

Summer 2006 Review of the New York Electricity Markets

David B. Patton, Ph.D. Potomac Economics

Independent Market Advisor

November 29, 2006





- The NYISO energy and ancillary services markets performed competitively during the summer of 2006.
 - There was little evidence of significant economic or physical withholding.
- Energy and ancillary services prices declined in most areas due to:
 - Lower fuel prices than the previous summer, including an 18 percent drop in natural gas;
 - ✓ 1 GW of newly installed combined cycle capacity in New York City, which reduced congestion, particularly into New York City; and
 - \checkmark Lower demand in most hours relative to the previous summer.



- Convergence between day-ahead and real-time energy prices was much better in 2006.
 - In 2005, a small number of real-time peak pricing events, not anticipated by the day-ahead market, were primarily responsible for the lack of convergence.
 - ✓ In 2006, convergence improved most likely because the additional experience of market participants under the SMD 2.0 markets has enabled them to more efficiently hedge real-time market prices.
- Convergence between day-ahead and real-time prices for reserves and regulation improved, although significant issues remained in 2006.
 - Eastern 10-minute reserves were significantly under-valued in the dayahead market.
 - ✓ State-wide 10-minute spinning reserves were over-valued in the dayahead market.



- The changes made in the SMD dispatch software improved the efficiency of energy and ancillary services pricing.
 - During the summer of 2005, real-time energy and reserves prices did not always reflect that the system was under shortage conditions – 43% of shortage intervals did not exhibit shortage pricing..
- Prior to the summer of 2006, a software change was made that better enables the real-time market model to set efficient prices.
 - ✓ In May 2006, the NYISO changed the pricing pass of RTD to ensure that gas turbine are consistent with their physical capabilities as reflected in the physical pass of RTD.
 - ✓ By improving the consistency of the two dispatch passes in RTD, only 15 percent of the intervals that exhibited physical shortages in 2006 did not result in shortage pricing.
 - Many of these shortages were slight and we recommend an additional change to make them less frequent.
 - ✓ In 97 percent of intervals that exhibited shortage pricing, the system was in a physical shortage.



- Non-local reliability uplift declined substantially due to less frequent commitment of uneconomic gas turbines by RTS.
 - ✓ The introduction of real-time line modeling for load pockets in NYC has enabled the real-time commitment software to anticipate better when gas turbines will be economic to relieve congestion.
- Balancing congestion shortfalls, which are uplifted to all customers in New York, declined as a result of:
 - ✓ Lower fuel prices, which reduces the costs of congestion;
 - ✓ Less frequent congestion within New York City; and
 - ✓ The introduction of real-time line modeling for load pockets in NYC.
- Local reliability commitments remained constant, but the resulting uplift increased because several generators raised their offers.
 - ✓ Local reliability uplift arises primarily from committing resources out-ofmerit for local issues, and thus, is allocated to the local TO.

Areas of Potential Improvement and Recommendations

- Additional improvements to made the pricing and physical dispatch passes of RTD are possible that would improve the efficiency of New York's real-time prices (particularly in shortages) and reduce uplift.
 - ✓ We recommend the NYISO re-calibrate the dispatch levels in the pricing pass for units that are not responding to dispatch signals.
 - RTC and RTD sometimes produce inconsistent results, which may result in inefficient commitment and scheduling. To improve this consistency, we recommend the NYISO evaluate:
 - ✓ Whether there is an alternative to RTC using the maximum of three fiveminute load forecasts.
 - ✓ Whether the assumptions in RTC and RTD regarding external transaction ramp can be made consistent to eliminate differences at the top of the hour.
 - ✓ Whether predictable adjustments to the RTD load forecast, which are made to minimize regulation deployment, can be reflected more quickly in the RTC load forecast.

Areas of Potential Improvement and Recommendations

- Convergence between day-ahead and real-time ancillary services prices remains poor.
 - The NYISO should consider introducing virtual trading of ancillary services.
 - ✓ This change would promote convergence of ancillary service prices and reduce physical suppliers' incentive to raise their offers.
 - ✓ However, it would need to be carefully studied to ensure it will not have unintended consequences on the day-ahead commitment.
 - Transmission constraint shadow prices can reach extremely high levels for brief periods when redispatch options are unavailable or relatively ineffective.
 - ✓ Transmission demand curves could be used to prevent costly redispatch in situations where there is little or no reliability benefit.
 - ✓ Therefore, we recommend that the NYISO continue to evaluate the reliability impacts of implementing transmission demand curves.

Areas of Potential Improvement and Recommendations

- Supplemental commitments through the local reliability pass of SCUC and the SRE process are often required to meet local requirements in New York City
- These commitments are not optimized with other commitments in the Day-Ahead market.
- We continue to recommend that:
 - ✓ The NYISO improve the modeling of local reliability rules and NOx constraints to include them in the initial SCUC commitment.
 - ► This change requires that the NYISO first work to revise the costallocation methodology for uplift associated with the local reliability requirements.



Market Prices and Outcomes





Day-Ahead Energy Prices and Natural Gas Prices

- The first figure presents average day-ahead energy prices by month during the summers of 2005 and 2006 in four regions of New York state.
- Electricity prices continue to vary substantially between regions due to the effects of transmission losses and congestion.
 - \checkmark These differences generally increased slightly from 2005 to 2006.
 - ✓ Long Island prices rose most significantly relative to other areas. The price difference between eastern up-state NY and Long Island rose from \$26/MWh in 2005 to \$38/MWh in 2006.
- Reductions in natural gas prices contributed to lower day-ahead electricity prices.
 - ✓ Natural gas prices decreased 18 percent in 2006
 - Day-ahead electricity prices in up-state areas and New York City decreased 13 to 18 percent.
 - ✓ New capacity added in NYC contributed to lower congestion into NYC.

