Resource Adequacy Mandates and Scarcity Pricing

("Belts and Suspenders")

William W. Hogan[i]

February 23, 2006

Introduction[1]

The resource adequacy mandate of the Reliability Pricing Model (RPM) arises in response to the defects of energy scarcity pricing rules and inadequate incentives for investment. The RPM initiative does not replace the need for improved scarcity pricing. Assuming the RPM proposal will go forward, better scarcity pricing would create no contradictions with the RPM design. Pursuit of the RPM should not be the occasion for further deferring attention to improved scarcity pricing. It would be better to implement better scarcity pricing sooner rather than later. Such a policy would provide its own benefits and would improve the chances of success for RPM. Adequate scarcity pricing would provide policy insurance that would pay off just when RPM was not enough, but when it would be too late to fix the problem by fixing RPM.

The origin of resource adequacy mandates and the limits on scarcity pricing

Providing consistent incentives and risk allocations that encourage the right investment in the right infrastructure, in the right places, and at the right times remains a persistent challenge in electricity market design. In addition to keeping the lights on, providing adequate investment is a fundamental test of using markets to improve performance in electricity systems.

The current RPM discussion in PJM is part of an evolution that identifies incentive problems in the existing electricity market design and seeks appropriate reforms. In the case of RPM the reforms target improved capacity market mandates. Although there are many important details in the RPM proposal, the basic proposition is simple. Market prices in the energy and ancillary services markets have on average been below those that would be needed to support efficient investment, and the gap has both a locational and operational character that must be addressed to meet reliability requirements. That is, the energy and ancillary service prices not only fail to support sufficient investment at the right locations, they also fail to provide the correct incentives for plants that offer quick-start, load-following and other features that are required for reliable operations. These are effects of the so-called "missing money" problem, which occurs when aggregate market revenues fail to provide either enough money to support resource adequacy investment goals or the correct incentives for reliable operations in

---

[1] This paper was prepared at the request of PPL Corporation in connection with the Federal Energy Regulatory Commission proceedings on the Reliability Pricing Model proposal for PJM Interconnection.

real time.[2] Hence, the underlying problem is not merely with poor investment incentives. The suppressed prices for energy and ancillary services also undermine the incentives generators need to take all steps necessary to ensure their units are available for dispatch when and where they are most needed, while signaling to loads that helpful demand reductions may not be warranted. The RPM is intended to help fill this gap by defining the locational and operational requirements and attempting to provide the missing money with greater certainty about the payments.

The RPM is a partial remedy for the various problems associated with the missing money. There is a complementary need to augment the RPM approach by focusing in parallel on the incentive problems created in the way energy and ancillary service are arranged and priced in the real-time and day-ahead markets. This dual "belts and suspenders" approach addresses not merely the symptoms but also the underlying causes of the investment and operational reliability problems in PJM.

The importance of the real-time energy and ancillary services markets

The energy market design in PJM is widely recognized as setting an example for other electricity markets. Built on the principles of a coordinated spot market with bid-based, security-constrained economic dispatch with nodal prices, the basic PJM energy market design reinforces reliability by using prices consistent with reliable dispatch. The use of consistent price signals simplifies the many tasks of the system operator, and provides a coherent framework for open access to the transmission grid. This core structure of the PJM design should be maintained and improved, because doing so will improve real-time reliability while laying a solid foundation to support the resource adequacy goals of RPM.

There are opportunities for improvement in the pricing of energy and ancillary services that will directly support RPM and its chances of success. The RPM proposal seeks to compensate for deficiencies in the existing market design for dispatching and pricing energy and ancillary services. The RPM does not address these underlying problems directly but instead seeks to address their symptoms. It is in public's interest that the RPM succeed and the market design does not result in a missing money problem. The best way to ensure success is to focus simultaneously on correcting these underlying problems in the basic markets, not just the symptoms of inadequate investments. As part of this process, it is important to give attention to continued identification and improvement in the real-time price signals provided through the basic market design. Accordingly, the energy and ancillary services market design should be improved to provide appropriate price signals reflecting the value of service provided.

---

[2] The term "missing money" describes the condition in which prices in the markets for energy and ancillary services are kept below market-clearing levels, especially in hours of scarcity, with one result being that the prices fail to cover the fixed costs of generators. As PJM notes in its RPM filing, "If a new unit is to recover all of its costs from the PJM markets in equilibrium, the unit needs to recover from the capacity market only those costs not recovered in the other PJM markets." Affidavit of Joseph E. Bowring, (hereafter, "Bowring Affidavit"), PJM RPM filing, Docket Nos. ER05-141-000 and EL05-148-000. August 31, 2005, Attachment G, at 2.

