

**Harvard Electricity Policy Group  
Twenty-Third Plenary Session**

The Four Seasons Hotel  
Boston, MA

Thursday, September 21 - Friday, September 22, 2000

**RAPPORTEUR'S SUMMARY\***

**Session One: Market Power and RTOs**

*Many regions are wrestling with market power issues. There are, of course, the traditional concerns about market power in energy markets in general. More specifically, however, RTOs are confronting concerns about market power in ancillary services, congestion markets, and other, more specifically defined niches in the market. Price caps, in various forms and time frames, have been imposed in various regions. While the caps have had their proponents and defenders, there are also critics who view the price caps as either undue interference in the marketplace or as the wrong pricing measure for addressing the problem. Are there really market power issues that are impeding the optimal functioning of regional markets? If so, precisely what are they and what substantive measures should be taken to remedy them? There are also institutional questions raised by the controversies. Who should raise, who should investigate, and who should decide how to remedy market power questions? What are the RTOs, particularly the ISOs, doing to resolve market power problems? Some argue that the ISOs have effectively become self-appointed, private regulators in addressing these issues. Is that true? Is such a role implicit in their missions.? Should ISOs be addressing market power issues at all, and, if so, how?*

**Speaker One**

What is market power? Economists define it as the ability to affect prices through either your output or your bidding decisions. We need to try to balance the cost of mitigating or

reducing the market power against the costs of the market power itself. There are efficiency costs from the exercise of market power in terms of cheap generation owned by large firms, perhaps being displaced by more expensive generation owned by smaller firms. The cost of reducing it to zero would be extreme. We have to think about ways in which we can reduce it to a manageable level, the options for doing

---

\*HEPG sessions are "off the record". The Rapporteur's Summary captures the ideas of the session without identifying the speakers.

so, and the costs of implementing those options.

Unilateral exercise of market power is, in general, not illegal in this country. Taking measures to extend or maintain a monopoly or a dominant position can be illegal. Explicit agreements between different competitors to raise prices can be illegal. Tacit collusion probably is illegal, but is unlikely to ever be punished. What we're talking about in the electricity industry is more of a regulatory than anti-trust problem.

The electricity industry is more vulnerable to market power. You can't store electricity cheaply unless you have a hydro facility. There are binding, short-run capacity constraints on production. We have transmission constraints. Transportation is also limited relative to most other commodities. All of these combine to make the supply response less elastic. You combine that with the demand side, where the price responsiveness is virtually non-existent since most customers don't know what the price is on an hour-by-hour basis, and we have extremely inelastic demand and, at times, extremely inelastic supply. All this combines to mean that raising prices under certain conditions of tight supply becomes fairly easy. Frequently repeated market interactions--hourly or daily bidding--also, some feel, creates more of a risk of tacit collusion.

Market power in electricity is most acute in periods of tight supply. The trick to market monitoring is to determine whether a price spike is determined by true scarcity, or whether we're close to scarcity, giving firms with any remaining capacity market power.

Given that we're going to have some market power, what do we do about it, if anything? There are three sets of policy tools that we might apply--structural solutions, regulatory solutions, and market rules.

Divestiture is the classic structural solution. It would ideally accomplish a decrease in concentration of generation. But the electricity industry in California is relatively unconcentrated. The standards for what levels of concentrations are prudent are borrowed from other industries. Those industries don't have electricity's problems in terms of lack of storage and inelastic demand. One would expect that we need a tighter standard in terms of concentration in the electricity industry than in other industries. Again, we need to think about the costs of the mitigation measure versus the benefits of reduced market power. There are times in which divestitures have been focused at the vertical market power problem. For example, the sale of San Diego's gas and electric generation assets, which were sold to an existing player in the Southern California market, actually increased the horizontal concentration.

Transmission expansion is another way of accomplishing a decrease in the concentration of supply by expanding the geographic market over which these suppliers are competing. In general, we need more transmission capacity in a deregulated regime than would have been optimal in a regulated regime.

Everyone is in favor of demand elasticity, although now in San Diego they're realizing what that means. Infrastructure needs to be put in place--

metering technology, a regulatory infrastructure that defines the responsibilities of retailers in a way that's more clear. It's worth considering requiring some level of these infrastructure elements before we move forward with market-based rates, at least in markets that haven't yet been granted it.

Regulatory solutions include incentive regulations directed at the revenues of returns earned by the companies selling power, and behavioral regulations aimed at the actual operations or decisions of the firm's selling power. Wholesale price caps are the most widely applied regulatory tool, at least in the markets that have restructured. There have been a lot of complaints about the uncertainty regarding what the price cap movements would be and what factors influence those movements. One tool that's been applied more internationally than in the U.S. is vesting contracts. Usually, this is a transitional tool in which the selling firm or regulator requires that a certain amount of the capacity of the portfolio being sold be contracted under a pre-negotiated or regulated rate. This shrinks the market that's unregulated or is competing on a short-run market-based process, and shrinks the size of the firms that have signed the contracts. The problem is that unless there are other things going on during the contract period, when the contract expires you're back where you started from.

Behavioral regulation is things like unit-specific bid caps, or the various ways in which we readjust bids in some markets. By doing this, we're producing a sort of dispatch in the system, a merit order, based on the regulation we're putting in. There are greater costs to applying this

more heavy-handed form of regulation. There are greater information requirements. We need to know a lot about the costs of specific units before we start redefining their bids to be based on their variable operating costs, for example.

Lastly, we have enforcement/punishment rules. The most infamous of these is the good behavior clause being debated in the U.K. The notion is that we can make it bad to try and raise prices beyond a certain level or to make too much money. This is very difficult, and there are a lot of costs in terms of monitoring and enforcement in trying to force firms to do things that are not in their economic interest.

The closer we get to competition, the less we have to apply these kinds of standards. They are Band-Aids to try and get us through transition periods or through periods of tight supply, rather than a foundation upon which we operate the market. I'm worried about the tendency in some areas to create a so-called deregulated market with an uncompetitive market structure and rely on these alternative forms of regulation to make that restructured market operate. We're not sure what the implications of many of these measures are.

Market rules can't eliminate market power, but there is a lot of evidence that they can exacerbate it. In bidding, for example, there is generally a tradeoff between providing flexibility to supply offers and also allowing them to manipulate different degrees of freedom or that flexibility to tailor those bids not just to their cost structure, but also to the market conditions. There are efficiency benefits to allowing them the flexibility,

but there also are potential market power consequences. Trying to empirically examine where those tradeoffs fall would be an interesting line of research.

In California, there's been a lot of attention focused on the uniform price auction of the Power Exchange. It is not the source of the market power problem. I don't think scrapping it is going to have any impact on outcomes.

Market power is a serious problem. Structural changes are needed and are definitely preferred. We probably need to think about concentration requirements stricter than those we have been applying. ISOs are becoming regulators of their market. The question is what kind of process they will follow. If we're uncomfortable with that kind of role, then we should think about the changes we can make that would limit the power of or the need for ISOs to apply those kinds of rules.

## **Speaker Two**

We have five real-time markets in New England--an energy market, three operating reserves, and automatic generation control. Structurally, there has been a lot of diversification. Most generation has been sold off, but while the names have changed, they're still one group controlling a huge block, which has created concerns with ownership in load pockets. Operating reserve margins have decreased.

New England has a centralized unit commitment, day-ahead process; the ISO performs a unit commitment for the next day based on the prices and operating characteristics submitted. The energy market currently is a single settlement market, though FERC has

approved a multi-settlement system. Transmission congestion costs currently are socialized across all of New England. Again, FERC has approved moving to a locational marginal pricing system.

As to unmitigated transmission costs in northeastern Massachusetts since the market went into operation in May 1999, the peak in December 1999 was a little over \$14 million in one month in one small area of New England. If you look at the energy uplift costs, prices, after being relatively stable until early this year, skyrocketed to over \$16 million in April and are still running quite high. What's that got to do with market power? Well, now we've got to figure out what part of those costs are due to scarcity rents and are really justified, and what part are due to people exercising market power.

New England has market rule 17, which governs market power monitoring and mitigation. It looks at two questions: Which sellers are exercising a pattern of behavior that is consistent with trying to raise the price? Have they been successful in actually raising the prices in New England? The monitoring process looks at three broad categories. First, unconstrained system operations. Is any party intentionally withholding capacity? Or is it being withheld just by economics? The second is more local. Is there a transmission constraint, and are there resources within that constraint that have market power? The third category is external energy purchases. This is in the latest draft of Rule 17. It will establish a cap on the price the ISO will pay for external purchases based on the price the ISO paid for emergency purchases from other power pools since

the ISO took over running the new markets.

What are some of the mitigation options in the unconstrained world? The ISO can negotiate voluntary mitigation arrangements. But it can also impose a variety of mitigation measures, including reducing the bidding flexibility of a resource, increasing the resource's reserve obligation, and replacing the resource's bid price with a reference price based on the historic bidding pattern of the resource when it was in merit.

In the constrained operation situation, Rule 17 tries to recognize the difference between the unit that pretty regularly runs in merit and just happens to now be in the load pocket because some transmission or other units are out of service, and the unit that's the true backyard peaking unit that is expected to run 10 hours a year and doesn't have a history of actually bidding and operating economically in the market.