Day-Ahead Electricity and Natural Gas Prices June to August, 2005 & 2006





Energy Prices

- The next two figures show how prices have changed in the last three years on an hourly basis.
- The first figure shows real-time price duration curves during the summers of 2004, 2005, and 2006.
 - These curves show the number of hours when the load-weighted price for New York State is greater than the level shown on the vertical axis.
- In 2006, prices were lower than in the previous year due to lower fuel prices and milder weather:
 - ✓ In 2006, there were 275 hours with prices above \$100, compared to 597 such hours in 2005.
 - ✓ In 2006, there were 61 hours with prices above \$200, compared to 80 such hours in 2005.
- The widespread nature of the price changes over the past three summers are primarily attributable to natural gas and oil price changes, which affect electricity prices in both high and low load conditions.



Price Duration Curves New York State Average Real-Time Price June to August, 2004 to 2006





Energy Prices

- To identify changes in electricity prices that are not driven by changes in natural gas prices, the second set of duration curves show the marginal heat rate that would be implied if natural gas were always on the margin.
 - ✓ Implied Heat Rate = (Real-Time Electricity Price) ÷ (Natural Gas Price)
- From 2004 to 2005, the dramatic increase in implied heat rates was due primarily to hotter weather that resulted in more frequent shortages:
 - ✓ Under SMD 2.0, the shortage pricing provisions led to approximately 20 hours of shortage prices in 2005 corresponding to reserve shortages.
 - ✓ Shortage pricing did not occur in 2004.
- Implied heat rates were comparable between 2005 and 2006.
 - ✓ Milder weather and one gigawatt of new capacity in New York City contributed to a modest reduction (10% decrease) in the number of hours when the implied heat rate exceeded 10 MMbtu per MWh.
 - ✓ These factors were partly offset by better recognition of reserves shortages by the real-time software, which led to a modest rise (8% increase) in hours when the implied heat rate exceeded 20 MMbtu per MWh.

Implied Heat Rate Duration Curves Based on New York State Average Real-Time Price June to August, 2004 to 2006



Load Profile

- The following figure shows how demand has changed across all hours during the past three summers.
 - ✓ The load durations 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.
- Between 2005 and 2006, load was most comparable under peak conditions (i.e. when load exceeded 30 GW).
 - ✓ In 2006, there were 60 hours when actual loads exceeded 30 GW, and 68 such hours in 2005.
- In the majority of the summer hours, load was higher in 2005 than in 2006.
 - ✓ In 2006, there were 136 hours when actual loads exceeded 28 GW, and 226 such hours in 2005.
 - ✓ In 2006, there were 293 hours when actual loads exceeded 26 GW, and 438 such hours in 2005.





Uplift Expenses from BPCG Payments

- The following figure summarizes uplift expenses from Bid Production Cost Guarantee ("BPCG") Payments during the past three summers.
 - ✓ These payments are made when a supplier does not receive sufficient revenue from energy and ancillary services to cover their as-bid costs.
- BPCG payments are categorized according to the following criteria:
 - ✓ Local Reliability BPCG payments are classified as local reliability when they result from out-of-merit commitment and dispatch by or on behalf of the local TO in order to manage a constraint not modeled by NYISO. This cost is allocated to the local TO.
 - ✓ Other These refer to all other BPCG payments that result when a generator is committed and dispatched in merit order but does not cover its commitment costs. This is allocated across all of New York state.
 - Day-ahead vs. Real-time This is based on whether the resource was scheduled in the day-ahead or real-time market.



Uplift Expenses from BPCG Payments

- Overall expenses for BPCG payments were similar in 2005 and 2006.
 - ✓ Sizable reductions in real-time non-local reliability uplift were offset by increased costs for local reliability uplift in the day-ahead and real-time.
 - Real-time non-local reliability uplift decreased 62 percent from 2004 to 2006 due to more efficient commitment and dispatch of gas turbines under SMD 2.0.
 - Real-time local reliability uplift arises primarily from commitments in the Supplemental Resource Evaluation ("SRE") after the day-ahead market.
 - ✓ The amount of SRE capacity decreased 28 percent from 2005 to 2006.
 - ✓ However, the uplift expenses rose because several generators in up-state New York raised their offer prices substantially above marginal costs, while staying under the applicable conduct thresholds for mitigation.
- Day-ahead local reliability uplift rose in 2006 as a result of more frequent commitments by the local reliability pass of the day-ahead model.



Uplift Expenses from BPCG Payments June to August, 2004 to 2006



Note: Real-time BPCG payment uplift does not reflect the application of RT BPCG mitigation described in the NYISO's filings in docket ER06-185 -20-



Balancing Congestion Shortfall

- The following figure shows the congestion revenue shortfall incurred in the balancing market on a monthly basis during 2005 and 2006.
 - Balancing congestion fell from \$146 million in the summer of 2005 to \$102 million in the summer of 2006.
- The primary cause of the balancing congestion shortfall is changes between the day-ahead and real-time markets in the amount of transfer capability associated with the transmission system.
 - ✓ When day-ahead schedules exceed real-time transmission capability, the NYISO must buy back the excess in real-time.
 - Reductions in real-time transmission capability from TSA operation contribute significantly to balancing congestion shortfalls.
- Several factors contributed to the decline in balancing congestion shortfall:
 - ✓ The introduction of new capacity in the New York City has reduced congestion, particularly within New York City.
 - ✓ Lower fuel costs have contributed to lower balancing congestion costs.
 - ✓ The introduction of line modeling to RTS in May 2005 has improved consistency between day-ahead and real-time transmission modeling.

