas turbines are the marginal source of supply, particularly in New York City, where gas turbines account for 34 percent of dispatchable capacity, and in the 138kV load pocket, where gas turbines account for 50 percent of dispatchable capacity. Thus, Hybrid Pricing is particularly important to setting efficient price signals in New York City load pockets.

The Hybrid Pricing methodology treats gas turbines as inflexible resources for the purpose of determining physical dispatch instructions and as flexible resources for pricing purposes. While this facilitates marginal cost pricing when gas turbines are deployed in merit order, it results in certain inconsistencies between the physical dispatch and the pricing dispatch. However, these inconsistencies should be limited such that: (i) prices reflect scarcity during physical shortage conditions, and (ii) high prices are only set when the system is physically in shortage of either energy or ancillary services. The analyses in this section assess whether these inconsistencies occurred frequently during the first eleven months under SMD 2.0.

The first analysis in this section assesses whether high prices have only been set when the system was physically short of a key reserves requirement. Figure 49 shows the amount of Eastern 10-minute reserves that were physically available during intervals of shortage pricing from February to December 2005.

Figure 49: Scheduling of 10-Minute Reserves in East New York During Shortage Pricing Intervals February to December, 2005

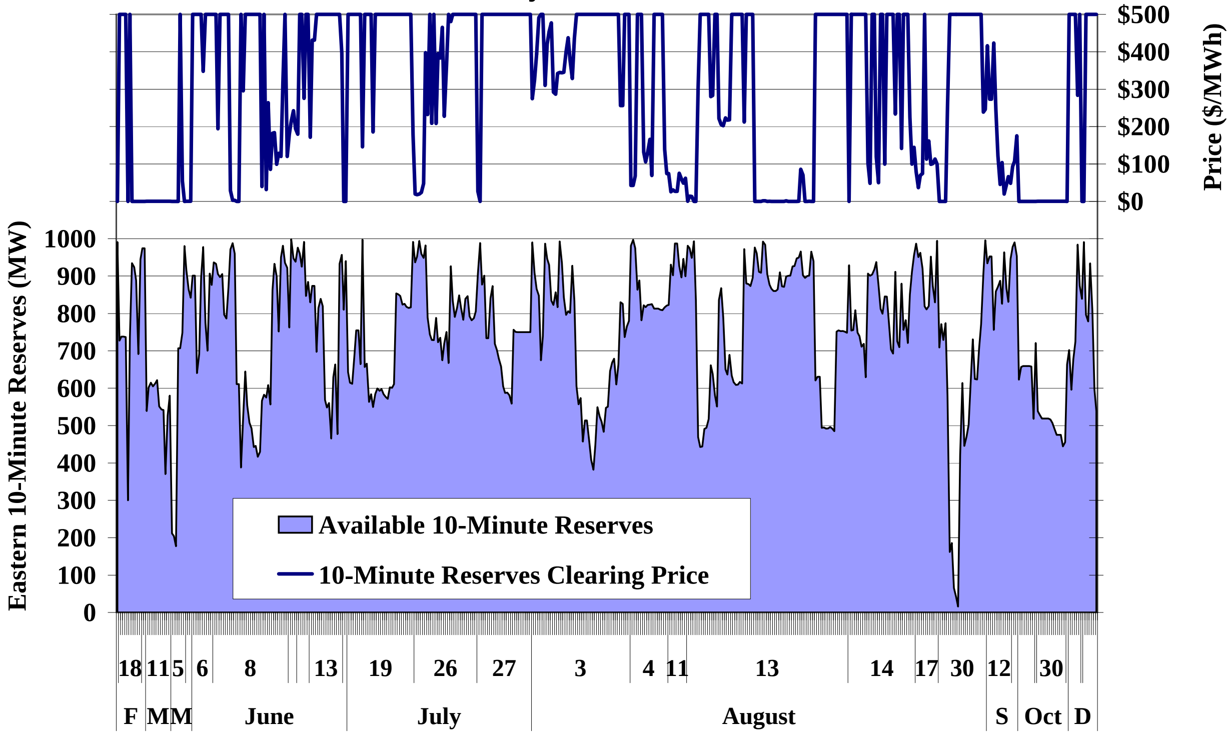


The figure shows 263 intervals with shortage pricing of Eastern 10-minute reserves during the study period. Based on the amount of physically available 10-minute reserves, Eastern New York was short in 89 percent of these intervals.[18] The results in the figure above indicate that the vast majority of shortage pricing intervals associated with the Eastern 10-minute reserves requirement occurred during authentic periods of physical shortage.

The second analysis in this section assesses how frequently physical shortages of Eastern 10-minute reserves are accompanied by shortage prices. The following figure shows the amount of available reserves during physical shortages of Eastern 10-minute reserves. It also shows a line indicating intervals with Eastern 10-minute reserves shortage pricing.

[18] In a previous assessment of the NYISO markets under SMD2, we showed a similar figure with a category of capacity that was available but not scheduled due to a design flaw. However, it was actually not scheduled for reserves in the physical dispatch due to a software error in RTD rather than flaw in the market design. In the following figure, the total quantity of available 10-minute reserves has been adjusted to include this capacity.

Figure 50: Scheduling and Pricing of 10-Minute Reserves in East New York During Physical Shortage Intervals February to December, 2005



Note: Eastern 10-Minute Non-Spin prices exceeding $500 per MWh are shown as $500 in the figure.

For Eastern 10-minute reserves, Figure 50 shows the real-time clearing prices and the available quantity during physical shortage intervals. This figure shows that 235, or approximately 50 percent, of the intervals with physical shortages were not accompanied by shortage pricing, although most of these shortages were relatively small. The shortage quantity was less than 100 MW in 43 percent of these intervals and less than 200 MW in 68 percent of these intervals. The average price in these intervals was $113 per MWh. Although there were a large number of intervals with physical shortages and correspondingly high reserves prices, there were also many instances with low prices during periods of physical shortage.

The efficiency of real-time energy and ancillary services pricing has greatly improved since RTD was implemented under SMD 2.0. It replaced the prior real-time market model, which did not consider how the provision of ancillary services affects the cost of energy. RTD has reduced production costs by reallocating energy and ancillary services every five minutes. Beginning August 16, 2005, enhancements were made to allow RTD to consider off-line quick-start gas turbines in the co-optimization of energy and 10-minute non-spinning reserves. Regarding the initial implementation of RTD in February 2005, there were software errors that resulted in inefficient physical dispatch instructions during periods of shortage. The NYISO has resolved these software issues.

The Hybrid Pricing approach generally enables the real-time software to calculate real-time prices more efficiently, especially in areas that are primarily served by gas turbines. Hybrid Pricing consists of two parts: a physical dispatch pass, which is used to determine physical dispatch instructions, and a pricing dispatch pass, which is designed to treat certain resources differently in order to set more efficient prices when gas turbines are on the margin. However, these differences can affect whether the pricing dispatch perceives a physical shortage of reserves. The next section examines some of the underlying reasons for the differences between the two passes of RTD.

2. Hybrid Pricing

Hybrid Pricing was specially designed to address the problems posed by gas turbines in a marginal cost pricing market. Hybrid Pricing treats gas turbines as inflexible resources for the purpose of determining physical dispatch instructions and as flexible resources for pricing purposes. For instance, in a case where the two most expensive on-line resources are a steam unit and a more expensive gas turbine, the steam unit is the most expensive unit that can be backed down in the physical dispatch pass, which leads the steam unit to be the marginal resource. If clearing prices were based on the incremental cost of the steam unit, the price would be lower than the running costs of the gas turbine. So RTD's pricing pass treats the gas turbine as capable of reducing its output, which allows it to be the marginal resource and set the clearing price. Under these circumstances, the steam unit will have a higher output level in the pricing pass than in the physical dispatch pass, while the gas turbine will have a correspondingly lower output level in the pricing pass than in the physical dispatch pass. This illustrates one of several reasons that output levels for individual resources are not consistent between the two passes of RTD.

In addition to the primary difference between the pricing and physical dispatch passes under Hybrid Pricing (associated with the on-line gas turbines), there are at least two other elements may account for the low clearing prices that sometimes accompany reserves shortages. The first additional element pertains to how ramp rate constraints are formulated in each pass of RTD. The physical dispatch pass constrains the instructed output level of each resource according to its ramp rate offer relative to its actual output level. In contrast, the pricing pass constrains the output level of each resource according to its ramp rate offer relative to its output level in the previous RTD interval's pricing pass. Although Hybrid Pricing was designed this way to facilitate treating gas turbines as flexible in the pricing pass, large inconsistencies can arise when a steam unit does not respond immediately to its physical dispatch instructions.

The second additional element that contributes to inconsistencies between the physical dispatch pass and the pricing pass of RTD has to do with the treatment of gas turbines that have reduced capability due to high ambient temperatures. Since the output capability of gas turbines is inversely related to ambient air temperatures, gas turbine capability tends to be lowest on hot summer days, especially during the afternoon. Given the difficulty of predicting weather, this creates uncertainty about how much capacity is available to RTD. To ensure that RTD matches load with the correct amount of generation, the physical dispatch pass of RTD uses a reduced upper operating limit for gas turbines that appear to be limited by the ambient temperature for several intervals after starting up. The pricing pass of RTD uses the upper operating limit submitted by the market participants, which is typically higher than the ambient temperature limit. The treatment of ambient temperature restrictions on output generally leads to a greater availability of supply in the pricing pass than in the physical dispatch pass.

