FERC STAFF PAPER ON
REGIONAL CHOICES FOR IMPLEMENTING
THE ELEMENTS OF THE WHITE PAPER
July 7, 2003

The White Paper sets out eight proposed elements of a Wholesale Market Platform. It offers options for regional choice regarding how to implement these elements. This paper summarizes the discretion offered in the White Paper for each element of the wholesale market platform, such as pricing of energy and ancillary services and allocation of transmission right.

The various elements of a regional market should work well together to produce an efficient, well-functioning wholesale market for the benefit of customers. A principal purpose of this paper is to describe how a regional choice for one element of a market design can affect or even limit the region's choices for another element. For example, there are important interrelationships between such wholesale market elements as the energy market design and the system for congestion management, and between the resource adequacy provisions and means for mitigating market power.

The views presented in the paper are based on experience with various electric market designs, mostly in the United States but also in other countries.

The eight elements of the Wholesale Market Platform are:

Element 1. Regional Independent Grid Operator

The White Paper would provide that independence of governance will be decided case by case. An RTO must meet the Order No. 2000 standard for Scope, but an ISO does not have to do so.

Although the White Paper offers regional flexibility on governance structure and on determination of scope, there is no independence or scope option presented in the White Paper. Thus, there are no specific options to be described for states or others to consider.

Element 2. Regional Transmission Planning Process

The White Paper would require each RTO or ISO to have a regional planning process. The process itself and the roles of various participants will be decided regionally.

Although the White Paper offers regional flexibility on the regional transmission planning process, there is no independence or scope option presented in the White Paper.

Thus, there are no specific options to be described for states to consider.

Element 3. Fair Transmission Cost Allocation

The White Paper would require each RTO or ISO to allocate the costs of existing and new transmission fairly. It would allow a variety of options for the recovery of existing transmission costs, the recovery of the costs of new transmission, and the recovery of the costs of transmission through more than one RTO or ISO. These three topics are taken up in turn.

Existing transmission cost recovery options

Under the White Paper approach, the costs of existing transmission (except for costs directly assigned) will continue to be recovered from rates paid by customers. The Regional State Committee (RSC) will decide whether there should be a uniform rate for transmission service throughout the region, called a postage stamp rate. Alternatively, the RSC may decide to maintain a different rate for each utility service area, called a license plate rate, which nevertheless allows customers in each area to purchase power from anywhere in the regions without paying multiple access charges.

If the objective is good economic pricing of transmission service, either a postage stamp rate or a license plate rate can be used because either type of rate collects existing capital costs (so-called "sunk" costs) separately from the cost of a particular transmission transaction. Many parties favor license plate rates, at least initially, because they avoid "transmission cost shifting," that is, raising transmission rates for some customers and lowering them for others in moving to a single regional average price for the access charge. Those who favor a single postage stamp rate often will accept an extended phase-in period to avoid a large shift in costs at any one time; they say that a postage stamp rate is, or at least appears to be, more equitable because all customers in the region pay the same rate for transmission service throughout one region.

New transmission cost recovery options

According to the White Paper, the costs of new transmission would be recovered in accordance with the regionally determined pricing policy, which will be informed by an RSC.

The two principal choices are rolled-in pricing and participant funding. Rolled-in pricing typically means adding ("rolling in") the cost of a new transmission facility into the total cost of the transmission facilities of an area, so that the cost of the new facility is included in that area's average rate paid by all its transmission customers. Under participant funding, which has traditionally been called direct assignment to a particular customer, the cost of the new facility would be borne only by the customer who agrees to pay for (or who benefits from) the new facility. (In fact, the "agrees to pay for" criterion and the "benefits from" criterion are somewhat different versions of participant funding with somewhat different advantages and disadvantages.)

Rolled-in pricing for new transmission costs may be applied to an entire RTO or ISO area, or to a single utility's license plate rate. For example, if a new transmission facility is needed to interconnect a new independent power producer to the transmission system, under rolled-in pricing all transmission customers of the transmission provider would pay a revised average rate, and the transmission provider may be a single utility or an RTO or ISO.

The new term, participant funding, may have different meanings in different contexts. For example, in an RTO or ISO some would refer to the direct assignment to one utility of the costs of a new transmission facility that benefits only that utility as participant funding, even though these directly assigned costs are then rolled into the rates of that utility's customers.

Most regions are likely to want a combination of rolled in pricing and participant funding. Typically a region may want to roll in backbone grid improvements needed for reliability because such improvements are essential and clearly benefit all parties, and to use participant funding for facilities that benefit only one party. A practical difficulty is that many grid enhancements do not fall clearly into one of these two categories. A grid enhancement may not only provide reliability benefits to all customers (by making the grid less vulnerable to the loss of a transmission facility and by providing more pathways for backup power to reach customers) but also provide specific economic benefits to some particular customers, not just to those customers willing to pay for the enhancements.

One advantage of rolling in all new transmission is that these difficult benefit determinations are avoided; even if a particular enhancement benefits some customers more than others, the benefits will average out when enhancements in various locations are rolled in over time. In addition, rolled-in pricing facilitates the development of needed transmission infrastructure development because the region does not have to determine a beneficiary or beneficiaries who are willing to sponsor each individual enhancement.

However, rolled-in pricing has potential disadvantages also. Rolled-in pricing may make new generators insensitive to where they locate because they themselves do not have to pay the full cost of the transmission upgrades needed for the new power to reach customers. There is a concern in particular with grid upgrades needed for new generation if that generation is expected to sell out of the region or if it is uncertain whether the generator will stay in business long enough to pay off its share of a new transmission facility that recovers its cost slowly over three or four decades. Further, the rolled-in facility may face siting difficulties unless local siting authorities find that the reliability and economic benefits to local customers who help pay for the facility outweigh its costs.

Participant funding offers the advantages of protecting customers that do not expect to benefit from a particular enhancement from having to pay for it, ensuring that each grid addition passes its own benefit-cost test as evidenced by a specific customer who is willing to pay for it, and avoiding building transmission that is not economically justified. However, some parties are concerned that many grid improvements needed to provide a stronger platform for wholesale competition simply will not be built if there is extensive reliance on participant funding.

