MARKET POWER MITIGATION: PRINCIPLES and PRACTICE

Larry E. Ruff

Charles River Associates

November 14, 2002

MARKET POWER MITIGATION: PRINCIPLES and PRACTICE

Table of Contents

MARKET POWER MITIGATION: PRINCIPLES and PRACTICE

Larry E. Ruff

Charles River Associates

November 14, 2002

1. INTRODUCTION

1.1 CONTEXT AND OBJECTIVES

Consumers and public officials are concerned that suppliers will use market power to drive prices too high, particularly during scarcity conditions when suppliers appear able to charge "whatever the market will bear." Suppliers are concerned that overreactions to such concerns will keep prices so low that existing assets will not recover their costs and future investments in electricity supplies will become unattractive. These opposing concerns both have some justification, and until they are dispelled everybody should be concerned about the reliability and continuity of their electricity supplies.

Current discussions of market power and its mitigation in electricity focus on the spot markets operated by independent transmission providers (ITPs) under the authority of the Federal Energy Regulatory Commission (FERC).[1] In particular, FERC's proposed Standard Market Design (SMD) includes a package of market power mitigation (MPM) measures based on limiting physical withholding and capping suppliers' bids in spot markets, with a resource adequacy or capacity requirement intended to give suppliers additional revenues to offset the lower spot prices. Some version of this general framework, with or without an effective resource requirement/payment, is in use or is being developed for functioning and emerging ITPs, including the Midwest ISO (MISO).

The objective of this paper is to provide an economic perspective on market power and its mitigation in electricity spot markets and to suggest policies that might reduce the concerns outlined above. The focus is on MPM as applied to suppliers, because this is the most important issue in practice. This paper has been commissioned by a group of electricity generating and marketing companies,[2] but the analysis and views are those of the author, an independent expert with extensive experience in the design and operation of competitive electricity markets worldwide. Individual sponsors of this paper do not necessarily share or endorse all the views expressed and recommendations made here.

[1] For the purposes here, an ITP is any system operator that uses spot market to manage and price physical operations. Both Regional Transmission Organizations (RTOs) and Independent System Operators (ISOs) are ITPs here.

[2] These companies are American Electric Power, Edison Mission Energy, Mirant, NRG Power Marketing, Inc. and PSEG Energy Resources & Trade, LLC.

1.2 OUTLINE AND SUMMARY

This paper consists of the four sections below in addition to this introductory section.

Section 2: The Economic Principles and Their Application

This section reviews some elementary economic principles as they apply to spot markets and discusses why and how competitive behavior in a specific market depends on the price-discovery process in that market. It then turns to electricity spot markets in particular, discussing why it is not so easy to know how a competitor "should" behave in an ITP's markets, what happens if spot scarcity prices are suppressed, capacity requirements/payments as alternatives to scarcity pricing, and demand response. The principal theme of this section is that MPM procedures of the type currently being implemented and proposed for ITPs are logically inconsistent and will produce inefficient outcomes unless ITPs use scarcity pricing methods that are more accurate than those currently used or likely to be implemented soon. Until ITPs implement such scarcity pricing methods, MPM procedures should expect and allow competitive suppliers to bid in ways that increase their chances of receiving efficient, market-clearing scarcity prices and recovering their reasonable costs.

Section 3: FERC's Market Power Mitigation Package

The objectives and the four MPM measures in FERC's SMD NOPR are described and analyzed at a general level consistent with the limited details provided by FERC. It is concluded that these measures will, in practice, suppress spot scarcity prices despite FERC's expressed desire not to let this happen. The Resource Adequacy Requirement (RAR) is discussed and shown to require much more thought and development before it can be regarded as a proposal for a workable or effective policy that will offset the effects of suppressed scarcity prices.

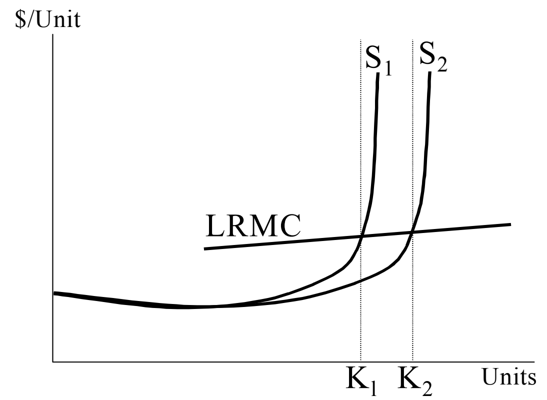
Section 4: The MISO Market Power Mitigation Proposal

The MPM proposal currently being considered by MISO is described and analyzed. It is shown that, even with some technical problems fixed, the proposed MPM procedure will result in the suppression of scarcity prices given the way the MISO (and other) spot markets determine such prices. Although some improvements in scarcity pricing that have been suggested for the MISO (and other ISOs) are steps in the right direction and should be encouraged, these are too limited to solve all the problems even if they are effectively implemented. The MPM procedures currently proposed for MISO would suppress scarcity prices and threaten the commercial viability of competitive and needed suppliers. They should be modified to apply less broadly and to allow suppliers to act in ways that will improve their commercial viability.

Section 5: A Suggested Approach to Market Power

The concluding section suggests five central objectives for a successful MPM policy: (1) Market power in spot markets should be put in perspective by focusing on overall market processes and outcomes; (2) spot scarcity pricing in ITP markets should be improved so that MPM can be both less distorting and more effective; (3) until spot pricing is much improved, MPM should be narrowly focused and light-handed; (4) because aggressive MPM procedures will suppress spot prices, as long as such procedures are in place there must effective capacity payment arrangements; and (5) electricity markets should quickly make the transition to full competition, which requires efficient spot scarcity pricing.

Figure 1:
Typical SRMC and LRMC Curves



2. THE ECONOMICS OF MARKET POWER MITIGATION

This section reviews some basic economic concepts as applied to spot markets and their implications for identifying and mitigating market power in spot markets.

2.1 THE ECONOMICS OF A COMPETITIVE SPOT MARKET

2.1.1 SHORT-RUN AND LONG-RUN MARGINAL COSTS

In economic textbooks, a perfectly competitive market is one in which no supplier is large enough to increase the market price by withholding or overpricing some or all of its potential supply. In most real markets, some supplier(s) could increase the market price a little for a while by withholding or overpricing supplies, but any supplier that did so would eventually lose more in profitable sales than it would gain from the higher prices as competitors increased their output to replace the withheld or overpriced supply. A market in which each supplier decides how much to supply at market prices that it cannot profitably affect for long is said to be workably competitive.

A supplier in a workably competitive market maximizes its profits (or minimizes its losses[3]) by selling up to the point where its short-run marginal cost (SRMC) equals the market price but not beyond the point where its SRMC begins to exceed the market price. For deciding how much to produce to maximize profits, a supplier properly considers its SRMC as the increase in present or future costs resulting from increasing output/sales by one (small) unit. A supplier's SRMC is not a single or easy-to-estimate number, but is a sometimes-hard-to-measure variable that depends on the supplier's current and expected future output, current and future (but not past) prices, the existing physical plant and staffing levels, etc.

Figure 1 illustrates typical SRMC curves for a supplier that produces a commodity using fixed physical assets and variable inputs such as fuel. S_1 is the SRMC curve corresponding to a particular set of fixed assets with a rated capacity of K_1. For output levels below K_1,

[3] This qualification is implied but not repeated whenever phrases such as "maximizing profits" are used here. A supplier's maximum profits may be negative – although if they are it cannot stay in business for long and is unlikely to be replaced by another supplier.

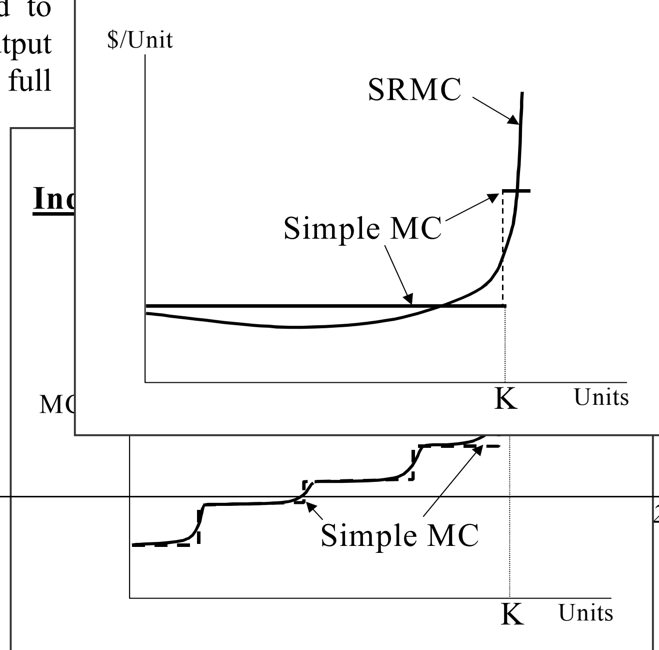
SRMC is the incremental cost of fuel and raw materials, maintenance and wear-and-tear on equipment, including any opportunity costs if producing more for this market now increases the costs of producing for some other or later market. Over this range, SRMC usually increases slowly with output beyond some point of maximum short-run technical efficiency and then becomes very steep near full capacity K_1.

The full capacity K_1 associated with some set of fixed assets can be defined as the steady-state output level at which SRMC is the same as long-run marginal cost (LRMC), defined as the minimum long-run unit cost of producing that output level taking into account all fixed asset costs as well as variable costs given those assets. Because it is always possible to operate above any such measure of full capacity, at least for a while, by paying overtime, sacrificing some technical efficiency, overstressing equipment or delaying maintenance at the risk of earlier or more costly repairs later, etc., the SRMC curve S_1 continues for some distance beyond K_1, becoming infinite where it really is impossible to get anything more from the existing facilities. Given time and money to expand the facilities, the SRMC curve itself can be shifted, as illustrated by the SRMC curve S_2 with higher full capacity K_2 and higher fixed asset costs.

It is critical for the analysis in this paper to understand that the proper or "real" SRMC that a competitive supplier must consider in deciding whether to provide an additional unit of output to the market is more complex than, and may far exceed, simpler and more conventional measures of marginal costs such as average fuel costs and variable operation and maintenance (O&M) costs. This is particularly true when a supplier is operating at or near its full output and may have to take extraordinary and very costly measures to increase output slightly. For clarity in this analysis, the term "SRMC" is used for the proper or complex SRMC and the term "simple MC" is used for simpler, more conventional measures of marginal cost.

Figure 2 illustrates the difference between a typical supplier's SRMC curve and the type of simple MC often assumed in electricity markets for illustrative purposes and – far more importantly – for operational and pricing purposes. Simple MCs are based on average SRMCs over wide ranges, and do not usually include even the kind of step illustrated in Figure 2, which at least acknowledges that marginal costs become high near maximum output. Even if simple MC is defined to include some such "sculpting" at high output levels, MC for a supplier operating at full capacity is often – but incorrectly – defined as the incremental cost of the last unit produced, not the much higher (or infinite) cost of the next unit that could (or could not) be produced.

Figure 2:
SRMC and "Simple MC" Curves



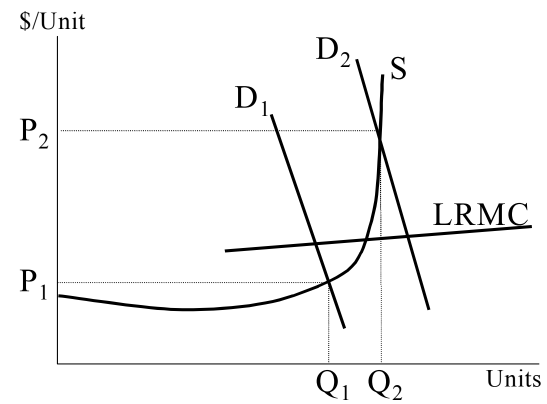
Because each supplier in a workably competitive market will produce up to the point where its own SRMC reaches or begins to exceed the market price, the SRMC curve for an individual facility is also the short-run supply curve for that facility. If all facilities face the same price for their output, the short-run supply curve for the market as a whole is the horizontal aggregation of the individual SRMC curves. Figure 3 illustrates an industry supply curve for an industry in which different suppliers have different SRMCs, and how such a supply curve might be approximated by using the simple MCs of different suppliers. Such simple MC curves are often incorrectly interpreted to mean that, when all suppliers and hence the industry are producing at full capacity, the SRMC of each supplier is the highest step in its simple MC curve and the SRMC of the industry is the highest of these, MC_{MAX} in Figure 3. The more correct interpretation is that the SRMC curve of each supplier and hence of suppliers as a whole becomes vertical at full output.