Monthly Balancing Congestion Uplift Expenses January 2005 to August 2006



-22-



Price Convergence



Day-Ahead and Real-Time Energy Prices

- Day-ahead to real-time price convergence is important because:
 - ✓ It helps ensure that the commitment of resources through the day-ahead market is efficient, and
 - Because most transactions settled by NYISO are through the day-ahead market.
- The following figure shows monthly average day-ahead and real-time energy prices during the summer of 2006.
- Overall price convergence was far better in 2006 than in 2005.
 - ✓ In 2005, the real-time price premium averaged 18 percent in New York City and 25 percent in Long Island.
 - ✓ In 2006, the real-time price premium averaged 1 percent in New York City and 8 percent in Long Island.
 - ✓ The Hudson Valley also exhibited relatively good price convergence.
- All four regions exhibited a real-time price premium during August, while day-ahead prices were higher during June and July.



Average Day-Ahead and Real-Time Energy Prices West Zone, Hudson Valley, New York City, & Long Island June to August, 2006



Day-Ahead to Real-Time Price Convergence

- The following three figures show the average real-time price premium on a daily basis during afternoon hours from June to August 2006.
- Day-ahead prices are higher than real-time prices on the majority of afternoons. The following figures show day-ahead prices were higher:
 - ✓ On 72 percent of afternoons in Hudson Valley,
 - ✓ On 69 percent of afternoons in New York City, and
 - ✓ On 51 percent of afternoons in Long Island.
- When the price difference is large, real-time prices generally exceed dayahead prices. The real-time price premium exceeded \$100/MWh:
 - ✓ On 8 afternoons in the Hudson Valley;
 - ✓ On 10 afternoons in New York City; and
 - ✓ On 12 afternoons in Long Island.
- Modest day-ahead premiums most days and large real-time premiums during shortages is consistent with efficient hedging of real-time price volatility.



Day-Ahead to Real-Time Price Convergence

- Large real-time price premiums occur when market participants in the dayahead market do not anticipate tight real-time operating conditions.
 - ✓ Tight conditions can be the result of high load, generation outages, or the result of the Thunderstorm Alerts ("TSAs").
 - ✓ TSAs can be difficult to predict and significantly reduce the transfer capability from the Capital region to down-state areas.
 - ✓ TSAs require double contingency operation of the ConEd overhead transmission system in real-time but not in the day-ahead market.
- TSAs were a factor in many of the real-time price spikes unforeseen by the day-ahead market.
 - ✓ TSA operation caused the Leeds-Pleasant Valley constraint shadow prices to be above \$1000/MWh in 182 intervals.
 - ✓ The most TSA-related transmission congestion was experienced on June 1st and 2nd and July 3rd, 11th, and 28th. Real-time price premiums exceeded \$100/MWh on all five afternoons.

Average Daily Real-Time Price Premium Hudson Valley – 1 p.m. to 7 p.m. Weekdays June to August, 2006



-28-

Average Daily Real-Time Price Premium New York City – 1 p.m. to 7 p.m. Weekdays June to August, 2006





Average Daily Real-Time Price Premium Long Island – 1 p.m. to 7 p.m. Weekdays June to August, 2006



-30-



Real-Time Transmission Price Spikes

- Real-time transmission price spikes occur when the re-dispatch costs necessary to resolve a transmission constraint reach extremely high levels, which contributed significantly to the severity of real-time energy price spikes.
 - ✓ During the summer of 2006, there were 814 intervals when shadow prices exceeded \$1,000/MWh on one or more constraints, and
 - ✓ 489 intervals when they exceeded \$2,000/MWh.
- These spikes typically occur for brief periods when there is not sufficient ramp capability within a constrained area.
 - ✓ This may result in large amounts of re-dispatch that provide little reliability benefit.
 - ✓ In some of these intervals, the real-time model cannot solve because of insufficient resources.
- Like ancillary services demand curves, transmission demand curves could be used to prevent costly re-dispatch when there is little reliability benefit. The NYISO has been evaluating the impact on reliability of using transmission demand curves.



Ancillary Services Price Convergence

- The following chart shows day-ahead and real-time eastern 10-minute reserves prices by hour of the day during the summers of 2005 and 2006.
- The NYISO requires 1,000 MW of 10-minute reserves east of the Central-East Interface. The market models include an economic demand curve value of \$500/MWh on meeting this requirement.
- Convergence between day-ahead and real-time prices has been poor since the creation of real-time ancillary services markets in 2005.
 - During afternoon hours, average day-ahead prices are a small fraction of average real-time prices.
 - During the morning and evening ramping hours, average day-ahead prices are substantially greater than average real-time prices.
- Convergence improved slightly in 2006 due to higher average day-ahead prices during the afternoon.
 - This is partly the result of rising day-ahead offer prices by 10-minute reserves providers in Eastern New York.

10-Minute Total Reserve Prices in East NY by Hour of Day June to August, 2005 & 2006





Ancillary Services Price Convergence

- The following figure shows day-ahead and real-time western 10-minute spin prices.
 - \checkmark Currently, the economic value of this requirement is set at \$500/MWh.
- Convergence between day-ahead and real-time spinning reserves prices has been poor since the creation of real-time ancillary services markets.
 - ✓ In 2005, average day-ahead prices substantially exceeded average real-time prices during morning and evening hours, while average real-time prices were considerably higher during the afternoon peak hours.
 - ✓ In 2006, convergence continued to be poor during the morning and evening hours, but exhibited improvement during the afternoon peak.
- Day-ahead spinning reserves prices depend on individual generator offers and the opportunity costs of providing reserves rather than energy.
 - ✓ Day-ahead offer prices for spinning reserves generally decreased in 2006 relative to the previous summer.
 - Thus, higher day-ahead prices are likely driven by higher opportunity costs from increased day-ahead load scheduling.