The first figure in this section summarizes major sources of deviations between the physical dispatch pass and pricing pass of RTD during the first 11 months of SMD 2.0. In particular, Figure 51 summarizes the effects of the following factors:

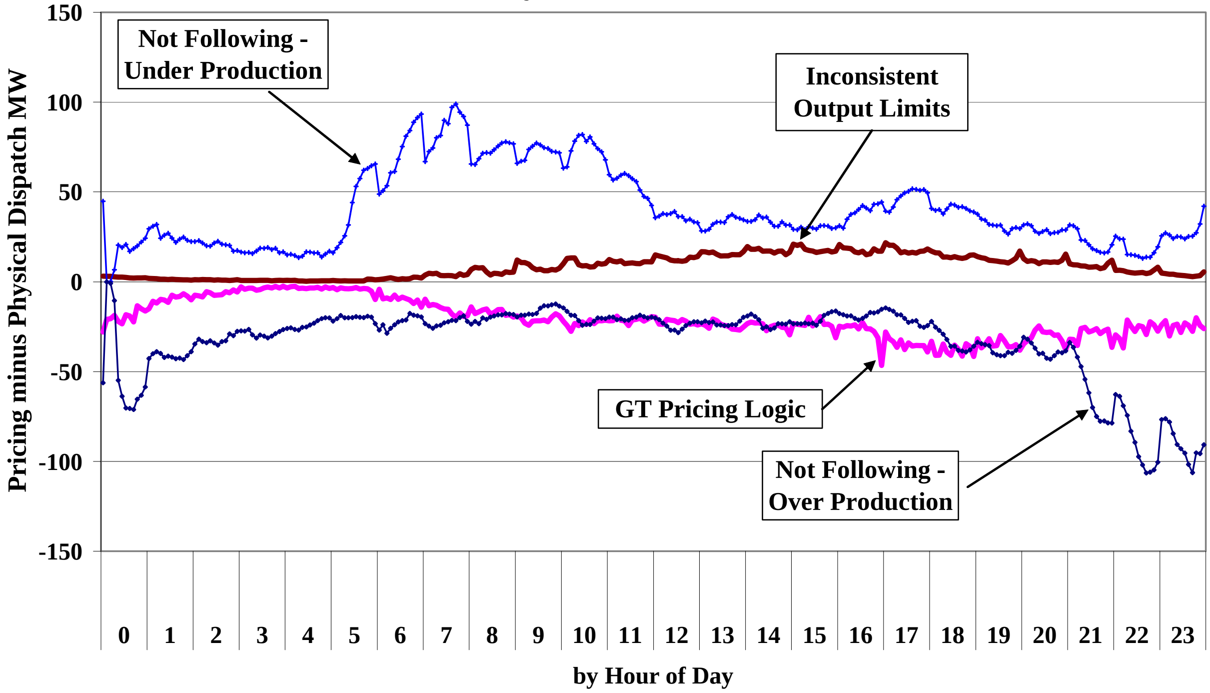
- GT Pricing Logic: RTD's pricing pass treats on-line gas turbines as flexible from zero to maximum output, while the physical dispatch pass assumes they produce at maximum output. Thus, the pricing pass may count less energy from these units.

- Not Following Dispatch: Physical dispatch instructions are "ramp-constrained" by the most recent observed physical output of the unit, whereas the output level in the pricing pass is constrained by the output level from the pricing pass of last RTD run. In the figure below, this group is separated into units that persistently under-produce and units that persistently over-produce as a result of not responding to physical dispatch instructions.

- Inconsistent Output Limits: Inconsistencies between the output capability offered and the actual production level can arise, particularly when high ambient temperatures reduce the maximum output level of gas turbines.

Figure 51 shows these three factors exhibited distinctive patterns according to the time of day from February to December 2005.

Figure 51: Sources of Deviation Between the Physical and Pricing Passes of RTD February to December, 2005



Note: Units are counted as not following dispatch if they lag behind their instruction for three consecutive ramp-constrained intervals.

Several conclusions may be drawn from Figure 51. First, the Hybrid Pricing logic is designed to allow gas turbines to set price when their energy is needed to satisfy load and resolve transmission constraints. This allows gas turbines to have lower output levels in the pricing pass than in the physical dispatch pass. This difference is correlated with the total amount of production from gas turbines, growing from less than 5 MW during the early morning to approximately 40 MW during the afternoon and early evening.

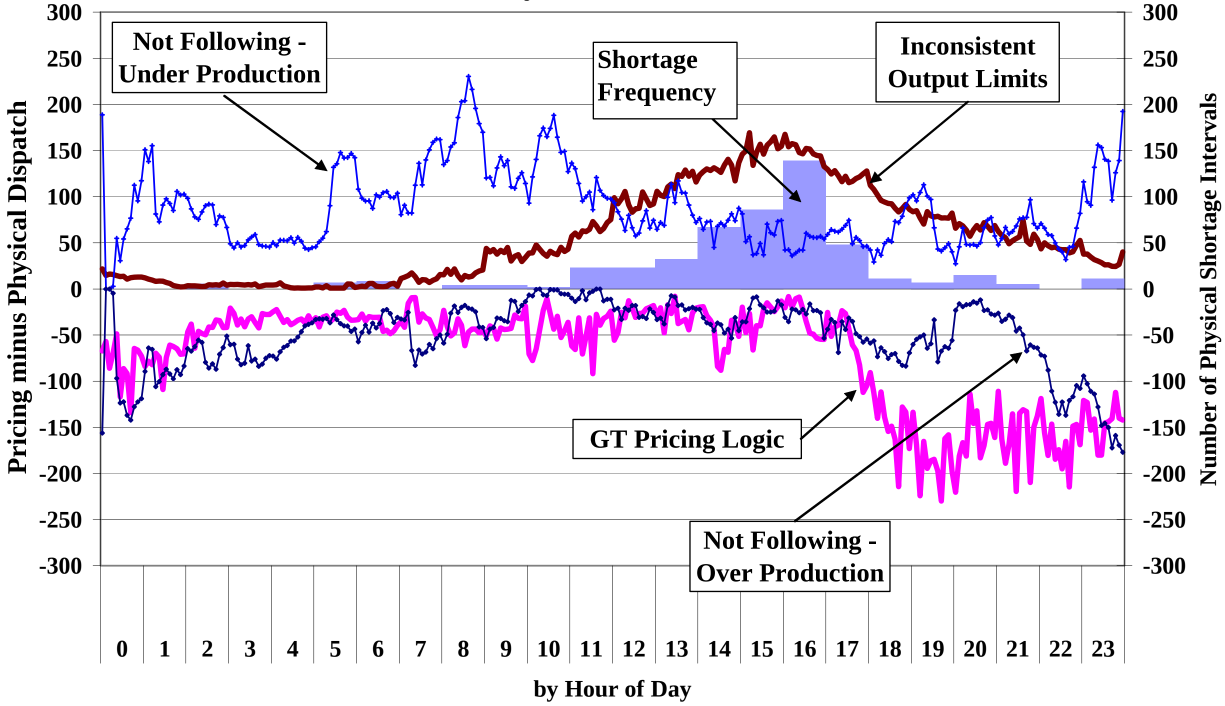
Second, there are several observations that can be made regarding non-gas turbines that do not follow their physical dispatch instructions. Generators that under-produce (i.e. fail to ramp up in response to physical dispatch instructions) may have output levels that are much higher in the pricing pass than in the physical pass. Under-production is largest during the morning ramp-up hours (i.e. 5 am to 11 am). Generators that over-produce (i.e. fail to ramp down in response to physical dispatch instructions) may have output levels that are much lower in the pricing pass than in the physical pass. Over-production is greatest in the evening ramp-down hours (i.e. 9pm to 1am).

Third, the inconsistent output limits associated with ambient temperature deratings of gas turbines are correlated with load, growing from around 0 MW during the early morning to approximately 20 MW during the afternoon. This is expected for two reasons: (i) the total capacity affected by ambient temperature restrictions is generally proportional to the amount of production by gas turbines, which increases with load, and (ii) deratings grow as a fraction of gas turbine capacity with ambient temperatures, which are correlated with load, especially during the summer.

Some inconsistencies between the pricing pass and the physical dispatch pass of RTD are necessary under the Hybrid Pricing methodology. Ideally, these differences should be limited to those that are needed to allow gas turbines to set energy prices in the real-time market. However, other differences are unnecessary and should be minimized. For example, unnecessary differences exist related to the treatment of non-gas turbines that do not obey dispatch instructions and gas turbines that are affected by ambient temperatures. Moreover, these differences may lead real-time energy prices to be inefficient under certain circumstances.

Figure 50 from the previous section indicated that significant differences may arise between the physical dispatch pass and the pricing pass of RTD, resulting in a substantial number of intervals when there was a shortage of Eastern 10-minute reserves while the Eastern 10-minute reserves price was low. Thus, the following analysis examines the major factors underlying differences between the two passes of RTD during high demand periods. Like the previous figure in this section, Figure 52 examines major sources of deviations the passes of RTD, although it focuses on the ten days with most frequent shortage conditions. The following figure also shows the frequency of Eastern 10-minute reserves shortages in each hour on these days.

Figure 52: Sources of Deviation Between the Physical and Pricing Passes of RTD Ten Days with Most Frequent 10-Minute Reserves Shortages in the East February to December, 2005



Note: Units are counted as not following dispatch if they lag behind their instruction for three consecutive ramp-constrained intervals.

These are the periods that are the most sensitive to differences between the physical and pricing treatment by RTD. The figure shows that most intervals with shortages of Eastern 10-minute reserves occur between 2 pm and 6 pm. Hour beginning 16, the hour from 4pm to 5pm, had the greatest number of shortage intervals, nearly 150. During these afternoons, deviations associated with resources not following dispatch instructions and the hybrid pricing of gas turbines were at a minimum.

The most significant contributor to deviations during the afternoons on these days was the inconsistent output limitations of gas turbines due to ambient temperature restrictions. The net effect of this inconsistency was that the pricing pass has an average of 150 MW of additional resources available to it during afternoons of the days with most frequent shortages. This inconsistency explains the two key observations from the previous section: (i) Figure 49 showed a small number of intervals with shortage pricing but no physical shortage of reserves, while (ii) Figure 50 showed a large number of intervals with physical reserves shortages but no shortage pricing.

Some differences between the pricing pass and the physical dispatch pass of RTD are necessary to implement the Hybrid Pricing regime. However, unnecessary differences will generally lead to inaccurate prices and increased uplift. Improving the consistency of the pricing pass and the physical dispatch pass of RTD will improve the efficiency of New York's energy and ancillary services pricing (particularly during shortages) and reduce uplift.

To address these issues, we have two recommendations. In the short-term, we recommend that the NYISO reduce of the upper operating limit of gas turbines according to ambient temperature restrictions in the pricing pass just as it does in the physical dispatch pass. This change, which the NYISO implemented on June 1, 2006, is expected to substantially address the problems with pricing efficiently during reserves shortages. We will evaluate the effect of this change after the summer of 2006. In the long-term, we recommend that the NYISO change the ramp rate constraint in the pricing pass of RTD for generators that do not respond to dispatch instructions.