2.1.2 Market-Clearing Prices and Price Volatility

A workably competitive market clears where demand equals supply, and the market price at this point will equal the SRMC for suppliers as a whole and for each individual supplier. This is true for a supplier producing at its maximum output because such a supplier can get another unit to sell if, but only if, it buys it at the market price; its SRMC curve at full output is vertical up to the market price, however high that is. The market price and hence the true SRMC of such a supplier may be far above any conventional measure of simple MC. If all suppliers are producing at or near full output the market price and hence all SRMCs can be far above the simple MC of any supplier.

Figure 4 illustrates a market in which suppliers have a fixed set of assets that result in the short-run supply curve S (assumed to remain unchanged over several short-run periods). As demand varies from one short-run period – which may be a year or a week or an hour, depending on the technical characteristics of the commodity – to another, the market equilibrium defined by the intersection of the demand and short-run supply curve will change and the short-run or spot market price will move along the SRMC curve – for example, from P_1 when demand is D_1, to P_2 when demand is D_2. The spot price or SRMC will be above LRMC at some times and below it at other times, but if there are no barriers to entry or exit and the market is approximately in a long-run equilibrium, the average price will be close to LRMC; if it is not, new facilities will be built or existing ones will be shut down until it is.

Figure 4:
Volatility in a Competitive Market



A scarcity[4] or peak period is defined as one in which market demand is high relative to the available supply, as illustrated by demand curve D_2 in Figure 4. At such times the market-clearing spot price can be much higher than it is usually or on average and much higher than LRMC. But the fact that spot prices are higher than LRMC during peak or scarcity periods does not mean they are "too" high during those periods or overall. Prices are too high during scarcity conditions only if a competitive market – one in which the market price is at the intersection of SRMC and demand – would clear at a lower price. Prices are too high overall only if spot prices averaged over a period in which entry is possible are above LRMC.

During scarcity conditions, most suppliers are operating on the steep parts of their supply curves where SRMCs depend on judgmental factors such as risks, and consumers may be taking or considering demand-reduction actions that have uncertain or unusual costs. Under these conditions, the cumulative effect of small and unpredictable changes in the judgments of individual suppliers and/or consumers can shift the supply or demand curves enough to have large effects on competitive scarcity prices. This makes it essentially impossible to forecast the levels of scarcity prices accurately or even to explain those levels after the fact based on easily observable factors that usually define simple MCs, such as fuel costs.

Because competitive scarcity prices can depend on so many complex and even judgmental factors, there is no reliable way to decide when market-determined scarcity prices are too high or to compel suppliers to act so as to produce Goldilocks prices: Not too high, not too low, but juuuust right. Any administrative procedure for controlling prices or market behavior will get it wrong much of the time. And, because any such procedure will target the highest prices that are the hardest to explain objectively or to tolerate politically, such a procedure almost inevitably suppresses scarcity prices below competitive levels.

[4] "Scarcity" is only a relative term. Economically speaking, anything that has a positive price is scarce, meaning that it would be good to have more of it. The term "scarcity price" generally refers to a price created when the demand curve intersects the supply curve on the (near) vertical section close to maximum output. Under these conditions, not only is the commodity itself scarce, but the assets needed to produce the quantity are scarce and are earning "scarcity rents" – the excess of revenue over average SRMCs or simple MCs need to recover fixed costs and encourage additional investment when needed. Real prices always include some scarcity rents for the lowest-cost producers, even in normal or "non-scarcity" periods.

Forcing spot scarcity prices below competitive market-clearing levels will reduce the efficiency of the market and increase total costs to consumers in the long run. In fact, even if scarcity prices are very high, reducing them may do consumers surprisingly little good, because other prices will eventually have to increase enough to make up for the loss of scarcity rents to suppliers. Given the difficulty of knowing when a spot price is too high, the high costs and risks of trying to reduce it, and the low payoff even if it is done well, it is usually better not to try unless there is clear evidence of harm to the larger market. This is discussed further below in the context of electricity spot markets.

2.1.3 Price-Discovery Processes and Supplier Offers

The logical conclusion that prices in a workably competitive market will tend to equal – more realistically, approximate – the SRMC of each supplier says nothing about the market process that determines the market prices. In particular, it does not say that each competitive supplier always will or should ***offer*** all of its supplies at its own SRMC. Whether or not and at what price a competitive supplier offers to sell anything in the market depend on the specific mechanics of that market.

In real, workably competitive markets for things as disparate as tomatoes and real estate, suppliers make offers by posting the prices at which they are willing to sell. But no seller in such a market decides what offer price to post by looking at its own SRMC curve. Instead, suppliers observe asking or transaction prices in the market yesterday or next door, consider any factors that may make things different today or here, make their best estimate of the market-clearing price here today, and then offer to sell at that price. A supplier may decide to offer more or less (or nothing) based on its estimate of the market-clearing price and its own SRMC curve, but its offer price is based on its estimate of the market-clearing price, not on its average SRMC much less on any simple MC.

Furthermore, because a supplier in such a market will ultimately sell at or near its own offer price, each supplier's offer price must frequently be above its average SRMC if it is to recover sufficient fixed costs to stay in business. Nobody expects a farmer to offer tomatoes at the cost of picking them and driving them to market, or a homeowner to offer its house at the cost of sprucing it up for sale. In most markets, no supplier is expected to offer to sell everything – or anything – at its own simple MC or is accused of trying to exercise market power if it does not.

The behavior of competitive suppliers expected in most real-world competitive markets described above is much different than what is usually expected in electricity spot markets. There is a widespread, but unexamined and often incorrect assumption that a competitive supplier in an ITP's spot market would always offer all its output at some average SRMC or simple MC, from which it seems to follow that the way to control or mitigate suppliers' market power is simply to require all suppliers to offer all their supplies at or near their simple MCs. What is the basis for this assumption and what are implications if that assumption is incorrect?

The proposition that a competitive supplier in an ITP's spot market would offer all it available output at its own simple MC is dependent on two critical assumptions: (1) that the supplier is allowed to submit sculpted bids so that its simple MC can become very high at high output levels to reflect its actual SRMC at those levels; and (2) that the ITP's pricing process will determine efficient, market-clearing prices that do not necessarily depend on, and may be much higher than, any supplier's bids. In particular, when demand is high relative to available supply, the ITP must base market prices, not on estimates of each supplier's average SRMC over its full output range, but on sculpted supplier bids that may become very high at high output levels, and/or on demand bids, import/export bids, and even on the implicit costs of such things as demand interruptions, low operating reserves or risky system operations. As discussed in sections 2.2.1 and 2.2.2 below, in theory ITP spot markets should operate this way but in practice they do not. And if they do not, there is no logical or practical reason to expect or to require suppliers in those market to operate as though they did. The implications of this commonsense observation are discussed further below.

2.2 The Economics of an Electricity Spot Market

Standard economic principles apply to electricity spot markets, but require some special twists and implications due to the physical and economic peculiarities of electricity.

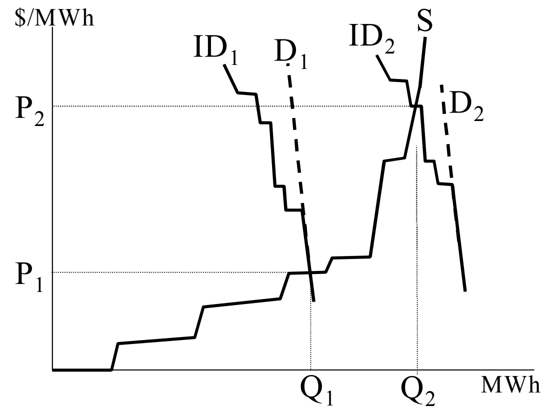
2.2.1 ITP Spot Pricing in Theory

In the "normal" market processes discussed above, a rational, competitive supplier bases its offer price on the expected market-clearing price, not at its own SRMC, and then offers to sell at that price the amount that maximizes its profits at that price. But an electricity system is so dynamic and complex that operations would be unnecessarily costly and even unreliable if each supplier had to estimate market-clearing prices and then offer to sell its profit-maximizing quantity at those prices without regard to what actually happens in real time. That is why an electricity market needs a central market-clearing and pricing process operated by an ITP.

The fundamental concept underlying ITP-operated energy spot markets is that all market participants should submit bids revealing the costs to suppliers of producing energy and the value to consumers of consuming it, and the ITP will use this information and everything else it knows about the system to determine spot prices that equal/approximate the real SRMC of meeting demand reliably at each time and place. In particular, under scarcity conditions when very costly actions are necessary to meet demand reliably, the market-clearing prices may be independent of and much higher than the bid prices of any suppliers. If an ITP spot market applies these concepts correctly, a workably competitive supplier will submit a bid curve with different prices at different quantities, in effect creating a simple MC curve that approximates its true SRMC curve, because this will result in the supplier selling the amounts that maximize its profits at the market prices, whatever these prices turn out to be. But it will do this only if it knows that the ITP will, if market conditions so indicate, determine a market-clearing price that is higher than even the highest step in its bid curve.[5]

The theoretical model of ITP spot pricing underlying the usual assumption about supplier bidding is illustrated in Figure 5. In theory, suppliers operating in the market submit to the ITP, not simple MC bids based on fuel costs, variable O&M, etc., but "sculpted" bids reflecting their true SRMCs, including the very high and sometimes judgmental SRMCs associated with operating near and beyond some measure of maximum output. The ITP uses these bids from "in-market" suppliers to construct a supply curve S, which will have many small steps increasing to very high levels near and beyond the full capacity of all suppliers.

Figure 5:
ISO Spot Pricing in Theory



The ITP also uses demand forecasts, estimates of demand elasticities and explicit demand bids[6] to construct a market demand curve for each period, illustrated by D_1 and D_2 in Figure 5. But the ITP knows that, if there is not enough supply to meet demand, it will not disconnect consumers as a first response. It will do many other things, such as buying imports from neighboring ITPs, canceling scheduled exports, calling interruptible load contracts, taking energy from reserves, overloading transmission facilities, letting frequency or voltage drop, and as a last resort curtailing some "firm" demand if necessary to keep the system from collapsing. Call these "out-of-market" or "OOM" actions, because that is how they are usually treated, even though the objective should be to get them into the market.

5 The discussion here is dealing only with the energy spot markets, but even in concept an ITP needs more than these to operate the system reliably. For one thing, reliability services such as operating reserves, voltage support or reactive energy, frequency control or regulation, etc., are not adequately managed or compensated in energy markets, so the ITP must acquire and pay for these in ancillary service markets. But even within the energy market itself, there can be situations where SRMC-based energy prices will not support energy suppliers that are needed for reliability purposes; in a typical example, a combustion turbine (CT) may be needed at some particular constrained location but runs at its full output so seldom that prices in energy and ancillary service markets will not cover its full costs. Such cases require special arrangements, such as a reliability-must-run (RMR) contract between the ITP and the CT. If the CT must support itself from energy market revenues it must be allowed to act in the energy market to produce prices that will cover its costs with a fair return.

6 A demand bid should be an offer to reduce energy purchases if the price increases above some level, not an offer to reduce energy purchases if paid to do so. Anybody should be free to sell back energy it does not consume if it has bought it (e.g.) in a forward market, but nobody should be paid for something it might have bought but didn't.

In theory, the ITP should estimate the costs of the various OOM actions it may take and then put these actions and their costs into its dispatch, market-clearing and pricing engines along with in-market supply and market demand. OOM actions could be treated as additions to supply for this purpose, but if they are treated as reductions in demand the ITP's total demand for in-market energy – ID_1 and ID_2 in Figure 5 – should have significant elasticity at high prices. When demand is low, as represented by ID_1, the market will clear at a low price P_1 with little or no demand response and no OOM actions. But when market demand is higher than in-market supply, as represented by ID_2, the ITP will have to move up the in-market supply and OOM-adjusted demand curves to clear the market at a very high price, P_2 in Figure 5, that is set by either an OOM action or a sculpted supplier bid.

If the ITP's dispatch and pricing process worked as the above theory says it should, competitive suppliers would submit sculpted bids to the ITP approximating their individual SRMCs. But suppliers would do so only because they would know that the ITP would use these sculpted SRMC bids, plus the bids or implicit costs of market-driven demand reductions and OOM actions, to determine efficient, market-clearing prices. In particular, under scarcity conditions the ITP would base the market price either on sculpted supplier bids that become very, even arbitrarily, high for the last few MWh offered, or on the bid or deemed costs of demand reductions and OOM actions that can be higher than any supplier bid prices. If the ITP does not live up to its side of the bargain and reliably determine scarcity prices based on this concept, there is no reason to expect or to compel competitive suppliers to ignore this reality and to continue bidding as though the ITP were doing what it is not doing.