Ancillary Services Price Convergence

- The following figure summarizes convergence between day-ahead and real-time prices for regulation.
- State-wide regulation requirements are generally of 275 MW during ramping hours and as low as 150 MW during other hours.
 - ✓ Currently, the economic value of this requirement is set at \$250/MW for the last 25 MW procured and \$300/MW for the balance of the requirement.
 - Day-ahead and real-time regulation prices are highly correlated across the day with real-time prices generally being higher.
- Price convergence improved in 2006 relative to the previous summer.
 - ✓ In 2005, real-time prices were approximately \$10/MWh higher on average than day-ahead prices.
 - ✓ In 2006, the average real-time price premium was reduced to \$4/MWh.
- The marked price increase from 2005-2006 is partly due to a rise in regulation offer prices. This was discussed in the 2005 State of the Market Report.


Ancillary Services Price Convergence Conclusions

- Convergence between day-ahead and real-time reserves prices improved modestly in 2006 relative the previous summer, but it is still poor in comparison to energy price convergence.
- Pervasive differences between day-ahead and real-time reserves prices give generators an incentive to adjust their day-ahead offer prices toward expected real-time reserves prices.
 - ✓ This reduces the efficiency of the day-ahead commitment to the extent that generators are not committed due to errors in projecting real-time prices.
- The NYISO should consider the feasibility and potential benefit of introducing virtual trading of ancillary services in the day-ahead market.
 - ✓ This would promote convergence of ancillary service prices and reduce physical suppliers' incentive to raise their offers above marginal cost.
 - ✓ However, it would need to be carefully studied to ensure it will not have unintended consequences on the day-ahead commitment.



Analysis of Bids and Offers





Analysis of Offer Patterns

- This section of the report analyzes the patterns of conduct that could indicate physical or economic withholding.
- This analysis evaluates the correlation of quantities of potential withholding to load levels.
 - ✓ Suppliers in a competitive market should increase offer quantities during higher load periods to sell more power at the higher peak prices;
 - Suppliers in markets that are not workably competitive will have the greatest incentive to withhold at peak load levels when the market impact is the largest.
 - ✓ Hence, this analysis allows one to discern quantities that may reflect attempts to withhold resources to raise prices.
- The first analysis is of potential physical withholding, analyzing total generation deratings (including planned forced outages, and partial deratings).



Analysis of Offer Patterns – Deratings

- The following two figures plot the total deratings and short-term deratings versus actual load in eastern NY during peak hours in the summer.
 - ✓ The figures focus on eastern NY because this area, which includes twothirds of the State's load, has limited import capability and is more vulnerable to the exercise of market power.
 - ✓ We focus this analysis on the summer to exclude the effects of planned outages that typically occur during off-peak seasons, and because market power is most likely during the higher load conditions in the summer.
 - The short-term deratings shown in the second figure include ones that last for fewer than 30 days. These are more likely to reflect attempts to physically withhold since it is more costly to withhold via long-term deratings or outages.
- These figures show that deratings did not substantially increase when load reached the highest levels on August 1st, 2nd, and 3rd, which is consistent with workable competition.

Relationship of Deratings to Actual Load Day-Ahead Market – East New York Peak Hours*, June to August, 2006



* Peak hours are defined as weekdays from 12 PM to 6 PM for purposes of this analysis.

Relationship of Short-Term Deratings to Actual Load Day-Ahead Market – East New York Peak Hours*, June to August, 2006



* Peak hours are defined as weekdays from 12 PM to 6 PM for purposes of this analysis.



Analysis of Offer Patterns – Output Gap

- The second analysis is intended to assess potential economic withholding, employing a measure called an "output gap".
- The output gap is the quantity of economic capacity that does not produce energy or ancillary services because a supplier submits an offer price well above a unit's reference level.
- The output gap:
 - ✓ Addresses all components of a supplier's offer, including start-up, minimum generation, and incremental energy offers.
 - ✓ Includes units that "set the price".
 - ✓ Excludes capacity scheduled to provide ancillary services.
- It is particularly notable that the output gap measured at the lower threshold declines and is very low during high load periods, because this conduct would not be subject to mitigation.



Analysis of Offer Patterns – Output Gap

- The following figures shows the real-time output gap in eastern New York during peak hours using:
 - \checkmark Standard conduct thresholds of \$100/MWh or 300% (whichever is lower).
 - ✓ Low thresholds, \$50/MWh or 100% (whichever is lower), and
- These figures both show that output gap generally decreases under the highest load conditions.
 - ✓ This is an important result because prices are most vulnerable to market power under peak load conditions.
 - ✓ These results indicate that economic withholding was not a significant concern during the summer of 2006.

Output Gap at Mitigation Threshold vs. Actual Load Real-Time Market – East New York Peak Hours*, June to August, 2006



* Peak hours are defined as weekdays from 12 PM to 6 PM for purposes of this analysis.

Output Gap at Low Threshold vs. Actual Load Real-Time Market – East New York Peak Hours*, June to August, 2006



* Peak hours are defined as weekdays from 12 PM to 6 PM for purposes of this analysis.



Summary of Day-Ahead Mitigation

- Local market power mitigation measures are triggered when constraints are binding into a load pocket to address market power in the NYC load pockets.
- The conduct and impact framework focus more effectively mitigates market power in the NYC load pockets and it also allows high prices to occur during legitimate periods of shortage.
- The following figure summarizes the frequency of mitigation in NYC during the summers of 2005 and 2006.
 - ✓ The line shows 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).
- Mitigation was most commonly associated with the constraints into New York City (i.e. Dunwoodie-South) and into the 138 kV system.
- Mitigation decreased substantially from 2005-2006, due in part to the new installation of one gigawatt of new combined cycle capacity in Astoria.