There is no inherent conflict in following this dual approach of pursuing RPM along with energy and ancillary market improvements; indeed PJM has already explained that improved energy and operating reserve pricing and its RPM proposal are fundamentally compatible.[3] Improvements in the energy and ancillary services markets should make the RPM easier to implement and more likely to succeed. Failure to correct deficiencies in the energy and ancillary services markets would only increase the burdens placed on the RPM process, adding to its complexity and expanding the role of regulators in place of markets, while increasing the chances that RPM, no matter how well designed, may fail to achieve its objectives.

Although the underlying problems that create the need for RPM arise primarily in the pricing of energy and operating reserves, the RPM proposal by itself does not correct any underlying deficiencies in the energy market. Rather the RPM proposal continues the practice of having a separate capacity requirement and creates a new procurement and payment procedure intended to change the outcomes of investment decisions from what they might have been under the current installed capacity (ICAP) mandate. The underlying market design deficiencies would remain. Even further, to the extent that the RPM is successful and produces incremental investment that expands capacity, the energy and ancillary services market design would tend to reduce market-clearing prices and exacerbate the incentive problems cited as giving rise to the RPM requirement.

Improvements in pricing for energy and ancillary services markets would reinforce the objectives of RPM. In the current market design, locational marginal pricing (LMP) is recognized as simplifying the task of transmission congestion management.[4] For similar reasons, improved energy scarcity pricing would provide better real-time incentives for reliability given whatever investment the RPM supports. Without adequate scarcity pricing, the system operator will be compelled to invoke inadequate substitutes for scarcity pricing through administrative penalties and continued reliance on special arrangements like reliability must run (RMR) contracts and "Maximum Emergency" out-of-market purchases.

There have been substantial discussions and some progress in moving in the direction of scarcity pricing in PJM.[5] However, there appears to be general agreement that more can and should be done. The most important first step was in adopting LMP. Under most circumstances, the simplest interpretation of the LMP pricing rules suffices. But in periods of capacity scarcity, application of the basic idea requires attention to the underlying purposes of market-clearing pricing. For example, although it is often true that efficient spot-pricing produces an LMP defined by the "highest running generator offer," this is not always the appropriate rule. When capacity is constrained, market-clearing prices should be determined (ideally by demand bids) at levels above generator

---

[3] PJM RPM filing, Docket Nos. ER05-1410-000 and EL05-148-000. August 31, 2005, Part 1, p. 15.

[4] Phillip G. Harris, "Relationship between Competitive Power Markets and Grid Reliability: The PJM RTO Experience," Issue Papers on Reliability and Competition, US Department of Energy and Natural Resources Canada, August 2005, pp. 4-5. (www.energetics.com/meetings/reliability/papers.html)

[5] For example, see Federal Energy Regulatory Commission, "Initial Comments Of The Commission Trial Staff In Support Of Offer Of Settlement," Docket Nos. EL03-236-006 and EL04-121-000, December 6, 2005.

offers, above even the higher offers allowed by the recent Settlement regarding scarcity pricing.[6] That is, genuine scarcity prices would be set by a demand curve, not simply by the supply offers of marginal generators. Further, the "demand curve" for power would reflect both the need for energy and the need for operating reserves, such that shortages in the level of operating reserves would tend to result in higher prices for both energy and operating reserves. The proper deviations from the simple "highest offer" pricing rule are most important precisely when scarcity in generation capacity applies. And it is artificially depressed pricing during these scarcity conditions that is cited as the principal source of the missing money problem.

The missing money is a symptom, so focusing solely on providing the missing money through RPM (or any other capacity payment mechanism) can fail to address the basic cause, while missing other aspects of the broader incentive problem. The litany of concerns is familiar. There is inadequate demand bidding and response in real time, and this creates its own problems by increasing incentives to exercise market power and the need for market power offer mitigation. But it is not just the absence of or disincentive for demand bidding that creates the problem. Nor is it just the existence of market power mitigation rules, which are likely to be necessary and appropriate in any event. Other operational and pricing rules may also suppress prices below appropriate clearing levels. Under current rules, the occasional conditions of scarcity, including maximum generation emergencies or voltage reductions, produce either no effect on LMPs, or perversely may cause calculated LMPs to decrease just at the time when other signals show that scarcity is increasing. The very existence of RMR contracts and out of market purchases could be prima facie evidence that scarcity pricing is defective and requires further reform.

