ELECTRICITY MARKET REFORM: APPA'S JOURNEY DOWN THE WRONG PATH

April 16, 2009

John Chandley
Principal at LECG

William W. Hogan
Raymond Plank Professor of Global Energy Policy
John F. Kennedy School of Government
Harvard University
Director at LECG

LECG

Contents

Executive Summary . . . 1

I. Introduction . . . 5
  A. The APPA's Search for RTO Reforms Has Been on the Wrong Path . . . 5
  B. APPA's Evolving Competitive Market Plan Is Still Seriously Flawed . . . 7

II. Effective Contracts Require A Consistent Market Design . . . 10
  A. Confusing Prices and Costs . . . 11
  B. Breaking the Logical Link Between Spot and Forward Markets . . . 13

III. APPA's Framework Would Not Reduce Long-Run Resource Adequacy Costs . . . 15

IV. Contract Scheduling, The Wrong Contract Path, Again . . . 18
  A. Reduced Grid Access and Trading . . . 18
  B. Unworkable Physical Transmission Rights . . . 19

V. The APPA Reforms Could Frustrate Federal and State Policies . . . 22

VI. The APPA Proposals Would Overburden Market Monitors But Not Reduce Prices . . . 24

VII. The RTO Approach Is A Proven, Superior Model . . . 27

VIII. Conclusion . . . 31

Appendix A: What Have We Learned About RTO Costs and Benefits? . . . 32

Appendix B: The Contract-Scheduling Structure Would Impose Major Costs on RTO Regions . . . 38

Appendix C: Experience From California to New England Shows The Risks of Restricting Access to RTO Spot Markets . . . 41

Appendix D: Estimated Costs of Reacquiring Generating Capacity . . . 44

Appendix E: Estimate of Increased Energy Costs within PJM . . . 50

Appendix F: APPA's Proposals Would Not Save Money On Capacity Costs . . . 54

Executive Summary

Regional Transmission Organizations (RTOs), currently manage much of the United States transmission system and provide reliable electricity service to more than two thirds of the nation's electricity consumers. Most RTOs also run day-ahead and real-time electricity markets to support reliable operations and capacity markets to support adequate investments; parties use the markets to buy and sell a variety of products and services. Those markets have evolved over time. The market model used in most of the United States is sound and has been endorsed by the International Energy Agency.

Various critics, including the American Public Power Association (APPA), propose major changes to RTO markets, but their reforms move in the wrong direction. The critics do not acknowledge the special characteristics of electricity that underpin these markets or the experience that led to the current designs. The markets should not be replaced in favor of discarded or untested alternative market models that ignore the many lessons we have already learned.

Proposals to unravel this successful market model threaten the investment required to maintain system reliability and promise to complicate the ability of independent market monitors to police against anti-competitive behavior and potential manipulation. Further, expectations that alternatives to this market model will result in lower prices for consumers are illusory. Rather, inconsistent and untested alternative market designs could cost consumers billions of dollars.

In the February 2009 publication, *Competitive Market Plan*, APPA offers another version of its evolving proposals to restructure organized electricity markets. The latest *Plan* follows APPA's *Consumers in Peril* (February 2008) and related papers maintaining that RTO markets cause electricity prices to be too high and do not lead to sufficient investments in new generation and transmission. But APPA's analyses are at odds with experience and reflect misunderstandings of how the RTOs and electricity markets work. The evidence shows that RTO market prices are not too high, and the studies APPA cites do not support a conclusion that the costs of RTO markets exceed the benefits they provide. Indeed the relevant evidence, much of it ignored by APPA, shows substantial benefits from RTO market design so obvious they appear to be "invisible in plain sight." (Appendix A)

Central to RTO markets are bid-based auctions, and a thrust of APPA's reforms is to limit or discourage use of these RTO auctions and instead somehow compel electricity suppliers (generators) to offer better contract terms to utilities and other load-serving entities (LSEs). There is nothing wrong with having long-term contracts, and RTOs are purposely structured to support voluntary contracting.

There is no evidence that contracting is failing in, let alone because of, RTO markets. Most of the trading done in RTOs today is through contracts easily accommodated by the RTO. But APPA asserts that suppliers will not contract at terms APPA deems acceptable because RTO markets offer too many choices. Much of the APPA critique flows from this premise. That buyers seek lower prices is neither unusual nor surprising. But this premise is hardly a sound basis for revising public policy.

The Proposal: Limit Spot Markets, Impose Long-term Contracts on Better Terms

Despite the unsupported diagnosis, APPA has proposed, in varying forms, two types of reforms; one to limit spot markets and another to impose long-term contracts on "better" terms. The details keep changing, sometimes in contradictory ways, and often in ways that would undermine how the RTO must operate the electricity system both to maintain reliability and provide all parties with open, non-discriminatory access to the transmission grid. Keeping the lights on at the lowest cost is an RTO's core function, and doing it while providing parties non-discriminatory grid access is a Federal mandate.