Two market price screens have been created. One is a structural screen the ISO looks at when a resource is in a load pocket: Was there competition for the supply of the necessary resources? It depends on the circumstances, but if there are three or five competitors to provide a service, then the resource is deemed to not have market power. If it doesn't pass the structural screen, they look at a price screen, which looks at how often the unit is being run out of merit and at its previous prices or the clearing price in the rest of the market.

If the price gets mitigated, it gets mitigated according to the same tables that are used to screen it. Essentially,

you can exercise a bit of market power for a short period of time, but as either the number of consecutive hours of operation increases, or the number of cumulative hours over the previous 90 days increased, then the percentage you can bid in over the prices that you bid when you were in merit decreases. The reference price is based on your in-merit operation over comparable hours, over the preceding 30 days. It weights near-term prices higher. The unit always has the option to negotiate a cost-based rate.

The ISO is required to report to the industry with monthly reports, quarterly reports to the regulators, and annual reviews. These reports identify general levels of mitigation activity, but don't provide specifics. The current NEPOOL information policy bars the ISO from publishing any information deemed to be commercially sensitive.

Concerns fall into several broad areas. There is a lack of transparency in the markets due to the lack of data that is actually published, so it is hard to tell if market power is being exercised. It is hard to tell if the ISO is actually running the market in accordance with the market rules. The information that is published is not published in a timely manner. FERC just ordered the ISO to publish bid data, but they are going to do it six months after the fact, with identities hidden, which makes it difficult to analyze portfolio bidding behavior. The tools and resources available to the ISO may not be adequate. Generators think there is too much intervention and the load thinks there's not enough.

Can lessons be learned from markets around the world? In England, Australia

and Alberta, they release a lot of data and they release it quickly. You can go to the NEMMCO website in Australia and find out what was bid in the previous day for prices, availability, re-declarations of availability--everything by unit. You can do the same thing in the UK, except you have to pay for the data. Parties in New England are pushing for release of specific data. Another concern is transparency of real-time operations. Is the ISO running the market according to the rules, dispatching units in economic merit order? And a move to locational pricing doesn't really address market power, but maybe makes it more visible.

In New England, a new market rule 17 is in the debating stages. The new rule will provide for an independent market advisor to advise the ISO board on market power issues and activities. The new rule has more specific trigger mechanisms as to how they're going to try to detect the exercise of market power when the operations are unconstrained by transmission constraints. They are trying to build on the experience of the last 18 months or so. There will be more specific remedies, although how specific they should be is a subject of debate.

### **Speaker Three**

What kinds of problems are best addressed by regulation as opposed to anti-trust types of enforcement? The highest level of generality says that regulation is good for markets or portions of markets that are permanently broken. We also look for permanent scarcity kinds of situations--lack of

storage in this industry, or if there's a permanent inelasticity of demand, maybe we have a market failure. If we do, then we've taken a long detour on the way back to regulation. Anti-trust is good for keeping markets competitive--for blocking mergers that will worsen things, and for policing for collusion and other bad behavior.

One thing we don't do in the U.S. very much is regulate commodities just because they're important. The point of regulation is not simply that we have an important product and the price is high. Usually under our system, we've looked to see if the market is flexible, and if it's not, then we try to impose the best kind of regulation that we can.

Which case is this? We have market power. But the problem is that market power is not necessarily a problem. Sometimes it is the result of competition. Someone can gain a very large share of a market because they happen to be providing the best product at the best price. How do we determine which kind of market power we're going to spend time addressing? In most markets, if somebody is getting a scarcity rent, their ability to collect it is policed by the magic of the market itself. It's called entry. Deregulated energy markets seem to be the proverbial party to which nobody came. We have the same people, about the same size resources, we may be having some marginal entry for new demand, but there hasn't been the entry that we expected was going to come in on its white horse, as it does in most industries, and save the day.

Why did we think there would be entry? First, we knew there was new technology that made it possible for

smaller plants to build and efficiently compete in the market. The smaller plants were easier to site and were faster to the market. Second, we thought that if we divested the transmission lines, we would have instant entry because everyone would hop on the transmission train and get that competition going. It didn't happen. The reasons include congestion in the system, the political difficulties of siting new facilities, and the fact that we have a transmission grid which is interconnected, which requires owners to get together and talk about transmission issues in an industry where we do not have complete divestiture and, therefore, there are people deciding on transmission rules and prices who have an interest in limiting entry.

Was our whole assumption about entry wrong? It's important to remember that the correct comparison of the present market to regulation is not to perfect regulation any more than it might be to compare it to perfect competition. Regulation in most of the U.S. was very, very broke. It did not induce efficient entry. Our entry assumptions have to be examined for each individual kind of energy product. Entry conditions are very different for wholesale than for ancillary services.

To have entry, first, we have to signal people that they'll be able to recover their costs. Limiting prices during a transition can have the effect of a never-ending transition. Looking at price caps in markets that you think are going to be competitive, as opposed to ones in which you think you're going to have a persistent market power problem--that is where you really have to do a cost-benefit analysis of what the effect will be on entry. Second, we have to make

sure that we have rates that will encourage investment in transmission. FERC has been looking at that seriously, and we want to continue to do that, to make sure that costs and business risk factor are reflected in those rates.

The hardest questions arise with ancillary service markets and load pockets. We have to look at them separately. Some load pockets happened because of a concentration of generating resources within transmission constraints. Then you can have a situation where there could be entry--we do have interesting and increasingly economic sources of smaller generation that can come in those circumstances. If we examine a load pocket and determine the problem is concentrated resources and ownership, there are two things you can do. One is to encourage entry, make sure there is a public will to site new facilities within the load pocket. If there is no public will to do it, then don't rely on entry to fix the problem. Second, you can decide that you only need one generator in this load pocket. Then maybe price caps would be good because you don't have to trade off much since you're not thinking that there's likely to be entry.

That's a different question from ancillary services. That will not necessarily be solved by entry since it frequently has to do with exactly where the generator is located. Under those circumstances, you may have to think of different regulatory answers. One of the things I don't think we can rely on, or that we have to examine very closely, is how much we can rely on transmission to provide entry for reserve products. I worry that we think we'll be able to import reserves. But if it's a hot day in the Northeast, why

do we think there's going to be power available in adjacent markets?

What should we do? You could shrink the market in which these unregulated sales are made by vesting contracts. But if you shrink markets, you make them easier to manipulate. Residual markets are thinner. It takes less generation to be able to manipulate price in a thin market. The good behavior rule is the worst of all possible worlds. It is after the fact, and gives you all the bad sides of regulation.

First, FERC has a role. It can look at transmission pricing to make sure that it is at a level that does not impede entry in the generation market or expansion of the grid. Second, we can watch for exclusionary collusive behavior. Third, there is a big role for legislation because all of us are hamstrung by the fact that many of these markets are concentrated, and ownership has been transferred, and the concentration has not been diluted. There are very few legal tools that clearly allow those ownership portfolios to be broken up absent legislation.

Fourth, there is certainly a role for the ISOs. One of the biggest roles that the ISO can play is to help identify whether these problems are problems of scarcity, of monopoly, and whether entry, given the local circumstances, would likely help solve the problem--and whether we therefore should look at competitive-type solutions, or whether it is a permanent problem where certain facilities will continue to be able to dominate a particular market in which maybe we need to look at market caps, etc. We should give ISOs a break. They are under a tremendous amount of pressure. Their tools are not clear, their

enforcement authority isn't clear, their independence is not clear, and their procedures are not public, so they're subject to after-the-fact criticism.

#### **Speaker Four**

The market structure in California has gone to full retail access. Scheduling coordinators provide balance schedules. The California ISO does not do unit commitment like the Eastern ISOs do; that's handled by scheduling coordinators in their own forward markets. The intended design was to have bilateral trades, block forward, day-ahead, hour-ahead and real-time in ascending order, with real-time being a very small market. In fact, we have some bilateral trading; we have blocks of energy being sold outside of California. We've been a heavily importing state for a long time. We end up in a situation where the real-time market is often 30 percent of our operating day.

We've had an enormous debate about price caps and their effectiveness in controlling market power and wholesale costs. In May and June we had \$750 price caps, in July we had \$500 price caps, and in August we had \$250 price caps. The total energy cost in August is higher than any of the previous months. Why did that happen?

The fundamental underlying issue is one of sufficiency of supply. I would argue that whether something is monopoly or scarcity is almost a meaningless distinction. In California--and this is true of much of the western U.S.--we haven't put any new generation in place for over 10 years. We're having load growth on the order of 1,000 MW per



year. In the Bay Area, in 1988 we saw peak load of about 8200 MW, in 1999 we saw 8450 MW, so we expected in 2000 to see peak load of about 8700. In fact, the peak, on June 14, was 9150 MW, which is probably three times the growth that we had expected. It was very hot, and the transmission constraints made it so you couldn't serve the load.

There has been huge load growth, not only in California but in all of the neighboring states. Typically, California imports 25 percent of its power. This summer I have seen, on average, a 2,000 MW decrease in available supply at any price. That is also reflective of the load growth in surrounding areas. So while the debate about market power and scarcity is important, the fundamental issue is how to facilitate entry, both of supply and of some demand elasticity.