Collectively, the several rules that combine to lower average prices in PJM produce the missing money problem. Although the conditions apply in relatively few hours, it would be precisely the scarcity prices in those few hours that would provide the needed compensation for the most efficient capacity investments. For electricity markets, the load curve and the marginal cost of supply combine to create a very volatile marginal cost that can be very high during shortage conditions, much higher than the current system offer cap of $1000/MWh, but such marginal cost is applicable in only a few hours per year.

Lower average prices will of course produce lower average investments. But in real time, the suppressed prices will also undermine real-time reliability, especially in those hours in which capacity resources are short, as indicated by lower levels of operating reserves. If energy and operating reserve prices do not rise to reflect these shortages, then it is unlikely that generators will see the incentives needed to encourage all reasonable efforts to make their plants available for dispatch. The changes in behavior may be subtle and difficult to trace to any individual pricing event, but collectively and over time they will adversely impact reliability. For example, generators may simply fail to purchase sufficient fuel in advance when shortage conditions might occur; or they might avoid investments that allow dual-fuel capability, or on-site fuel storage; or they

---

[6] "PJM will not cap offers from any generation in the region while scarcity pricing is in effect, although such generation will remain subject to PJM's overall cap of $1,000 per megawatt-hour." Settlement Comments, p. 5.

may change maintenance practices and schedules. Collectively, these decisions will reduce the probability that various units will be available when and where they are most needed. For the same reasons, investments that would tend to make the units more flexible, such as the ability to perform quick starts, to recycle more frequently, and to follow load more easily may not be made, because the incentive for such reliability-enhancing investments is simply not provided by the real-time prices.

The introduction of improved scarcity pricing in the PJM energy and operating reserve markets would address these concerns directly, thus complementing the investment goals of RPM, which seeks to target specific kinds of investments in those operational features that improve real-time availability and dispatch flexibility. The more important point, however, is that the real-time prices for energy and operating reserves would tend to encourage an efficient level of investment in these operating characteristics, thus making it less important whether PJM had correctly anticipated the exact level of investment needed for these operational features in the four-year ahead RPM auctions.

The importance of the demand curve for operating reserves

To go beyond the description of symptoms of the scarcity pricing problems, the diagnosis and prescription should be targeted precisely at these infrequent conditions when scarcity arises. For example, in most hours there will be adequate spinning or quick start reserves. However, in some hours when the system is capacity constrained, there will be increasing steps to acquire more expensive sources of operating reserves, or reduce the available reserve protection, until we reach the point where load curtailment becomes necessary to ensure security. It is generally recognized that this process should address the increasing scarcity through rising prices of both reserves and energy. An important contribution of this scarcity pricing would be to recognize that as operating reserves are reduced, the reduced reserves increase the probability of load curtailment. By allowing energy and reserve prices to rise when this occurs, the enhanced incentives of scarcity pricing encourage both greater supply availability and increased demand response, thus reducing the need for involuntary load curtailments.

One approach to scarcity pricing is to use not a fixed operating reserve requirement but a sloped operating reserve demand curve above a minimum level to reflect higher prices determined simultaneously for energy and reserves.[7] Although the simplification of an operating reserve demand curve may not capture the full picture, it would remove many of the pricing anomalies that appear with the current fixed requirement model, a model that is equivalent to a vertical operating reserve demand curve that cannot always be honored.

---

[7] William W. Hogan, "On An 'Energy Only' Electricity Market Design For Resource Adequacy," Center for Business and Government, John F. Kennedy School of Government, Harvard University, September 23, 2005, pp. 11-14, (available at www.whogan.com). For a related discussion of the importance of an operating reserve demand curve, see ISO New England, "2006 Wholesale Markets Plan," September 2005, pp. 16-17.

An effective operating reserves demand curve with simultaneous determination of energy and reserve prices could provide several advantages in addition to sending better signals for investment. An operating reserve demand curve would provide an effective administrative tool for incorporating scarcity prices that would be necessary to provide adequate incentives for increased energy demand bidding. Hence, this could help "jump start" the demand side of the market. Further, by introducing some "elasticity" in total energy and reserve demand, the operating reserve demand curve would make a contribution similar to demand bidding in lessening the incentives to exercise market power.

The scarcity prices implied by an operating reserve demand curve would promote reliability by providing greater and timely incentives for market participants to make special efforts to provide capacity during periods of scarcity. And a well defined operating reserve demand curve would provide greater clarity for PJM system operators and transparency in their decisions about the need to obtain additional expensive capacity to meet reserve requirements.