Differing regional views on the appropriate rate incentives for building new transmission and different regional approaches for determining the beneficiaries of new facilities can result in each region having a different mix of rolled-in and participant funding for new transmission facilities. Because of the difficulty and inevitable subjectivity of determining how various market participants benefit from a particular new facility, the Commission, while allowing for regional differences, would require that this determination be made by an entity that is independent of market participants.

Options for recovering transmission costs across RTOs

Under the White Paper, RTOs and ISOs should work to eliminate the payment of multiple charges for transmission through more than one RTO or ISO, unless an RTO has a notable imbalance between imports and exports, in which case the exporter may apply for an export rate. An option for the RTO or ISO in a region is to determine whether to work out with its neighbors an agreement for reciprocal waiver of access charges or whether to negotiate an export rate agreement. Reciprocal waiver of access charges gives customers in all the participating regions a wider range of supply choices, reduces market concentration, and promotes wholesale competition. An export fee, however, recovers the cost of transmission that is constructed to export to customers in another region but that is not used much for imports to serve local customers who may otherwise have to pay for that transmission.

Linkages among these cost allocation policies

Decision makers should be aware of the linkages among these transmission cost allocation policies. For example, the policy chosen for recovering the cost of existing facilities can affect the policy choice for recovering the cost of new facilities, as these examples indicate:

- *Postage stamp with rolled-in*. A postage stamp rate for existing transmission fits well with rolled-in pricing for new transmission because the first assumes common regional benefits for existing facilities and the second assumes common regional benefits for new additions.

- *License plate with participant funding*. Similarly, license plate pricing fits well with participant funding because both assume that subregional rates or rates for specific customers should reflect the costs—whether historical or new—that are incurred to meet the specific needs of each subregion or customer.

- *Postage stamp with participant funding*. Mixing a postage stamp rate for existing facilities with participant funding for new facilities is possible. However, it means that each service area has the same "base average postage stamp rate" with a different surcharge for each area's participant funded upgrades. Over time, the accumulation of different surcharges in different service areas could become the equivalent of different license plate rates.

- *License plate with rolled-in*. It is also possible to mix a license plate rate policy with a regional roll-in of the costs of new facilities. The result is to have a common surcharge on each area's license plate rate, and as older facilities are retired from rate base over time, every area in the region would gradually come to have the same postage stamp rate. This could provide a natural, though prolonged, transition from license plate to postage stamp rates.

- *Export fee with postage stamp*. Similarly, an export fee paid by an importing RTO or ISO is more readily spread to all customers in the RTO or ISO if it has a postage stamp rate.

- *Export fee with license plate.* An export fee that is directly assigned to just one utility in the RTO or ISO because that utility is the only explicit importer is more readily accommodated with a license plate rate.

Further, there appears to be an important linkage between these policies and the Commission's preference for each state to determine the transmission rate component used for bundled retail service. If a region relies exclusively on license plate rates for existing transmission and participant funding for new transmission, then there is no difficulty. However, suppose the RSC can decide to have a postage stamp rate for the region and assume this is decided by a super-majority vote over the objections of one or a few states with costs that would increase. Then the authority of the RSC to make this decision appears to be at odds with the authority of the individual state to determine its own bundled retail transmission rate. The same apparent conflict between state and RSC responsibility occurs if the region rolls in even some new facility costs to the rates of all customers in the region, or if the region spreads among all customers in the region the payment of an export fee charged by a neighboring region.

Element 4. Market Power Mitigation and Monitoring

An essential element of a competitive wholesale electric power market, as set out in the White Paper, is the provision for market power mitigation and market monitoring. The White Paper aims to protect customers against high prices that come from the exercise of market power while not suppressing market prices below the level necessary to attract new investment in efficient generation, transmission expansion, and demand response. The type of mitigation measure chosen, and in particular the level of any price cap or bid cap, can affect the region's approach to assuring that there are adequate resources. Further, the choice of market power mitigation measure may affect or be affected by whether the region has a day-ahead market for energy, as explained further below.

Three terms should be explained by way of background. Market power is the ability to raise prices above competitive levels by withholding generating capacity from the market, whether by "physical" withholding or "economic" withholding of generating capacity. Physical withholding means that a supplier says that its generator is not available to run or that it otherwise does not make the generator's capacity physically available to the market for reasons other than those approved by the RTO or ISO—such as an approved planned outage or an environmental restriction. Economic withholding means that a supplier intentionally bids too high to avoid being selected to run. There may be circumstances in which a generator legitimately submits a high bid for scheduling purposes, and distinguishing such a bid from the exercise of market power is a job for the independent market monitor. One rule adopted in several regions is that withholding that does not affect the market price is not subject to mitigation; here, the distinction is between conduct (bidding too high) and impact (affecting the market).

The White Paper stresses that the Commission will consider alternative methods of market power mitigation, subject to the assessment that they are effective and will not create additional problems for the market, especially at the boundaries with neighboring RTO markets. Under the detailed rules proposed in the original SMD NOPR, which follow recent Commission precedent, the RTO or ISO must address economic withholding in the spot markets by implementing market rules that act to limit spot market bids, as necessary, before settlement prices take effect. This is preferable to letting spot markets work initially without restraint, then trying to detect bad behavior after the fact, and later making amends by trying to determine what the correct market prices should have been.