In all of the investigations in California, demand elasticity has been touted as a silver bullet. While I think that it is extremely important, I struggle with relying on it too much. We have some demand programs that we've tried to put in place. We had an RFP for this summer and only got 60 MW. We need 2,000 or 3,000. If I have a new 50 MW gas turbine and I'm only going to run it 30 hours in a summer, and it has to recover all of its annual cost in those 30 hours, the costs are much higher. A number of businesses indicate that they need between \$7,000 and \$8,000 per MWh to shut off. In the residential area, shutting off your average air conditioning load of one-and-a-half to two kW is only worth \$4 for that hour for that customer. So the most you could pay a customer to shut off his air conditioner all afternoon is \$20.

We have had underscheduling, in which, instead of scheduling power in the forward markets, because people are doing cost-minimizing and profit-maximizing behavior, they tend to drive to the real-time market.

What do we do in the meantime? We operate with a social contract that electricity is a right, that it should be available to everyone at a reasonable cost, that price spikes are not okay, and that it's not okay to shut the lights off. Sixty-four percent of California's generation is over 30 years old and a good portion is over 40 years old. No generation owner has more than a nine percent share of total generation capacity. Yet it has been concluded by all of our market monitoring groups that significant market power has been exercised. So concentration of ownership is an insufficient measure of this capability. In San Diego, the retail rate structure was not in place this summer for customers to respond effectively to prices.

The actions I believe are needed are to accelerate permitting and siting of new generation; develop load-responsive programs, to the extent they will work; and ensure that the distribution companies have enough incentives to hedge forward to lower the risks of wholesale price volatility. All regulatory barriers to retail competition have to come down. Generation is proposed to come on-line in particular places, but we could use 10,000 MW in addition to that, and begin to retire existing fleets or modernize fleets of aging generation.

I leave you with four questions. One, California has only 200 hours of loads above 40,000 MW. Do we want to create a separate peaking market, and have that handled outside of the normal market functionality, so that we can have less volatility in prices? Two, how much can we depend on demand elasticity to moderate high prices? Three, what is an appropriate transition? While there are clearly design factors that need to be fixed, some of them will take time, and entry will take time. So what is an appropriate transition to avoid political outrage, backlash, slowdown of the process? Four is market power: How much market power is too much? How much do you need as a price signal to get the kind of entry you need? And where do we cross the threshold between having enough to stimulate entry, and wealth transfer and political backlash?

## **Discussion**

*Question:* What do you mean by a structural solution to transmission? Some type of additional central planning, or a mandate for facilities to be built?

*First Response:* I don't know if we've really sorted out whether building new transmission is going to be centralized or mandated. We need more transmission capacity in a deregulated framework. The responsibility has increasingly fallen on the ISO to take the lead, but the money still is going to come from distribution companies, at least for now. Measures can be taken to upgrade the grid by eliminating the need for certain must-run status for many of their units. These kinds of improvements can get at permanent reliability-related problems and can also lessen the concentration of

ownership by making that market bigger.

*Second Response:* Disaggregation in the natural gas industry may give us a hint about transmission planning. It took a long time for people who ran pipes to stop thinking of themselves as people who served load and instead as people who sold to people who were going to serve load. They therefore wanted lots of flexibility, and started developing market hubs and integrated storage. All of those kinds of things followed from a fundamental philosophical change in management. We're nowhere near that point in electricity.

*Question:* A lot of learning occurs politically and business-wise out of these crises, and hopefully we can start to see more rapid decisions. I think we see signs that the politics of siting and building new facilities is significantly improved. If the decisionmakers in California are aggressive, can those timetables of getting new plants, and some of those investments on the transmission side, be accelerated?

*Response:* It is possible to move the social paradigm to a place where you have an outage or two on hot summer afternoons. But what are the economic and societal consequences? California has an RFP out for 3,000 MW of peaking capacity to be put in place by next summer. I believe there is a critical place for demand-side response. But some large customers are well-suited for that and others aren't. Many are tired of being interrupted. Some said this summer, "When I signed up for this, I didn't think you were ever actually going to use it." This goes back to questions of the timing we need to transition and the

appropriate mechanism to keep the volatility at a politically acceptable level, and still get the economics done.

*Response:* On demand side, Sydney Electric in Australia has 1,200 MW of controllable demand. I get daily reports that sometimes show that when the demand is ramped in, right about the same time, one of the generators is being ramped off. And they can interrupt the pot lines for up to two hours to respond to contingencies.

*Question:* Why have the August prices in California been so flat at all load levels?

*Response:* Gas prices are a factor; people who use gas should have bought forward at cheaper prices. Emission credit prices have gone up. Price caps have come down, changing people's bidding strategies. We have a lot less water available. Market power could also contribute. Do these factors fully explain it? No. We tend to look at prices versus ISO demand, and we should probably think about it as the price relationship to ISO demand, less imports, maybe less hydroproduction. Costs were higher in August than in May or June, when the price cap was \$750, and that's a big driver of why the average costs in August were higher.

*Question:* You can't identify the best fixes for the problem unless you've got adequate information. It seems that we don't have that right now. Do we need a national information access policy?

*First Response:* We need to release data, we need to release it quickly, and we need to release it by unit. I'm not so worried about a single unit exercising

market power. I'm more worried about a portfolio of units bidding in a strategic manner that can exercise market power. Let's get the information out there. Some will say that putting all that information out there will make it easier for entities to collude, but shining light on it will be better in the long run.

*Second Response:* Is this information that the market participants have anyway? That is a real concern, because people who participate in some of these markets, because of their position in the markets, are both customers and competitors. They exchange information in the guise of transmission, in which they are not competitors, that tells them about how to bid their generation where they are competing. It is a difficult question, but one of the national information policies is the anti-trust laws; there are certain things that you can't tell your competitors. So we have to look at how the information exchange is taking place, but it doesn't make sense to pretend that it is not going on unofficially.

*Question:* It was stated that California had only about 200 hours where load exceeded 40,000 MW, and it was asked whether the market for peaking units should be separate from the rest of the market. Have you thought about doing that so that it doesn't skew the entire market?

*First Response:* California has talked about it and may look at, one, does it make sense to do it, and two, if you did it, what would it look like? I don't think it's going to be as easy as you might expect because of the nature of electricity and since you're talking about something that's going to happen during

high demand periods. I think we're back to demand-side response being a big potential contributor during those periods.

*Second Response:* I think there is more skepticism than there ought to be that there will be entry for peak periods. I have seen market studies showing that

generators as large as 100 MW can be planned to function at fewer than 50 hours annually and still have a three-year payout. There are a lot of problems with expecting that to happen soon, but it doesn't necessarily mean that we should always assume that building peakers isn't an answer.

## **Session Two: Multi-Settlement Systems: Consistency and Efficiency**

*Coordination of electricity markets through real-time adjustments is necessary for reliability and can support competitive markets in unbundled generation. A principal purpose of coordination is to deal with the interactions of balancing requirements and transmission constraints. In order to gain more time for commitment and operational decisions, as well as to provide a greater degree of price certainty, electricity market designs have included day-ahead markets with a separate financial settlement. For example, California included such a day-ahead market in its original design; New York launched with such a design in November; and PJM introduced a day-ahead market in June. The operation of day-ahead,*

*hour-ahead and real-time markets with binding financial commitments and settlements raises questions of consistency and efficiency. What is the theory of the market interactions? How do differences in rules (such as the application of price caps) affect the interaction of the markets? How do multi-settlement systems affect market power mitigation policies? How has the introduction of a day-ahead market affected the behavior of market participants? How do the markets affect the valuation of financial transmission rights? How do pricing rules affect these markets? What has been the practical experience? What are the pros and cons of extending the formal coordination process to periods before real time?*

### **Speaker One**

PJM calls its multi-settlement a two-settlement system because it has two settlements, a day-ahead and a real-time. PJM's market design is fairly flexible. It allows a variety of trading options. Similar to the California market, it supports balanced bilateral transactions with no spot market backup, although few participants are actually using that type of transaction right now. The majority of transactions are bilateral with implicit spot market backup, and then generation is able to submit a generation offer, which consists of a three-part bid, and demands can participate in the day-ahead market.

With the two-settlement system, suppliers and customers bid into the day-ahead market both the prices they're willing to pay and the quantities. That's the first settlement. Then the suppliers offer into the real-time market to supply incremental requirements.

PJM supports a variety of financial contracts that are separate from the physical spot market. It has several different types of forward energy markets. PJM uses nodal pricing with overlaying transmission zones and commercial trading hubs. It has a NYMEX futures contract. It has the

day-ahead market, which is the two-settlement system, and financial transmission rights. Financial energy contracts allow internal bilaterals to be scheduled between participants within the PJM system. That does not change the scheduling of energy; it's just a transfer of megawatt ownership within the PJM control area. Participants that use these internal bilaterals can determine among themselves whether they want to settle at day-ahead or real-time prices.

So what is the two-settlement system? It allows PJM market participants to participate in a forward market for energy. There are two markets, and there are separate settlements performed for each of the markets. In the day-ahead market, PJM is developing a day-ahead schedule using least-cost security-constrained unit commitment and security-constrained economic dispatch programs. That means taking in all the generation offers in the day-ahead market, then matching that up with the load that's bid in.