Summary of Real-Time Mitigation

- While the previous figure summarizes mitigation in the day-ahead market in New York City, the following figure summarizes real-time mitigation.
 - ✓ GTs mitigated at start-up are shown in the "MinGen/StartUp" category.
 - ✓ GTs mitigated after being on-line are shown in the "Energy" category.
- In 2005, real-time mitigation was more commonly associated with the subload pockets inside the 138 kV system than day-ahead mitigation, which was generally done for the larger load pockets.
 - This was because the real-time market had significantly more congestion between areas inside New York City than the day-ahead market.
- In 2006, real-time mitigation was much less frequent than in 2005.
 - ✓ The installation of new capacity that has significantly reduced congestion into the sub-load pocket areas inside the 138 kV system, thereby reducing the need for real-time mitigation.
 - ✓ The introduction of detailed line modeling has allowed greater utilization of transmission into the load pockets, which has reduced the effect of generators' offer prices on LBMPs in the load pockets.





Analysis of Load Bidding Patterns

- The following figures show day-ahead load schedules and offers as a fraction of real-time load during the summers of 2005 and 2006 at various locations in New York.
 - Virtual supply effectively nets out an equivalent amount of scheduled load, thus it is shown as a negative quantity.
 - Net scheduled load = Physical Bilaterals + Fixed Load + Price-Capped Load + Virtual Load - Virtual Supply
- Load is generally over-scheduled in New York City and Long Island and under-scheduled in up-state New York.
 - This implies a higher level of imports to constrained areas in the day-ahead market than in real time.
 - This pattern of scheduling is consistent with prior years and generally contributes to better price convergence.
 - \checkmark This pattern was more pronounced in the summer of 2006.

Composition of Day Ahead Load Schedules as a Proportion of Actual Load in New York City June to August, 2005 & 2006



Composition of Day Ahead Load Schedules as a Proportion of Actual Load in East Up-State New York June to August, 2005 & 2006



-54-

Composition of Day Ahead Load Schedules as a Proportion of Actual Load in West Up-State New York June to August, 2005 & 2006



-55-



Market Operations – Real Time Scheduling and Shortage Pricing





- RTD co-optimizes procurement of energy and ancillary services. This has several advantages:
 - The software efficiently allocates resources to provide energy and ancillary services every five minutes.
 - ✓ This incorporates the costs of maintaining reserves into the price of energy, whereas these costs were not considered prior to SMD 2.0.
 - Demand curves rationalize the pricing of energy and reserves during shortage periods by setting limits on the costs that can be incurred to maintain reserves.
- This section evaluates the consistency between Eastern 10-minute reserves pricing done by the new software and the actual physical scarcity of Eastern 10-minute reserves.
 - ✓ The real-time software maintains 1000 MW of 10-minute reserves inside Eastern New York up to a cost of \$500/MWh.
 - ✓ The Eastern 10-minute reserves requirement has been the most costly to maintain since the introduction of real-time ancillary services markets.



- Under SMD2, co-optimization of energy and reserves is integrated with the Hybrid Pricing approach. Hybrid Pricing of gas turbines has been a key element of the real-time market software since 2002.
 - The inflexibility of gas turbines creates challenges for pricing energy efficiently when the gas turbines are the marginal source of supply.
 - ✓ 28 percent of dispatchable capacity in New York City and 42 percent of the dispatchable capacity in the 138kV load pocket are gas turbines.
 - Thus, Hybrid-Pricing is particularly important to setting efficient price signals in NYC.
- Hybrid Pricing treats gas turbines as flexible resources for pricing purposes, which results on certain inconsistencies between the pricing dispatch and the physical dispatch of the system. However, these inconsistencies should be limited such that:
 - ✓ Under physical shortage conditions, prices should reflect scarcity; and
 - \checkmark High prices are only set when the system is physically in shortage.



- The following chart shows the amount of Eastern 10-minute reserves that were physically scheduled during shortage pricing intervals in the summer of 2006.
 - The figure shows 318 intervals with shortage pricing of Eastern 10-minute reserves.
 - ✓ Based on the amount of physically available 10-minute reserves, Eastern New York was in a physical shortage in 97 percent of these intervals.
 - This is an improvement over the previous summer when 93 percent of shortage pricing intervals occurred during periods of physical shortage.
- The following figure shows very good consistency between the pricing dispatch and physical dispatch passes of RTD during periods when shortage pricing was invoked.
 - Thus, shortage pricing in Eastern New York has occurred during true shortages, and these shortages have been accurately reflected in the real-time prices of energy and reserves.



Scheduling of 10-Minute Reserves in the East During Shortage Pricing Intervals – June to August, 2006



-60-

٠ ٠

- The following figure shows available reserves during physical shortages of Eastern 10-minute reserves as well as a line indicating intervals with Eastern 10-minute reserves shortage pricing.
- There were 57 intervals with physical reserves shortages but no Eastern 10minute reserves shortage pricing.
 - ✓ This represents just 15 percent of the intervals with physical shortages;
 - ✓ The shortage was less than 100 MW in 77 percent of these intervals; and
 - ✓ The average Eastern 10-minute reserves price was \$185/MWh during these intervals.
- The following figure shows dramatic improvement in the consistency between the pricing dispatch and physical dispatch passes of RTD during periods when the East is short of 10-minute reserves.
 - ✓ During the summer of 2005, just 43 percent of the shortage intervals did not result in shortage prices (compared to 15 percent in 2006).
 - ✓ This improvement is the result of two software changes made after the initial implementation of SMD 2.0. These are discussed below.