3. Transmission Shortages and Congestion Price Spikes

Real-time transmission price spikes occur when the re-dispatch costs necessary to resolve a transmission constraint reach extremely high levels. During 2005, there were 562 intervals when shadow prices exceeded $1,000 per MWh on one or more constraints and 238 intervals when they exceeded $2,000 per MWh. The shadow price of a transmission constraint indicates the marginal cost to the system of resolving the constraint. High shadow prices during these intervals contributed significantly to the severity of real-time energy and reserves price spikes.

Spikes in the shadow prices of transmission constraints typically occur for brief periods when there is not sufficient ramp capability within a transmission-constrained area. This may result in large amounts of re-dispatch that provide little reliability benefit when only remote generators are available to be re-dispatched. For instance, there are cases where RTD must re-dispatch 100 MW or more in order to provide one megawatt of relief to a transmission constraint. In such cases, relieving the transmission constraint by re-dispatching hundreds of megawatts may cause shortages of operating reserves. Therefore, the actions taken to maintain reliability by resolving a transmission constraint may actually undermine reliability by causing a shortage of reserves.

In some intervals, the RTD cannot fully solve because there are not sufficient resources available to resolve a transmission constraint. For instance, flow on an interface with a transfer capability of 1,000 MW might only be reduced to 1,020 MW. Brief violations of the transfer limit are not uncommon and do not significantly undermine reliability when they are brief in duration or small in magnitude. Under these circumstances, RTD cannot always set clearing prices that reflect marginal costs, which undermines the efficiency of price signals.

Similar to ancillary services demand curves, transmission demand curves could be used to prevent costly re-dispatch when there is little reliability benefit. Use of transmission demand curves could result in a more reliable prioritization of transmission constraints and reserves requirements when it is not possible to meet all requirements and constraints simultaneously. Under the current operation of the system, the lack of transmission demand curves results in transmission constraints always having priority over reserves requirements. Use of transmission demand curves would also result in more efficient price signals during periods of shortage. Therefore, we recommend that the NYISO continue to evaluate the effect on reliability of using transmission demand curves.

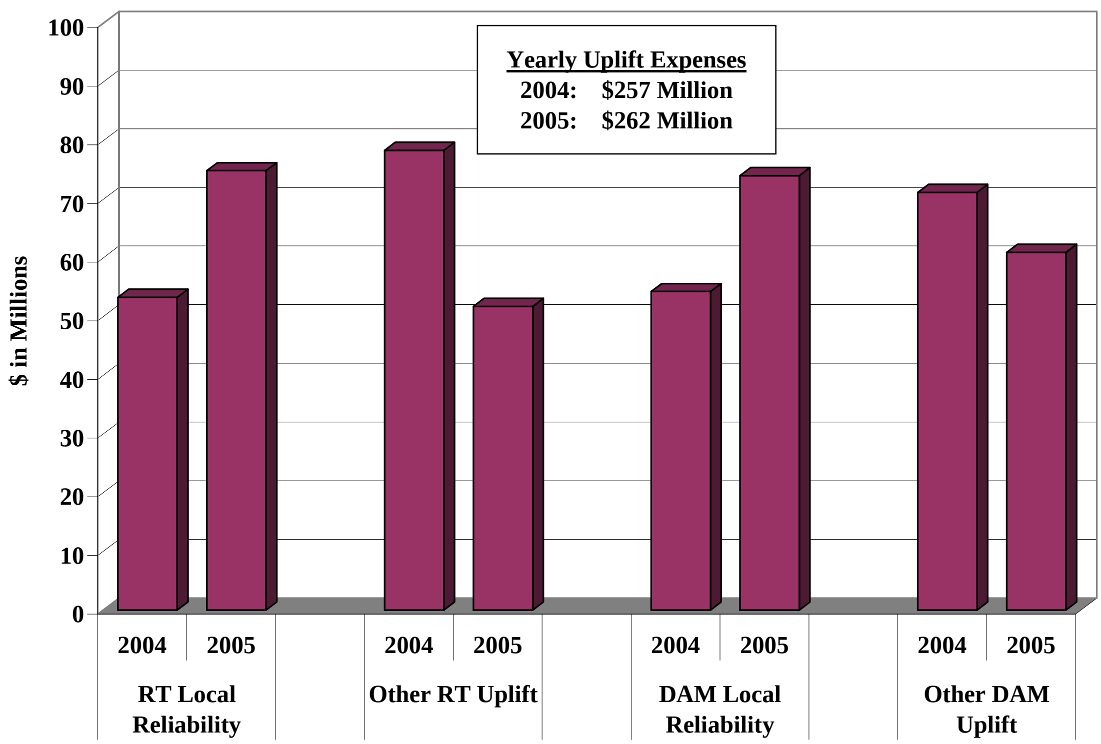
D. Uplift and Out-of-Merit Commitment/Dispatch

In this section of the report, we evaluate patterns of uplift and out-of-merit actions that occurred in 2005. This evaluation is an important component of our overall assessment of the performance of the NYISO's markets because it indicates the extent to which the markets satisfy New York's operational requirements. The first analysis presented in Figure 53 shows the magnitude of uplift costs for BPCG payments in the last two years. The figure shows uplift costs for four categories of BPCG payments, which include payments for local reliability and non-local reliability operation in both the day-ahead market and the real-time market.

There are two categories of day-ahead market uplift, which is paid to units committed by SCUC that do not recoup their as-bid costs from the day-ahead clearing prices. First, uplift is paid to units committed economically that do not receive sufficient revenue to cover start-up costs and minimum generation costs over their entire run time. Second, uplift is also paid to generators committed in the local reliability pass of SCUC, which commits generators out-of-merit in New York City to protect against second contingencies. These commitments for local reliability by SCUC have a tendency to decrease day-ahead prices. As a result of lower prices, more uplift is paid to generators committed before the local reliability pass. Only uplift paid to units committed in the local reliability pass is allocated to the local area, while approximately half of the day-ahead market uplift is assessed market-wide.

There are two categories of real-time market uplift, which is paid to units committed by the NYISO after the day-ahead market that do not recoup their as-bid costs from the real-time clearing prices. First, uplift is paid to generators committed and/or re-dispatched for local reliability reasons. While this can occur for a variety of reasons, the majority of this uplift is related to Supplemental Resource Evaluation ("SRE") commitments. Second, the remaining real-time uplift for BPCG payments goes to units committed economically by RTC and RTD that do not receive sufficient revenue to cover start-up costs and other running costs over their entire run time.

Figure 53: Day-Ahead and Real-Time Uplift Expenses 2004 – 2005



Note: These uplift figures do not reflect all adjustments to real-time local reliability uplift associated with mitigation of OOM and SRE units.

Figure 53 shows that day-ahead and real-time local reliability uplift increased approximately 38 percent, while the total non-local reliability uplift actually decreased 25 percent. Overall, total expenses for BPCG payments remained relatively flat from 2004 to 2005, although the allocation of these expenses has changed substantially, because local reliability uplift is charged to specific transmission owners or load serving entities while non-local reliability uplift is allocated to all loads in the New York Control Area.

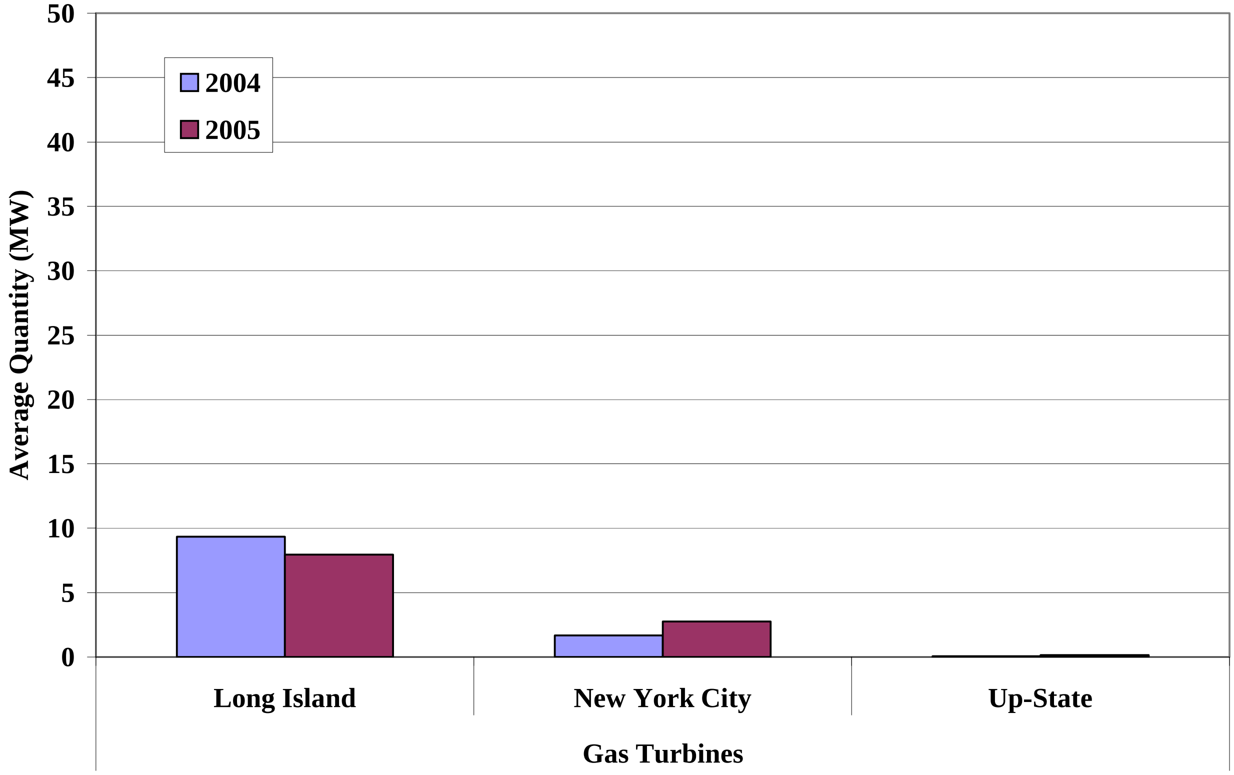
1. Real-Time Out-of-Merit Dispatch

A resource is out-of-merit ("OOM") when it is dispatched by the ISO even though its energy offer exceeds the price at its location. This can be caused by the physical parameters of the unit (e.g., minimum run-time that requires the unit to run after it has become uneconomic) or by operator action. OOM actions are generally taken to ensure reliability and resolve congestion. Actions to ensure reliability in the day-ahead market to ensure enough capacity is committed for the real-time market results in OOM commitment, as discussed in the next subsection. OOM dispatch in real-time can also be used to manage network constraints that are not included in the model.