The load consists of fixed load, price-sensitive load that's cleared, "inc" offers that have cleared--and those are just financial instruments. It also includes "dec" bids, which are virtual loads. A load is considered both a physical and a financial load, and PJM matches its generation resources up with that. If it

encounters a constraint, it uses the same process as in real-time, which is to have units used to control the constraint, and you see different locational prices during those hours. PJM is calculating an hourly price for every bus within the system and posting it on its website. All this takes place between noon and 4:00. So all bids have to be in at noon, and between noon and 4 the ISO runs its technical software. At 4:00 it posts its results to the market participants. Then there is a two-hour rebidding period during which any generation resource that was not selected in the day-ahead market can change its bid.

In terms of the settlements, PJM calculates a locational price and a schedule for every load and every generator, and that's the binding financial commitment for the day-ahead market. In the real-time market, it's based on actual deviations. So if the load has 100 MW that it cleared in the day-ahead market and it takes 105 in the real-time market, it will pay the real-time price for that 5 MW of deviation. Similarly, generators that either don't perform or are asked to provide additional MW will be paid the real-time price for that deviation.

In terms of the day-ahead market mechanisms, the generation resources submit offers. It is a three-part bid. There are startup, no-load, and incremental energy costs. A unit can participate in the day-ahead market by saying it is just going to be dispatchable at PJM's discretion, or it can self-schedule itself in the day-ahead market. Demand can put in fixed demand bids, meaning it will just be a price-taker and take whatever the price ends up being, or

it can put in a price at which it doesn't want to take that load. Transactions into and out of the PJM control area can also put in fixed transactions and can specify a maximum amount of congestion they're willing to pay, currently capped at \$25. Two financial mechanisms in the day-ahead markets allow participants to go into the real-time market with a short or long position--incremental offers, which are imaginary generators, and decrement bids, which are virtual loads.

Why did PJM implement the two-settlement system? It was high on the list of what PJM's participants were looking for it to do. It creates a more robust and competitive marketplace, and adds price certainty in terms of allowing generation resources to obtain a financial schedule in the day-ahead. Loads are now able to submit price-sensitive demand bids, and participants are able to identify a maximum price at which they don't want to pay to participate in the day-ahead market. The increment offers and decrement bids give participants a better ability to do the business they need to do to cover their positions. PJM is looking at what it would take to allow the technical software to clear both an inc offer and dec bid, basically an internal bilateral transaction.

In terms of fixed transmission rates, when PJM had only one settlement, real-time prices determined the economic value of the FTR. Now, the day-ahead locational prices are being used to determine the FTR value. This provides protection against day-ahead congestion charges. Participants asked, "Why can't we have FTRs in both the day-ahead and the real-time market?" You can't sell the same service twice. All of the real-time

market pricing is based on deviation. So as long as you flow the MW you said you were going to in the day-ahead market, you're indifferent to prices in real-time.

There has been broad participation in the two-settlement system. Over 900 participants have been trained. There are over 1,000 registered users on the interface that participants use to enter their data. There are continuing working group meetings in which participants tell PJM what they like about the system, what they're struggling with, what enhancements they would like to see to the technical software and to the design of the screens and the database. The day-ahead and real-time average LMPs track each other closely.

Prior to implementation of the market-based regulation product, there was insufficient regulation available in some situations. Post-implementation, there is sufficient regulation available. There is always a sufficient supply. The purchase price remained about the same, and there has been a significant improvement in system control that has allowed PJM to reduce its transaction notification times. PJM is evaluating its regulation requirements to see if they can be reduced.

### **Speaker Two**

Although the New York system is called a two-settlement system, it really is a multi-settlement system. There is an installed capacity requirement, and load-serving entities are required to contract with the suppliers to meet that requirement. When you become an installed capacity supplier, you take on

the obligation to bid into the ISO's day-ahead market. This allows the ISO to ensure that it has availability for tomorrow while at the same time giving the market the most flexibility in making decisions as to how that market should work. That is tracked, and when suppliers are not bidding into that market, the ISO contacts them to find out why and making sure that there are valid reasons.

The hour-ahead is not really a market. It's really a vehicle for the ISO to enter into the scheduling process with its neighbors on an hourly basis and accept new contracts.

Finally, we have the real-time market, which is based on a security-constrained economic dispatch algorithm, very similar to the PJM system.

There was mention of the need for a transitional mechanism. New York has at least in part addressed that with the installed capacity market. The ISO wasn't comfortable that all generation would bid in. And it was a reliability issue. I think the market will tell when it is no longer needed--hopefully, as generation suppliers begin to build, there will be sufficient supply and the value of installed capacity may drop.

In terms of how the markets interact, installed capacity providers and other suppliers bid into the day-ahead market. The market closes rather early, at 5:00 a.m.; there is discussion as to whether to keep it then or align it more closely with the other ISOs. By 11:00 a.m., the process produces a set of hourly clearing prices for the next day. Those represent the ISO's contracts to buy and sell energy and, if you've submitted

bilaterals a day ahead, you're locking in that congestion cost. Therefore, as long as your transaction flows in real time at the same values, you would incur no additional congestion costs.

The balancing market evaluation uses the same software platform as for day-ahead to do the unit commitment, but it's a tighter window, three hours, and starting that analysis 90 minutes ahead of the hour. If a participant wants to modify any schedules that they submitted a day ahead, they can do that 90 minutes in advance of the hour.

This all flows into the security-constrained dispatch. NYISO uses the same software for real-time dispatch as it used under power pool operation, developed in the early '80s. Having the day-ahead market there was a very important element since 95 percent of the business is being done in the day-ahead market. Most implementation problems have been with the real-time market--price spikes, some of the calculation errors, a lot of the price correction errors. Day-ahead prices have been very stable.

Scheduling day-ahead, there is bilateral transaction power; scheduling where you basically come to the ISO and say, I want to buy 1,000 MW tomorrow regardless of the price; and price cap loads, which represent additional energy that you can schedule a day ahead based on price. You'd say, well, I'll take an additional 100 as long as the price is no more than \$60.

Once you have scheduled your load day-ahead, we go into the real-time market. One of three things can happen. One,

your load could come in exactly at the value that you scheduled day-ahead. That doesn't happen often. Two, your actual load is higher than what you scheduled day-ahead, in which case you would buy the additional amount at LBMP. You could also schedule a bilateral transaction in the hourly market to cover it if you see it coming. Three, the load won't come in--you won't use as much as you said you would--in which case you simply sell it back to the ISO at the real-time price. It balances.

There are several passes in security-constrained unit commitment. First, the ISO looks at the load that people have indicated they want to buy day-ahead. It schedules the generation it needs to meet that load. It then looks at the forecast load. Generally that's somewhat more than what people wanted to schedule day-ahead. Additional generators might be committed in this pass. There is then the local reliability pass, where I might have trouble keeping the lights on in Binghamton if I don't run this particular generator because it's holding the voltage up. His bid may not have been selected in the first pass because it wasn't economic, but for reliability needs, that I need to have that generator on. Then the ISO, to determine prices, goes back and says, If we took this set of generators and dispatch that against the load that people said they wanted to buy a day-ahead, how does that work out? That calculation is where it gets the prices for the forward contracts.

Some comments about the two-settlement system, about 11 months into it. The first settlement, which is those day-ahead prices, has been reasonably stable. Parties are using those prices on



a secondary market basis to establish contracts. The second settlement, real-time, can be volatile for a lot of reasons. Real-time never turns out the way you thought it was going to day-ahead. That sends an incentive for people to schedule day-ahead. Ninety-five percent of the market is scheduling day-ahead. About 45 percent of that 95 percent is buying in at the day-ahead LBMP price.

Although there have been convergence problems, it's maybe not as bad as people say. The lines cross, and although some supply shortages made hour-ahead estimates high, that is really a function of the manner in which some of the bidding rules have been implemented.

To conclude, the two-settlement system is a key feature of the New York market, and I don't believe the market could have been started when it was without it. It has allowed the ISO to produce a reliable day-ahead unit commitment, and that process has worked very well. It provides the market-based incentives to follow the instructions of the ISO, and that's key because when you shrink the commitment window to 24 hours, all of the money is on that commitment and the system operator has to be sure that the operating plan for tomorrow is correct.

There are discussions with market participants on whether or not there should be a third settlement. That would be in the hourly area.

### **Speaker Three**

Why should FERC get into these markets? Market failures, public goods,

mitigation of market power, dealing with externalities. Reliability is our number one public good; it trumps markets. So you have to deal with that in a collective way. Market power mitigation is also a public good, meaning that everybody can participate. If someone has market power and is making a lot of money, it's only rational for that person to invest in keeping it. And it is very important to have information. In ISOs today you get real-time prices and quantities, but you may need to know more about the system in order to get better.

The markets are incomplete. One reason is a long history of management by regulators, where no one cared much about the price. There was a very passive demand side. And there's no institutional history of markets. The markets are asymmetric; if you let the suppliers bid and you put the demand in as a vertical line, that's not a very good market. There are bidding non-convexities.