Introduction of an operating reserve demand curve would not necessarily require changes in operating practice and should only enhance reliability. Even with no changes in the dispatch rules, the operating reserve demand curve should play a role in the determination of consistent energy and ancillary services prices. Just as with LMP determination, all prices should be calculated to be consistent with the dispatch assuming that the dispatch was optimal given the bids. This dispatch-based pricing principle maintains the primacy of reliable dispatch while using prices to provide consistent incentives. And just as experienced with LMP, there should be a convergence where any gaps between reliable dispatch and consistent prices would be both small and easy to explain.

The difference from present practice would be substantial during periods of reserve scarcity. The simple logic of scarcity pricing says that the reserve and energy price should approach the value of lost load (VOLL) when the level of reserves reaches the security minimum and load curtailment is required. The precise estimate of VOLL is controversial, but what is not controversial is that the VOLL is far above the de facto price cap of approximately $1000/MWh in PJM. A VOLL of $10,000/MWh is closer to the mark, and prices should sometimes approach or reach this level.

Customers would be hedged against the risks of scarcity prices

For most customers, the few hours of high prices would be managed through hedging programs like the Basic Generation Service in New Jersey or other default service rules in other states in PJM. A key point is that the degree of hedging, and the choice of whether to make such hedges mandatory for most consumers, or voluntary for some, would be left to the States to decide, just as they do today.

In effect, the few hours of high prices would be anticipated in the hedging contracts and spread over all the hours of load. This is the same effect that applies for the missing money recouped through the RPM payments that would also reflect the value of reliability and be spread over all hours of consumption. The important incentives at the

margin, high scarcity prices that encourage availability of capacity, would still apply to market participants, even though average prices would be largely hedged and not subject to the same volatility.

Market power mitigation would remain

Better scarcity pricing would not eliminate the need for market power mitigation, although it would not necessarily make the market power problem worse.[8] The principal tool for mitigation, offer caps for generators that might have an incentive and ability to exercise market power, would remain. Although the short-hand descriptions of the problems of scarcity pricing often make reference to the effects of offer caps, it is not the offer caps per se that are the problem. To the contrary, the theory of offer caps is quite consistent with the principles of competitive markets and appropriate scarcity prices. The problem is the combination of the offer caps and other rules in dispatch and pricing. For example, with an appropriate operating reserve demand curve and offer caps on generators, market rules could provide that scarcity prices would still rise as PJM became short of operating reserves, and price could rise to the VOLL if the available capacity were not sufficient to meet the minimum operating reserve requirements.

Scarcity pricing complements RPM and other resource adequacy goals

Improved pricing in energy and ancillary services would provide other benefits. It is possible that adequate scarcity pricing, combined with hedging and long-term contracts, would be sufficient to ensure the right investment incentives and procedures. However, it is likely that better scarcity pricing would leave some problems that would be beyond the reach of prices and markets alone. The usual example cited would be some condition where lumpy investment decisions have a material effect on prices and create free-rider problems.[9] Adequate scarcity pricing alone would not be sufficient to deal with all these conditions. However, improvements in scarcity pricing would isolate the special conditions that needed further regulatory treatment and allow for better diagnosis and prescription.[10] This would provide a principled basis for regulatory decisions and provide the tools that would support regulatory restraint. Absent such a principled argument, the market could slowly unravel as each regulatory treatment

---

[8] See William W. Hogan, "On An 'Energy Only' Electricity Market Design For Resource Adequacy," Center for Business and Government, John F. Kennedy School of Government, Harvard University, September 23, 2005, pp. 24-25. (available at www.whogan.com)

[9] For example, a large plant in a load pocket might depress prices enough to prevent it from recouping its annual fixed costs. But if the plant closes and no new capacity were to enter, prices could rise more than enough to justify keeping the plant in place. Market participants would be better off with the plant open, but there may not be an adequate market mechanism to capture these benefits.

[10] Dispatch-based pricing to provide consistent ex post LMP and reserve prices can reduce some of the gap associated with short-term fixed costs and lumpy decisions. For a further discussion, see William W. Hogan and Brendan J. Ring, "On Minimum-Uplift Pricing For Electricity Markets," Center for Business and Government, Harvard University, March 19, 2003, (www.whogan.com)

creates perverse results that add to the symptoms of market failure and in turn create appeals for further regulatory treatment.

For those who argue that adequate scarcity pricing is all that is required, implementation along with RPM provides an advantage in giving credibility to the sustainability of scarcity pricing. Absent sufficient capacity and adequate long-term contracts that provide viable hedging instruments, there would be a legitimate concern that a regulatory promise to support scarcity pricing could not be honored in the face of the political pressure that would appear precisely when scarcity emerged. But with parallel programs that provide the best hedging and capacity investment that the regulator can craft, scarcity pricing to support real-time operations should be seen as a solution and not part of the problem. As experience grows, the hope that good market design can reduce or eliminate the need for programs like RPM might be realized.