Before-the-fact market power mitigation in spot markets may consist of either or both of:

(1) **Bid caps, bid thresholds, contracts, or other limits that are for specific generators, locations, or conditions.** A bid cap is a maximum price that a generator may bid that is set in advance. In contrast, a bid threshold is the price of a bid that triggers application of a conduct or impact test to determine whether to limit the price bid; the threshold price can vary or apply only under certain conditions. Typically, the bid thresholds that apply inside a "load pocket," an area in which import capacity is very limited, are more stringent than those outside a load pocket. The approach to before-the-fact mitigation in each region is somewhat different, particularly for generators that must run for reliability reasons. New York and New England have automated mitigation procedures (called AMP) that substitute a so-called reference price for a generator's bid if the bid exceeds a threshold and also fails a conduct and market impact test. In PJM, a generator that is dispatched out of merit order because it must run for reliability reasons is subject to a bid cap of marginal cost plus 10 percent. In other regions, reliability "must run" (RMR) contracts with some needed generators set predetermined prices for those generators that provide for recovery of their fixed costs.

(2) A **safety net bid cap** is a bid cap that applies to all generators at all times.

For a spot market design in which the various elements work well together to benefit customers, the choice of whether to implement one or both of these types of bid cap measures and the stringency of the bid caps in particular locations can be affected by the region's approach to assuring that there are adequate resources, as discussed below. Also, the choice of market power mitigation measure may be affected by whether a region has a day-ahead market for energy, also discussed below.

In addition, in some regions the Commission has approved other market mitigation measures that are not mentioned in the White Paper. For example, in a spot market design that allows a generator to make a bid to recover in the energy market its cost of starting up, which is separate from its hourly cost of running, there can be a restriction on how often a generator can change its start-up costs. Also, where a generator is allowed to state that, if it is called on at all, it must run at least a specified number of consecutive hours, there can be a restriction on how often the generator can change this minimum run time. Those designing spot markets should also consider how these mitigation measures fit with the other elements of the market design.

Design of Bid Thresholds and Bid Caps

Two major spot market design decisions are:

(i) The geographical coverage; that is, should generators in all areas or only some areas be subject to a bid cap?

(ii) The conditions under which the bid caps or thresholds are operative and the level at which they are set.

Various combinations of geographical coverage are possible. Several combinations have already been approved for spot markets by the Commission for some regions:

(1) **Safety net bid cap and a bid threshold for *all* generators**. This approach typically would substitute a reference price for a generator's bid when (a) the market clearing price reaches a critical level and (b) a generator's bid violates the conduct and market impact tests. A reference price is established for each generator as its expected bid under competitive conditions. It could be based on prior bidding history, cost data, or consultation. A threshold bid price is set for each generator at either a fixed percentage or a fixed dollar amount above that reference price. These threshold prices are set for all generators throughout the RTO region. Under this approach, any generator's bids will be subject to a conduct and market impact test when the market clearing price reaches a critical level, with more restrictive thresholds applying in the load pockets. The safety net bid cap, even if lower than the threshold price, is the highest price that a generator is allowed to bid. For example, using a New York-style AMP approach, a generation owner may be allowed to make very high bids from some of its generators located outside load pockets—reference price plus 100 percent or reference price plus $300/MWh—before it is subject to a conduct and market impact test. However, if there is also a safety net bid cap such as $1,000 per MWh (the current safety net bid cap in the northeast RTO or ISO markets), no bid could exceed $1,000 even if thresholds were higher. This type of full coverage of the market is currently used in New York when market clearing prices reach $150 per megawatt-hour.

(2) **Safety net bid cap for *all* generators and a bid threshold for *some* generators.** In this option, the safety net bid cap is again applicable to all generators throughout the region, but the generator-specific threshold mitigation method is limited to generators in certain locations. In some regions, tests for market power apply only to a generator located in an area designated as a "load pocket." This is an area found to be highly susceptible to market power, generally an area with few suppliers of generation and limited transmission to import power from suppliers outside the area. A

safety net bid cap applies throughout the region, both inside and outside the load pockets. This combination of market power mitigation measures is being applied in New England. PJM caps bids of certain generators at marginal cost plus ten percent if certain transmission constraints require that they be dispatched out of merit order for reliability reasons.

(3) **No safety net bid cap anywhere in the region and a bid threshold for *some* generators.** This combination of market power mitigation measures is the same as for option (2) above but with no safety-net bid cap. MISO has proposed this type of option, but the details are still under development.

(4) **Safety net bid cap only.** As long as a bid does not exceed the safety-net bid cap, there are no restrictions on a generator's bid. This approach has not been used in the United States. There is no market yet reviewed by the Commission that does not have at least some generators with persistent locational market power.

The level of the safety net bid cap varies from one RTO or ISO region to another, reflecting regional differences in market conditions. For example, each Northeast ISO or RTO has a safety net bid cap of $1,000 per MWh, while the California ISO currently has a more stringent safety net bid cap of $250 per MWh.

In load pockets, regions use various thresholds to determine when a bid must be evaluated for its impact on spot market prices. For example, a bid is evaluated for market power if the

- bid exceeds 2% of Average Area Price × (8,760 ÷ number of constrained hours) (New York ISO),

- bid is out of merit order and needed to relieve a reliability constraint (PJM RTO),

- bid for high cost, seldom run units in designated congestion areas is above a certain level (such as the proxy price of a new gas combustion turbine) (New England ISO), or

- bid exceeds the sum of a reference bid and the capital cost adder for a new peaker (Midwest ISO).

Both the coverage and stringency of market power tests are important market design decisions. They affect the incentives facing investors in new infrastructure for generation, transmission and demand-side response. They also affect revenue recovery by most existing generators. These market design components are generally decided through assessment of the following factors:

(1) **The level of market concentration in the region under different system conditions**. In general, the more concentrated the market or the greater the likelihood of extreme scarcity, the more restrictive market power tests are likely to be. However, in some areas such as New England market power tests for selected generators in load pockets are temporarily looser, allowing higher threshold prices, to encourage new generators to locate in the load pockets or to address concerns about the market's limited ability to price scarcity.

(2) **The interaction of energy market mitigation and a resource adequacy provision**. Depending on how they are designed, the coverage and stringency of mitigation measures may affect recovery of investment for existing and new generators and therefore affect the entry of new resources with an adequacy obligation. A resource adequacy provision and its level of reserve adequacy also affect investment recovery and the entry of resources with an adequacy obligation. How safety net bid caps and a resource adequacy provision—or lack of one—act together to affect the incentives or disincentives for new investment should be carefully designed.