Reserve Shortages and Shortage Pricing Conclusions

- The dispatch software implemented under SMD 2.0 has significantly improved the efficiency of energy and ancillary services pricing.
 - It replaced software that did not consider how ancillary services affect the cost of energy.
 - It reduces system costs by re-allocating ancillary services every five minutes.
- During the summer of 2005, real-time energy and reserves prices sometimes did not fully reflect that the system was under shortage conditions.
- Prior to the summer of 2006, two software changes were made that better enable the real-time market model to set efficient clearing prices.
 - ✓ In mid-August 2005, enhancements were made to allow off-line quickstart GTs to be co-optimized by RTD for providing energy and reserves.
 - ✓ In May 2006, a change was made to allow the physical and pricing passes of RTD to be more consistent regarding the ratings of gas turbines in high ambient temperature conditions. This is explained below in greater detail.



- Hybrid Pricing generally enables the real-time software to calculate efficient prices, especially in areas that are primarily served by GTs.
- Hybrid Pricing utilizes a pricing dispatch and a physical dispatch that can differ significantly, which can affect whether the pricing dispatch perceives a shortage in 10-minute reserves.
 - ✓ The Hybrid Pricing approach allows the pricing dispatch to treat on-line GTs as flexible from zero to maximum, while the physical dispatch always includes them at their maximum output level.
 - Thus, the pricing dispatch may count less energy from these units, but only when it is not economically in-merit, which is generally not the case during reserves shortages.

- Two other factors have contributed to deviations between pricing and physical.
 - ✓ Units Not Following Dispatch: In general, physical dispatch instructions are "ramp-constrained" by the expected physical output of the unit plus or minus what can be ramped in one interval, whereas the pricing dispatch level is ramp constrained by the last pricing dispatch level plus or minus the ramp limit.
 - Thus, the pricing dispatch may count *more* energy from units that persistently *under*-produce.
 - And, the pricing dispatch may count *less* energy from units that persistently *over*-produce.
 - ✓ *Inconsistent Output Limits for GTs*: Inconsistencies between the offer amount and the actual production level can arise when high ambient temperatures reduce the maximum output level of GTs.
 - The physical dispatch uses the actual production level while, until May 2006, the pricing dispatch used the offer quantity.
 - The physical dispatch and pricing dispatch currently use the same value.
 - Until this software change, the pricing dispatch generally counted more production from GTs than the physical dispatch.



- In May 2006, inconsistencies were eliminated between the pricing and physical dispatches in the output limits of GTs.
 - ✓ GTs that fail to reach their as-bid maximum output level after three intervals are treated as having a derated maximum output level.
 - ✓ Now, the physical dispatch and the pricing dispatch both assume the maximum output level is equal to the telemetered output level.
 - ✓ Once the reduction in capability is recognized by RTD, it is also fed back to RTC, which takes it into account when making commitment decisions.
- Consistent ratings of GTs under high ambient temperatures has greatly improved the efficiency of prices during reserves shortages.
- The following analysis examines the effects of inconsistent treatment of units not following dispatch instructions on Eastern 10-minute reserve prices during intervals when a) there was a physical shortage and b) no shortage pricing.

- The following figure summarizes the effect of units persistently not following dispatch instructions on Eastern 10-minute reserves prices during the 57 intervals when there was a physical shortage and no shortage pricing.
 - \checkmark The bars indicate the shortage quantity in the physical dispatch pass of RTD.
 - The line indicates the additional energy and 10-minute reserves available in the ideal dispatch pass due to inconsistencies in the treatment of units not following dispatch instructions.
 - The additional supply available to the ideal dispatch pass was greater than the physical shortage quantity:
 - \checkmark in 12 of the 57 intervals shown; and
 - \checkmark In 3 of the 13 intervals when the shortage exceeded 100 MW.
- The inconsistent treatment of units not following dispatch instructions explains a modest share of the instances when the physical dispatch pass perceived a shortages of reserves while the pricing dispatch pass did not.



Impact of Units Not Following Dispatch Instructions During Periods of Shortage and No Shortage Pricing June to August, 2006



-68-



Hybrid Pricing – Conclusions

- Some differences between the pricing and physical dispatches in RTD are necessary to implement the hybrid pricing regime. However, unnecessary differences will generally lead to inaccurate prices and increased uplift.
- The consistent treatment of GTs under ambient temperature restrictions, which was implemented in May 2006, has greatly improved the efficiency of prices during Eastern 10-minute reserves shortages.
- Additional improvements to the consistency of the pricing and physical dispatch passes of RTD should lead to more efficient pricing of energy and ancillary services (particularly during shortages) and reduce uplift.
 - ✓ We recommend the NYISO assess the feasibility of re-calibrating the dispatch levels in the pricing pass for units that are not following dispatch signals.



Market Operations – Real-Time Commitment

- The NYISO upgraded its real-time commitment model as part of the SMD 2.0 implementation:
 - ✓ The RTC model commits gas turbines, and schedules generation, ancillary services, and external transactions. It runs every 15 minutes and is a significant improvement over its predecessor, the hourly BME model.
- Convergence between RTC and actual real-time dispatch is a substantial concern because a lack of convergence can result in:
 - ✓ Uneconomic commitment of generation, primarily gas turbines; and
 - ✓ Inefficient scheduling of external transactions.
- When excess resources are committed or scheduled, the results are increased uplift costs and depressed real-time prices.
 - Alternatively, committing insufficient resources leads to unnecessary scarcity and price spikes.
- This section includes several analyses that evaluate the consistency between RTC and actual real-time outcomes.





Efficiency of Gas Turbine Commitment

- The following figure measures the efficiency of GT commitment during the past three summers by comparing the offer price (energy plus start-up) to the real-time LBMP over the initial commitment period.
 - The left panel shows the average volume of gas turbines being started whose energy + start-up costs (amortized across the commitment period) are:
 - (a) < LBMP (clearly economic);
 - (b) > LBMP by up to 25 percent;
 - (c) > LBMP by 25 to 50 percent; and
 - (d) > LBMP by more than 50 percent.
 - The right panel shows the quantity gas turbines that were likely economic, but not started (i.e. the LBMP > Energy plus start-up offer).
- Some of the GTs with offers greater than the LBMP in the left panel are also economic, because GTs that are started efficiently may sometimes not recover their start-up costs.