2.2.2 ITP Scarcity Pricing in Practice

In practice, ITP spot markets do not operate as the above theory says they should. In fact, ITP pricing software usually mechanically calculates market prices from the offer prices of on-system generators even when the ITP must use demand reductions, imports, energy from reserves and even higher-cost OOM actions to clear the market, and even when it must use emergency actions to keep the system from collapsing.

Figure 6:
ISO Spot Pricing in Practice

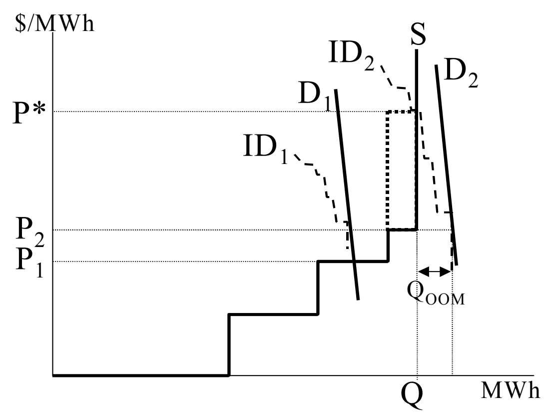


Figure 6 illustrates an ITP dispatch and spot pricing process that more closely resembles current practices than the theoretical process illustrated in Figure 5. In this process, the ITP constructs a supply curve S using bids from suppliers, but now expects or requires these bids to be relatively simple, with no more than a few large steps at levels reflecting fuel and variable O&M costs and at most some "reasonable" allowance for risks and opportunity costs. The ITP also creates a market demand for each period, D_1 and D_2 in Figure 6, perhaps reflecting some price elasticity and/or demand bids. The ITP also has a set of OOM actions it will take if necessary to keep the lights on, but regards these as truly out-of-market actions in the sense that their actual or deemed costs will not be used in computing market prices. The ITP will presumably have in mind some implicit demand curve for in-market energy indicating the order in which it will take the OOM actions if necessary and the implicit costs of doing so – ID_1 and ID_2 in Figure 6 – but does not use these for pricing purposes.

When market demand is D_1 in Figure 6, the ITP dispatches the supply curve S to meet demand at the market-clearing price P_1 with no need for the OOM actions represented by the ITP demand curve ID_1. But when market demand is D_2, there is not enough in-market supply to meet market demand, so the ITP must use the OOM actions represented by the ITP demand curve ID_2. When the ITP has dispatched Q_{OOM} of OOM energy to close the gap, the implicit – and correct – market-clearing price is P*. But the price-determination process ignores the OOM actions and sets the market price at P_2, the highest in-market supply bid taken.

This ITP market-clearing pricing process is very different than the theoretical one described in the preceding section – the one that would motivate competitive suppliers to offer all their energy at sculpted bid prices approximating their individual SRMCs. In fact, this ITP market process is much more like the processes in most other markets, in which suppliers will not get the market-clearing price unless they estimate it themselves and then offer to sell only at that price. In particular, if all suppliers offered all their energy at their simple MCs, none of them would get the market-clearing price under scarcity conditions and some of them would soon be out of business.[7]

So how would competitive suppliers, if unconstrained by such things as market mitigation procedures, conduct themselves in the market given this ITP price-formation process? The textbook concept of "perfect" competition is not helpful here,[8] but examples from real workably competitive markets are. When sellers of tomatoes or houses know they will not get the competitive market-clearing price unless they ask for it, they estimate it and ask for it. It is reasonable to conclude, therefore, that when an ITP market will set the market price below market-clearing levels unless at least some suppliers bid at the market-clearing price, some suppliers will estimate the market-clearing price and bid at that level, even if the market is highly competitive. For competitive suppliers to do otherwise – to offer all their supplies at some estimate of their individual simple MCs knowing that the ITP's software will use such offers to compute and pay them a price that is much less than the real market-clearing price – would be irrational and even commercially suicidal.

---

7 More accurately, those "peakers" who make most of their money from scarcity prices would go out of business unless they could qualify for the ITP's OOM payments or capacity payments. "Baseload" suppliers who can make money at normal times would survive, but only if they received higher average prices in normal times and/or capacity payments.

8 In the textbook definition, a "perfect" competitor does not make a price offer, but passively observes the market price and then sells as much or as little as it chooses at that price. In any real market, suppliers must make pricing decisions that will affect what they are paid.

This reasoning leads directly to the conclusion that, in an ITP market that uses the kind of scarcity pricing process illustrated in Figure 6, workably competitive suppliers would not all offer all their supplies at simple MCs approximating their individual SRMCs. At least some of these suppliers would – and should, if the objective is to produce reasonable scarcity prices – offer some significant amount of their output at levels reflecting their estimate of market-clearing prices. In Figure 6, for example, if the marginal generator with SRMC of P_2 increases its bid price to P* the ITP will set the market price at P* and all generators will get the efficient, competitive, market-clearing price. In fact, there is no other way for such a market to produce competitive scarcity prices.

An electricity market that requires suppliers to predict and bid the market-clearing prices – a so-called pay-as-bid market – can be inefficient under normal conditions, because suppliers will not be able to predict market clearing prices accurately and errors in their individual predictions will result in inefficient dispatch and operations. During scarcity conditions, however, essentially all suppliers will be running at or near full capacity, so such dispatch inefficiencies are less likely to be important, particularly if it is the highest-cost suppliers who are bidding at their estimates of the market-clearing price.

If an ITP market will not produce market-clearing scarcity prices unless some suppliers bid such prices, competitive suppliers should be expected, allowed and even encouraged to do what unconstrained competitive suppliers would do: bid at their estimates of market-clearing prices even if this means bidding well above some estimate of their simple MCs. Such bidding conduct should not be regarded as an exercise of market power just because it increases prices up to competitive levels, as long as it does not push prices higher than competitive levels. And a market in which suppliers can, do and must behave this way in order to get market-clearing scarcity prices should be regarded as workably competitive as long as such behavior does not result in average prices above competitive levels.

The main problem with these suggested definitions of the exercise of market power and a workably competitive market is that they require difficult and contentious judgments about what competitive prices really are. But this is just as true of the approach to MPM currently favored by ITPs and FERC, which implicitly assumes that it is easy to determine the competitive market price in an ITP market: simply plug some estimate of suppliers' simple MCs into the ITP's pricing engine and turn the crank. As discussed in the next section, there is no logical basis for this assumption given the way ITPs currently determine scarcity prices.

The most relevant measure of spot market prices for judging when a market is workably competitive at some location is not the spot price during a few scarcity hours, but the average spot price over some period such as a year. If such average spot prices in some region are above estimated supplier LRMCs when there is no shortage of supply, there is good reason to investigate to see if supplier market power – as opposed, for example, to unusually high local LRMCs – is the problem and if it is to do something about it. But if average spot prices are below LRMCs, particularly when there is no capacity surplus, the conclusion should be just the opposite: the interaction of the ITP's computer and suppliers' bidding conduct is keeping prices too low. Indeed, if a MPM procedure discovers that supplier bidding is significantly increasing prices during scarcity periods, and yet average prices are below LRMCs, it is merely demonstrating how inadequate the ITP's scarcity pricing really is and why bids and prices should not be mitigated further.

If spot prices averaged over (say) a year are higher than locally relevant LRMCs, or if some suppliers are consistently bidding at levels above any plausible estimate of true market-clearing prices, it may be necessary to do something to mitigate market power. Even here, however, the best solution may be to stimulate additional supply or to negotiate contracts between the ITP and local suppliers, not to try to control prices or even bidding behavior directly. Even if it is decided to control bidding behavior, suppliers should not be required to bid at some low estimate of their simple MCs that will be used in the ITP's pricing process to keep prices below market-clearing levels during scarcity conditions. Something more flexible is required.

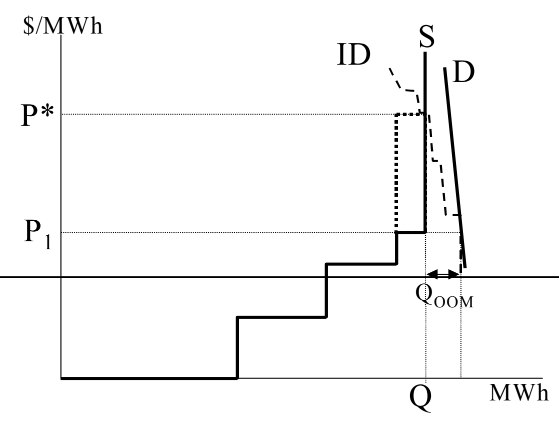
2.2.3 THE LOGICAL INCONSISTENCY IN AUTOMATED MPM PROCEDURES

Concerns about supplier market power in spot markets, particularly during scarcity conditions when unconstrained suppliers seem able to drive prices to arbitrarily high levels, have led to the development of automated MPM procedures (AMPs). The basic, if unstated, assumption behind these procedures is that it is easy to know how truly competitive suppliers would conduct themselves – they would all bid at some easy-to-determine estimate of their individual simple MCs – and to know what the resulting competitive spot prices would be – whatever comes out of the ITP's pricing engine when simple MCs are plugged in. Given this assumption, it is perhaps reasonable to say that any supplier conduct that departs significantly from the assumed competitive conduct and that causes a significant increase in the ITP-computed price is a successful exercise of market power, and as such should be prohibited or mitigated. Unfortunately for those looking for a simple solution to the problem of market power, this approach involves a fundamental logical inconsistency.

Figure 7 illustrates a scarcity situation in which market demand is D, the ITP's demand for in-market energy is ID, and the highest simple MC of any in-market supplier – call it a "peaker" – is P_1. An ITP's Independent Market Monitor (IMM) using a typical AMP procedure will set the peaker's reference bid price at P_1, on the assumption that this is what a workably competitive peaker would bid. But this assumption about the behavior of a competitive peaker is valid only if the peaker knows that the ITP will set the market price at the market-clearing level P* even if the peaker bids all its energy at price P_1. If the ITP's market does not use such a scarcity pricing process but instead responds to a peaker bid of P_1 by setting the price at P_1 and meeting demand with OOM energy, there is no reason to expect the peaker to bid P_1; in this case, the peaker would – indeed must, as a matter of commercial survival – bid well above P_1 under scarcity conditions. If the market is workably competitive in the sense that no supplier can profit much for long by increasing prices above competitive levels, the peaker would bid up to but not much beyond P*.

An AMP procedure does not ask how workable competitive suppliers would bid

<u>Figure 7:</u>
<u>The Logical Flaw in AMP Procedures</u>



given the actual ITP pricing process, but sets the peaker's reference price at P_1 because this is what the peaker would bid in the theoretical ITP pricing process. The AMP procedure then monitors the conduct of the peaker in the real market in which a competitive but rational peaker will bid at P* because that is what it must do to get the market-clearing price. If P* exceeds P_1 by more than some significant but arbitrary amount such as $100/MWh, the AMP procedure will view this bid as a violation of the peaker's conduct threshold and will then test whether the increase in the peaker's bid from P_1 to P* would increase the market price significantly. If the impact test were applied using the theoretical ITP pricing process it would find no market impact, because good ITP scarcity pricing would result in a scarcity price of P* whether the peaker bid P_1 or P*. But the AMP procedure tests for market impact using the real ITP pricing process that sets the market price at the peaker's bid and hence finds that the peakers bid has a significant impact on the market price. So the AMP procedure mitigates the peaker's bid down to P_1, the ITP sets the market price at P_1, and all suppliers are paid less than the market-clearing price they would have received if the ITP really did set scarcity prices the way the AMP procedure assumed when it set reference prices.

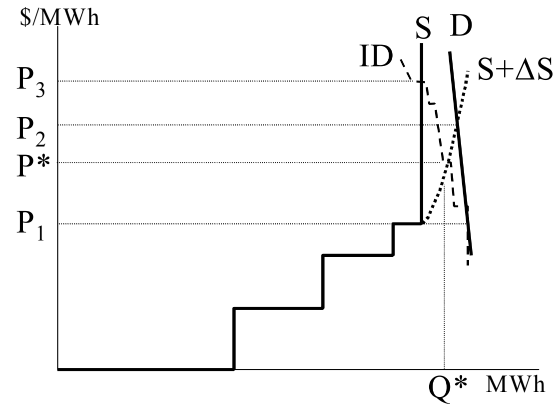
Setting reference bids on the assumption that the ITP uses good scarcity pricing and then mitigating bids to those reference levels even though the ITP does not use good scarcity pricing is logically inconsistent and has the practical effect of virtually guaranteeing that scarcity prices will be suppressed. The best way to fix this logical problem and the serious market distortions it creates is to require ITPs to implement much better scarcity pricing as discussed in the next section, and to allow suppliers to bid in ways that will produce scarcity prices near real market-clearing levels as long as that that is the only way to get such prices. This may be more difficult than simply applying mechanical but illogical AMP procedures, but – as in the old joke about the drunk looking for his keys under the street light – at least offers some hope of finding what market designers and regulators should be looking for.