Mitigation with an Administratively Set Spot Market Price

One of the central problems in market design is whether and how, in the absence of market power, to allow prices to rise sufficiently to reduce demand when supplies are scarce. There are several ways of addressing this in overall market design.

Currently, certain generators may bid high for some of their capacity—typically generating capacity with a high risk of equipment failure that might be used only when supply is very tight. However, the market price set by such bids may be unrelated to the price needed to reflect the level of scarcity and to reduce demand to the level of supply. Although this helps rarely-used generators to recover some of their fixed costs, a shortage of reserves can occur despite the higher prices.

Another method is to allow a high spot market price when supplies are low, even through there are constraints on bids. Although market power mitigation remains in place to block economic withholding through high bids, the highest mitigated bid does not have to set the market clearing price. The RTO or ISO could administratively set a market clearing price that increases as demand tends to outstrip available supply. Both New York and New England have proposed so-called "scarcity pricing" to adjust market prices upward to better reflect scarcity when demand response programs are triggered or when energy must be produced from reserves. While these "scarcity prices" would be set administratively, they should be determined through a market analysis, for example by establishing a hypothetical demand curve based on prices that reflect the values that customers place on avoiding loss of electric service. In addition, a measure referred to as a "demand curve for capacity" may provide higher or lower prices for capacity that satisfies a need for long-term planning reserves; it would compensate generator owners satisfying this reliability need at a level that accurately reflects supply and demand conditions for such capacity. Reasonable capacity prices provide the extra revenues needed by some generators to remain in or enter the market.

Scarcity pricing and a demand curve for capacity help support resource adequacy provisions while market power mitigation measures help assure competitive bidding behavior. Each could play an important role in a well-designed wholesale electricity market.

Detractors of scarcity pricing and a demand curve for capacity may see high prices during a shortage as “price gouging,” or they may argue that an administratively set price is not the same as a true market price. Further, although mitigation may limit economic withholding under this approach, detecting physical withholding to bring about a shortage and high prices could remain difficult.

Mitigation With and Without a Day-Ahead Market

The mitigation measures proposed in the SMD NOPR act through the day-ahead market. As discussed in the next section, although the White Paper's Wholesale Market Platform would require the RTO or ISO to establish a day-ahead spot market eventually, it does not have to begin with one if the market is not ready for it. As a result, the mitigation measures of the SMD NOPR may not be usable, unless modified appropriately, in a market that starts up without a day-ahead market. A good overall market design—whether for initial or eventual use—should have market power mitigation provisions and some sort of day-ahead processes that are compatible.

Each type of market power mitigation measure discussed above can be designed to be compatible with the various spot market design options, which are discussed below, including those designs without a day-ahead market. But the interaction between these Wholesale Market Platform elements must be understood and made complementary to avoid a bad market design outcome that could disadvantage customers.

Although a region may start with a real-time market without a day-ahead market, the RTO or ISO will have some process that precedes the real-time market to ensure reliability in the real-time market. Absent a day-ahead market, this would mean either that most supply and load must be matched up ahead of time in a known way such as through a so-called "balanced schedule" requirement, or that the RTO or ISO would establish a pre-real-time scheduling process.

Both physical and economic withholding are easier to control if there is a day-ahead market. With only a real-time market and no pre-real-time process for reliability assurance, the RTO or ISO must identify physical withholding in real time, and it is difficult for the RTO or ISO to detect and react to physical withholding within the short time available before the real-time market clears. If capacity is tight, this may give the RTO or ISO little time to find alternative supply. It may also force the RTO or ISO to call on any available units to run with little warning, possibly increasing generator start-up costs and making it difficult to recover such start-up costs efficiently. Further, the RTO or ISO is more likely to have to deal with the market price consequences of physical withholding after the fact, creating difficulties regarding what the market prices would have been without such withholding.

A day-ahead market reduces most of these problems by allowing the RTO or ISO to determine in advance if there is supply that is not made available to the market. Efforts to react to physical withholding are most effective when there is a day-ahead market and, somewhat less effective, when there is at least a day-ahead submission of real-time market bids.

Market Monitoring

The White Paper would require the RTO or ISO to have an independent market monitor to monitor the markets operated by the RTO or ISO. The options and tools for market monitoring would have to be designed to suit the markets and market elements for that region. In particular, if a region designs and the Commission approves a combination of wholesale market features that is somewhat different from any previously experienced, the market monitor should analyze carefully the interaction of market elements for problems not experienced before.

Element 5. Spot Energy and Ancillary Service Markets

Under the White Paper's Wholesale Market Platform, most energy would continue to be bought and sold under long-term arrangements; however, the White Paper would require that the RTO or ISO make available to customers spot markets for residual energy sales and purchases and for ancillary services. The Commission has approved various combinations of design features for energy and ancillary service markets in RTOs and ISOs. RTOs, ISOs, and the Commission have learned from experience about the market pricing problems associated with some poor combinations of design features. (A discussion of some such poor combinations is in Appendix C of the SMD NOPR.) Here, we examine some possible advantages and disadvantages of various design decisions in view of the regional market design discretion allowed in the White Paper. Although an RTO or ISO has discretion to propose spot market features designed to suit its region, such market design and the interaction of its features must still be reviewed and approved by the Commission.

Real-Time Spot Market

The RTO or ISO must conduct a bid-based, real-time spot market for energy, sometimes called a market for balancing energy or a balancing market. Real-time here refers to the actual operating day, so that real time decisions are decisions about market operations and reliability made in the same day (typically from one midnight to the next midnight) as the hour or set of hours in which they will be implemented, as opposed decisions made to be implemented the next day or longer out. Bids must be submitted at a reasonable time prior to each real-time hour in question and can be submitted for multiple hours during that day. Although not required, the RTO or ISO may allow start-up bids and no-load bids in its design in addition to energy bids.