Efficiency of Gas Turbine Commitment

- The left panel of the following figure indicates that the fraction of GT commitments that are clearly economic has risen dramatically over the past three summers.
 - ✓ For 10-minute GTs, the fraction rose from 33 percent in 2004 to 24 percent in 2005 to 59 percent in 2006; and
 - ✓ For 30-minute GTs, the fraction rose from 11 percent in 2004 to 47 percent in 2005 to 70 percent in 2006; and
- The reduction in uneconomic commitment of GTs had led to substantial decreases in the amount of real-time non-local reliability uplift. The first section of this review indicates that this category of uplift was:
 - ✓ \$27 million during the summer of 2004;
 - ✓ \$20 million during the summer of 2005;
 - ✓ \$10 million during the summer of 2006;
- The reduction in uneconomic commitment must be weighed against the additional capacity that would likely have been economic if it had been started.
 - ✓ However, this category (shown in the right panel) has remained small relative to the volume of economic starts.

Efficiency of Gas Turbine Commitment Comparison of SMD and SMD2 June to August, 2004 to 2006



Efficiency of Gas Turbine Commitment – Improvements Under SMD 2.0

- While the introduction of SMD 2.0 led to more efficient commitment of gas turbines, the previous analysis indicates significant efficiency improvements since 2005.
- In May 2006, RTD and RTC began to model transmission constraints in New York City using a detailed representation of the network of transmission lines, replacing the use of simplified interface constraints.
 - ✓ When constraints are binding, this allows RTD to re-dispatch generators more efficiently to optimize flows into each load pocket.
 - RTC must frequently commit generation before constraints are actually binding. Thus, the detailed line model of New York City enables RTC to better anticipate congestion, which leads to more efficient commitment of gas turbines.
- Discrepancies between RTC and RTD have likely been reduced by better consistency between the physical and pricing dispatch passes of RTD.





Comparison of RTC and RTD Prices

- The following analyses in this section examine the reasons for differences between RTC and RTD prices.
 - RTC runs every 15 minutes, and each RTC run produces advisory prices at 15 minute intervals over a 2 hour and 30 minute horizon.
 - ✓ The following analyses compare RTD prices with the RTC prices for the interval that is closest to the time when RTC runs.
- The comparison of RTC and RTD prices provides a general indication of convergence between RTC and RTD. Inflated RTC prices can lead to:
 - ✓ Uneconomic commitment of generation, primarily gas turbines; and
 - ✓ Inefficient scheduling of external transactions.
- Excess commitment and scheduling results in increased uplift costs and depressed real-time prices.
 - Alternatively, failing to commit economic resources leads to unnecessary scarcity and price spikes.



Comparison of RTC and RTD Inputs

- The following figure shows the differences between RTC and RTD in loads, net exports, and prices at 15-minute intervals during the day.
- Loads and net exports are inputs which jointly determine the quantity of internal resources that must be scheduled by RTC and RTD.
 - Thus, increasing load and net exports requires additional internal generation, which leads to higher prices.
 - ✓ Net exports and loads are stacked in the figure to show their cumulative effect.
- RTC load is consistently higher than RTD load during the morning ramp period, which leads to correspondingly higher RTC prices.
 - RTC schedules resources at time *t* using the highest of the load forecasts of time *t*, *t* plus five minutes, and *t* plus ten minutes.
 - As a result, RTC load is approximately ten minutes ahead of the load forecast during the morning ramp period.
- The difference between RTC and RTD prices briefly spikes high and low at specific times during the day.
 - ✓ RTC prices are *higher* on average by at least \$15/MWh at 6:00, 7:00, and 12:30.
 - ✓ RTC prices are *lower* on average by at least \$20/MWh at 17:00, 17:15, 21:00, 23:00, and midnight.



Comparison of RTC and RTD Inputs

- Systematic differences between RTC and RTD prices are correlated with differences between RTC and RTD values of load and net exports.
 - ✓ At the top of each hour, RTC and RTD do not expect the same level of exports. RTD assumes that each interface "ramps" at a constant rate from five minutes before the top of the hour to five minutes after, whereas RTC assumes that each interface meets the next hour schedule at the top of the hour.
 - Reasons for the 15-minute variations in the differences between RTC and RTD load are discussed in the next analysis.
 - At specific times of day, systematic differences between RTC and RTD prices seem to be explained by differences between RTC and RTD values of load and exports.
 - \checkmark From 5:15 to 11:00, there is a strong correlation;
 - ✓ Likewise, from 20:30 to midnight, there is a strong correlation; and
 - The afternoon and early evening do not exhibit an obvious correlation. Thus, other factors, such as transmission constraints and locational reserves shortages, become increasingly important.
- The analysis suggests that differences between RTC and RTD values of load and exports play a significant role during ramping hours.





Comparison of RTC and RTD Inputs

- The following analysis compares differences between the load forecasts used by RTC and RTD to the net estimated regulation deployment by time of day.
- There is a strong correlation between variations in regulation deployment and the difference between the load forecasts used by RTD and RTC.
- For example, at 5:15:
 - ✓ Regulating units are usually being instructed to increase output relative to 5:00.
 - The difference between the RTD load forecast and the RTC load forecast shifts in the positive direction.
 - The additional load scheduled by RTD reduces the amount of regulation that must ultimately be deployed.
- To minimize regulation deployment, the operators make incremental adjustments to the load forecast, which reduces the need for regulation.
 - ✓ Lower regulation requirements lead to lower regulation procurement costs.
 - Reduced deployment of regulation results in less economically out-of-merit generation.
- Because RTC looks further into the future than RTD, adjustments to the load forecast are reflected "sooner" in RTD than in RTC.