2.2.4 Improving ITP Scarcity Pricing

Part of the solution to the problem of getting reasonable scarcity pricing must be to allow real suppliers to bid the way even "perfectly" competitive suppliers would bid in a centralized market-clearing process: by sculpting their bids to reflect the fact that the SRMCs of incremental supplies become very high near and beyond the rated full capacity of a generating unit. Such incremental sculpting does not require a supplier to guess the market-clearing price and then offer significant quantities at that price – a practice that is hard to distinguish from economic withholding – but only to come up with plausible estimates of its own SRMC for the last few percent of its available capacity. Such sculpted SRMC bids, perhaps subject to guidelines defining the quantities that can be offered at prices far above easy-to-estimate variable costs, could produce reasonable scarcity prices even if the ITP uses only supply bids to determine prices.

It is not easy to know just how fast and how far SRMC increases near and beyond "maximum" output, because this depends on judgmental factors such as the effects of stressing equipment and personnel, the value tomorrow of fuel or reservoir water not used or a boiler tube not replaced today, etc. Under normal, non-scarcity conditions it may not be worth worrying about such things, so the last increments of output may simply not be

Figure 8:
Sculpted SRMC Supplier Bidding



offered, or may be offered but not delivered if dispatched at a too-low price. But when scarcity conditions are likely, suppliers should be allowed, even encouraged, to submit sculpted bids reflecting their judgments about how their SRMCs increase at high output levels and these judgments should not be second-guessed as long as they are within specified guidelines.

For example, a generator with a nameplate capacity of 100 MW might offer 95 MW at its average fuel and variable operating costs of $30/MWh, but offer the next 5 MW at its full average cost or LRMC of (say) $100/MWh and each of an additional 5 MW as emergency energy at increasing prices reaching some high bid cap such as $5,000/MWh. Only the lower prices in such a sculpted bid could be easily related to objective measures of cost, but the higher prices for incremental output, while judgmental, are probably closer to real SRMCs than some simple measure of fuel costs would be. And offering incremental output at very high prices is better than not offering it at all – i.e., offering it at a price of infinity – which is usually the realistic alternative.[9]

Figure 8 illustrates how an ITP market could clear with sculpted supply bids. In addition to the in-market supply curve S representing the supplies available at "normal" supplier SRMCs, there would an incremental supply curve ΔS representing incremental amounts offered at higher prices. If the ITP used such sculpted supplier bids in the natural way, under normal conditions little or nothing would change, because the market would clear at low prices with no need for the ΔS supplies. But under scarcity conditions, as represented by the demand curve D, the ΔS supplies would be used and, if there was no other response, would set the price at P_2.

Even if suppliers are providing sculpted in-market supply bids the ITP should still use demand bids and/or OOM actions when these are cheaper. The ITP should create its demand curve for in-market energy, ID in Figure 8, call the indicated actions when they are in merit – and then set the market price at the bid or implicit cost of the marginal supply, demand reduction or OOM action taken. In Figure 8, if the ITP used OOM actions without sculpted

9 Some generators, particularly combustion turbines (CTs), have little flexibility in operations but are essentially either on or off. This creates serious logical and practical problems for the ITP's dispatch and pricing process and hence for the owners of CTs. For example, when a CT comes on some other generator with lower SRMC may have to reduce its output; in such situations there is no simple definition of the system's SRMC or the market-clearing price. Special pricing and payment rules must be created to deal with such situations. These are very important, particularly for owners of CTs, but are beyond the scope of this paper.

supplier bidding, the market-clearing price would be P_3; if it used sculpted supply bids without taking OOM actions the market-clearing price would be P_2; but if it did both the market-clearing price would be P*, which is necessarily less (or at least no greater) than either P_2 or P_3.

One of the principal advantages of allowing sculpted bidding by suppliers is that it would stimulate the ITP to look for lower-cost ways to clear the market, i.e., to develop a demand curve for in-market supplier energy that is more elastic. Of course, the ITP could allow high-priced bids for incremental or emergency supplies but then treat these as OOM energy that cannot set the market price; this would give the ITP more OOM options and hence would make the actual dispatch more efficient and reliable, but would do little to improve the efficiency of market pricing.

Letting sculpted SRMC bids from suppliers set prices under scarcity conditions is conceptually no different than letting demand bids from consumers or LSEs set prices – everybody's favorite solution to the scarcity pricing problem. If and when demand bidding is used to set scarcity prices, consumers will not be expected to have objective justifications for the specific prices at which they will turn off air conditioners, shut down industrial processes or run back-up generators; it is understood that these price levels are based on complex and often subjective judgments about costs, benefits and risks. In principle, suppliers should have much the same freedom when it comes to offering incremental supplies near and beyond full capacity, although in practice concern about supplier market power may require limits on how much sculpting is allowed.

2.2.5 The Effects of Suppressing Scarcity Prices

Suppressing scarcity prices below competitive, market-clearing levels does consumers no good and even harms them in a long-run, expected-value sense, because demand and supply must be matched somehow and all the non-market ways of doing so are more costly than letting the market work. In fact, even if prices during scarcity periods are somewhat too high, reducing them to just "the right" levels will do consumers little good in the long run as long as entry or the threat of it will keep average spot prices close to LRMCs. Given the impossibility of knowing when and by how much scarcity prices are too high and the low pay-off and high risks involved in trying to reduce them just enough, it is probably better not to try unless there is clear evidence that average prices are materially too high.

If scarcity prices are kept below competitive market clearing levels, the ITP has two basic options for matching supply to demand during scarcity periods: (1) close the supply-demand gap with OOM actions that do not set spot market prices and recover its costs with an uplift or tax on all consumers; or (2) compel or subsidize excess capacity so that spot markets will clear at lower prices, recognizing that the costs of this excess capacity will ultimately be paid by consumers. The capacity requirement/payment option is discussed in section 2.2.7 below. The option of subsidizing OOM energy in the spot market is discussed here.

If there are no significant barriers to entry into or exit from generation – and no capacity payments, as assumed throughout this section – in a long-run, expected value sense average spot scarcity prices must equal the LRMC of peakers and the time-average of scarcity and normal (i.e., non-scarcity) prices must equal the LRMC of non-peakers.[10] Notice that spot prices averaged over the relevant time periods must equal suppliers' LRMCs, not their average SRMCs or simple MCs. In theory, a proper ITP pricing process will produce LRMC-level prices when there is an optimal amount and mix of capacity, because sculpted bidding, demand bids, OOM actions, etc., will be used to set scarcity prices above suppliers' bids.

If an electricity system begins with a reasonably economical amount and mix of supply resources but the ITP's pricing process does not produce efficient and compensatory scarcity prices, all suppliers, not just peakers, will lose money and will stop maintaining or adding supply. Scarcity will increase until eventually even flawed ITP processes will produce more energy, ancillary service and OOM revenues for the remaining suppliers. Total costs will be higher and reliability may suffer because prices are so distorted, but unless the electricity sector is allowed to disappear or is subsidized, in the long run consumers' total bills will have to cover total costs.[11]

If the ITP uses OOM actions to keep spot prices below market-clearing levels during scarcity periods, the reduction in spot scarcity prices will be less than the uplift needed to pay for OOM actions and consumers will be better off during scarcity periods, at least in the short run. Peakers will not be commercially viable based on market revenues, so they will either become OOM suppliers or be replaced by OOM actions (perhaps including load interruptions). Non-peakers are not getting OOM payments but in the long run must still cover their total costs, so prices in non-scarcity periods must be high enough to make ***average*** spot prices equal the LRMC of non-peakers – just as they would be if the ITP were not suppressing scarcity prices. In a long run, expected value sense, suppressing spot scarcity prices by subsidizing OOM energy during scarcity periods leaves average (wholesale) prices essentially unchanged and requires an uplift or tax on consumers; on balance, consumers are worse off.

Even if market power during scarcity periods is consistently pushing spot prices during such periods too high – i.e., above the LRMC of peakers – the benefit to consumers of reducing the too-high scarcity prices is small as long as entry by non-peakers can be counted on to keep ***average*** prices near the LRMC of non-peakers. Reducing prices during scarcity periods to the LRMC of peakers will reduce the margins of non-peakers and make investment in

---

[10] A peaker here is any generator whose SRMC and even simple MC is so high that it earns net revenue or operating profits from energy sales only during scarcity periods, although it may earn significant amounts by selling ancillary services and reserves at other times. A non-peaker is any generator that earns significant net revenue or operating profits – or scarcity rents – outside scarcity periods.

[11] Whether the electricity sector just has costs that are a little too high or experiences massive supplier bankruptcies during the transition and on-going reliability problems depends on how bad the ITP's pricing processes are. If an ITP market works too badly it will presumably be fixed – or the whole idea of competitive electricity markets will be scrapped – before the whole system declines into some "long-run equilibrium" with high costs and poor reliability.

them unprofitable or even force some of them to shut down. As demand grows, non-peakers will not be built until increasing scarcity during non-peak periods brings higher-cost suppliers to the margin more often and pushes time-average prices back to the LRMC of non-peakers. During the transition consumers may see lower average prices for awhile and suppliers will have financial problems, but in the long run average prices will be right back where they were before scarcity prices were reduced.[12]

Given all this, and the difficulties of finding and maintaining Goldilocks prices, an aggressive policy of capping or mitigating spot prices whenever market power is suspected can easily do more harm than good. Unless average spot prices are staying above LRMCs even when new supply is not needed, and particularly if average spot prices are below LRMCs when new supply is needed – as seems to be the case in ITP markets these days – suppliers are not benefiting from any exercise of market power and consumers would not gain in the long run from any mitigation of supplier behavior. If MPM procedures discover that some suppliers sometimes bid above their simple MCs and that doing so increases some scarcity prices, this demonstrates only that such bidding is necessary to get prices to competitive market-clearing levels given the way the ITP market actually operates. There is no reason to mitigate such bidding behavior in a real ITP market just because such behavior in a theoretically perfect ITP market might – or might not – be an exercise of market power.

2.2.6 The Problem of Load Pockets

The analysis above concludes that, where there are no significant barriers to entry, it is probably better to err on the side of under-mitigating occasional high spot prices than to create an aggressive MPM program with its associated costs, uncertainties, distortions and probably suppressed scarcity prices. But in a transmission-constrained load pocket there may – or may not – be barriers to entry so there may – or may not – be better reasons to worry about and to mitigate the exercise of market power.

A load pocket is defined here as an electrical area within which (or perhaps a single node at which) prices are or could be higher than they are outside the load pocket because of transmission congestion.[13] The mere existence of a load pocket does not mean that market power within the load pocket is necessarily a problem or should be mitigated, because there may be adequate competition within the load pocket to keep spot prices at competitive levels – even if those competitive spot prices are unusually high during constrained periods or even on average over time.

---

[12] The rebalancing of prices may yield some efficiency gains. For example, lowering scarcity prices to suppliers' SRMCs will allow incremental supply to replace more-costly demand reductions during scarcity periods, and increasing non-scarcity prices may prevent some waste of energy during non-scarcity periods.

[13] A load pocket is sometimes defined as an area within which there is or is likely to be market power when transmission constraints are binding. The definition used here is more general and is used to stress the fact that things other than market power that can cause high local prices.

If there are no significant barriers to new suppliers entering a particular load pocket, there is no more reason to worry about market power or its mitigation in occasional scarcity periods here than anywhere else. Even if suppliers within the load pocket could somehow keep prices "too high" during constrained periods indefinitely, entry or the threat of it would drive prices in unconstrained periods down until time-average prices equaled LRMCs. In this case there would be some efficiency losses because prices were too high during constrained periods and too low in other periods, but these costs would probably be less than the costs of an aggressive MPM procedure – particularly given that such a procedure may create the very disincentives to entry that should be eliminated so that prices can adjust to competitive levels in the long run.