Many characteristics of the real-time market depend in important ways on the features of the day-ahead market, day-ahead procedures, or other prior-to-real-time process that precedes the real-time market. For this reason, it is better to begin with consideration of these prior-to-real-time options.

Prior Day Procedures

Prior to real time, typically on the prior day, the RTO or ISO has to follow some reliability procedures—with or without a formal bid-based day-ahead market—to ensure that sufficient supply will be available in real-time to meet the forecast load and to provide resources for operating reserves and for real-time energy balancing. Under the day-ahead procedures, the RTO or ISO does the following:

(a) establish a next-day forecast of load for each hour,

(b) evaluate the sufficiency of scheduled energy and bid-in energy a day ahead to meet the forecast load, and elicit additional supply if necessary, and

(c) establish a next-day schedule with consideration of transmission constraints as well as any generator constraints based on generator start-up costs or minimum run times.

Experience with markets suggests the following choices for day-ahead procedures:

(1) The RTO or ISO could require that bids and schedules for the real-time market be submitted a day ahead. The RTO or ISO then uses this information to establish a provisional schedule for the next day. Although this schedule is based on day-ahead price bids, it is not financially binding on the parties; that is, the parties in real time may deviate from this schedule without financial consequences. This approach was used in PJM and New England for a period.

(2) The RTO or ISO could require that each buyer in the market enter real time with a balanced schedule, and limit deviations from that schedule to some percentage (often 5%). Deviations are priced in the real-time balancing market. Penalties are applied for deviations beyond the percentage allowed. Versions of this were used in California and ERCOT (and in the England-Wales market).

(3) The RTO or ISO can establish a voluntary day-ahead, bid-based market. No one would be required to bid into or buy from this market. As in option (1), bids for this market would typically be submitted in the morning of the prior day, and the resulting schedule is released that afternoon. However, bids accepted in this market are financially binding at day-ahead prices. Parties may deviate from the day-ahead schedule, but in this case there is a financial consequence: deviations are paid for at the price set in the real time market. This approach is used in New England, New York and PJM and proposed for MISO.

Advantages and Disadvantages of the Spot Market Choices

Having only a real-time market with either option (1) or (2) above for the day-ahead procedures has advantages (+) and disadvantages (-):

(+) Having only a real-time market costs somewhat less to implement than having both real-time and day-ahead markets because of the extra costs of software and the costs of financially settling two markets

(-) Having only a real-time market does not let transmission users that are not hedged with an FTR "lock-in" congestion charges prior to real time. This was a concern expressed by traders in PJM when it had only a real-time market with LMP.

If there is only a real-time market that is combined with the option (1) day-ahead procedures,

(-) Suppliers are less likely to follow the day-ahead schedule because there is no financial consequence for deviating from the provisional schedule.

If there is only a real-time market that is combined with option (2) day-ahead procedures,

(-) Balanced schedules with penalties for deviations may result in significant penalty charges to transmission customers because an inaccurate weather forecast or other unanticipated or uncontrollable factor may produce significant deviations from the schedule set in the prior day.

(-) The balanced schedules may not be feasible. In the former California market, the balanced day-ahead schedules were often not feasible because transmission constraints could not be adequately considered on the prior day and because the start-up costs of generators were not permitted to be expressed in the day-ahead bids. This required the ISO to compensate for the infeasible day-ahead schedules by purchasing large amounts of real-time balancing energy at the last minute at high cost, spreading these costs to all grid users.

A voluntary, bid-based, day-ahead market has the advantages on ensuring reliability without the problems listed above. However, it is more costly to implement. PJM and New England started without a day-ahead market and later added a day-ahead market. ERCOT is evaluating having a day-ahead market. Some may assert that "going up the learning curve" may be a good reason for starting with only a real-time market. That is a permissible choice under the White Paper. Market designer should be aware, however, that there may be a cost to customers associated with the problems above during the learning period in each region.

Element 6. Efficient Grid Congestion Management

The White Paper would require each RTO and ISO to manage transmission congestion with an approach that protects against market manipulation, uses the grid efficiently, and promotes the use of lowest cost generation. Several ways of managing congestion could be compatible with this element of the Wholesale Market Platform, including:

(1) **Locational Marginal Pricing**. Under locational marginal pricing, or LMP, the RTO or ISO uses price bids to calculate the lowest cost way of meeting

an increase in load at each location on the high voltage network, taking transmission limits into account. The Commission prefers use of LMP because it is an efficient congestion pricing system that results in the lowest cost to customers, given grid constraints. The RTOs and ISOs in the Eastern Interconnection use or are proposing to use LMP, and two of the three RTOs forming in the Western Interconnection have proposed to use some form of locational pricing to manage congestion in real time. ERCOT is currently considering the use of LMP.

In implementing LMP according to the White Paper, a region could choose to charge the transmission customer either the marginal cost of line losses or the average cost. Use of marginal losses conveys economically correct information about the marginal cost of each transaction, but it tends to overcollect revenues for losses. The RTO or ISO must then devise a system for disbursing these revenues that does not undermine the price signals that result in use of the lowest cost generation. Use of average losses lacks both the advantages and disadvantages of marginal losses. The New York and New England ISOs use marginal losses, and MISO plans to use marginal losses. Most others use average losses, but some are considering switching to marginal losses. A decision to use average losses requires further choices. For example, should the charge for losses depend on distance? Should counterflow transactions that actually reduce the total energy lost in transmission pay for losses at the average rate, or should such transactions be paid for their reduction of the amount of energy lost?

(2) **Market Redispatch Program**. Several years ago, the Commission directed public utility transmission providers to work through the North American Electric Reliability Council (NERC) to develop a market-based method of managing congestion, as an alternative to relying exclusively (outside the ISOs) on NERC's emergency procedures for handling grid overloads (known as transmission loading relief, or TLR). This approach solicits bids for redispatch in a decentralized bilateral market and relies on sufficient offers being available to manage congestion efficiently. All congestion that is not managed through bids falls back on being subject to TLR procedures. In principle, the more liquid such a market for redispatch becomes, the closer it would approximate the prices and congestion charges associated with an LMP system. In practice, this approach has had little success to date.