Some groups support privatizing everything. Some think the electric system is the natural gas system. A group which I'm closer in affinity to is the efficiency group, which says, let's try to maximize the efficiency, not the amount of off-RTO trading. Fortunately, the law's on our side. When you look at what the law says, it's consumer protection and efficient allocation of scarce resources. That's the interpretation of what just and reasonable means: intervene to correct market failures; make sure there's choice. Very importantly, because you want complete markets, you want to accommodate the off-RTO markets via self-scheduling. I think that FERC has

spent a lot of time in its orders making sure that the off-RTO markets are accommodated.

The off-RTO markets are a different settlement system, and we want to make sure that they are part of this. The original design was for ISOs and RTOs to be market operators. They tend to be market participants. We want to eliminate the bias between being in the off-RTO and the RTO markets. We want people to choose which one to be in to fit their needs the best. But if the off-RTO markets were so good, why would we be clearing trades in the RTO markets? And why would we be limiting RTO trades? Again, the idea is to have lots of this trading.

One of the biggest arguments in California is over too much load showing up in the real-time market. It is very important that, before you start making penalties for real-time market participation, you make sure that there aren't already biases to get into the real-time market. Otherwise, it is very difficult to figure out what's happening.

I think New York didn't think they would be able to get the real-time market working without the day-ahead market. There seems to be a fear on the part of the operators of having too much unknown supply or demand showing up in real time. What the two-settlement market does is help you control the reliability of the system by keeping too much of the system from showing up accidentally. And if you get too far off from what you think your point is going to be you may have a lot of very strange problems taking place, because a lot of the AC load flow issues are very local in

nature. Again, the first thing to do is make sure you've eliminated the biases to be in the real-time market, and then you can think about incentives for driving people out of the real-time market, if you think it's important.

The day-ahead market offers some additional opportunities. FERC is looking for large RTOs, and there is at least some thought about, in the Northeast area, having a day-ahead schedule that encompasses three or four of the current ISOs. This gives you a chance to balance the system over a much larger region. It can help deal with the seams problem because you are dealing with it day-ahead instead of real-time, and that gives you some time to manage the system.

These are multi-settlement systems, and in the pre-day-ahead markets we sell FTR, a lot of different things. We also do market mitigation. There is another settlement that says if you have market power, in California we declare you an RMR unit and we give you another special contract. These are different settlement systems. Another proposal is that we deal with market power via forward contracts, options contracts or vesting contracts. It is a crude market and these long-term contracts are probably going to fall back into some kind of regulatory scheme, but that may be the best solution other than divestiture. Allow self-scheduling. FERC asked PJM to allow bilateral trades to hedge their congestion, and it did what FERC asked.

To conclude, the idea is to combine the strengths of the bilateral and the coordinated markets with governance,

and let people implement these markets as the market participants see fit. It's very important to manage the transition, and what you want to focus on is giving people choices and options, not a forced ideology.

#### **Speaker Four**

When California started in April 1998 with its restructured market, it worked pretty much as expected. There were relatively low prices, lower than New York. In the second year, these price trends continued. There was a summer price hike, but it wasn't until two plus years into the market that things really went crazy. In the two years before May 2000, there had never been an hourly price above \$250. There had never been a day above about \$90.

In May, June, July, and August 2000, prices went as high as \$750. In June, a price cap of \$500 was imposed, moved down to \$250 in August. But there were days that averaged over \$350. The average price for 90 days of the summer was almost twice the highest previous day we had ever seen. So there was a significant shock.

One of the most mysterious things is that we don't know why this is all happening. You've heard supply and demand, gas prices are up, there's less hydro, there's less imports available. But none of those explains everything that's happening. For '99, as the load grows, you get a predictable supply/demand increase in costs. In 2000, somewhat predictably, the costs are higher because the gas price has gone up. But in 2000, with demand significantly lower than in '99, prices are up. If this was supply and

demand, there would be some sort of curve instead of this shotgun pattern.

Price caps did have an impact. Prices came down 40 percent when the \$250 price cap was imposed. But there was the strange effect that prices also flattened out. I don't understand why simply being unable to charge a certain amount on peak turns into being able to charge whatever you want off-peak. The other interesting thing is that prices also came down in neighboring states. Perhaps there was some competition between California and neighboring states.

There's very little difference between what the three utilities are experiencing in terms of what they are paying for commodities. Compared with a regular supply/demand curve--where as demand goes up, the price goes up, when demand is low off-peak, the price is low, and when demand is high on peak, the price is high--there's something very different going on in the California market, because prices can be high when demand is low and low when demand is high.

There are some rules in place in California that are sort of odd. For instance, when the ISO goes to cure congestion in California, it does so through what's called a minimum shift dispatch. It makes the minimum shift in the generation that was previously bid in. It doesn't do what New York and PJM do, a security-constrained economic dispatch. That is a design pillar of California's that is inefficient and may well cause some of what's going on here. Another issue is that California has separated the power

exchange, the commodity market, from the ISO, the control of the transmission. In real time, that's not a very sensible thing. In real time, when you go to solve congestion, you need to do it economically. The separation of these two markets has created an ISO that seeks reliability without relation to cost. Its mantra is reliability at what the market wants to charge us. And that is also a significant difference from what you see in PJM and New York.

Finally, California has a zonal pricing scheme, and no zonal pricing scheme has worked yet. What it does in California is sweep up large areas and take congestion within those large areas and socialize it, as opposed to charging locational prices. It doesn't give very good locational price signals, and allows for some gaming of intra-zonal congestion.

### **Session One's Speaker One, Redux**

I'm doing some work on the relationship between the day-ahead and real-time price movements in California. We looked at the difference between the schedules as of the close of day-ahead and each increment. What we see is a lot more stability than expected. The real-time market has been running around three percent for most of the period of operation, and then as we got to the spring, it shot up. It's gone down a little since then, to about 4.3 percent in July.

An alternative way of thinking about this is the percentage of time in which what's traded in real time means a net increase in supply. So if we were thinking of the real-time market as a true imbalance,

where we're trying to schedule everything that we think is going to happen and random things cause us to need to buy extra supply because demand is lower than we want, the amount of time that we're increasing or decreasing should be around 50 percent. Early on, it was pretty high, but more recently California is consistently buying additional supply in the real-time market.

The ISO trades were hovering around zero for a lot of the period. You can see a trend more recently where most of the volume tends to be an increase in acquisition at any moment. There were a couple of spikes in the 10,000 MW range, when total load was around 40,000, so there were quite a few hours in which a quarter or more of the volume was being traded in the real-time market.

The ISO is worried that this is a threat to reliability. I think the solution would be, if there is a social cost to showing up in real time with either your load or your supply, if we can quantify that social cost, we can apply a trading charge to real-time. If you buy or sell power in the real-time energy market, you pay, say, \$5 per MWh to do that transaction.

### **Discussion**

*Question:* Is it imposition of the price cap that has provided the incentive to underschedule, or is it a problem of the California system prior to the cap?

*Response:* There is probably a structural issue such that the underscheduling is rational economic behavior. If that's a problem, the problem is in market design because it's designed to produce that

kind of behavior. You know the maximum you're going to have to pay in real time, because there are caps. So you're not going to bid anymore than that in the day-ahead, and if the day-ahead doesn't clear, it goes over to real time. There are also incentives for supply, because they can earn replacement reserve on top of their energy.

*Question:* My understanding is that, in New York, and also PJM, a load-serving entity can nominate a certain amount of load or even less than they think they're really going to need, and buy the balance in the real-time market, but that generators are required to submit matched schedules. So if they submit generation they have to have a place. That doesn't make sense.

*First Response:* Capacity resources within PJM must submit an offer into the day-ahead market. That offer can be that they are going to self-schedule in both the day-ahead and in the balancing market. They have to take a forced outage if they don't do one of those two things. A load-serving entity can either submit a fixed demand bid or put in a price-sensitive demand bid. Anyone can put in an increment offer or a decrement bid, which is a financial instrument without physical generation or load assigned to it. So I don't think one side is more leveraged than the other. It levelizes the playing field. The only requirement is the capacity resources.

*Second Response:* New York does not have the ability to utilize the financial increments and decrements. If you are an energy service company that has one MW of load that you have signed up and

are bidding into the market, there's no rule that you can't buy 10 or 100 MW day-ahead. However, entities that are not qualified to serve load cannot take that same position. The ISO recognizes that that is a problem and is developing the rules and the software specifications to make those changes. The price cap load bid--the load that parties can purchase day-ahead at a price--is also limited in that only one entity can do it at a particular bus. That needs to be available to all parties.

*Question:* Why does a multi-settlement system create a more efficient market than a single-settlement system?

*Response:* The two-settlement system, from an operations perspective, gives the right incentives. It is in the ISO's interest to know day-ahead, when it's developing its operations plan for the next day, what resources are going to be available and needed. The two-settlement system does that. It also provides a hedging mechanism for market participants who may be adverse to the volatility of real time. Some want to play in the real-time market and some would rather not. The two-settlement system allows those parties to sort that out and bid accordingly.

*Question:* In PJM over the summer, day-ahead prices averaged higher than real-time prices, yet the day-ahead load has generally been lower than the real-time prices, which suggests it's the exact opposite of California. Any reason why?