If it is unusually expensive to build generation within a load pocket because of fuel supply, siting or environmental problems, LRMCs are higher in the load pocket than elsewhere and hence average market prices and SRMCs should be higher, at least until constraints on transmission, fuel, sites or environmental impact are relieved. There may be many, fiercely competing suppliers in and free entry of electricity suppliers into the load pocket, and yet average spot prices may stay above the full cost of new supplies elsewhere or even above the apparent cost of new supplies in the load pocket for a long time. But it is the scarcity of sites, permits, etc., not supplier market power, that is to blame, and the only real solution is to increase the supply of these scarce assets. If existing suppliers in such a situation appear to be making "too much" money without attracting entry, it is because their excess profits are more properly regarded as scarcity rents on the sites, fuel sources or environmental permits they own, not because they are exercising market power in the spot market.

When it comes to spot pricing, a load pocket is just like anywhere else except that scarcity prices arise more frequently and perhaps rise to higher levels. The best way to get reasonable scarcity prices in a load pocket or anywhere else is for the ITP to allow sculpted supplier bids, price-responsive demand and the deemed bids of OOM actions to set market prices, as outlined above. Until such ITP scarcity pricing is working effectively, suppliers should be expected and allowed to bid some quantities at their estimates of scarcity prices within some limits, even when this is well above their simple MCs, because that is the only mechanism the market has for getting reasonable scarcity prices. Bidding to get prices to competitive levels is not an exercise of market power or a sign the market is not workably competitive unless average spot prices remain above LRMCs when more capacity is not needed.

The only situation in which a load pocket creates serious potential market power problems is where there are only a few independent suppliers within a load pocket, or a single supplier at a node, and entry is difficult because the local market is just not big enough there to support more competitors. In such situations the incumbent supplier(s) may be able to set local prices much of the time and, indeed, may have to bid well above their own simple MCs much of the time just to make enough money to cover their costs. Even here, however, the best solution is to negotiate or impose contracts that assure the needed supplier(s) that they will cover their costs and that remove the incentives to create or maintain artificial scarcity, not to try to keep spot prices from reflecting actual scarcity.

2.2.7 CAPACITY PAYMENTS AS SUBSTITUTES FOR HIGH SCARCITY PRICES

Despite the economic arguments for letting, indeed assuring that, spot prices rise to market-clearing levels even when these levels are very high, doing so is difficult both technically and politically. As a result, most ITP markets in the United States[14] use some form of installed capacity/capability (ICAP) requirement to try to reduce reliance on spot prices, particularly during scarcity periods. FERC's SMD NOPR proposes, as one of its mandatory MPM measures, a resource adequacy requirement (RAR) that, in effect, tries to find a "third way" between reliance on spot scarcity pricing and ICAP-like mechanisms. FERC's RAR proposal is discussed in detail in section 3.2.3 below. This section discusses the basic options for and implications of any such mechanism.

The basic mechanism of any resource[15] payment mechanism is a requirement that somebody – usually LSEs in the U.S. context – contract with or otherwise pay enough resources to meet peak demand plus some reserve margin. If such a requirement is effective, resources will receive resource payments in addition to payments for energy and ancillary services in spot markets. These additional payments will stimulate more resources to be available than would otherwise be there – in effect, shifting the SRMC curve outward along the LRMC curve in Figures 1 and 2 – and hence spot prices will be lower, especially at times of scarcity. Unless the resource payments and resulting excess capacity are very large, scarcity pricing will still be needed at times and hence the ITP should develop some version of the scarcity pricing process outlined above. But the market should now be able to clear more easily and at lower prices during scarcity periods.

There are three principal choices to be made in designing a capacity payment mechanism: (1) should the ITP require LSEs to contract bilaterally for resources or do the contracting itself and allocate its costs to LSEs; (2) where on the spectrum between "short term" (e.g., one day) and "long term" (e.g., five years) should the time horizon be chosen; and (3) should enforcement be *ex ante* based on forecasts or *ex post* based on what actually happens? Not all combinations of these choices are logical or workable, but any workable resource requirement/payment regime uses some combination of these options. The numerical parameters of the requirement or payment can be set so that the non-spot-market payments are higher or lower, resulting in more or less additional capacity and impact on spot prices.

If each LSE is required to contract bilaterally for resources, the ITP must define the MW requirement for each LSE, define criteria for resources that can meet this requirement, and enforce the requirements and criteria. If the ITP contracts directly with the resources itself, it must pick the specific resources, negotiate and enforce contracts with these resources, and allocate the costs to LSEs. Most existing ICAP programs and FERC's RAR rely on bilateral

[14] Capacity requirements are much less common outside the United States. For example, there is no such requirement in the ITP-like markets in England and Wales, Australia, New Zealand, Norway, Alberta or Ontario.

[15] The term "resources" includes both generation capacity and demand-reduction capability.

contracting by LSEs rather than centralized procurement by the ITP, but the latter is being considered in some ITP markets.

Existing ICAP programs put a short-term resource obligation on LSEs with both definition and enforcement of the obligation *ex post*. Each LSE's MW ICAP obligation is based on its peak demand during (say) a month as determined at the end of the month and any ICAP deficiencies or surpluses are priced in an ICAP market at the close of the month.[16] Defining and enforcing the obligation monthly and *ex post* makes it compatible with retail competition but may do little for investors or planners looking for assurance of supplies and revenues years ahead, or for consumers looking for long-term stability in monthly bills.

Long-term assurance of supply and long-term stability of monthly bills require multi-year resource commitments. In principle, the ITP could define the total resources needed (say) three years in advance and then require each LSE to contract or pay the ITP for its assigned share of that total three years in advance; but this is probably not realistic in a world of retail competition, where LSEs cannot forecast their loads years in advance.

With any of these (or other) alternatives, the size of non-spot resource payments and hence the impact on spot prices will depend on parameters such as the level of reserve margins and the enforcement penalties. These parameters could be adjusted to increase reliability and reduce spot scarcity prices as much as anybody could ever want – or more. For example, a monthly ICAP requirement with a 100 percent reserve margin would induce suppliers to make a lot of capacity available, even though they could lose money if they collectively built so much of it that both spot energy prices and monthly ICAP prices fell. Reliability would be high and spot prices would be low, but total costs would be high and the ITP's criteria, enforcement and cost allocations would be a major force in the market. The monthly reserve margin could be adjusted until an acceptable balance were found between high reliability and low spot prices on the one hand, and high costs and a strong ITP role on the other.

The fact that any effective resource requirement/payment creates some problems and inefficiencies does not mean that such a resource requirement/payment should never be used. The real world is full of unpleasant choices, and if high scarcity prices are politically unacceptable or are unlikely because of excessive MPM, some form of resource requirement/payment is required. The point here is that nobody should think that any such scheme is a costless way to keep scarcity prices down or to make up for suppressing them.

16 *Ex post* ICAP markets have proven to have highly volatile prices and have been accused of being susceptible to the exercise of market power. The basic problem is that the demand for ICAP in such a market is a simple multiple of peak demand in the month just ended while the supply of ICAP in the market may depend totally on equipment ratings and tests months earlier. With both demand and supply curves essentially vertical, the price is either zero or whatever price cap is imposed.

2.2.8 THE ROLE OF DEMAND RESPONSE

There is widespread agreement that electricity demand should and could be more price-elastic in the short run than it is, that ITPs should incorporate demand bidding into their dispatch and pricing processes, that there should be more use of time-of-use meters and demand management technologies, and that more consumers should be exposed to spot prices on the margin. But the fact that demand is not elastic "enough" is no reason not to let spot markets clear at scarcity levels, particularly when there is no way to know how much demand elasticity is "enough" and there are ways to clear spot markets even if market demand is not highly price-elastic. In fact, until spot prices reflect the true supply-side costs of meeting inelastic demand there will be too little incentive to make demand more price elastic.

Even in the absence of explicit demand bids from LSEs and large consumers, at least some demand will respond to expected spot prices, and the ITP should estimate and use this elasticity in forecasting demand in its dispatch and pricing processes. Interruptible load contracts imply some incremental cost of calling for an interruption,[17] and emergency actions such as taking energy from operating reserves, letting frequency or voltage drop or even shedding some load are logically high-cost demand reductions. In principle, all such actions should be given deemed bid prices and treated as demand bids – or supply bids; it can be done correctly either way – in the ITP's dispatch and pricing processes, along with sculpted supplier bids and the deemed costs of OOM actions.

3. FERC'S MARKET POWER MITIGATION PACKAGE

In the Notice of Proposed Rulemaking (NOPR) describing the SMD to be implemented by ITPs, FERC proposes a package of four MPM measures, three mandatory and one voluntary. The NOPR provides few details about these measures, but the general MPM framework proposed there will have a strong influence on the development of MPM procedures in all electricity markets, even those that may not adopt the full SMD.

3.1 OBJECTIVES OF THE FERC MPM PACKAGE

The SMD NOPR says that "the development of structurally competitive markets is the Commission's long-term goal," but that "at this stage of the industry's evolution" wholesale markets are not yet structurally competitive primarily because of "two significant structural flaws" in electricity markets: (1) "the lack of price responsive demand;" and (2) "generation concentration in transmission-constrained load pockets." [SMD NOPR, para. 390] This statement is significant primarily because it suggests both that FERC would be more relaxed about market power if ITPs were to take some of the steps suggested above to make demand

[17] A typical interruptible load contract may not specify an explicit cost to be paid when the load is interrupted but limits the number and/or cumulative effect of interruptions. Such a contract implies that interrupting load now has an opportunity cost, analogous to taking water from a reservoir.

for in-market energy more price responsive, and that FERC's current concerns about market power are limited to load pockets in which generation is so concentrated that workable competition is threatened.

The SMD NOPR also says: "the challenge for market power mitigation on the supply side is to assure that it allows long-term competitive prices, which allows the opportunity to recover the fixed costs of the investment as well as the short-term variable costs … If some degree of scarcity pricing is not allowed, … then some generators needed for reliability could fail to recover their full costs and may be retired. Worse yet, prices could be held so low that investors decline to invest … because they do not see a reasonable expectation of recovering their costs." [SMD NOPR, para 393] Thus, "the market power mitigation plan should be calibrated so that it does not inefficiently suppress prices, or mask scarcity prices, providing the wrong economic signals for efficient investment or demand response." [SMD NOPR, para. 397] These quotes imply that MPM should not suppress scarcity prices so much that average spot prices are less than LRMCs when and where investment is needed.

FERC's MPM mitigation package includes a mandatory resource adequacy requirement (RAR), justified with the statement that "the spot market does not yet work well to produce long-term reliability investment, even without price mitigation," because it takes a long time to build power plants and demand does not respond to price. FERC also justifies the RAR on the grounds that "market power mitigation may tend to suppress the scarcity price that would otherwise stimulate new resource development [and] as a result, investors may not develop adequate infrastructure …unless there is a provision for resource adequacy." [SMD NOPR, para. 468].

The most logical interpretation of the quotes above is that FERC does not want MPM to suppress scarcity prices below competitive levels but recognizes that this could happen and hence proposes the RAR to assure long-term resource adequacy given FERC's belief that the spot market is unlikely to do so. This is fine as far as it goes, but does not go far enough to explain how competitive prices should be defined and assured, or how the RAR can be made workable and effective.

3.2 Description of FERC's MPM Package

FERC's MPM package consists of the following four measures, the first three mandatory and the fourth voluntary.

3.2.1 Local Market Power Mitigation

Mandatory generator operating agreements (GOAs) between generators and the ITP will be used to mitigate local market power created by either "persistent and foreseeable" or "sporadic" transmission congestion. [SMD NOPR, para. 406] Each GOA will define the transmission system conditions under which that generator will be deemed to have local market power, and in those conditions the generator must offer all its available energy to the market either by scheduling it under bilateral contracts or by offering to sell it in the spot markets subject to a unit-specific bid cap.

3.2.2 SAFETY-NET BID CAP

No supplier will be allowed to offer energy to ITP spot markets at prices exceeding an ITP-specified safety-net bid cap, such as $1,000/MWh, said to be necessary because of inadequate price-responsive demand. FERC says that imports offered at higher prices (subject to some higher bid cap) could set the market price above the safety-net bid cap, but also acknowledges that some OOM suppliers may receive higher payments without setting the market price. FERC does not say whether higher demand bids or (for example) the deemed cost of taking energy from operating reserves could set prices above the safety-net bid cap, but does not rule this out and even includes just such a mechanism in determining the spot-market compliance penalty in its RAR.

3.2.3 RESOURCE ADEQUACY REQUIREMENT (RAR)

The RAR is described and analyzed in section 3.3.2 below. FERC regards its RAR as a MPM measure for two reasons: (1) Because the RAR is intended to induce LSEs to contract with generators for their full load and if they do generators will have less incentive to exercise market power in the spot market; and (2) because FERC recognizes that MPM may suppress scarcity prices so some other source of revenue for generators is required. As discussed below, something very different than the RAR outlined by FERC will have to be developed to accomplish what FERC wants.