(3) **Congestion Management by a large regional ITC**. An independent transmission company that both owns and operates the grid for a large region would schedule transmission and charge users an efficient congestion charge.

However, the ITC would be subject to incentive-based regulation to reduce congestion in the system. A form of this approach has been used in England and Wales, and some urge its adoption in the United States. (This is an alternative to the prevailing U.S. approach of relying on an independent RTO or ISO that does not own the grid to use bid-based markets, locational pricing, and regional planning to reduce congestion.)

The following congestion management systems are generally not compatible with the White Paper's market-based congestion management element (although an RTO or ISO could still propose to manage congestion with one of the following methods if it first demonstrated to the Commission that the costs of a White paper's market-based approach would exceed the benefits in its region):

(4) **Fixed Zonal Congestion Pricing**. Under fixed zonal congestion pricing, the RTO or ISO would identify zones of the grid that are anticipated to have little or no congestion within the zone but are likely to experience congestion going from one zone to the next. Fixed zonal congestion pricing would apply congestion charges only between zones. Experience has shown that such zones are difficult or impossible to identify for extended periods in most systems. Thus, the fixed zones do not reflect market conditions. To the extent that zones are not fixed and can change or be subdivided to reflect evolving congestion patterns, this type of flexible zone could be expected to approximate the LMP approach over time. A practical disadvantage of this is that the zone-to-zone transmission rights would have to be re-established every time the zones change.

(5) **Transmission Line-Loading Relief (as the primary congestion management method)**. TLRs are used to manage congestion on an administrative basis, an inefficient method that is not designed to support a competitive market. In fact, the TLR system was not designed to manage congestion at all, but to deal with grid emergencies such as a downed power line. Under TLR, non-firm transmission schedules are subject to curtailment, with no reference to the value of the transaction. The market redispatch program mentioned above was intended to substitute for TLRs as the primary congestion management tool until LMP became more widely used. (Retaining use of TLRs as a fallback curtailment measure for use when necessary for true grid emergencies is appropriate if it is not the primary congestion management tool.)

Advantages and Disadvantages of Alternative Congestion Management Systems

The market-based congestion management methods each have advantages (+) and disadvantages (-):

Locational Marginal Pricing

(+) LMP is an efficient way to calculate the least cost dispatch consistent with congestion. It provides clear and transparent information regarding the costs of congestion. The quality of the price signal is independent of changes in patterns of congestion (in contrast to fixed zonal pricing, in which the zonal price can suppress information about changes in intra-zonal congestion).

(+) LMP has a several-year track record of successful use. Various types of firm transmission rights (FTRs) can be used together with LMP to protect customers from high congestion prices. FTRs are discussed in the next section.

(+) LMP can be simplified for buyers through zonal averaging and for buyers and sellers by hub pricing.

(+) LMP is compatible with pricing line losses at either marginal or average cost.

(-) Because LMP depends on having a centralized RTO or ISO to operate a regional electricity spot market, this approach may have significant start-up costs.

(-) LMP is perceived as being a complex pricing system and also perceived as hard to adjust to fit the scheduling requirements, and perhaps other requirements, of a hydroelectric region.

(-) For market participants conducting short-term transactions (e.g., under one month), LMP may be difficult to hedge with FTRs (although flowgate rights may mitigate this problem).

Market Redispatch Programs

(+) Market redispatch programs are compatible with decentralized operation of regional markets and, in principle, have smaller start-up costs than LMP.

(-) The market redispatch program conducted by NERC did not solicit sufficient bids from market participants and is generally considered unsuccessful to date. It has been difficult to coordinate bilateral bids in real-time without a regional market. And there is dissatisfaction with the method of determining which bilateral transactions are selected for curtailment and hence responsible for redispatch costs. More software development is needed to continue the NERC-led program, and market participants that are committed to LMP are reluctant to fund continued investment in this program.

(-) Not tested for marginal loss pricing.

(-) Congestion charge hedging has not been attempted or examined.

Congestion Management by a large regional ITC

(+) A central entity is responsible for transmission upgrades to reduce congestion management

(-) Not tested in markets with locational pricing. A form of this approach is used in the England/Wales market, where there is significant excess supply and no transmission rights, conditions that do not readily fit the current U.S. situation.

(-) Requires a large entity that owns all the transmission in a region that is large enough for effective congestion management, which is an institutional development that is difficult to bring about in the near term in the United States because of the current diverse transmission ownership arrangements and the possible tax consequences of divestiture of transmission to a large ITC.

Element 7. Firm Transmission Rights

Under the White Paper proposal, RTO or ISOs that manage congestion with locational marginal pricing (LMP) must make firm transmission rights (FTRs) available to transmission customers. With FTRs, transmission customers can protect themselves from congestion charges. The White Paper would have each region design transmission rights that are compatible with its approach to congestion management. Firm transmission rights can be designed and allocated in several ways, as discussed next.

Types of FTRs

Firm transmission rights can be obligation rights, option rights, or flowgate rights. An obligation right not only permits the holder of the right to be compensated for congestion from one point to a second on the grid, but it also obligates the holder to make payments for congestion that occurs in the opposite direction from the second point to the first. An option right, on the other hand, does not obligate the right holder to make such congestion payments. Both obligation and option rights are referred to below as point-to-point rights. A flowgate right, in contrast, is the right to collect congestion revenues associated with congestion on a particular grid facility.

A point-to-point option is the transmission right that is most like a familiar point-to-point contract. The eastern ISOs began with obligation rights. Some parties object to a requirement to offer obligation rights. (The financial exposure associated with obligation rights appears not to be well understood outside the areas that have experience with them.)

Under the White Paper proposal, an RTO or ISO that manages congestion with LMP can offer any of the following combinations of types of FTRs:

(1) point-to-point obligation rights only,
(2) point-to-point option rights only,
(3) flowgate rights only,
(4) point-to-point obligation rights + point-to-point option rights,
(5) point-to-point obligation rights + flowgate rights,
(6) point-to-point obligation rights + point-to-point option rights + flowgate rights.