*Response:* PJM believes that supply is withholding from the day-ahead market for various reasons, whether the

transactions are deciding that they don't want to participate in the day-ahead market and just want to flow energy real time, or whether they're taking positions in other areas and waiting to see what happens to the day-ahead prices. Some generation resources perhaps are engaged in transactions where they don't want to be scheduled in the day-ahead market, so they're pricing themselves out of it.

*Question:* Do you believe that ISOs and the markets they run have failed inasmuch as prices for consumers are high? Is it the responsibility of ISOs to keep end-user prices low?

*First Response:* We have to verify that the prices we calculate are correct and consistent with the tariff we're operating under and, presumably, the rules that we have set up and the markets that we've designed were designed in a manner to ensure an efficient outcome. So we do not routinely go through and look for ways to lower high prices if those high prices are calculated correctly. We can deal with market power, but not retroactively.

*Second Response:* From FERC's perspective, the goal was to make the ISO a market operator, not a market participant. The problem is that they also have responsibility for reliability. When they have a pretty good idea that they're short so many MW, they have to do something about it because we're not going to let the system fail. I think the goal is to ultimately get the ISO out of the market and make it a market operator, not a player. In California, they seem to be getting sucked into a game where they're going to be more of

a participant.

*Question:* What is the reaction in Oregon to the California prices?

*Response:* People are pointing to California and saying, I told you so, you shouldn't have passed legislation to deregulate. It's mostly at the sweeping generalities stage than at any sort of real analysis. At the same time, we are struggling with the effect of California on prices in the Northwest. We're seeing much more of an integrated effect.

*Question:* There is a perception in the East that the PX/ISO separation in California is a big part of the problem. Do you agree?

*First Response:* I'm not sure there is any implicit or explicit value from joining the two institutions or that it would lead to a different result for the end user's bill. In fact, I'd argue the opposite, that New York has failed because they won't allow virtual bidding.

*Second Response:* I think the real problem in California is the requirement that load serving entities buy out of the PX. That has forced much of the market into the day-ahead market, which creates the volatility that is the problem.

*Comment:* Something that hasn't been discussed a lot is the function of the supply coordinator, who can be an aggregator of supply and of load and bidding into the PX and ISO markets. FERC can find that a generator doesn't have market power, but that generator may end up having his supply coordinated through a supply coordinator with someone else's supply

and someone else's load.

*Question:* What is the role of price volatility, and is it a good thing or a bad thing as we transition? That every customer has to have a regulated, fixed price alternative can't be right.

*Response:* The residential customer wants some degree of stability, but it doesn't have to be a regulated price. It's a question of the interaction between the marketer and the customer. If a marketer says, I'll pass the wholesale price on to you, that's probably not a good way to get the customer, but it's different if the marketer offers a product the customer can buy. What happened in California is that, in the way they structured stranded costs, they didn't create any head room for anybody to come into the residential market and then, when the stranded costs were out and all of a sudden the wholesale price

got passed on, there isn't anybody there who's already made the investments or set up the mechanisms to come in and respond. And then you have this immediate political response, which messes things up even more. Pennsylvania created a situation where there is room for marketers to come in and make the necessary investments. That is a better approach than California's, which now has the wholesale price pass through directly to the consumer without any retail competitors in place except for a few customers.

### **Session Three: Demand-Side Participation: An Essential Part of the Reliability Equation**

*In considering issues of electric system reliability, the focus has been on supply. However, with growing concern about shortages, demand-side participation comes into focus as an essential component of the supply-demand equation. ISO New England is operating a summer demand-side bidding program; separately, it has filed a proposal for a future, permanent program. The California ISO, just for this summer, is inviting loads to bid for non-spinning and replacement reserve (up to 400 MW) and supplemental energy (up to 1,000 MW). PJM and New York ISO have discussed, but have not taken action on, this front. Few retail customers are exposed to market prices. Large industrial customers in many states that have implemented retail competition can choose the standard offer, and are therefore insulated from market fluctuations. Clearly, there is great potential for progress in this area. What experience has there been with customers responding to spot price movements? How can demand-side participation be increased? How effective were the demand-side programs that were put in place for this summer? What about other policies that can enhance (or deter) demand-side participation? What is the best avenue for making large customers more price-sensitive? How can programs and technology geared to increasing small consumers' price*

*responsiveness move ahead?*

### **Speaker One**

The question is, how do we reconcile these two themes--harnessing market forces, while also protecting the public good? Often when the subject of the demand side comes up, the response is, Oh, that sounds like that horrible central planning, integrated resource planning stuff that we're all trying to get away from; that's why we're going to markets. I would offer the following observations.

Demand management did not cause the nuclear cost overruns of the 1970s that led to later reforms. Demand management didn't cause the high-priced QF contracts of the 1980s. Demand management didn't cause the price spikes and reliability problems that we've experienced around the country in the summers of 1999 and 2000. In other words, demand management isn't the problem. Demand management is part

of the solution to those problems. It was a decade ago, and it should be in this decade as well. We're going to do significant harm to our emerging electricity markets, as well as to the public good, if we don't figure out the best mechanisms to harness the demand side in these emerging power markets.

What do we now face? A lot of the authority that state regulators had with respect to vertically integrated franchises has been lost in the move to regional wholesale markets, which has both strengths and weaknesses. We're seeing reliability problems, market power, and price spikes across regions that makes it tough for state regulators to control. So

there's an emerging tension. Retail markets and customer relationships are state-regulated, but a lot of the problems are in wholesale markets. Problems at the retail level will make it difficult for wholesale markets to operate properly and vice versa. So we need to think about both retail and wholesale solutions at the same time.

What do the markets look like? Peak loads vary by, say, 50 percent from one week to the next over the course of a year. But spot prices vary by multiples of that. There are three public policy problems. First, price spikes, which are almost always associated with thin margins, no matter what the load is. Second, increasing market power is associated with these thin margins and the ability to run up prices when there's high demand or margins are thin for other reasons. The forced outage rate in New England jumped from 7 or 8 percent to 19 to 20 percent since the market opened in this region. Everybody is being careful not to say that's evidence of strategic withholding or other market manipulation, but there's no other conclusion to draw at this point.

Third, reliability is a problem. Across the country, we see a consistent, persistent increase in electrical demand. Sales have risen 31 percent in the past decade, and that pace is expected to continue for the next decade. Peak loads are rising even faster. As a nation, we are facing the prospect of trying to build on the generation side the entire electrical equivalent of the nations of Japan and Germany over the next 15 to 20 years. It's going to be a significant challenge if we're going to try to meet all



of our demand on the supply side.

Looking at load duration for New England, there are just a few hours at the end that are causing the problem. Over the past year and half since our market opened, one percent of the hours accounts for 16 percent of the dollars traded in the spot market. About one percent of the hours accounts for the last nine percent of the total system demand in a capacity sense and for about 16 percent of the total dollars traded. That tells us, again, something about where we might want to look in terms of demand-side responses.

Besides those one percent of the hours, we ought to remember that day in and day out efficiency of electricity is a proven resource that ought not to be shunted aside as we move to competitive electric markets. Whatever one thinks about utility, IRP programs, and even demand side management programs over the 1980s up to the mid-'90s, they were pretty successful. We saved a lot of energy under those programs at a system cost of less than three cents, and we reduced our need for new capacity by about 30,000 MW. We need to figure out a way to maintain that progress.

When we did energy efficiency in the utility IRP program, we wondered about the "no losers" test: Who was going to have to pay more on account of the fact that the utility was investing in efficiency in some customer locations? In the current market environment, it's almost the flip. Now the customer who invests in efficiency, thereby lowering the total demand in the market, lowers the market clearing price marginally for everyone else who is taking out of the spot or paying a price related to the spot.

The public value of efficiency in California over the past year or so is significantly greater than the private gain to the customers who actually invested in the efficient end-uses.

A market without a demand curve is not an efficient market. Reliability managers often see demand on any given day as fixed. So they are forced to go into the market and clear it at whatever it takes to meet that demand curve. If, however, we had a more fully functioning demand-side portion in the market, we could reveal the implicit demand curve that customers actually have and clear the market at a lower price and a lower quantity.

In terms of solutions, the multi-settlement system has numerous advantages, including creating a market where you settle the market on a forward-looking basis, which creates an entitlement that aggregators and customers can then sell back into the market to reveal their personal demand curves, ultimately giving you the potential for new businesses coming into place. Demand-side bidding is just one of many things that we need to be talking about. Congestion management pricing also has the effect of bringing more economic rationality to the market.

How do we price and pay for reliability? What's lurking in the general category of uplift charges? In most instances we create reliability charges and uplift charges to pay for wires and turbines without thinking much about it. But if someone suggests using those same mechanisms to pay for demand management measures that might achieve the same ends more cost-effectively, it's a problem.

There are at least three ways that you can bring the demand side into the market. The first is where the economists often start, simply by saying that if you have posted prices, there's a demand response. That works to a degree, but has problems. A second is market rules that allow load-serving entities to bid variable amounts of demand with a price attachment to them, a price cap for each portion of their quantity. The third is what could be called dispatchable load, where system managers are enabled through an economic system to go back into the market and dispatch load to meet reliability and clear the market, just as they can reach into the market and purchase supply-side ancillary services.