The system operator's task would be simplified with better scarcity pricing. The concern that reliability cannot be met unless the right investment occurs motivates the interest in the longer horizon and locational and operational features requirements of RPM. However, the same logic implies that whatever the investment, in the end the key is to make sure that the plant is available, fuel acquired, and any extraordinary measures needed are forthcoming and creative. It would be very hard to write operating rules and procedures to mandate cooperation when the prices don't reflect opportunity costs. But if prices in real time reflect the best estimate of opportunity cost, behavior would more naturally align with the needs of reliability. In this sense, adequate scarcity pricing provides policy insurance that would pay off just when RPM was not enough, but when it would be too late to fix the problem by fixing RPM.

The RPM proposal would be quite compatible with improved scarcity pricing in the energy and ancillary services markets. For example, the current RPM proposal includes setting the variable resource requirement curve to net out an estimate of the likely scarcity profits in energy and ancillary service markets. However, the procedure does not require RPM resources to rebate the actual scarcity payments that appear. This preserves incentives in real time and is an important attractive feature of the RPM design. Of course, bidders in the RPM auctions will anticipate the expected value of the scarcity payments and these estimates will reduce the missing money that must be handled through the RPM auction.

Hence, the logic of the RPM treatment of scarcity rents is compatible with implementing the RPM auction and improving scarcity pricing. Simply launching a credible effort to reform energy market pricing would have a beneficial effect on the RPM auction. Conversely, a de facto decision to defer reforms in energy market pricing would have a detrimental effect on RPM's prices and its success in achieving its stated goals.

Implementing resource adequacy mandates and scarcity pricing

PJM should implement better and more comprehensive scarcity pricing at least as fast as the schedule to implement RPM. Since it is important that proper energy pricing signals accompany the implementation of RPM, it would be an unnecessary risk for PJM

to first implement RPM and only then examine necessary energy and ancillary market improvements. If a choice is required, the priority should be on improved scarcity pricing. Otherwise, do both and provide both belts and suspenders.

There is, of course, still much to do, including developing a proper approach for implementing co-optimization of energy and operating reserves, defining an operating reserve demand curve that allows demand to set price at the VOLL during extreme scarcity conditions, analyzing the interactions between scarcity pricing and locational problems that have given rise to greater reliance on RMR contracts, and addressing the scarcity pricing implications for other ancillary services and constraints. However, this development work should begin promptly and should proceed with some sense of urgency. We should not rely on a false sense of security that RPM provides an easy fix that will avoid the need to treat the underlying problems.

---

[i] William W. Hogan is the Raymond Plank Professor of Global Energy Policy, John F. Kennedy School of Government, Harvard University and a Director of LECG, LLC. This paper draws on work for the Harvard Electricity Policy Group and the Harvard-Japan Project on Energy and the Environment. The author is or has been a consultant on electric market reform and transmission issues for Allegheny Electric Global Market, American Electric Power, American National Power, Australian Gas Light Company, Avista Energy, Brazil Power Exchange Administrator (ASMAE), British National Grid Company, California Independent Energy Producers Association, California Independent System Operator, Calpine Corporation, Central Maine Power Company, Comision Reguladora De Energia (CRE, Mexico), Commonwealth Edison Company, Conectiv, Constellation Power Source, Coral Power, Detroit Edison Company, Duquesne Light Company, Dynegy, Edison Electric Institute, Edison Mission Energy, Electricity Corporation of New Zealand, Electric Power Supply Association, El Paso Electric, GPU Inc. (and the Supporting Companies of PJM), GPU PowerNet Pty Ltd., GWF Energy, Independent Energy Producers Assn, ISO New England, Luz del Sur, Maine Public Advocate, Maine Public Utilities Commission, Midwest ISO, Mirant Corporation, Morgan Stanley Capital Group, National Independent Energy Producers, New England Power Company, New York Independent System Operator, New York Power Pool, New York Utilities Collaborative, Niagara Mohawk Corporation, NRG Energy, Inc., Ontario IMO, Pepco, Pinpoint Power, PJM Office of Interconnection, PPL Corporation, Public Service Electric & Gas Company, Reliant Energy, Rhode Island Public Utilities Commission, San Diego Gas & Electric Corporation, Sempra Energy, SPP, Texas Utilities Co, TransÉnergie, Transpower of New Zealand, Westbrook Power, Western Power Trading Forum, Williams Energy Group, and Wisconsin Electric Power Company. The views presented here are not necessarily attributable to any of those mentioned, and any remaining errors are solely the responsibility of the author. (Related papers can be found on the web the web at www.whogan.com).