In contrast to the White paper, in the SMD NOPR the Commission had proposed that regions should at least begin with option (1) and then consider options 4, 5, or 6 on the basis of user requests, with option (6) as an end goal. This is like the approach taken in the Northeast ISO markets, all of which began with option (1) and are now considering option (4). RTO West proposed and the Commission approved a variant of option (2). ERCOT has implemented a version of option (3), focused on what were presumed to be major flowgates only. The Midwest ISO has proposed a variant of option (6), in which point-to-point obligation rights are allocated initially, but point-to-point options and flowgate rights can be sold in subsequent auctions.

Allocating FTRs to Customers

There are two main issues with allocating FTRs to customers. First, there is (A) the issue of whether the initial allocation ensures that all pre-existing physical rights are converted into equivalent FTRs. Second, there is (B) the issue of whether the FTRs should be assigned directly to eligible customers or whether there should be an auction of FTRs in which the customers eligible for rights have a type of preferential bidding that allows them to retain their rights and in which such customers receive the revenues from any rights sold at auction. Options for resolving these issues, A and B, are set out next.

A. Converting Existing Rights to FTRs. The conversion of existing rights to FTRs is important because customers with current rights want assurance that these rights will not be diminished in the conversion process. The White Paper offers such assurance but leaves to the regions, particularly the RSCs, to devise the method.

Some customers ask why, if all existing rights are met today, there might not be enough FTRs to satisfy all rights tomorrow. This is in part because the grid operator understands that various customers use the grid at different times, but they cannot all receive FTRs with the right to use the same capacity all the time. One approach to resolving this could be to give out rights to time periods, such as seasonal rights. Another reason that there may not be enough rights to cover all existing firm uses of the grid is that the grid operator understands that customer flows in opposite directions today at least partly cancel one another, but the operator cannot give out rights to flows in one direction unless another customer has an obligation to provide flow in the opposite direction, often called counterflow. There are several ways that a region could convert existing rights to FTRs without diminishing existing rights. Here are several examples in two categories:

(1) *All* pre-existing physical rights would be honored regardless of their initial physical simultaneous feasibility. To do this, a method for ensuring feasibility is needed. One such method is to create new FTRs that do not correspond to any customer's rights, such as rights to "counterflow" against the normal direction of power flow. These additional FTRs create new rights to flows so that all the flows taken together are feasible, including the flows corresponding to initial rights. The new rights are, in effect, "rights to pay" to relieve congestion through counterflow, allowing all the initial rights to be honored. However, the holder of these new rights has to pay for counterflows, that is, pay for generator redispatch to make all rights feasible. As no one is likely to want to acquire these new rights, which are a cost instead of a benefit, the RTO or ISO would create and hold the rights so as to be able to honor all existing firm rights and would allocate the costs of holding these rights to (probably all) transmission customers. A similar approach is for the RTO or ISO to pay a generator to run under a long-term contract so as to create counterflow to make up any customer's deficiency in historical rights. The cost of either the new FTRs or the long-term contract would be a so-called "uplift charge" assessed to all transmission customers. This would be treated as a necessary cost of protecting all existing firm customers from congestion charges during the transition. Either approach would be a transition mechanism that would automatically phase out as existing contract rights expire. Another approach, taken in RTO West, is to catalogue pre-existing rights and continue to honor them as option point-to-point FTRs.

(2) *Some* pre-existing physical rights would remain as physical rights and be honored without conversion to FTRs. Under this approach, a selected set of pre-existing physical rights is grandfathered as physical, and all other rights are converted to FTRs. In effect, some transmission capacity is left out of the FTR modeling. A disadvantage of doing this for significant amounts of transmission capacity is that the amount of the grid that can be managed efficiently by the RTO or ISO is reduced. For example, the grid operator cannot arrange counterflows over the grandfathered capacity to increase the number of beneficial trades that regional market participants can make.

Either of these general approaches is consistent with the White Paper proposal. Another approach, that does not necessarily ensure that all existing rights are honored, has been used by New England, New York, and PJM:

(3) *Only* simultaneously feasible physical rights are converted to FTRs. Under this approach, utilities serving customers are offered FTRs up to some amount (typically their peak load from the prior year) and allowed to prioritize the locations that they value most. However, the RTO or ISO makes the final allocation, possibly giving transmission customers only a pro-rata share of their existing rights.

The Commission would not require these regions to alter their existing approach, and would consider requests from other regions to follow such an approach.

B. Assignment or Auction of FTRs. Once the existing customer rights are converted to FTRs, there are two approaches for initial allocation of these FTRs. The White Paper would allow the region to decide between these two approaches:

1. Direct Allocation. All those eligible for FTRs would directly receive their FTRs. They can then keep them or voluntarily sell or re-configure FTRs in monthly auctions or sell them in secondary markets. The advantage (+) and disadvantage (-) of direct allocation are:

(+) Allows entities that do not want to develop a bid strategy in an initial auction (as discussed next) to directly receive FTRs

(-) Encourages entities that receive FTRs to be conservative in deciding to sell those rights, possibly limiting unnecessarily the FTRs available to others.

2. Initial Auction. All those eligible for FTRs would receive the *revenues* from an initial auction. These rights to revenues are typically linked to the FTR locations for the

FTRs that would have been allocated directly; the regions that have used this approach have used somewhat different formulas for allocating these revenues. Hence, each right holder is able to bid high enough to ensure it can buy back in the auction all the FTRs that it would have been assigned. New York and New England chose the auction approach. The advantage (+) and disadvantage (-) of an initial auction are:

(+) Encourages entities to evaluate the financial benefits of bidding to retain their FTRs or selling some excess FTRs. Any extra selling would benefit both FTR sellers and buyers and should foster a more liquid market for FTRs.

(-) Entities not familiar with FTRs could make mistakes in the auction, such as selling the rights for less than their true worth.