There are a lot of reasons why load is not yet responding to real-time system costs. Customers see average prices, and they see them long after consumption. It just doesn't help very much to say to a customer, Oh, by the way, we had a peak last month, it's rolled into your bill. We're not telling you what hours it was for or what you could have done about it. We have very few customers on interval or real-time meters. We have created a system of default service prices for most customers that make the demand-side markets almost unworkable for two reasons. First, when customers are given an artificially low default price, they have no incentive to manage their load in response to what's going on in the real market. Second, the way we've created default service systems in this country, we have made it very difficult for new entrants to come into those markets, capture those customers, and offer them a package that includes load

management that is an attractive package to them economically.

Further, we don't have rate designs for default service providers that would reward them if they actually went into the market and worked with customers to lower demand cost effectively. There is no incentive for wires companies to lower their customers' demands. Through-put is where they make their money. If you have something like California does, with just a pass-through of the PX rate, there is no incentive for the default service provider to do anything to manage costs in the wholesale market. A rate cap doesn't do any better. The only mechanism that would do better would be a revenue cap.

If you're in a pool that is assigning wholesale responsibilities to a load-serving entity on the basis of a set of load profiles, consider its incentives to manage load in response to real-time prices. You can buy back some expensive peak load and lower the market-clearing price for everybody in the pool or help the pool resolve itself more reliably. What do you get for it? When your cost assignment comes around from the pool or ISO for your customers during that time period, you're assigned costs according to the load profile just as though you hadn't done a thing. If you can't change the load profiles when you improve market performance, you have no incentive to do it, and you're penalized for making that change. If you can change the load profile or meter the customer whose load you affected and prove it, maybe you can get that benefit.

We repeatedly see system managers concerned about whether demand

management is real and whether it can deliver reliability benefits in real time. We have to be more creative in figuring out ways to let dispatchable load play on a level playing floor with peakers or imports to manage the system when there are reliability needs. That would include also providing dispatchable load as an ancillary service on the same terms as other reserves.

Efficiency and load management can improve reliability at all levels. Improving efficiency in load management lowers demand at the distribution and transmission levels, as well as at the generation adequacy level. When margins are thin, even small actions to manage load can have significant effects on market power, prices, and reliability. The supply curve is steep, and there are a lot of options to improve things if we can bring the demand side more fully into the picture.

## **Speaker Two**

Looking at the electricity market today, demand is completely inelastic, so there is a tremendous incentive for generators to withhold or push bids upwards.

People respond to prices. But you can only respond to prices if you see them. And right now, most people do not see prices. So real-time metering is essential in the near term. It is beginning to happen in industry, but is slow to happen in the residential sphere. Demand-side is not only one way to handle market power, it is the best, most effective way. With demand-side, as a generator, I have a lower incentive to withhold because I'm not getting as large a price increase for each MWh that I withhold.

Even a modest demand response works wonders. The market performs well much of the time, but performs poorly when we get near capacity. Having a small fraction of dispatchable load--say, five to 10 percent--would do wonders to enhance system reliability and avoid excessive prices.

What can be done? One thing is load response programs, such as the NEPOOL program implemented this summer. People sign up for so much load curtailment--200 MWh blocks at \$500, \$750 or \$1,000. If I sign up for \$1,000, if the price reaches \$1,000, I get curtailed. This is a band-aid, not the ultimate solution.

The ultimate solution is to make more load dispatchable. Last time I checked, NEPOOL had 47 MW of dispatchable load. Out of over 2,000 MW, 47 isn't going to do the job. You don't need a lot, but you need more than 47 MW.

Another solution is to introduce true retail competition. This has been decimated by poor political decisions that make it impossible for power marketers to come in and compete. The innovation we want to see on the demand side is going to come from retail competition, once we have it. There will be incentives for real-time metering and power marketers will promote smart appliances, which we already have. The technology has been around for a long time and continually becomes cheaper and better. Ultimately, everything in the house will be controllable to optimize energy consumption based on input a computer receives, the real-time price.

Some people say that what consumers want is one rate; they want to pay a

certain amount per kWh. I disagree. A lot of consumers recognize that energy price is highly volatile, and in those times would like to cut back their use. I would be happy to pay a low price most of the time, then on occasion reduce my demand. Contrast this with telephone "one rate" service, like 10 cents a minute anytime, that has become popular. The economics of telephone service are radically different: there is endless capacity, marginal cost can be zero. In electricity, marginal cost changes enormously at different times of day and in different seasons. In telephone, it is appropriate for consumers to demand simplicity because the phone company doesn't face different costs at different times of day. In electricity, the reality is that the marginal cost varies greatly. Once consumers face that, they will be responsive.

### **Speaker Three**

I propose an approach that deals with the underlying problem, not the symptom. The underlying problem is an overall low system capacity factor in the electricity business. The solution is to lower the peaks and to boost the off-peak load systematically through consistent long-term change in the system, so we wind up with an electricity system which has a much higher capacity factor. To a certain extent, we have seen that in the telecomm and airline industries as one of the causes of lowered unit cost there.

Unfortunately, in our business, the infrastructure is in relatively poor shape to try the kinds of load reshaping that we need to do. I studied an oil-fired heat pump for Volkswagen that could lower energy bills for a typical oil-fired home

about 75 percent. But we couldn't find anybody in the U.S. that could service it. There's a standing joke in our business that you can go into almost any large commercial building that has spent hundreds of thousands, perhaps millions of dollars on an energy management system, and it doesn't work. It's not calibrated. There are no trend logs. It's very difficult to get people to pay attention to technologies that are widely used in the rest of the world.

But the big issue is that customers are not, for the most part, real-time energy consumers. Many will change their load shape given appropriate signals over some period of time, but their response time is not hours or minutes, it's months, years. A typical retrofit project on a major facility is 12 to 18 months. To convert people who have that kind of time frame to responding on an hourly basis is very difficult.

There was a proceeding in New Jersey where the major electric utilities submitted a report describing how much cost-effective energy conservation was available in the state. Their view was that roughly 30 percent of the entire load could be substituted with cost-effective energy conservation.

What has been the public policy response to this? We've cut demand-side management funding dramatically during the '90s. We've shifted DSM programs from resource acquisition to market transformation programs. In some states, notably California, we have a continuous planning exercise as opposed to programs. Building codes are a problem. There is no peak load factor in building codes, so it's all based on annual energy use. It's difficult to

site any kind of combined heat and power plant. Utilities still dominate all of these programs. It's a conflict of interest; in what other industry do we ask people to take the lead in telling people not to buy their product? That's exacerbated in situations where there's stranded cost recovery based on kWh sales. Finally, there's a misunderstanding of ESCOs. ESCOs have margins that are like the construction business. They're not in a position to push innovations. They have little marketing or R&D expertise or budgets.

We've got to solve the technical problems. We need massive marketing and education. A good portion of that probably has to come out of public funding. We need to be able to do heat storage. Domestic hot water should all be done off-peak. We need electric storage space heating in the winter. We need thermal ice storage in the summer. We need building operator training for EMS systems. We need monitoring and targeting in industrial institutions.

Any good marketing company will say that marketing should be five to ten percent of your expected sales. We're looking at \$1-2 billion a year in electricity. So we see no alternative other than some sort of public mechanism for funding. There's essentially no venture capital available for ESCOs. There's zero interest in the venture community in any sort of energy service venture. It's much too complicated a business.

There is a lot of DSM available for long-term load shifting and load reshaping. We need to focus on building the system capacity factor, not on real-time pricing

mechanisms. We need a significant effort to solve these very solvable technology problems. We're probably going to need some public funding for this effort.

#### **Speaker Four**

We have the reliability curtailment programs that most utilities do. We also have an appliance cycling, air conditioning, and water heat cycling program. But those are not really based on economics--they're more on overall capacity. When we started to see the price spikes, particularly in 1999, we said, We should look at some load reduction programs different than our traditional programs.

2000 turned out to be a very mild weather year. But if you go back to 1999, you see a jump in the daily prices when the load starts to get up towards the peak. May 8 was a hot day and limited generation was available, so you see a price spike. What those curves do, and particularly what the May 8 price did, is drive the forward market up. In July and August, power was trading up in the \$125 range. You can see that it has grown steadily over the three years.

Since our customers have capped rates, we cannot impose a load reduction program on them. We said, let's try a voluntary program where we'll actually pay them instead of paying a generator to reduce power. We targeted our larger customers, those who had advanced interval metering. We set up a term of June 1 to September 30, although we got started late. We based it on economic energy. We told them right up front, we're doing this for economics, not for reliability. One significant thing is that

we were going to utilize the day-ahead bidding settlement system within PJM.

This is how we designed the program. We came up with an on-peak period from 12:00 to 8:00. One of my problems with the five by 16 block that marketers love to sell is that the 16 hours is not what I need. I really need the eight hours in the middle. We said as a starting point that we would split any savings 50/50 with the customer. We signed a contract with a customer just the way we would sign a contract with a marketer. The customer pledges an amount and we compare it a profile in order to see that he's actually reduced. There are checks for gaming and penalties if the amount of the reduction is less than 90 percent of what was pledged.

We signed up with a small consulting firm that provided us the web-based product that allowed the customers to go onto the web to see what the prices are. It also allowed us to monitor. They have a statistical program that develops the baseline and helps with the measurement.