3.2.4 MITIGATION TRIGGERED BY MARKET CONDITIONS

This fourth, voluntary MPM measure would presumably be some version of the automated mitigation (AMP) mechanisms "such as those approved by the Commission in New York ISO and California" [SMD NOPR, para. 231]. The MISO MPM proposal discussed in detail in section 4 below is the sort of thing FERC seems to have in mind.

FERC says this "measure, if needed, would apply to unanticipated and sustained market conditions that would give the ability and the incentive to exercise market power. For example, extreme supply or demand conditions to which the market cannot quickly adapt, such as … drought, or … a major transmission line outage. These kinds of events, which are not transitory, can provide opportunities to exercise market power even in a market that is normally workably competitive." [SMD NOPR, para. 415] Furthermore, "since this form of market power mitigation is for temporary market conditions, it will be equally important … to indicate the criteria to determine when the market has returned to normal competitive conditions and this market power mitigation method will be suspended." [SMD NOPR, para 416]

The quotes above would appear to limit the application of any AMP-like mechanism to unexpected, sustained but temporary periods of shortage, such as occurred in western power markets in 2000-2001. However, FERC also says that "it may be appropriate for other conditions to trigger this mechanism" and invites comment on what these triggers should be, suggesting that an ITP could ask FERC to apply such measures more broadly. [SMD NOPR, para. 415]

As to what this fourth measure is intended to accomplish, FERC says: "Although market-clearing prices would be expected to rise in these situations [to which this measure might apply], and perhaps sharply and significantly, … the market [may want] … the assurance that the price increases are attributable to the extreme circumstances and not to the exercise of market power." [SMD NOPR, para. 415] Again, the clear intent seems to be to assure that prices do not go significantly above competitive scarcity levels, but without suppressing prices below competitive scarcity levels in the short run and LRMC levels in the long run. Just how any MPM procedure is supposed to find this fine line, or how successfully it will do so, is not discussed.

3.3 Analysis of the FERC Proposals

The FERC MPM proposals are not defined in enough detail to allow detailed analysis of them, and until FERC decides what it will and will not approve it is not possible to say much about the likely effects. However, based on the general analysis above and on existing MPM procedures, including those approved by FERC in existing ITPs, the following general observations can be made.

3.3.1 The Effects on Scarcity Pricing

The safety net bid cap would apply at all times and places and would probably be set at something like $1,000/MWh. This level is low enough that it will suppress prices below competitive market-clearing levels during serious scarcity conditions, at least until ITPs implement scarcity pricing methods that can produce prices higher than any supplier's bid. FERC suggests that imports could set market prices higher than this bid cap but does not strongly endorse the development of other scarcity pricing mechanisms.

The measures for mitigating local market power included in generator operating agreements would be applied wherever "persistent and foreseeable" or "sporadic" transmission congestion creates local market power, which could be almost anywhere. The voluntary, AMP-like measure could also be interpreted to require widespread, even universal testing of market conduct in order to determine when and where some special market conditions are creating "unanticipated," "sustained" but "temporary" market power. Given that existing MPM procedures often misdiagnose and mitigate as exercises of market power the competitive market behavior that is needed to produce reasonable scarcity prices in today's ITP markets, widespread application of these tests and mitigation measures in search of market power to mitigate will tend to suppress scarcity prices.

3.3.2 The Resource Adequacy Requirement

In its introduction to the RAR proposal in the SMD NOPR, FERC asserts that spot prices – particularly when mitigated but even when not – cannot be relied upon to assure long-term resource adequacy, and criticizes existing ICAP arrangements, partly because of "concern about the[ir] effectiveness" and partly because they require the ITP to "play a strong role … that may not suit regions without a history of tightly coordinated reserve sharing." [SMD NOPR, para. 483] FERC then proposes a RAR that appears to be trying to find a "third way" that is neither scarcity pricing nor ICAP when, in fact, some versions of these are the only two logical alternatives.

The essential features of FERC's RAR proposal are the following:

- **Long-Term Resource Planning:** Each ITP, in cooperation with a Regional State Advisory Committee, use long-term – FERC says three-to-five years, but call it three years for discussion purposes – regional demand forecasts and resource projections, and a reserve margin of at least 12 percent, to "establish the appropriate level of resource adequacy for the region." [SMD NOPR, para. 491]

- **Allocation of the RAR Among LSEs:** The ITP will allocate the regional RAR among LSEs based on their individual peak demands; whether these should be actual past demands or forecast demands is one of the things on which FERC invites comments.

- **Resources That Can Satisfy the RAR:** Only contracts that meet specified criteria and require physical performance by real, identifiable resources can satisfy the RAR. The ITP will establish the criteria for resource performance and ISO contracts.

- **Enforcement of the RAR:** No LSE will be penalized for not contracting in advance for all (or any) of its allocated resource requirement. Instead, an LSE that does not meet its RAR in one year will be penalized three years later if it buys anything in the ITP's spot markets. The penalties will include a punitive surcharge over and above the spot price on spot purchases, plus an attempt to curtail the customers of deficient LSEs first if demand curtailments are necessary.

If FERC's RAR proposal were implemented as now proposed, it might create a long-term planning process but would provide no mechanism for implementing the resulting plan because the enforcement provisions are hollow – which would make the long-term planning process largely meaningless. Consider the enforcement provisions:

- **Curtailment of Customers:** The threat to curtail first the customers of LSEs who did not contract "enough" three years ago would be very difficult even to define and impractical to implement. With reasonably efficient wholesale markets the probability of involuntary curtailments is very small and with retail competition there is no practical way to target curtailments on the customers of specific LSEs.
- **Spot Market Penalties.** The threat to impose penalties on today's spot market purchases by LSEs who were deficient three years ago is very easy to avoid but creates real-time inefficiencies. Even if no LSEs contracted for anything three years ago, they can all contract enough sometime in the intervening three years (including yesterday) to avoid buying and paying penalties in the ITPs's spot markets. But whatever their contract portfolio of long-term and short-term contracts, each LSE must avoid taking either more

or less than the specified contract quantities because contract imbalances are penalized.[18] The effect would be a form of contract dispatch that distorts real-time dispatch and consumption, increasing system and LSE costs.

The RAR as proposed in the SMD NOPR does little to assure long-term resource adequacy, because its enforcement is all based on what happens in short-term and real-time markets. Its penalty on spot purchases could be regarded as a way to get effective spot prices above some *de jure* or *de facto* cap on spot market prices, except that the spot-price-plus-penalty applies only to any LSEs who were/are deficient both three years ago and now, not to other LSEs or to suppliers. Any LSEs unlucky or incompetent enough to be in that situation will face incentives to reduce demand that are much higher than the demand-reduction incentives faced by other LSEs and much higher than the prices being paid to suppliers.

The most interesting part of the RAR proposal is that the penalty on spot purchases "would increase in stages as the shortage becomes more severe. For example, the penalty price could be $500 (in addition to the spot market energy price) when operating reserves are just below the minimum level, $600 when operating reserves are more than below 1 percent below [*sic*] the minimum level, $700 when operating reserves are more than 2 percent below the minimum level, and so on." [SMD NOPR, para. 530] Something similar should be used to determine a market price that applies to all LSEs and suppliers, not just a penalty that inefficiently and unfairly applies only to a few LSEs.

Long-term resource adequacy in electricity requires either efficient spot pricing, including scarcity prices that can be very high and are not subject to stringent caps or aggressive mitigation, or an effective resource requirement/payment mechanism to make up for the loss of scarcity prices due to price caps and mitigation. FERC has disparaged both of these options, but its proposed third way does not work; it will have to decide what combination of the two disparaged options it will support. Several such combinations are feasible, but will require accepting the reality of high scarcity prices, an effective capacity requirement/payment defined and enforced by a strong ITP, or some combination of the two.

4. THE MISO MARKET POWER MITIGATION PROPOSAL

The MISO is considering a MPM proposal being developed by Dr. David Patton, the MISO Independent Market Monitor (IMM) and the Independent Market Advisor to the New York ITP. This MPM proposal is similar to the AMP process used in New York and elsewhere, and is a relatively detailed example of the fourth, voluntary type of MPM measure proposed in FERC's SMD NOPR.

This section discusses the MISO MPM proposal, as described in a draft dated October 3, 2002 (the "MISO Draft"),[19] in terms of the economic principles outlined in section 2 above.

[18] It is unclear whether contract imbalances would be assessed on each of a LSE's individual contracts or on its portfolio. The former increases the operational inefficiencies and the latter gives an uneconomic competitive advantage to large LSEs.

The focus is on the definition and enforcement of the proposed prohibitions on physical withholding and bid caps for suppliers, because these are the most important features of the proposal. Silence on or cursory treatment of any feature of the MISO MPM proposal here means only that such feature has not been analyzed in detail.

This section also discusses a proposal by Dr. Patton that the energy price in the MISO spot market be very high when the ITP must take energy from operating reserves. This proposal and the reasons given for it, as well as similar recommendations that Dr. Patton has made to the New York ISO, are consistent with the ITP scarcity pricing recommendations made in section 2.2.4 above and are encouraging steps in the right direction, but would not solve all the problems with ITP pricing.

4.1 OBJECTIVES OF THE MISO MPM PROCEDURES

The MISO Draft stresses that the intent of the proposed MPM procedures is to “mitigate the market effects of any conduct that would substantially distort competitive outcomes” in the MISO markets, “while avoiding unnecessary interference with competitive price signals” [MISO Draft, sec. 1(a)] and while allowing “prices to rise efficiently to reflect legitimate supply shortages.” [MISO Draft, sec. 1(b)] These assurances that MPM will not prevent prices from rising to scarcity levels are particularly critical given that there are apparently no plans to include a capacity requirement or payment in the MISO market, at least initially. Unfortunately, the MPM procedures outlined in the MISO Draft are unlikely to deliver on this promise.

The operational objective of the IMM in the MISO MPM proposal is to “remedy conduct that: (1) is significantly inconsistent with competitive conduct; and (2) would result in a substantial change in one or more prices” or production guarantee payments[20] in a MISO market [MISO Draft, sec. 2.3(a)]. This objective is noteworthy primarily for its focus on conduct that is “inconsistent with competitive conduct” as defined by the procedure itself and that causes “substantial change[s]” in prices in MISO spot markets however flawed the MISO pricing processes may be. As discussed in section 4.4.2 below, this inward focus on the procedure itself should be (at least) supplemented by efforts to evaluate and calibrate the MPM procedure in terms of its broader impacts on market outcomes, i.e., on average spot prices compared to LRMCs.

---

[19] Attached as Appendix A.

[20] Production guarantee payments are made to generators to cover the non-energy start up and other costs of a dispatched generator when spot market payments are inadequate to cover all such costs. The MISO proposal includes both conduct and impact thresholds related to these non-energy payments, but these are not discussed here.

4.2 Description of the MISO Proposal

The basic approach to MPM in the MISO Draft is the same approach used in functioning ITPs, endorsed by FERC and outlined in section 2.2.3 above. The specific definitions and procedures proposed for the MISO are summarized in this section.

4.2.1 Transmission Constrained Areas

The IMM will, at least once a year, identify any Narrow Constrained Areas (NCAs) within the MISO. A NCA is defined as an electrical area in which transmission congestion, however often it arises, can be "relieved" only by resources owned or controlled by (initially) "less than three" suppliers. [MISO Draft, sec. 2.5.1] This language suggests that the IMM will be able to draw lines around well-defined areas that are sometimes isolated from the rest of the grid by transmission constraints and count the suppliers within that area, but in practice things are seldom that simple. Many NCAs are likely to be a single node at which a single generator can sometimes increase the LMP with its own supply bid.

A Broad Constrained Area (BCA) is defined as any electrical area in which the conduct and impact thresholds defined below are violated. This language suggests that there is some distinction between a BCA and "the rest" of the grid, but in fact suppliers everywhere will be subjected to the same conduct and impact thresholds – except those in NCAs, which are subject to more stringent thresholds – and those who violate their thresholds in any given day or hour will have their bids mitigated. The intent and effect of the MISO MPM proposal would be clearer if the term "Broad Constrained Area" were dropped and replaced with some version of "anywhere" or "everywhere" wherever it appears in the MISO Draft.

4.2.2 Bidding Reference Levels:

Reference levels will be set in advance for each component of a generator's bids, such as the energy price ($/MWh), start-up costs ($/start-up), minimum run times (hours), etc., at levels "intended to reflect a resource's marginal costs, including legitimate risk and opportunity costs." [MISO Draft, sec. 3.1.4(a)] Reference levels for each generating unit will be determined by the IMM using one or more of several methods: past bids or prices during periods when the unit is presumed to have been bidding its "true" SRMC, i.e., when bids or prices were low and the unit was dispatched; negotiations with the unit operator; IMM estimates of costs; or comparison with competitive bids from similar units.