OOM actions tend to depress spot market prices, particularly during peak demand conditions when prices are most sensitive to small changes in the quantity of load or supply. This is because OOM units are ineligible to set prices and when they are added to the supply stack, the result is to supplant higher-offer units on the margin and depress prices, causing a divergence between the spot price and the actual marginal cost of meeting load.[19] The use of OOM units to maintain reliability also creates a need to make supplemental payments to the OOM units because the spot price is not sufficient to pay the OOM units' offer costs. The costs of these payments are recovered through uplift charges. Figure 54 shows the average quantities of OOM dispatch of gas turbines committed for local reliability reasons in 2004 and 2005.

19 While OOM resources are ineligible to set energy prices, in many cases these resources turn out to be economic (i.e., in merit). Only units that are economically OOM will affect prices.

Figure 54: Average Out-of-Merit Dispatch Quantities
Gas Turbines Called for Local Reliability
2004 – 2005



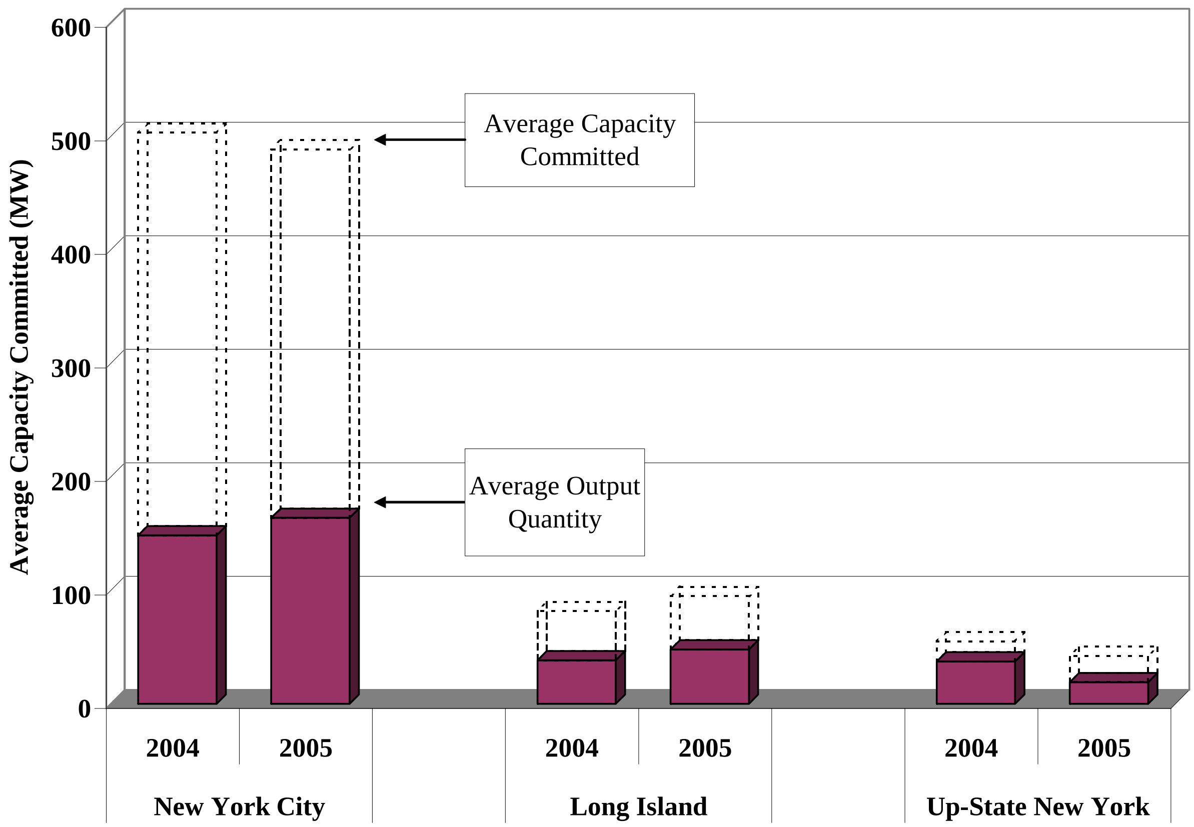
Prior to changes in the modeling load pockets in New York City in June 2002, OOM dispatch for local constraints in New York City accounted for most of the resources dispatched OOM in the real-time market. Because this demand for OOM dispatch has been substantially eliminated, Long Island units now account for most of the OOM dispatch of gas turbines. OOM dispatch quantities are generally very low across the state.

2. OOM Commitment

There are two types of OOM commitment: local reliability commitment by the day-ahead model and the SRE commitment. Day-ahead local reliability commitment is a form of out-of-merit commitment that takes place during the day-ahead market process, as opposed to the SRE that occurs after the day-ahead market closes. The day-ahead local reliability commitment is an element of the SCUC market process whereby some units that are not committed economically may be committed to meet certain specific reliability requirements, particularly second-contingency requirements in New York City. The SRE is a process by which additional resources are committed after the day-ahead market closes in order to meet reliability requirements not included in the SCUC.

Our first analysis in this section, Figure 55, shows the average quantity of SRE commitments made in 2004 and 2005 in New York City, Long Island, and upstate New York.

Figure 55: Supplemental Resource Evaluation 2004 – 2005



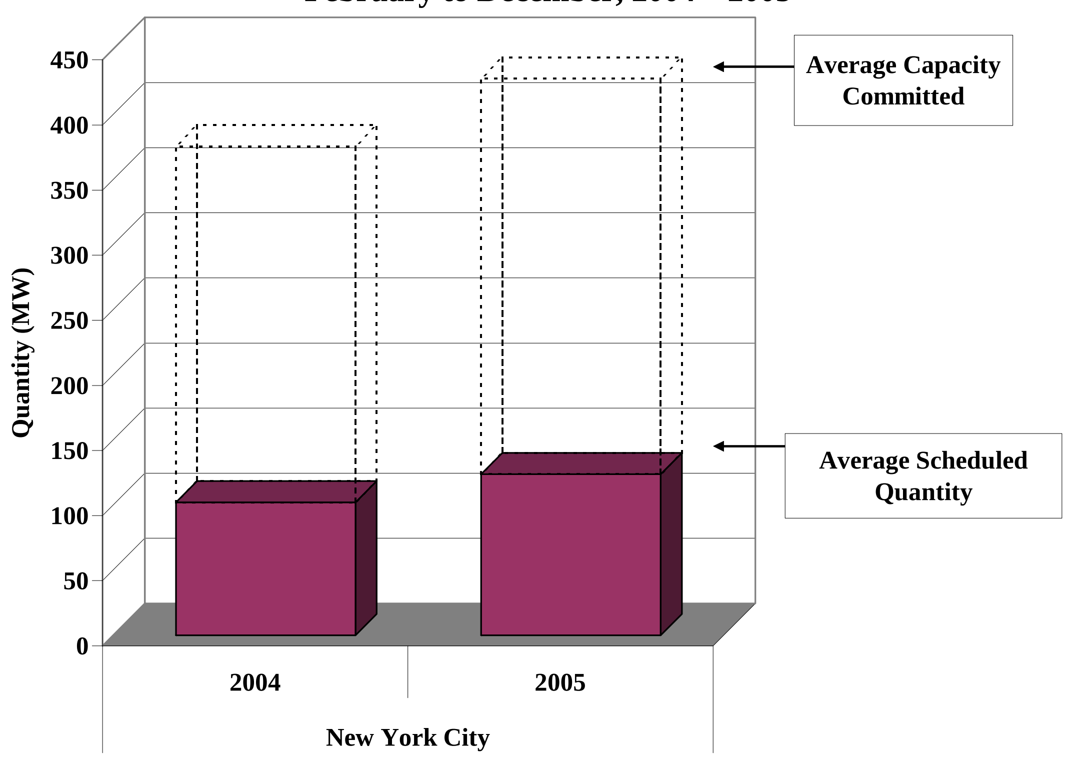
When the operators undertake SRE commitments these actions are logged and reported on the NYISO website. Such supplemental commitments do not directly affect the day-ahead prices, but instead make additional resources available in real-time, and, therefore, may reduce real-time prices as a result of additional units operating at their minimum generation levels. Because of the potential for price distortion as a result of these actions, it is important to evaluate the SRE process and its effects.

One reason for the SREs in New York City is nitrous oxides (NOx) emission limits that require certain baseload units to turn-on in order for gas turbines to operate. The SRE commitments in the City are generally made to satisfy the generators' NOx requirements, which restrict the average emissions (per MWh of output) from a generator's portfolio. Because gas turbines emit NOx at a much higher rate per MWh generated, each supplier must have a steam unit committed to provide the capability to dispatch the gas turbines, if necessary (to keep average emissions below permitted levels). Hence, certain steam units in the City are committed through the SRE process when they are not committed by SCUC. The uplift associated with the SREs constitutes the majority of RT Local Reliability Uplift and is allocated to the local area.

Figure 55 also shows that most of the units committed through the SRE process are dispatched at close to their minimum generation levels (i.e., 25 to 35 percent of the maximum capacity). Hence, although nearly 500 MW of capacity is committed in the City, only 150 MW of additional energy is produced due to these commitments on average. This reduces the effects of these commitments on the NYISO energy markets.

The next analysis focuses on local reliability commitments made in the day-ahead market (i.e., by SCUC). Figure 56 shows the average capacity committed in the day-ahead market for local reliability and the day-ahead scheduled quantity from February to December in 2004 and 2005.