-81-

Comparison of RTC and RTD – Conclusions

- Currently, three factors undermine convergence during ramping hours:
 - ✓ RTC schedules resources at time *t* using the highest of the load forecasts at time *t*, *t* plus five minutes, and *t* plus ten minutes. This practice consistently leads RTC prices to be higher than RTD prices during the morning ramp period.
 - ✓ RTC and RTD use different assumptions about the level of expected exports. RTD assumes that each interface "ramps" at a constant rate from five minutes before the top of the hour to five minutes after, whereas RTC assumes that each interface meets the next hour schedule at the top of the hour.
 - ✓ The load forecast is adjusted in real-time to reduce the need for regulation deployment, which results in differences between RTC and RTD load.
- We recommend the NYISO evaluate whether:
 - ✓ There is an alternative to RTC using the highest of three five-minute load forecasts;
 - The assumptions about external transaction ramp can be made consistent to eliminate differences at the top of each hour; and
 - ✓ Predictable adjustments to the RTD load forecast, which are made to minimize regulation deployment, can be reflected more quickly in the RTC load forecast.

Market Operations – Supplemental Commitment POTOMAC ECONOMICS



Supplemental Commitment

- The last section of this review evaluates supplemental commitments during the summer of 2006.
- Supplemental commitment occurs when a generator is not committed in the economic pass of the day-ahead market but is needed for local reliability. Supplemental commitment primarily occurs in two ways:
 - ✓ The Day-Ahead Local Reliability Pass of SCUC commits generators after the economic commitment but before clearing prices are determined.
 - ✓ The Supplemental Resource Evaluation ("SRE") process is used to commit generators after the day-ahead market.
- In the first section of this review, we reported increased uplift expenses for both day-ahead and real-time local reliability.
 - Day-ahead local reliability uplift arises entirely from commitments by the local reliability pass of the day-ahead model.
 - ✓ Real-time local reliability uplift arises primarily from SRE commitments.



Supplemental Resource Evaluation

- The following figure summarizes supplemental commitments made by the NYISO after the day-ahead market.
 - ✓ They are important because they influence the real-time market results.
 - ✓ To the extent that they are anticipated by the day-ahead market, they will also influence day-ahead market results.
- The average quantity of capacity committed through SRE in New York City decreased significantly in the summer of 2006.
- SREs are called by individual TOs, so the resulting uplift is allocated to the local area. SREs are the primary source of RT Local Reliability Uplift.
 - Although SREs decreased significantly from the summer of 2005 to 2006, RT Local Reliability Uplift increased from \$23 million to \$26 million.
 - ✓ Uplift expenses rose primarily because several generators in up-state New York raised their offer prices substantially above marginal costs, while staying under the applicable conduct thresholds for mitigation.

Supplemental Resource Evaluation Commitment June to August, 2004 to 2006





Day-Ahead Local Reliability Commitment

- The next analysis focuses on commitments made in the day-ahead market (i.e., by SCUC) to meet local reliability requirements.
- These commitments are not made because they are economic to serve dayahead load. However, they are important because they tend to:
 - Reduce prices from levels that would result from a purely economic dispatch; and
 - Can increase non-local reliability uplift a portion of the uplift caused by these commitments is incurred to make guarantee payments to other generators that will not cover their as-bid costs at the reduced price levels.
- The following figure shows the average quantity of these commitments.
 - ✓ The increase in day-ahead local reliability commitment in 2006 is consistent to the decrease in SRE commitment.
 - ✓ Day-ahead local reliability uplift increased from approximately \$7 million in the summer of 2005 to \$13 million in the summer of 2006.





Units Frequently Committed for Local Reliability

- To further evaluate both the local reliability and SRE commitments, we analyze them at the individual unit level.
- The following figure shows seven units committed most frequently for local reliability or through the SRE process in NYC and Long Island.
 - ✓ The values shown are the hours that each unit is committed as a percent of the hours that the unit is available to the day-ahead market (i.e., not on outage).
 - ✓ Six of these units are in NYC and one is on Long Island.
- When these units were available but not committed economically, they were committed in the local reliability pass of SCUC or through SRE a large share of the time.
 - Supplemental commitments can cause units committed in the economic pass to be uneconomic, thereby increasing uplift and depressing energy prices.

✓ It would be more efficient for these units to be committed within the economic pass of SCUC.

Units Most Frequently Committed by SRE and the Local Reliability Pass of SCUC in NYC and Long Island June to August, 2006





Supplemental Commitment Conclusions

- Supplemental commitments have a number of significant market effects:
 - ✓ Inefficiently reducing prices in the day-ahead and real-time markets;
 - ✓ When they occur in a constrained area, they will inefficiently dampen the apparent congestion into the area; and
 - Increasing uplift as units committed economically will be less likely to recover their full offer production costs;
- In the summer of 2006, a substantial share of local reliability commitments shifted from the SRE process to the day-ahead process, although the total amount did not change significantly.
- To reduce the inefficiency and uplift associated the supplemental commitments we recommend:
 - ✓ In the short-run, that the ISO allow operators to pre-commit units needed for NOx compliance or other predictable local reliability needs; and
 - ✓ In the long-run, that the local reliability and NOx constraints be included in the initial economic commitment pass of SCUC.



Supplemental Commitment Conclusions

- Both of these recommendations will 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 associated with payments made to units supplementally committed to meet the requirements are allocated locally.
 - Payments made to other units due to the price changes caused by the supplemental comments are allocated throughout NYCA.
 - ✓ When the recommendations are implemented, a methodology would need to be developed to identify units due to the local reliability requirements.