Reference levels for energy prices ($/MWh) can vary with energy output "over the output range of the resource," but there is no indication that sculpted SRMC bids reflecting a generator's judgment about the very high costs and risks of the last few MW are contemplated or would be allowed. The IMM will inform each market participant of the reference levels applicable to its units and market participants may request changes in reference levels. The MISO will have dispute resolution procedures and market participants can ultimately appeal to the FERC if they disagree with the reference levels determined by the IMM.

4.2.3 Conduct Thresholds

The bid parameters and offered capacity of all suppliers will be tested to see if they violate conduct thresholds for economic withholding or physical withholding. There are minimally stringent conduct thresholds that apply everywhere and additional thresholds that apply to a supplier in a NCA when they more stringent than the minimally stringent thresholds. The conduct thresholds relevant for the discussion here are:

- **Conduct threshold for economic withholding**: A supplier violates its economic withholding threshold if its bid parameters exceed its reference levels by more than:
  - For an energy bid (energy bids less than $25/MWh are exempt): The lesser of 300 percent or $100/MWh or (for a supplier in a NCA when constraints are binding) a "NCA Threshold" discussed further below and defined as:

    NCA Threshold = (Net Annual Fixed Cost)/(Constrained Hours)
  - For a start-up cost bid: 200 percent, reduced to 50 percent for a supplier in a NCA
  - For time-based parameters (e.g., minimum down-time): 3 hours for any single parameter or 6 hours in total for all parameters
  - For bid parameters expressed in units other than time or dollars: A 100 percent increase for parameters that are minimum values, or a 50 percent decrease for parameters that are maximum values.
- **Conduct threshold for physical withholding:** Physical withholding is defined to include not-offering a unit's "capability" without good reason, economic withholding as defined above, and operating a unit at less than a dispatched level. For suppliers within a NCA, any physical withholding, however small, is a violation of the threshold. A supplier anywhere outside a NCA violates its physical withholding thresholds if it:
  - Withholds more than the lesser of 10 percent or 100 MW of a unit's capability;
  - Withholds more than the lesser of 5 percent or 200 MW of its total capability; or
  - Operates at less than 90 percent of a dispatched output level

4.2.4 Market Impact Thresholds:

If a supplier violates any applicable conduct threshold above, the impact of the violating behavior on market outcomes will be tested, presumably using procedures similar to those used elsewhere by Dr. Patton. In these procedures, the IMM will use the MISO's market rules and models to determine what market outcomes would have been if everything in the market remained the same except that all bid parameters that violate a conduct threshold are simultaneously[21] mitigated to their reference levels, e.g., to simple MCs for energy bids. A supplier that is in violation of any of its conduct thresholds is also in violation of the market

[21] An alternative interpretation is that each bid parameter that violates a conduct threshold would be tested individually for its market impact, leaving all other such bid parameters at their as-submitted, unmitigated levels. This interpretation would significantly reduce the impact of MPM but seems unlikely.

impact threshold if this test finds that, as a result of all suppliers' conduct violating thresholds:

- o Any price in a MISO market increases by more than the lesser of 200 percent, $100/MWh or (for a supplier in a NCA) the NCA Threshold defined above and discussed below.
- o A daily guarantee payment to the supplier increases by more than 200 percent.

4.2.5 Prospective Bid Mitigation

When the conduct of a supplier violates any applicable conduct threshold and the market impact of all such conduct violates an impact threshold, the following mitigation measures are applied.

- **Price Mitigation:** If the conduct that violates a conduct threshold relates to a bid parameter, the MISO will substitute for the violating bid parameter a default bid parameter equal to the reference level for that bid parameter. This substitution will be done prior to determining market prices, which means it must almost certainly be done in a computerized AMP procedure; prices will not be mitigated *ex post* except as may be specifically authorized by FERC. The resulting mitigated prices shall apply to all market participants, whether or not their bids have been mitigated.

- **Financial Sanctions:** If the conduct that violates a threshold was physical withholding or any other activity that cannot realistically be prevented or mitigated before the fact, the violating market participant will be subject to a financial penalty equal to the LMP at the most relevant location(s), multiplied by the MW withheld, multiplied by a penalty factor that increases the more often the market participants is guilty of withholding.

The MISO Draft also proposes provisions intended to remedy and discourage actions that are inconsistent with competition and that create undesirable divergence between day-ahead and real-time prices, but these are not discussed here.

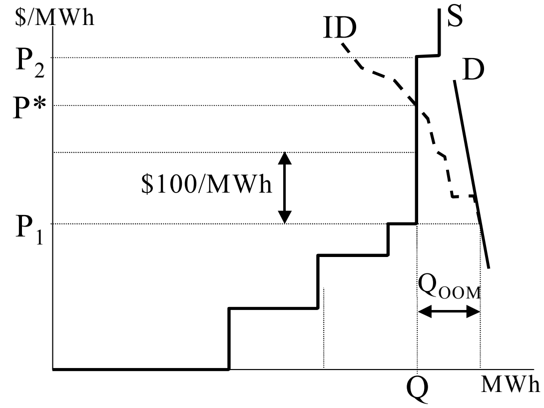
4.3 Effects of the MISO Proposal

This section discusses the likely effects of the MISO proposal and Dr. Patton's suggestion to improve scarcity pricing in the MISO spot market.

4.3.1 The Suppression of Scarcity Prices

The MPM procedure proposed for MISO (and used elsewhere) suffers from the logical inconsistency discussed in section 2.2.3 above: Reference prices are set based on how competitive suppliers would bid if the ITP determined scarcity prices properly, but then suppliers are (in effect) required to bid that way even though the ITP does not determine scarcity prices properly. The inevitable result is a suppression of scarcity prices, with all this implies about short-run efficiency, the effects on investment incentives, the need for a resource requirement/payment, etc.

Figure 9:
Suppression of Scarcity Prices



This is illustrated in Figure 9, where it is assumed that the marginal in-market generator – call it a "peaker" – has a simple MC or reference price of P_1 and a conduct threshold of $100/MWh. If the market demand curve for energy is D and the MISO's implicit demand curve for in-market energy is ID, the true scarcity prices is P*. If the peaker bids its reference price of P_1, the MISO will set the market price at this level and will call on Q_{OOM} MWh of OOM supplies to meet demand, but does not let the high costs of these OOM supplies set the market price. The result is inefficient, because high-cost supplies are being used to meet demands that – according to the market demand curve – value the energy at less than it costs. The result is also unfair and, if it happens very often, unsustainable, because the peaker and all other generators are being paid less than the true competitive value of their energy and less than they should get as a contribution to their fixed costs.

If the marginal generator increases its bid price to P* – and this bid is not mitigated – the MISO will set the price at the market-clearing level P*. This would be a good thing for all concerned – even consumers in the long run – because it would allow some in-market demand response and other market-drive actions to replace some higher-cost OOM energy, and because it would allow the marginal generator and all others to earn a competitive operating profit or rent that helps pay fixed costs.

Under the MISO MPM procedure, however, a peaker bid price of P* would violate the conduct threshold, and (as drawn in Figure 9) then the impact threshold, and hence would be mitigated down – not just to the allowed maximum bid/price impact of P_1+$100/MWh, but all the way down to P_1. The ITP would then set the price at P_1 and use high-cost OOM energy to meet the excess demand at that price.

This type of mitigation will effectively limit bids to levels that will not violate conduct thresholds, because bidders will not want to risk being mitigated all the way back to their reference bids, and hence will often result in prices that are far below market-clearing levels. One change in the mitigation rules that might help in situations such as this would be to mitigate an energy bid down to the conduct threshold, not to the reference bid, when both conduct and impact thresholds are violated. This would allow the peaker to bid more aggressively when it expects the real scarcity price to be very high, with less risk of having its bid mitigated down to the reference level and losing everything.

The MISO Draft recognizes that the proposed procedure will suppress scarcity prices, so it proposes that suppliers in NCAs be allowed to bid above their simple MCs or reference prices by an amount that is apparently intended to give them a chance to recover their fixed costs. Furthermore, Dr. Patton has recommended some changes to the MISO pricing rules that could, under some conditions , result in scarcity prices that exceed any supplier bids. As discussed in the next two sections, however, these proposals, as welcome as they are, are not enough to eliminate, or perhaps even significantly to reduce, the tendency of this procedure to suppress scarcity prices.

4.3.2 Treatment of Suppliers within Narrow Constrained Areas

In the MISO Draft, a supplier within a NCA is, under some conditions and to some extent, allowed to bid energy prices that exceed its simple MC by an amount related to its fixed costs without such bidding violating its energy bid conduct threshold. This is noteworthy because it implicitly acknowledges that the MISO spot markets will not reliably produce scarcity or compensatory prices unless suppliers bid such prices. However, the provision as it stands will not necessarily accomplish its apparent objective for both theoretical and practical reasons.

The theoretical problem is in the definition of the "NCA Threshold" that limits both the (maximum) amount by which the supplier is allowed to bid above its SRMC and the amount by which the market price is allowed to increase. The NCA Threshold is defined as follows:

$$\text{NCA Threshold} = \frac{\text{(Net Annual Fixed Cost)}}{\text{(Constrained Hours)}}$$

where:

Net Annual Fixed Cost = "Annual fixed costs of a new peaking generator per MW, including recovery of capital costs, minus appropriate credits for Net Revenue the new peaking generator would receive from the MISO electricity market." [MISO Draft, sec. 3.1.2(c)(1)]

Constrained Hours = The number of hours over the prior 12 months in which imports into the NCA were constrained, but not more than 2,000 hours

The problem is that it is theoretically incorrect to use a peaker's annual fixed costs in the definition of the NCA Threshold unless the generator subject to that threshold is itself a peaker. If a generator is to recover its total costs with a price that has two components, one related to energy costs and the other related to fixed costs, the two components must relate to the same kind of plant. As both a logical and a practical matter, a generator in a NCA should be allowed to bid its own energy costs plus its own fixed costs (or perhaps the fixed costs of a proxy plant of the same type). It is mixing apples and oranges to say that a non-peaker generator can recover its own energy costs plus a peaker's fixed costs. For example, if the generator at issue is a baseload plant, prices equal to its own low fuel costs plus the low fixed costs of a peaker will not cover its full costs.

Other, less theoretical problems with this provision may be more important in practice, particularly if the generator in question is a peaker. Even if the NCA Threshold is defined correctly, a generator may not be able to bid at the implied level because a generator's conduct threshold is the ***minimum*** of the NCA Threshold, $100/MWh or 300 percent of the generator's reference energy price, so if the NCA Threshold is very high the generator cannot bid it. For example, if the Net Annual Fixed Cost of a Peaker is $80/kW-year and there were 500 constrained hours last year, the NCA Threshold is $160/MWh, which is greater than $100/MWh, so the effective conduct threshold is $100/MWh, not the NCA Threshold. If the purpose of the NCA Threshold is to give supplier in a NCA a better chance to recover its fixed costs, the effective threshold should be the NCA Threshold, period.

Furthermore, even if the generator can bid the NCA Threshold in all constrained hours, it is unlikely that doing so would result in a price at this level in all hours. If the market price were lower than the price defined by the NCA Threshold in some of the constrained hours, a generator under this bidding constraint would be unlikely to recover its fixed costs even if the average market-clearing price P* over constrained hours would do so.

4.3.3 THE PATTON RESERVE PRICING PROPOSAL

In a memorandum to the MISO Operating Reserves Task Force dated October 29, 2002 (the "Patton Memo"), Dr. Patton has recommended a mechanism for "establishing economically efficient energy prices during capacity shortages (i.e., when there is insufficient capacity to meet both the energy and operating reserve requirements of the system)." The reasons given in the Patton Memo for this recommendation are similar to the arguments in this paper: "[I]f the market rules are developed without an explicit economic relationship between the reserves and energy markets, spot energy prices that are determined during the capacity shortage are likely to not reflect the full value of energy. By muting energy prices during shortages, the market will not send appropriate signals in the short-run for resources to enter from other regions nor send efficient signals over the longer-term for new investment." [Patton Memo, p. 1]

Dr. Patton's recommendation and arguments here, and similar things he has proposed for the New York ISO, are consistent with the recommendations and arguments in this paper and are certainly a step in the right direction, But they are not enough to remedy the deficiencies in ITP pricing even if implemented, largely because they would add only one large and seldom relevant step to the supply curve, when what is needed is a whole series of steps that can routinely set scarcity prices.