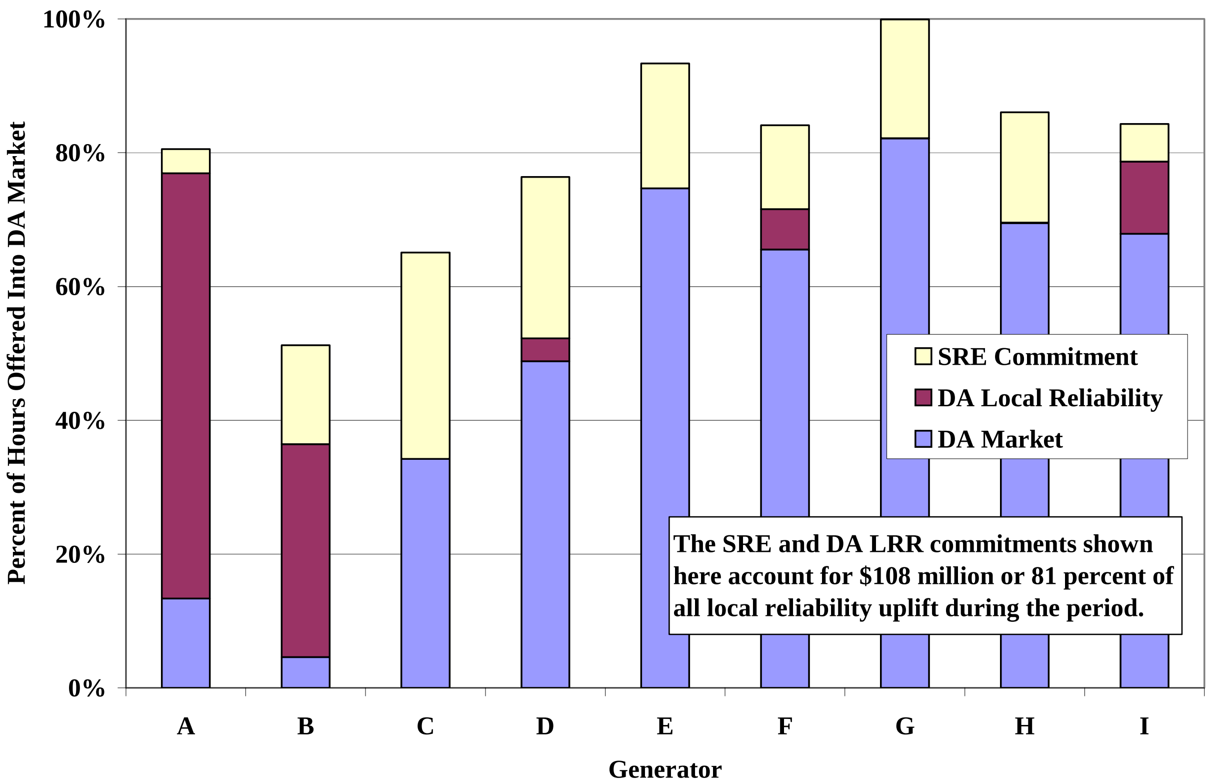
Figure 56: SCUC Local Reliability Pass Commitment
February to December, 2004 – 2005



The commitments shown in Figure 56 tend to reduce prices from levels that would result from a purely economic dispatch; and can increase uplift incurred to make guarantee payments to other generators that will not cover their as-bid costs at the reduced price levels. The figure shows that the average capacity committed for local reliability was approximately 420 MW for the period shown in 2005, which is a modest increase from 2004. Virtually all of the local reliability commitments made by SCUC involved three units in New York City. These units received much lower day-ahead schedules, indicating they are generally scheduled at their minimum generation level. This is the quantity of energy that will affect the day-ahead prices.

In our final analysis of the OOM commitment, we evaluate the frequency of these commitments at the individual unit level. Figure 57 shows nine units that were frequently committed for local reliability by the day-ahead model or through the SRE process. The values shown are the hours that each unit is committed as a percent of the hours that the unit is available (i.e., not on outage). The units in the figure accounted for 81 percent of the uplift costs for local reliability in the day-ahead and real-time. Six of these units are in NYC and three are on Long Island.

Figure 57: Units Most Frequently Committed for Local Reliability and SRE in 2005



Note: DA Market Based included periods when the unit is committed economically in the day-ahead market.

The figure above shows that these nine units were committed for local reliability in a large share of the hours when they were not otherwise committed economically. For instance, Generator A was committed economically in the day-ahead market in 13 percent of hours. However, this generator was committed in 77 percent of the remaining hours by the local reliability pass of SCUC or through the SRE process. In hours when these nine units were available but not committed economically, they were committed in the local reliability pass of SCUC or through SRE at least half the time. Based on this analysis, it seems clear that under certain operating conditions certain generators are predictably needed for local reliability, and will be committed for reliability if they are not economically committed.

It would be more efficient for these units to be committed before or within the economic pass of day-ahead market model. Committing additional units after the economic pass of the day-ahead market model causes some economically committed units to no longer be economic. This leads to excess capacity, depressed clearing prices, and additional uplift.

3. Out-of-Merit Dispatch and Commitment -- Conclusions

Out-of-merit dispatch and commitment have significant market effects. Primarily, they inefficiently reduce prices in both the day-ahead market and real-time market. When this occurs in a constrained area, it will inefficiently dampen the apparent congestion into the area. OOM commitments also may increase uplift payments as units committed economically will be less likely to recover their full bid production costs in the spot market.

Out-of-merit commitment by the local reliability pass of the day-ahead market model increased in 2005, because the resources needed in New York City were committed less frequently on an economic basis. In the long-run, it would be superior to include local reliability constraints into the initial economic commitment pass of SCUC. In the short-run, we recommend that the ISO consider the feasibility and benefits of allowing operators to pre-commit units needed for NOx compliance and other predictable conditions.

Both of these recommendations would require the NYISO to work with participants to revise the cost allocation methodology for uplift associated with the local reliability requirements. Currently, the uplift costs for BPCG payments to units committed for local reliability are allocated locally, while BPCG payments to other units are allocated throughout NYCA. If the recommendations were implemented, a methodology would need to be developed to identify units that were committed as a result of the local reliability requirements.

VI. CAPACITY MARKET

A. Background

The capacity market is intended to ensure that sufficient capacity is available to meet New York's electricity demands reliably. This market provides economic signals that supplement the signals provided by the NYISO's energy and operating reserve markets. In this section we assess the design and competitive performance of the capacity market.

Since 2001, capacity payments have been made for Unforced Capacity ("UCAP") rather than Installed Capacity ("ICAP"). UCAP is a measure of resource availability adjusted to reflect forced outages. Thus, a unit with a high probability of a forced outage would not be able to sell as much UCAP as a reliable unit of the same installed capacity. For example, a unit with 100 MW of nameplate capacity and a forced outage probability of seven percent would be able to sell 93 MW of UCAP. This creates a mechanism that attaches an explicit value to investments in reliability and gives suppliers a strong incentive to maintain their units for reliable performance.

The New York Reliability Council has recommended certain installed capacity margins for the NYISO in order to achieve NERC's one-day-in-ten-years outage standard. Since these recommendations are stipulated in terms of ICAP, the NYISO uses a control area-wide forced outage rate to convert this recommendation into UCAP terms. Likewise, suppliers sell capacity from each of their units on a similarly adjusted basis. An LSE could contract for capacity bilaterally, self-schedule, or rely on the NYISO-run auctions to fulfill its UCAP requirements. The NYISO conducts 3 UCAP auctions: a strip auction that includes each of the upcoming months in the capability period, a forward monthly auction, and a spot monthly action. The two forward markets are voluntary. However, all requirements must be satisfied at the conclusion of the spot market immediately prior to each month.

Starting in June 2003, the New York state-wide ICAP purchases were no longer fixed at 118 percent of peak load. Instead, it varies depending on the market price for ICAP, which is determined using an ICAP Demand Curve in the spot capacity auction that occurs each month. Thus, the ICAP Demand Curve replaced what was effectively a vertical demand curve with a sloped demand curve. In addition, the fixed deficiency charge was replaced with a variable charge equal to the ICAP price that results from the spot auction. For the state-wide capacity requirement, the ICAP Demand Curve was set so that at a capacity of 118 percent of peak load (or the UCAP equivalent in the UCAP deficiency auction), the demand price would be set equal to the annualized cost of a new peaking unit. The demand price would reach zero at 132 percent of peak load, and rise to a maximum of twice the annualized cost of the new peaking unit if capacity declines below the 118 percent.

The ICAP Demand Curves for Long Island and New York City work in a similar manner, but they are adapted to the specific requirements for native generation in those areas. The ICAP Demand Curve for Long Island goes from the annualized cost of a peaking unit at 99 percent of peak load to zero at 117 percent of peak load. The ICAP Demand Curve for New York City goes from the annualized cost of a peaking unit at 80 percent of peak load to zero at 94 percent of peak load. In the unlikely event that the sales of ICAP in New York City were to exceed 94 percent, the New York City UCAP price would be equal to the UCAP price in the rest of New York State.

The results of the monthly spot UCAP auction using the Demand Curve determines each LSE’s obligation for the following month. The aggregate UCAP requirement and the associated UCAP price are established at the point where the supply curve of offers crosses the Demand Curve. All UCAP resources accepted in the auction, including resources offered by LSEs, are paid the applicable market-clearing UCAP price, and all LSEs pay the applicable market-clearing UCAP price for their UCAP requirement.

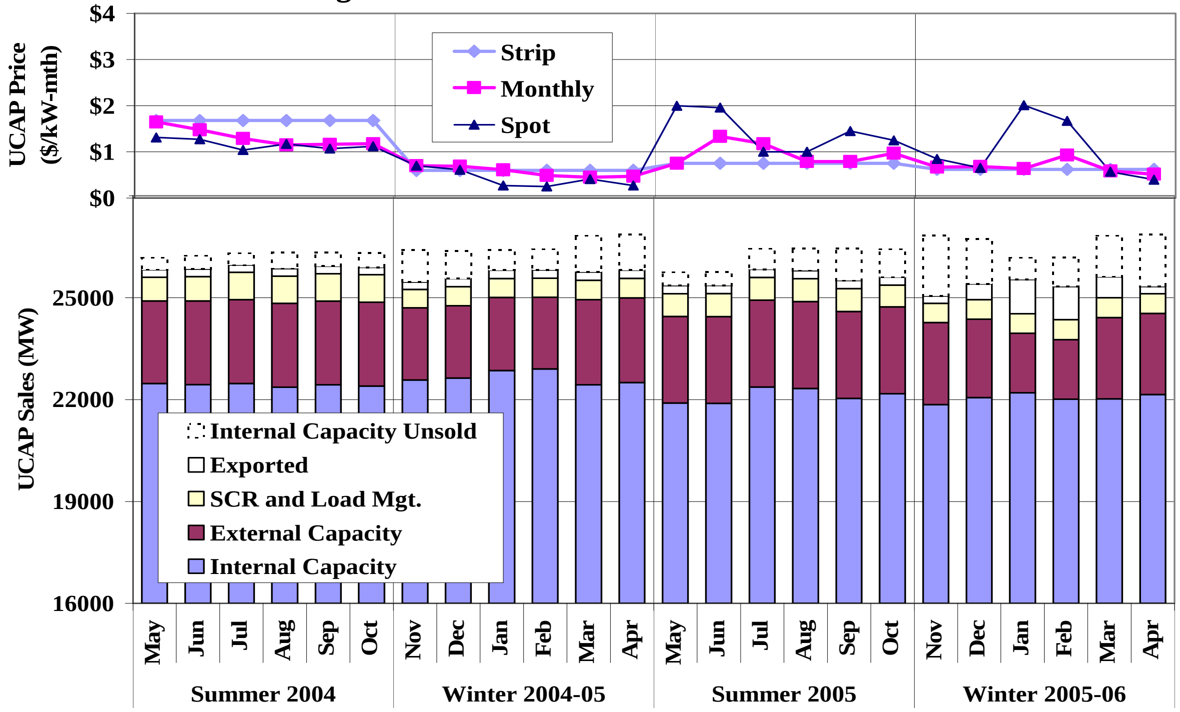
B. Capacity Market Results

To evaluate the performance of the capacity market, Figure 58 and Figure 59 show capacity market results during the two-year period from May 2004 through April 2006. This includes four six-month capability periods from the Summer 2004 capability period through the Winter 2005-06 capability period. These figures show the sources of UCAP supply and the quantities purchased in each month.