Here's an example, from what turned out to be a cool summer. We take the forecast risk, so we forecasted a day-ahead rate of \$200 for these hours. We posted that on the website so that each of the customers that signed up could look at it. We sent that signal out by 9:00 a.m. We get a response back from the customer by 11:00 a.m. In this case, he decided to pledge a reduction for one hour. We accept the pledge and then factor it into our day-ahead settlement system with PJM by noon. By 4:00 p.m., we know what the prices are.

It turns out that our forecast wasn't very good and the price came in cheaper than what we expected. In this case we got kind of a benefit out of PJM, either in avoided spot purchases or if you want to think of this as a virtual generator. We transmitted data to the consulting firm so they could calculate the actual load reduction for the customer, which turned out to be 1.3 MW. He pledged 1.3 and delivered 1.33, so there was no penalty. So we paid the customer \$133. We probably lost money because we didn't get the revenues for that 1.33 MW.

We started the program late and found that it requires quite a bit of education, so we only got seven companies to sign up, but they were a variety of companies. We ended up with 15 MW, ranging in size. We only ended up calling it four days. We had set for ourselves an objective to only invoke this voluntary load reduction when temperatures were forecasted to be above 90 degrees and we thought PJM's load was going to hit 50,000 MW and that LMPs would be between \$150 and \$1,000.

The mechanics of the system do seem to work. Customer interest was high. I guess they liked the fact that it was voluntary as opposed to traditional curtailment programs where the utility interrupts them with two hours notice or so. In this program they got to volunteer when they were going to do it and they knew about it at 11:00 a.m. the day before.

Selling this requires face-to-face contact because it's a complicated subject. People want to see how it works and know how easy or hard it's going to be for them to make it work. A lot of these people are building managers or energy

managers. If they save a few bucks here and there but one day they really blow it, they view that as damaging to their careers. The program's penalties created some confusion. Using a third party to do the measurement was useful because they want to feel that it's not the utility controlling all of the measures.

We probably need to rethink the split of savings because when you look at it in the example, 1.3 MW, you could ask whether \$49 in an hour is worth all the effort. That's still a big question mark. Is it enough money to make it worthwhile? At least for this type of program, advanced meters and preferably ones that we can communicate with by telephone are preferred. We were a little concerned when PJM came out with their program; how do we make sure we don't get people who signed up for both? We think that some real-time customer energy analysis system and total energy solution might be a better approach. And try to make it simpler. I think we also need to consider whether customers have a preference for day ahead versus real-time. Is the day ahead where we want these kinds of programs to be? Do we want to also consider options where people look at real-time prices and perhaps only reduce when they see real-time prices in the \$500 to \$1000 range?

## **Discussion**

*Comment:* Two thoughts. First, we should recognize that customers really aren't that interested in playing in the real-time market. That means that we need to create wholesale markets and public policy rules at the retail level that permit energy service companies and load-serving entities to operate in the

real-time markets and make money by shifting load around within their customer base. Second, I want to draw attention to a point that Speaker Four made about how it looked to him like his company was going to lose money on their load management operation because as a wires company under the current default service pricing rules, they were going to lose on the lost sales. That ought to ring some warning bells. When we're counting on the wires company to help manage load for the benefit of the entire system, but the signal we give it is, if you do this well you lose money, we are headed for a crater on that one.

*First Response:* There is some evidence that people want to play around in the real-time market, not in a hands-on way but in an automated way. The very successful air-conditioning programs show that consumers, if they reap some of the benefit of the cost savings, will participate. For a large power marketer, the cost savings could be dramatic if some segment of their load was willing to be price-responsive. And the cost of the necessary device declines exponentially with time and with the size of the market. Something that is \$2,500 could be \$2.50 if you give it five years and enough demand.

*Second Response:* Our experience is that residential consumers in particular are not as price-sensitive as you would expect. A \$5/month savings did not get a lot of people to switch energy suppliers. There's a lot of inertia in customer choice.

*Question:* In terms of what consumers want, my experience is that different sets of consumers on the commercial side are

interested in different issues. For some, it's a question of reliability. For others, it's not seeing fluctuations in price, because they have to budget. What these programs do is enable different consumers with different needs to get some of their needs met. An interesting area is hospitals, where they've been very focused on reliability, and suddenly they have an opportunity to use their on-site generation. Do you see a distinction between small-scale on-site generation and general load reduction programs? Are there barriers that need to be addressed on the distribution level, particularly for on-site generation? Are there any environmental issues that will come up as we see more on-site generation?

*First Response:* As to the fact that different customers want different things, we all need to understand that better. I hope as we actuate the demand side we will find different customer niches developing. As for distributed generation, we need a market in which there's a neutral trading floor, so that whether someone is bringing generation or demand reduction to the market, we are operating by the same rules. There is the associate pollution burden with on-site generation. I would recommend that a standard emissions profile apply to distributed generation on an output basis so that they can all compete freely.

*Second Response:* The example of load reduction versus distributed generation is an illustration of the market really being niche markets. Some people are very interested in distributed generation. Others wouldn't touch it. We need to understand who those people are and why. There is typically a proceeding in each state examining the technical and

economic issues of distributed generation. People in the policy world need to understand that those kinds of proceedings are very difficult for ESCOs to participate in. They're expensive and time-consuming, and ESCOs tend to have limited staff. So there's a real imbalance. The environmental issues are significant. These are diesel engines that would be running on the days when the air quality is already the worst.

*Question:* In a lot of states that have restructured, there are legislatively imposed price caps or rate caps. As fuel and other costs go up, a lot of costs are being deferred. What would the default suppliers do if they weren't allowed to defer those costs that would otherwise take them above the price cap? Would they move more aggressively in DSM? And what is the relationship between these rather arbitrary price caps and how the default supplier plays in the marketplace?

*First Response:* There are three dominant options. There's the straight pass-through, like California. Then there's the price cap option and maybe the pass-through and deferral. The third option would be something like a revenue cap on a revenue per customer basis, or a different kind of performance-based regulation for the default service provider. They have different effects on the companies' attitudes towards energy efficiency and cost-effective load management. A wires company that is making money from additional sales, but that loses money in high-cost hours, will be interested in changing the load curve and spreading consumption to off-peak hours. A company that can pass through both doesn't care about either load management or energy efficiency. And



a company on a revenue cap would care about investing in both energy efficiency and cost-effective load management.

*Second Response:* The ability of utilities to respond in any creative way has been hindered by the history and current practice of demand-side management. One, demand-side management and energy conservation have since the beginning been mandated, so utilities tend not to view it as a voluntary part of their business. Two, if you look at the system benefit charge programs that have been instituted as part of deregulation, they are, if anything, more tightly constrained and regulated than previous programs. So we haven't yet seen what utilities would do in a business sense on these issues, because they don't view it as a business.

*Comment:* I'm part of a company that serves aggregation load; our supplier is Enron. We have about 350 MW of customers, of whom about 50 to 70 MW are contracted with Enron on an interruptible basis, created solely to deal with the problems that were anticipated this summer. On the infamous May 8, we shut down an entire mill for the bulk of a day to take advantage of that rate. I think that kind of solution has a lot more probability of success than does the ISO's interruptible rate solution, because the company trusted us. Customers don't want to shop for electricity? Our national sport is shopping. Our kids grow up in malls. What's getting in the way when it comes to electricity? We have conditioned behavior on the part of consumers over a long period of time. And we have to change that. The problem is that we have a centralized planning system that the customers know has not gone away. They don't

trust us. We've done a great job for utilities. We need to start worrying about consumers.

*Question:* You need metering and you need an information system--a settlement system and an infrastructure--that allows this data to pass back and forth. What will get people interested in the metering and in paying for the data systems?

*First Response:* Perhaps those investments ought to be supported by the system benefit charges or the uplift charges that the pools are paying for, in order to accelerate the deployment of meters and metering systems that would get us these more robust markets.

*Second Response:* I would take the approach that there is a subset of the customer base that's interested in this kind of stuff, and I would work on trying to identify what that subset is and how to get at them. One of the difficulties with many utilities is they don't have a good understanding of their customers in terms of what motivates them to take actions.

*Comment:* Instead of trying to figure out how utilities should do rate design or analyze their customers, why not push for competitive markets and allow players like marketers to make the right decisions, given their experience, information and capital.

*Question:* You mentioned that we need public funding. For what and to whom?

*First Response:* If we try to reshape loads on an aggregate basis and move to a squarer or rectangular load shape, a tremendous public education campaign

has to be undertaken. The industry won't be motivated to spend huge amounts of marketing dollars to do it. The UK has a carbon tax that's backed by a huge government-sponsored marketing campaign, where they're systematically trying to change behavior in the energy field.

*Second Response:* The system benefit charges that most states have adopted, in part to support energy efficiency, are well-supported by the fact that there are significant market barriers to--and significant public goods benefits from--the deployment of energy efficiency. Given that the incentives utilities face don't support success in delivering energy efficiency, we should be careful about who gets the money. A whole host of charges are collected at the pool or RTO level under various terms, sometimes called uplift, sometimes a reliability investment. We ought to be looking carefully at those investments to make sure we're making them on a least-cost basis. That might include replacing some reliability-motivated investments on the supply side or for wires with investments on the demand side.