Simply stated, Dr. Patton recommends that activation of reserves by the MISO be treated as an addition to market supply with an energy bid equal to or above the level of FERC's safety-net bid cap and that this reserve energy bid be used to determine prices. The justification given for setting the minimum price of reserve energy at the safety-net bid cap is that reserves should not be activated until all supply bids are taken. In principle, such a rule for reserve activation is not required or necessarily optimal; it might be better to release small amounts of reserves at prices that are high but less than the safety-net bid cap. Be that as it may, the Patton Memo recommends putting a bid floor on operating reserves at the safety-net-bid cap, call it $1,000/MWh for concreteness.

Given that the marginal suppliers are likely to be CTs with a reference energy bid price of no more than about $100/MWh and a conduct threshold of no more than $100/MWh (or less, for a peaker in an NCA with a lower NCA Threshold), the effective limit on energy offer prices from suppliers will be about \$200/MWh – 20 cents/kWh – leaving a very big gap between the highest supplier bids and the minimum bid price for reserve activation. For example, if P_2 in Figure 9 represents the safety-net bid cap/minimum reserve bid and P_1 is the highest simple MC of any supplier, there is a large gap between P_1 and P_2 and there will many hours when the market-clearing price is somewhere in between. If suppliers are allowed to submit sculpted bids for small amounts of energy at high prices, there may be some supplier bids in that range, but there is no guarantee this will be the case everywhere, e.g., in a small area or single node designated a NCA.

A critical issue is what happens if/when the market-clearing price is above the highest supply bid but below the reserve activation bid at (e.g.) \$1,000/MWh, which is likely to be common in non-emergency scarcity conditions. If demand expects a price of P_1 there will be excess demand at that price, which the MISO will have to meet using OOM actions other than activating reserves. If these actions do not set the price the pricing problem has not been solved by putting in the very high and rarely used bid at P_2.

One possible solution to the "large gap" problem would be to develop a formula that would set the market price at a weighted average of P_1 and P_2, with the weights depending on how much OOM energy is needed to close the gap. This would not be necessary if the MISO's implicit demand curve for in-market energy ID were really known, but in practice it probably will not be. What the MISO will know is whether, after taking all the supplier energy available at price P_1, it had to use a lot or a little OOM energy to meet demand and how close it came to activating reserves. The market price should be a continuous function of such operational factors between P_1 and P_2.

Again, if all this sounds too hard and *ad hoc* to be realistic, the implication is that the MISO will not really have a very good scarcity pricing mechanism, and the only way to get reasonable approximations of scarcity prices is to let suppliers bid at least some non-trivial amount of their energy at their estimate of the scarcity price without it being regarded as an exercise of market power that must be mitigated. As long as such bidding behavior does not result in time-average spot prices that are significantly above the supplier's total costs or the LRMC of similar suppliers, the supplier is not benefiting from any exercise of market power but is simply bidding as it must to get the prices it deserves given the deficiencies of the MISO's pricing process.

4.4 Evaluation of the MISO Proposal

The proposed MISO MPM procedure is similar to the AMP procedures currently used in ISOs/ITPs and endorsed by FERC in its SMD NOPR. As such, it is worth evaluating the MISO proposal in terms of its apparent consistency with FERC's SMD objectives and the logic of its basic approach.

4.4.1 Consistency With FERC's SMD NOPR

The treatment of Narrow Constrained Areas in the MISO proposal corresponds closely to the local market power mitigation measure mandated in the SMD NOPR, although FERC proposes that this measure be applied through Generator Operating Agreements. MISO's proposed treatment of Broad Constrained Areas is harder to put into the FERC framework and, in fact, appears to be inconsistent with at the least the spirit of FERC's proposals, because MISO would apply the MPM thresholds and, if so indicated, the mitigation measures everywhere and at all times, not just in special, clearly defined circumstances.

FERC says that its voluntary AMP-like measure is to be applied only in "unanticipated and sustained" but "temporary" market conditions such as a drought or major transmission line outage, and with clear "criteria" for determining when such special conditions exist and when they have ended. It might be argued that the threshold tests are these "criteria," and that they must be applied at all times and places because there is no other way to know where and when they are being violated. But this would seem to be inconsistent with FERC's clear intent to leave the market alone except under unusual circumstances.

Universal application of the thresholds might be defended as a way, or even the only practical way, to determine when and where "sporadic" transmission congestion is creating local market power. But this is casting a very large and heavy net just in case there are a few fish out there and without regard for the effects on the innocent dolphins. This, too, would appear to be inconsistent with FERC's intent to leave the market alone unless there are clear indications of a problem that needs fixing.

4.4.2 Circularity and Self-Verification

The MISO proposal contains many numbers – 200 percent here, 300 percent there; \$100/MWh here, \$25/MWh there – that, taken together, will determine whether the MPM procedure will be ineffectual or draconian or somewhere in between. None of these numbers has been or realistically could be chosen based on either sound theory or quantitative analysis. At best they reflect judgments about what is "reasonable" based on very little experience.

Any new policy or procedure is implemented without knowing precisely what its effects will be and with the understanding that it may have to be adjusted based on experience with it. But for experience to be useful there must be clear guidelines stating what the procedure is intended to accomplish, how its performance will be measured and evaluated, and how it might be changed in response to experience. The MISO Draft does state the objective of the MPM procedure: to "allow prices to rise efficiently to reflect legitimate supply shortages while effectively mitigating inflated prices …" [MISO Draft, sec. 1(b)] But without some objective way to determine when mitigated or unmitigated price levels are high enough "to reflect legitimate supply shortages" without being "inflated" there is no way to know when or if the procedure is accomplishing its stated objective.

The MISO MPM procedure, like all AMP-like procedures, implicitly assumes that the right competitive price in an ITP's market is the price resulting from application of the MPM procedure in that ITP's market, and hence here is no need – or even no way – to compare the prices resulting from application of the MPM procedure to some external estimate of competitive prices. Even if AMP-like procedures did not suffer from the serious logical problems discussed in section 2.2.3 above, such circular self-verification would be illogical and potentially dangerous.

Before adopting the proposed (or any) MPM procedure, the MISO and FERC should require the development of a plan for evaluating the impact of the procedure, calibrating it so that it does what it is supposed to and nothing else, and modifying or deactivating it as appropriate. Such an evaluation/calibration plan should look beyond specific hourly prices and outside itself for evidence that it is or is not doing what is intended. Evaluation should be based primarily on comparisons of time-averaged spot prices to the LRMCs of needed generation. The calibration procedure should adjust the parameters to assure that the interactions of the MPM procedure and the realities of the spot market do not suppress average spot prices below LRMCs and should deactivate the whole procedure unless there is clear evidence that without it average spot prices would be too high.

5. A SUGGESTED APPROACH TO MARKET POWER

The analysis of economic principles and current practice in the preceding sections suggests the following general approach to the identification and mitigation of market power in electricity spot markets.

5.1 KEEP SPOT MARKET POWER IN PERSPECTIVE

Most consumers and in fact many suppliers and middlemen do not care what the spot price is in any specific hour or how much it may change from hour to hour, because they will not or cannot respond to hourly spot prices and because they (at least should) do most of their business under contracts and price-averaging arrangements; what these players care about are average spot prices over periods ranging from a day to a year. But for those who do care about prices over shorter periods because they can respond to them, it is important that the prices to which they respond reflect the real value of energy at that time. Artificially suppressing scarcity prices increases costs for everyone and does no good for anyone in the long run.

Given the difficulty of knowing precisely how a competitive supplier should behave in a spot market or what the “right” competitive spot price is, it is a serious mistake to try to identify and control market power by observing either behavior or outcomes in individual spot market periods, i.e., hours in the case of electricity. If suppliers are effectively exercising market power in spot markets, it will show up in time-averaged prices over periods such as a year or more. If such average spot prices are not above the LRMC levels needed to stimulate needed investment, there is no good reason to implement comprehensive, intrusive and potentially distorting procedures for identifying and mitigating market power in spot markets and many good reasons not to do so.

The evidence from functioning ITP spot markets is that suppliers are not making too much money in those markets, and in fact may be making too little. For example, Dr. Patton says that spot prices in New York, even after ICAP payments, “would not likely support new investment in GTs” [gas turbines] outside New York City “with significant uncertainty regarding GTs within" New York City."[22] A similar analysis for gas-fired combined cycles plants in New York and New England indicates that average spot prices were below LRMCs even during 2000-2002 when the supply-demand balance was very tight.[23] Such results do not suggest there is a compelling need to implement aggressive or comprehensive MPM procedures at this time.

5.2 Improve ITP Spot Scarcity Pricing

Much of the behavior that is viewed as an exercise in market power in ITP markets can be explained at least as well as the necessary behavior of workably competitive suppliers given that the ITP markets do not have the kind of scarcity pricing procedures they should have – and that they are likely to be assumed to have when reference levels are set in AMP-like MPM procedures. If ITP scarcity pricing is improved, there will be less need for competitive suppliers to act in ways that are commonly, if incorrectly, interpreted as an exercise of market power, and it will become less difficult to identify and remedy true exercises of market power.

ITP pricing processes can be improved by basing scarcity prices on sculpted supplier bids, demand bids, the costs of imports and recalled exports, and the implicit costs of what are now treated as out-of-market actions. Dr. Patton's suggestions along these lines for the New York ITP and for MISO are a step in the right direction that should be adopted and extended; but even if these specific steps are implemented, ITP markets will not establish efficient scarcity prices if all suppliers are required to bid at their simple MCs.

5.3 Keep MPM Focused and Light-Handed

Given that ITP markets do not have good scarcity pricing and that the market evidence does not provide compelling or even a plausible case for aggressive and comprehensive MPM procedures at this time, efforts to control market power should focus on identifying specific, localized problems and tailoring specific programs for these. Even where specific problems are identified, it should not be assumed that AMP-like procedures are the only or the best solution. Indeed, the preferred approach should be structural and contractual remedies, such as contracts that reduce the incentive to drive up scarcity prices, rather than intrusive efforts to control market behavior and outcomes directly.

Where comprehensive MPM procedures are implemented, they should be carefully calibrated so that they mitigate egregious behavior and outcomes without interfering with the normal efforts of suppliers in complex and imperfect ITP markets to remain commercially viable. MPM procedures should not apply mechanical measures of what competitive suppliers

---

[22] David B. Patton, PhD, "Summer 2002 Review of the New York Electricity Markets," presentation to the New York ITP Board of Directors and Management Committee, October 15, 2002.

[23] See Klein, Abram, "Scarcity Pricing in Northeast ISOs: An Assessment of Market Performance," Presentation for NECA Wholesale Markets Conference, November 14, 2002, Boston, MA.

would do in a perfect ITP market and then compel such behavior in real ITP markets where it produces much lower prices than it would in a perfect ITP market. Instead, MPM procedures should ask what workably competitive suppliers would, indeed must, do to survive in real ITP markets, and then not prevent such commercially sensible and necessary behavior until there is some better way to assure the commercial viability of needed suppliers.

5.4 Use Capacity Payments To Offset Aggressive MPM

Aggressive MPM procedures will inevitably suppress scarcity prices. At least if and while such aggressive MPM procedures are in place, some effective form of resource requirement or payment is required. The basic choice to be made here is whether the time horizon for such a program should be short-term or long-term. A short time horizon – e.g., a month or so, like most ICAP programs – can reduce and stabilize spot market prices but will not give much assurance to those who want to see long-term commitments and long-term price stability. A long time horizon – e.g., three-to-five years, as proposed for FERC's Resource Adequacy Requirement (RAR) – can produce long-term commitments and stability, but only if the ITP makes long-term plans and then either makes itself or effectively compels LSEs to make long-term commitments.

FERC appears to be trying to avoid the fundamental logical choice between efficient spot markets with good scarcity pricing and an effective resource requirement/payment mechanism. Its RAR proposal is neither of these nor even a logical combination. FERC will have to develop something different and much better, particularly if FERC endorses and approves aggressive MPM procedures.

5.5 Move Quickly To Full Competition with Good Scarcity Pricing

The long-term objective for competitive electricity markets should be, as FERC says, "the development of structurally competitive markets." An important step in accomplishing this objective is good scarcity pricing in spot markets, because such pricing is necessary to reduce concerns about market power and the need for MPM procedures, and to reduce the need for the ITP to make decisions that should be made in the market.

Conversely, and just as importantly, if intrusive and distorting MPM prevents or delays good ITP scarcity pricing, concerns about market power will be perpetuated, there will be continued demand for aggressive regulatory interventions and ITP decision-making that distort market dynamics, and the development of structurally competitive markets will be delayed. Good pricing by ITPs is not a substitute for structurally competitive markets or *vice versa*; the two go hand in hand or not at all.