The amounts shown in Figure 58 include all capacity offered by New York capacity suppliers outside New York City and Long Island into the New York capacity market. The hollow portion of each bar represents the in-State capacity not sold in New York or in any adjacent market.

Figure 58 shows UCAP prices in the "rest-of-state" area (i.e., the price applicable to capacity outside New York City and Long Island).

Figure 58: UCAP Sales – Rest of State



The figure shows that most of the capacity requirement is satisfied by internal generation, although external suppliers (in the rest-of-state area) and alternative capacity suppliers (including special case resources and load management) each provide a significant amount of capacity in this market. In up-state New York, the amount of available UCAP has fluctuated for several reasons. Approximately 600 MW was retired, although this has been replaced by 700 MW of new capacity that came online in July 2005. The level of imports from other control areas has varied between 1,750 MW and 2,550 MW during the period. The UCAP that can be provided from individual resources varies with changes in the effective forced outage rate of the resource. Although New York is a net importer of capacity, exports from New York rose in January and February 2006, leading to a rise in the spot auction price.

Figure 59 shows that capacity purchases in New York City over the last two years have risen modestly, consistent with the rise in peak load requirements in the City.

Figure 59: UCAP Sales – New York City

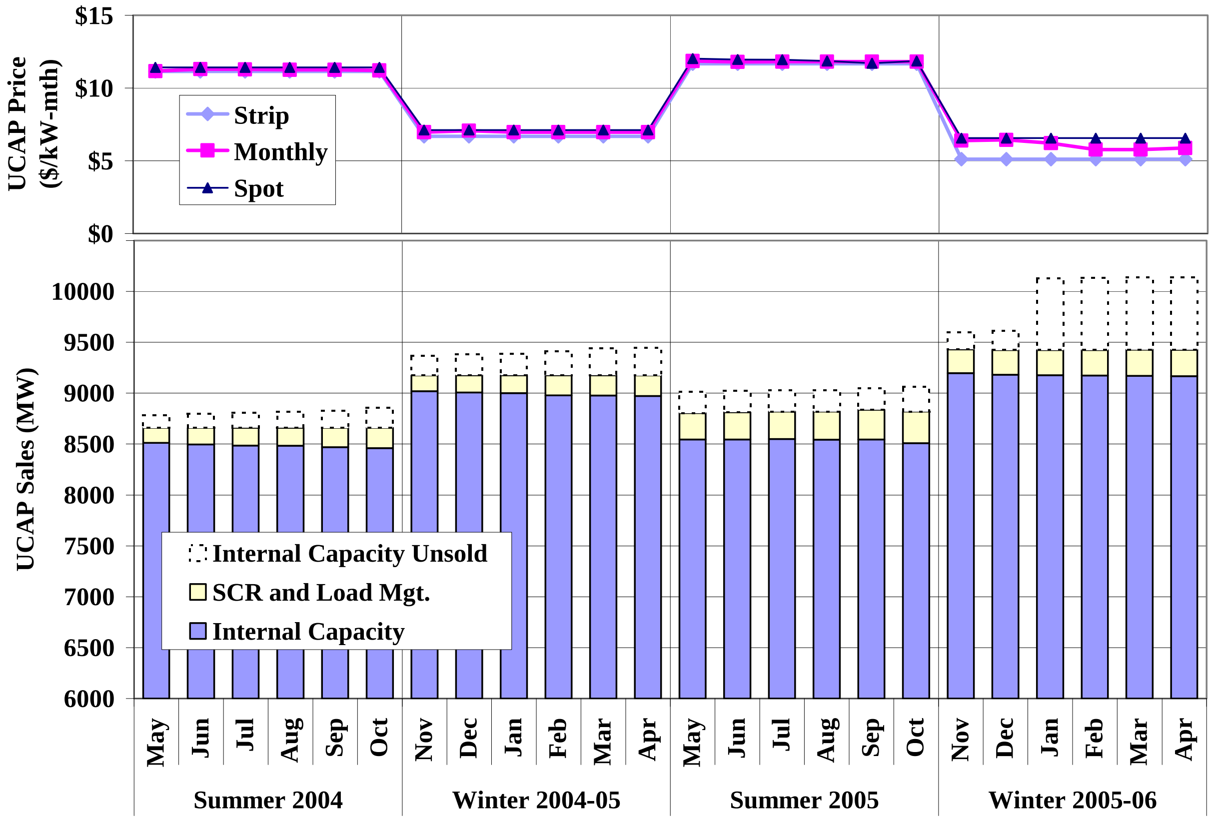


Figure 59 shows that new supply became available in New York City during this period. Approximately 275 MW of new capacity came on-line before the Summer 2005 Capability Period, although this was largely to replace retiring capacity. The amount of capacity satisfied by demand response has increased by approximately 150 MW over the 24 months shown. However, the most notable addition of capacity occurred in January 2006, when approximately 500 MW came into service.

Prior to January 2006, virtually all of the capacity in New York City was sold. However, after the addition of new capacity in January 2006, there was virtually no increase in the amount of scheduled capacity and, thus, no reduction in clearing prices from the In-City suppliers' price cap. After examining the data on capacity and energy outcomes, we found that the unsold capacity participated in the energy market. The results shown in Figure 59 raise concerns regarding economic withholding from the capacity market. The lack of additional sales after the installation of new capacity in January 2006 had a substantial effect on clearing prices in the

New York City UCAP market. The unsold capacity in New York City also raised Rest-of-State capacity prices.

The capacity market in New York City is highly concentrated and these results are consistent with one or more suppliers having market power. Most of the capacity in New York City is owned by several Divested Generation Owners ("DGOs") that purchased the capacity from ConEd when it was required to divest itself of most of its generation in 1998. Regulators foresaw the potential for market power and imposed market power mitigation measures at that time. These primarily consisted of caps on the revenue that DGOs could earn from the capacity market, and a requirement to offer the capacity in the NYISO's market at a price no higher than the cap. Although this provision was intended to mitigate the DGO's market power, it allows the DGOs to raise prices substantially above competitive level under conditions when NYC has surplus capacity. It is also important to recognize that the price that the DGOs paid for the in-city units were likely substantially higher than they would have been if more effective market power mitigation measures had been required. Hence, modifying the mitigation measures to more effectively mitigate the market power in NYC would raise significant equity considerations.

VII. External Transactions

This section evaluates the extent to which prices have been efficiently arbitraged between New York and adjacent regions by analyzing the price differences between the markets and the use of the interfaces. Although several market design improvements have been made in recent years to improve the efficiency of flows between adjacent markets, the interfaces are still not fully-utilized. There are additional changes that should be made to improve the efficient price convergence at these "seams" between New York and the adjacent markets.

In particular, we encourage the NYISO to continue working with ISO New England to develop the external scheduling provisions to enable the two markets to realize many of the benefits of a larger control area. In April 2005, PJM and the Midwest ISO implemented the Joint Operating Agreement ("JOA") to coordinate congestion management in the two markets. Under the JOA, the dispatch software of each market incorporates transmission constraint information from the other market in real-time, allowing for more efficient congestion management and pricing in the two markets. The JOA could serve as a model for future coordination between New York and adjacent markets. However, in the near term, it is reasonable for New York to focus on implementing external scheduling provisions with New England to improve the price convergence between the markets.

Price convergence occurs when the energy prices at the border are equal in the absence of transmission congestion. In real-time, it has proven difficult for the adjacent markets to achieve price convergence by relying on transactions scheduled by market participants. Uncertainty, imperfect information, and offer submittal lead times limit the ability of participants to capitalize on real-time arbitrage opportunities. This failure of real-time arbitrage gives rise to market inefficiencies that could be remedied if the ISOs were to coordinate interchange to reduce or eliminate the price differences.

A. Price Convergence between New York and Other Markets

The performance of the wholesale electricity markets depends not only on the efficient use of internal resources, but also the efficient use of transmission interfaces between New York and other areas. Absent transmission constraints, trading should occur between neighboring markets to cause prices to converge. When the interfaces are efficiently utilized, one would expect that the hourly prices in adjacent areas would not differ greatly except when the interface capability is fully used (the interface constraint is binding). In other words, when prices are higher in New England than in New York, exports to New England should continue until the interface is fully scheduled or until prices have converged and no economically-viable exports remain.

The series of scatter plots/charts in Figure 60 show the hourly difference in real-time prices between New York and neighboring markets relative to net exports during hours when transmission constraints are not binding.