Element 8. Resource Adequacy Approaches

The eighth and last element of the White Paper's Wholesale Market Platform is that each RTO or ISO must have a regional approach to assessing resource adequacy. The approach itself and the roles of various participants would be decided regionally.

Because the White Paper offers complete state and regional flexibility on the regional resource adequacy approach, there are no specific resource adequacy options presented in the White Paper. Nevertheless, there is interest in the range of possible options available to a region. The following discussion presents example of possible approaches without attempting to present an exhaustive list of resource adequacy approaches allowed under the White Paper.

Resource adequacy measures may be implemented through a central capacity market for generating capacity that may be enforced through the RTO or ISO tariff. It may be implemented through traditional state requirements for a traditional utility to satisfy a reserve margin or loss of load probability standard of adequate service; such a requirement is typically enforced by state regulation of these utilities, or by comparable regulation of the service adequacy of public power by government bodies and of cooperative utilities by their governing boards. A resource adequacy requirement may be implemented through some combination of these measures as, for example, in a region with a mix of states with and without retail choice programs. Or it may be implemented through other measures not discussed here, such as the current regional planning council mechanism of the Pacific Northwest.

One region may prefer to state its requirement in term of loss-of-load probability, and another in terms of reserve or capacity margin. Each region would be free to choose the measure of adequacy that best fits the region's needs and resource mix. Further, a region may choose to implement its measure by allowing one utility to express its measure as loss of load probability, and another as reserve margin, as long as they meet a comparable level of reliability in such a way that no one utility is in effect planning to lean on the resources of another.

Each region would set its own level of adequacy, such as a loss-of-load probability of one day in ten years or a reserve margin of 18 percent. Both the measure and the numerical value would be determined by each region.

Each region should consider the time it takes to develop new supply and demand response infrastructure in the region and how this should affect the time frame for resource planning. Also, the region may consider whether to make uniform what resources qualify as satisfying the regional needs, such as demand response and transmission upgrades as substitutes for generation capacity, and any requirements for the assuring that generation claimed as a resource is deliverable to the location that claims it as a resource.

However, a regional resource adequacy plan that allows a load serving entity to satisfy its share of a regional requirement by relying on spot market purchases from uncommitted supply resources during peak demand periods would not appear to be appropriate if the purpose of the plan is to assure that new investment is in place in advance of the potentially high prices that occur during shortage periods. In a purely spot market approach to resource adequacy, new infrastructure investment is motivated by the expectation of future high prices during the peak periods. But there is no guarantee that these investments for uncommited capacity will in fact be made in advance. Needed investments may be made only after the shortage begins and prices have risen. However, a purely spot market approach to resource adequacy presumes no price caps or other market power mitigation measures; any price caps or other market power mitigation rules would have to be designed to provide the opportunity for these uncommitted peak spot resources to recover their costs through the spot market. If the market rules do not provide this opportunity, some existing resources could exit the market and some new resources may not be developed. Further, although the expectation of high prices during shortages should give load serving entities a natural market incentive to build in advance or contract forward to avoid high prices during these shortages, an expectation of limits on prices should have the opposite effect, suggesting the need for a regional forward build-or-contract requirement with mitigation. Note too that suppliers of resources have less ability to exercise market power in a forward market that plans several years ahead because new entrants can compete with existing resources, disciplining the forward market price.

Reliance on State and Other Local Regulation of Resource Adequacy

To meet a region's resource adequacy requirement each state could rely on its own state authority (or municipal city council or cooperative utility board, as appropriate) to ensure that each utility or other load serving entity owns or contracts for enough reserves, whether generation or demand response, to satisfy its share of the regional need. A state may choose to require that each load serving entity either own all necessary generation resources or rely on forward contracts or call contracts with other suppliers—or a combination of the two—to meet the region's adequacy requirement.

Reliance on a Regionwide Capacity Requirement for Resource Adequacy

With a regionwide capacity requirement, each load serving entity in the RTO or ISO market would be required to acquire its share of the capacity or other resources needed to ensure that the total resources available to the region are sufficient to meet the regions' reliability and market needs. As mentioned, the appropriate capacity requirement would be determined by each region. Further, generators designated as capacity resources could have specific obligations to supply energy into the market, or power customers could have an obligation to provide demonstrable demand response. A region may choose in addition to create a centralized market to trade in such resources. The RTO or ISO could then ensure that resources that seek to qualify as capacity resources meet the necessary deliverability and availability criteria to maintain system reliability.

Relationship of Resource Adequacy to Other Market Design Elements

The White Paper emphasizes that the approach to resource adequacy must be designed to work together with other elements of the regional market design: market power mitigation measures, demand response programs, and any scarcity pricing measures. Those designing elements of a regional market must assess how the various regional choices work together. This is because the Commission is responsible for just and reasonable wholesale prices, and wholesale market prices depend on having enough resources available for the market to function effectively.

Importantly, investment in new generation and other infrastructure is needed to keep supply and growing demand in balance. To invest in such infrastructure, investors must find that the combination of mitigation measures, resource adequacy provisions, and scarcity pricing provisions—taken together—provide a reasonable opportunity to recover the costs of their investments. For example, a region with little mitigation or mitigation that permits prices to rise quite high to reflect scarcity may not need to have a strong administrative approach to resource adequacy. But a region with a low safety net bid cap and no scarcity pricing to hold demand in check is unlikely to have a market with prices that attract new supply, unless such a region also has an additional approach to resource adequacy that provides additional assurance of capital cost recovery for new investment in the region.

Other Areas of Flexibility

The White Paper contains two other areas of flexibility not presented above because they are not elements of the Wholesale Market Platform and they do not lend themselves to options analysis. One is the offer to opt out of any market element (of the eight above) with a demonstration to the Commission that its cost exceeds its benefit. This is not properly characterized as a regional choice because it depends on the outcome of an analytical cost-benefit analysis. The other is timing flexibility for implementation, which depends on the outcome of discussions begun at the regional conferences.