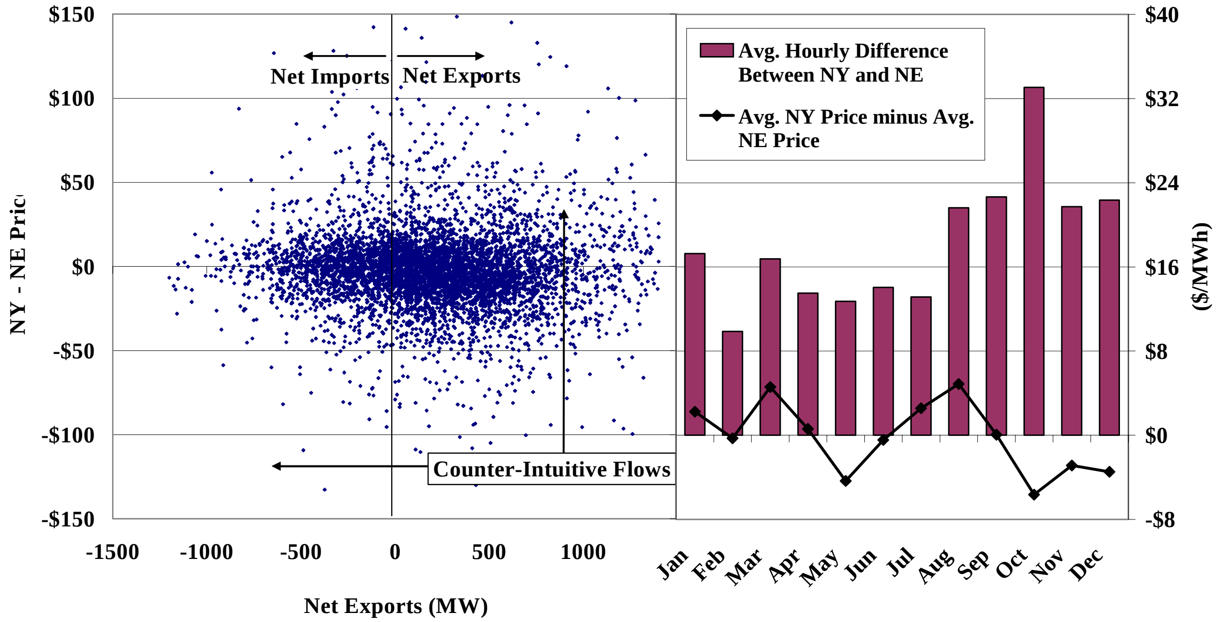
On the left side of the figures:

- The price differences plotted against the left axis are always computed by subtracting the external price from the New York price (i.e., positive price differences mean prices are higher inside New York). The top half of each scatter diagram, therefore, reflects hours when the price in New York was higher than the price in the neighboring region.
- The net exports are shown on the x-axis with positive values reflecting net exports from New York and negative values representing net imports.
- Two "counter-intuitive" quadrants are shown where power is scheduled from the higher priced market to the lower priced market.

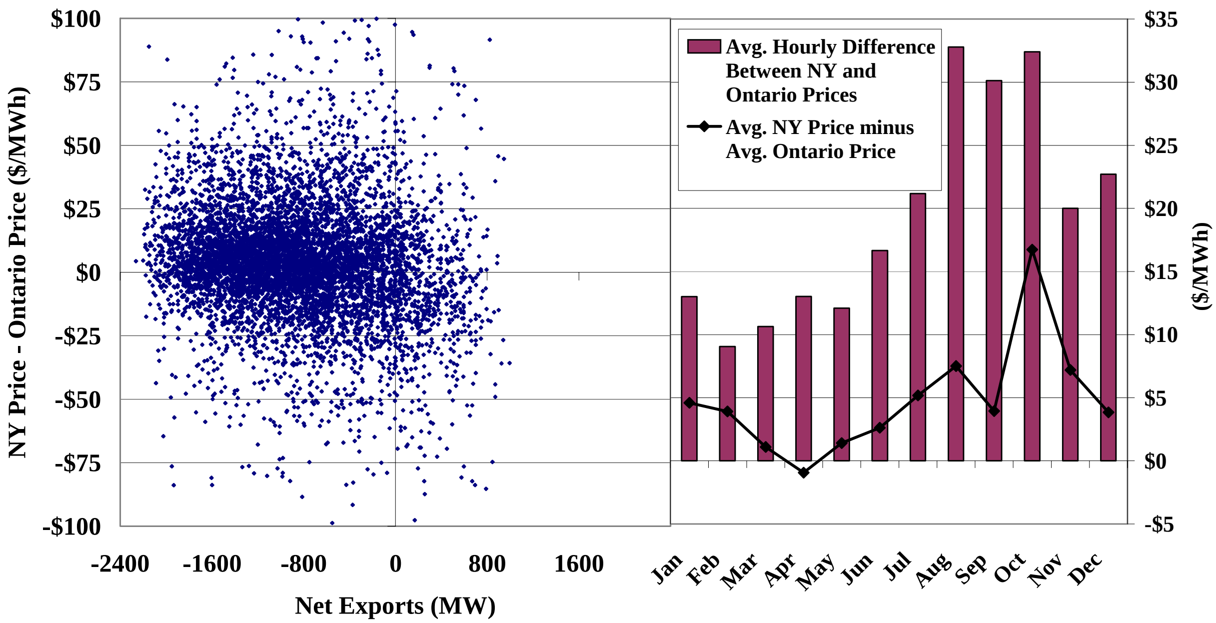
On the right side of these figures, the monthly average price differences between New York and the adjacent market are shown.

If transactions were scheduled efficiently between regions, it is expected that the points in each of the charts would be relatively closely clustered around the horizontal line at $0 – indicating little or no price difference between New York and the adjacent region in the absence of a physical transmission constraint (quantities of imports or exports can vary widely, but without transmission constraints power flows should continue in one direction or another until prices differences were arbitraged away). Moreover, one would not expect net exports to occur when the New York price substantially exceeds the price in the neighboring region. Likewise, one would not expect net imports to occur when the New York price is substantially less than the price in a neighboring region.

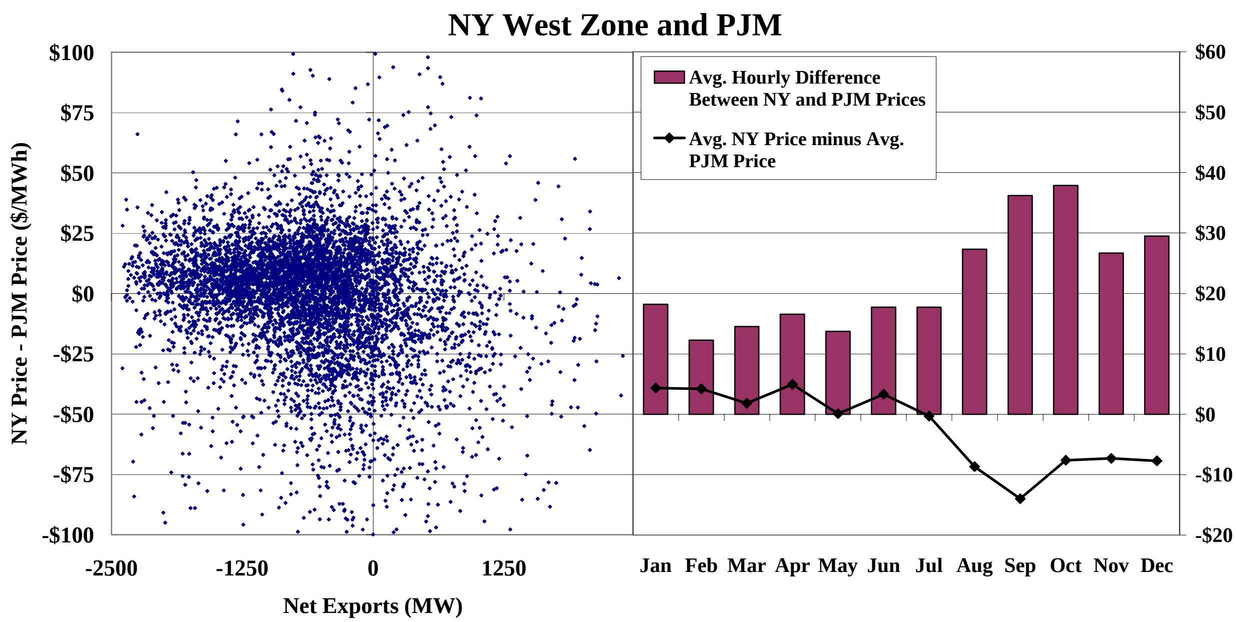
Figure 60: Real Time Prices and Interface Schedules
Eastern NY and New England



West NY and Ontario*



* Price difference measured in US dollars



These figures show the real-time markets are still not being arbitraged efficiently by participants. The dispersion in prices during unconstrained hours is shown to be considerable. In a significant number of hours for each interface, power is scheduled from the high-priced market to the lower-priced market. These results are similar to results presented in prior years.

Several factors prevent real-time prices from being fully arbitraged between New York and adjacent regions. First, market participants do not operate with perfect foresight of future market conditions at the time that transaction bids must be submitted. Without explicit coordination between the markets by the ISOs, complete arbitrage will not be possible. Second, differences in scheduling procedures and timing in the markets serve as barriers to full arbitrage. Third, there are transaction costs associated with scheduling imports and exports that diminish the returns from arbitrage. Participants would not be expected to schedule additional power between regions unless they expect a price difference greater than these costs. Last, risks associated with curtailment and congestion will reduce participants' incentives to engage in external transactions at small price differences.

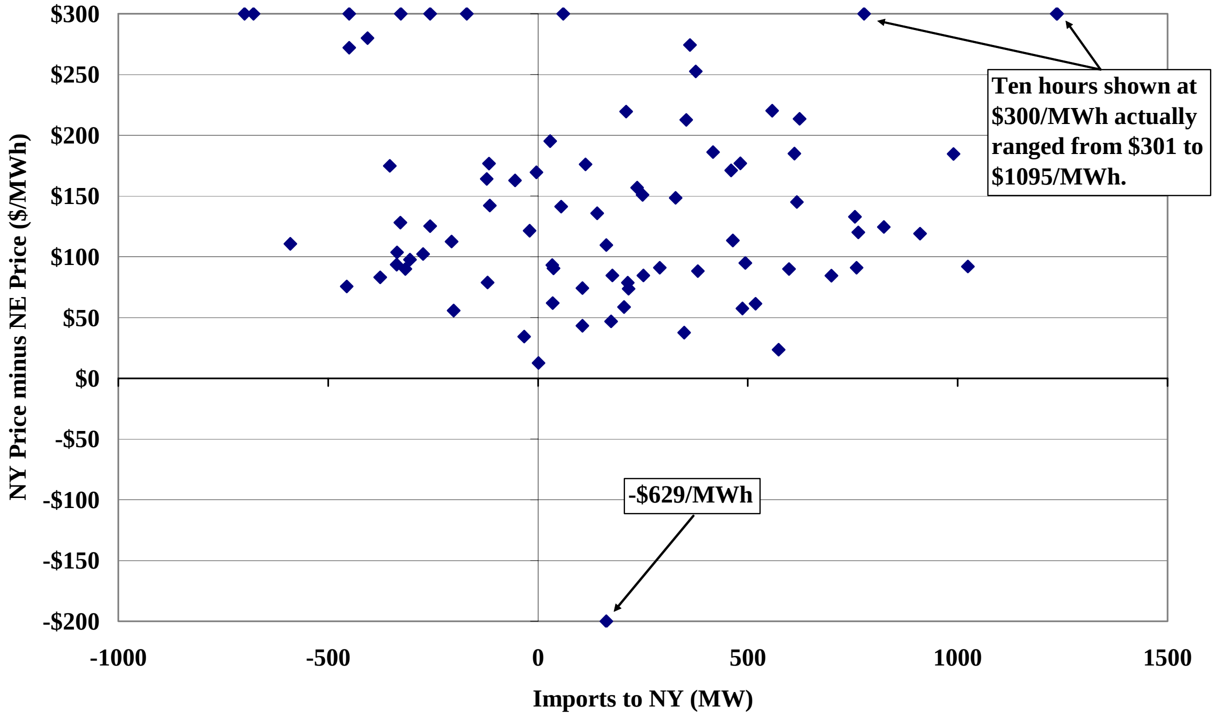
Uncertainty, imperfect information, and required offer lead times limit the ability of participants to capitalize on real-time arbitrage opportunities. These results reinforce the importance of intrahour transaction scheduling provisions that are being developed to improve real-time interchange between New York and New England. This will be particularly important when the capacity surpluses in the Northeast are eliminated – when optimizing the flow between areas will have larger economic and reliability consequences.

Fees assessed to transactions between control areas tend to inhibit convergence. In 2005, export fees between New York and New England were eliminated, facilitating arbitrage of the adjacent markets. However, several other charges are still imposed on transactions between New York and New England. Exports from New York and New England scheduled after the day-ahead market continue to be allocated charges for certain ISO/RTO operating costs. Prior to the fall of 2005, the ISO-NE allocated these charges based on the largest transaction scheduled during a month. This method resulted in very large charges (on a per MWh basis) for some market participants. In the fall of 2005, the ISO-NE addressed this problem by allowing market participants to choose an alternative method which allocates on a per MWh basis. Transactions from New York to New England scheduled after the day-ahead market continue to be allocated uplift for certain types of supplemental commitment by both ISOs. However, neither ISO assesses these charges to transactions that flow from New England to New York.

Figure 60 shows that although the difference in average prices for New York, New England, and PJM is relatively low, the dispersion of prices in the absence of congestion is substantial. The figure shows that the typical price difference between New York and these adjacent markets ranged from $10 per MWh to $38 per MWh on a monthly average basis. This indicates that significant seams issues remain that continue to prevent efficient interchange between the market areas.

During peak demand conditions, it is especially important to efficiently schedule flows between control areas. The following chart, Figure 61, examines the difference between New York and New England real-time border prices in unconstrained hours when the Capital Zone price exceeded $200/MWh.

Figure 61: Interchange and Price Differences Between New York and New England Unconstrained Price Spike Hours* 2005



* Includes hours when the real-time Capital Zone price exceeded $200 per MWh

Figure 61 indicates that price convergence between adjacent markets has been especially poor during peak demand conditions. In 18 of 82 hours, the New York price was higher than the New England price by $200 per MWh or more. Likewise, the New York price was higher than the New England price by $100 per MWh or more in 50 of 82 hours. In 29 of the hours shown, power was flowing out of New York even though the New York price was higher. Frequently during peak demand conditions, a small amount of additional imports can substantially reduce the magnitude of a price spike. This underscores the potential benefits of ITS (Intra-hour Transaction Scheduling) especially during peak demand periods.

B. Inter-regional Dispatch Coordination

The results shown in the previous section indicate that price convergence between New York and adjacent markets continue to be relatively poor. This implies that more efficient scheduling of flows between markets would result in substantial production cost savings and benefits to consumers. In this section, we estimate the potential benefits of perfect coordination between the

New York and New England markets where we have been recommending that the ISO's coordinate the physical interchange.

In each hour of 2004 and 2005 when transmission constraints were not binding between New York and New England, we performed a simulation of the re-dispatch that would likely have occurred if real-time flows were scheduled efficiently between markets. This simulation used actual real-time generator offers and took into account marginal losses, congestion patterns, the flow limits between markets, and operating reserves constraints in Eastern New York. We assumed that lower average real-time prices would benefit consumers by leading to lower average prices in the day-ahead and other forward markets for purchases by LSEs. We also assumed that reduced congestion would lead to correspondingly reduced payments to transmission owners from the TCC auctions, leading to higher Transmission Service Charges for consumers. The results of these simulations are summarized in the following figure.

Figure 62: Estimated Benefits from Optimizing Flows Between Control Areas
New England Interface
2005

	2004	2005
Production Cost Savings (in millions of $)	**$13**	**$21**
Consumer Savings (in millions of $):		
New York Consumers	**$35**	**$124**
New England Consumers	**$13**	**$31**

This analysis indicates that optimizing the interchange between markets would bring very significant benefits. Particularly in 2005, there were many occasions when a price spike occurred in Eastern New York while conditions were moderate in New England and the interface was not fully utilized. During many of these hours, a small change in flows across the interface would have greatly reduced or eliminated the price spike. Based on our analysis, more than 50 percent of the consumer benefit shown in the figure above would have come from mitigating price spike events. The fact that the results indicate that New York consumers would have benefited more than New England consumers is consistent with the larger number of price spike events New York.

C. Conclusions and Recommendations

Over the past several years, modeling improvements and rule changes have led to improved convergence between control areas during non-transmission-constrained hours. While the external transaction scheduling process is functioning properly, significant price differences remain between markets in hours when no congestion is present. The economic consequences of these issues were minimized in 2003 and 2004, because demand was relatively mild and there were no instances of shortage. However, the economic effects of the seams issues were larger in 2005 when the market experienced a substantial number of shortages, some of which could have been avoided if the external interfaces were fully utilized.

These results reinforce the importance of addressing remaining seams issues. We continue to encourage New York and New England to develop and implement new scheduling procedures, such as Intra-hour Transaction Scheduling ("ITS"). Intra-hour transaction scheduling is a process that would allow the physical interchange to be adjusted within an hour when prices diverge at the interface between the two markets. These adjustments would ensure that the interchange levels are efficient, eliminating the price distortions and other inefficiencies caused by poor market arbitrage. This will lead to less volatility and more predictability in the New York to New England prices. Likewise, we recommend that the NYISO work with PJM to eliminate export fees and improve scheduling procedures.

VIII. DEMAND RESPONSE PROGRAMS

The New York ISO has some of the most effective demand response programs in the country. There are currently three demand response programs in New York State:

- Day-Ahead Demand Response Program (DADRP) – This program schedules physical demand reductions for the following day, allowing resources with curtailable load to offer into the day-ahead market like any supply resource. If the offer clears in the day-ahead market, the resource must curtail its load in accordance with the accepted offers and is paid day-ahead clearing price for each MW of curtailed load.

- Special Case Resources (SCR) – These are loads that must curtail within two hours. They are called when operators forecast a reserve deficiency. These resources may sell capacity in the capacity market corresponding to their commitment to curtail load.

- Emergency Demand Response Program (EDRP) – The emergency demand response program pays loads that curtail on two hours notice the higher of $500/MWh or the real-time clearing price. SCRs receive this payment as well.

The EDRP and SCR programs have been effective in achieving actual load reductions during peak conditions. The total registered quantity of more than 2000 MW is larger than most comparable programs in other ISOs.

The success of these programs is largely due to incentives provided by the programs. EDRP participants are paid the higher of $500 per MWh or the LBMP for voluntary load reductions (i.e., they have no obligation to respond), which is the only source of revenue for the EDRP resources. SCR resources can sell their curtailable load in the capacity market in exchange for an obligation to respond when called. SCR participants are paid the higher of a strike price that they bid (limited to be less than $500 per MWh) and the LBMP.[20]

This payment structure satisfies two critical objectives. First, it results in payments to participants that are close to or exceed $500 per MWh, which allow them to be paid an amount that covers their marginal value of consumption during peak periods. Hence, it would provide an adequate incentive for loads to respond, even though most are served under regulated or

[20] The NYISO will provide a 24-hour notice if it anticipates a need to make curtailments under the SCR program to meet reserve requirements. These curtailments may or may not ultimately be called. However, there is a two-hour notice given when the NYISO determines that the load should be curtailed. EDRP also provides the NYISO with resources to meet potential reserve shortfalls. These curtailable load resources are given two-hours notice prior to being asked to curtail.

otherwise fixed rates that cause them not to incur the wholesale price of electricity.[21] Second, during times when EDRP and SCR are the marginal sources of supply in the market that allow the system to satisfy its reserve requirements, the LBMP typically will be set at $500/MWh. This price is in a range that is consistent with the marginal value of reserves to the system. Hence, these payments and the associated pricing provisions contribute to efficient pricing during shortage (or near-shortage) conditions.

The EDRP and the SCR programs can contribute substantial demand-side resources to the market. Special Case Resources are qualified to sell into the capacity market, and by adding to the total supply, help reduce capacity prices. In the summer of 2005, the quantity of SCR/ICAP subscribers that sold capacity were 279 MW in NYC, 111 MW in Long Island, and 667 MW in upstate New York. These demand response resources have had a substantial effect on prices in the capacity market. They also help increase the competitiveness of the capacity market in New York City where ownership of generation is relatively concentrated.

In 2005, there was one afternoon, July 27th, when the NYISO called upon SCR and EDRP resources in five zones to curtail. The NYISO estimates these resources provided 819 MWh of response during four hours when the average real-time price exceeded $600 per MWh. Nearly 90 percent of the response came from resources located in New York City and Long Island. The day-ahead demand response program has provided considerably less valuable demand reduction than the EDRP and SCR programs. Although approximately 2000 MWh of day-ahead demand response was scheduled in 2005, this was dispersed across many hours with moderate prices.

21 While the average regulated rate paid by load is much lower than $500/MWh, the value of power at peak times is typically much higher than the average. Therefore, in the absence of the NYISO's payments for EDRP and SCR load reductions, load that is interrupted would save only the regulated rate. This rate does not reflect the marginal system cost of serving the load as embodied in the wholesale LBMPs.