First, APPA seeks to prevent parties from relying too much on RTO spot markets. Under today's rules, parties are free to use RTO markets as much or as little as they want or need. Parties can contract to cover their loads and use RTO spot markets to deal with imbalances, or they can use RTO spot markets to supply their loads, or they can use a combination of contracting and spot market transactions. Parties are free to determine the mix between contracting and spot transactions. APPA would somehow restrict that choice, attempting to compel all parties to rely heavily on bilateral contracts.

To accomplish this, APPA would seek to artificially suppress spot market prices. The RTOs organize bid-based day-ahead and real-time spot markets to dispatch generation and ensure reliable operations. APPA would replace these with a real-time "optimization market" and some unspecified means for dealing with day-ahead commitments. In the latest version (*Competitive Market Plan*), supplier participation in the "optimization market" would be mandatory; all generators would be required to submit supply offers to the RTO for the real time optimization market at prices approved in advance by the RTO market monitors. Offer prices would be individually set at an estimate of short-run marginal costs. The proposal would essentially re-regulate all generation on a less-than-cost-of-service basis.

The first effect, and apparent hope, is that spot prices would be artificially suppressed. But unintended consequences would undermine short-run reliability and long-run resource adequacy. Suppressing spot prices would reduce incentives for resources to be available during shortages and keep total revenues below levels needed for adequate investment.

Second, APPA would require the RTO to implement (impose) a requirement that LSEs and suppliers trade almost exclusively through contracts, preferably long term, to cover any loads not met by the LSEs' own generation. The mandate for forward contracts would then buttress the goal of limiting use of the RTO spot markets.

APPA is silent on how the RTO would enforce this regime. Somehow, LSEs would be required not to plan on RTO spot markets to meet any portion of their needs except for inadvertent imbalances, even when spot market prices seemed attractive. And somehow suppliers would be compelled to contract with LSEs at terms the LSEs preferred. How this would be achieved is not explained. If everyone had been compelled to sign contracts at the higher prices existing a year ago, would this have been better than taking advantage of lower spot prices that exist today?

The Unfinished Evolution in APPA's Reforms

The preferred contracting framework has undergone significant evolution since *Consumers in Peril*. In one sense, such an evolution might be viewed as progress as the APPA confronts the realities of electricity systems that others have learned and embodied in the current RTO market design. However, a continuing missing chapter in the APPA analysis is any forthright description of the special characteristics of electricity systems that underpin the current RTO market structure. The several elements of bid-based auctions, economic dispatch, security constraints, locational prices, unit commitment, long-term contracts and capacity markets all work together to solve the complicated coordination problems that come hand-in-hand with an integrated transmission grid. The RTO market design elements are there for a good reason, and the lessons about missing pieces were learned at great cost. The APPA continues to sidestep the issues or give new labels to old ideas ("optimization market") that obscure the message and ignore the lessons of the past.

For example, in its early version, and in a November 2008 article in the *Energy Law Journal*, APPA officials described a market design best described as "contract scheduling." The model has a history, and it is not encouraging. With its limits on spot markets, the contract-scheduling model contains features originally proposed by Enron and others in the initial restructuring debates in California, PJM and New York. Early experiments with these features turned out to be costly policy mistakes, as we describe in Appendix C.

The contract-scheduling model is also familiar in non-RTO regions. It describes how utilities operate the transmission system where there is no independent RTO. The utility system operators are not required to use their ability to redispatch generation to provide third parties the same, non-discriminatory access to the transmission grid they provide to themselves. Most notably, third parties do not have equivalent access to economic dispatch.

The core features of contract scheduling are limited access to the dispatch (and related spot market) and requirements to obtain physical transmission rights to match bilateral contract schedules. These features have been shown to reduce grid utilization while increasing the risks of schedule curtailments. That means less economic trading across the grid and thus higher costs. A separate study performed by Ventyx (Appendix E), estimates that such restrictions, if implemented, could increase PJM customers' energy costs by about $13 billion over the next decade.

Moreover, the regions with limited spot market access and the contract-scheduling framework generally involve large vertically integrated utilities that operate the transmission system. This is key to making this model work, in that one entity owns the generation and transmission, provides ancillary services, and controls the dispatch. Recreating this framework in RTO regions would be extremely costly because RTOs do not own or control generation, and if bid based markets were constrained then the RTO (or delivery utilities) would have to acquire or contract with multiple generators in order to obtain scheduling rights to be able to operate the system reliably. We estimate (Appendix D) that the cost of reacquiring previously divested capacity for this purpose would cost utilities in the PJM region as much as $130 billion.

Most RTOs abandoned this approach (limited spot markets, physical rights, contract-scheduling) years ago. The superior approach used in RTOs today makes their bid-based spot markets and associated dispatch open to all parties on a non-discriminatory basis. Parties rely on the dispatch for balancing and use the open spot market to buy and sell energy to any degree they find beneficial. The RTO arranges the dispatch to keep the system balanced at the lowest as-bid cost; it adjusts the dispatch at the lowest as-bid cost to change electricity flows to manage congestion, so that no transmission line exceeds safe operating limits.

APPA's Formula for Shortages

Parties naturally tend to sign forward contracts at prices that reflect their expectations of what spot prices would be over the forward period. Deliberately suppressing spot market prices would logically lead to suppressed contract prices as well. That is apparently what APPA hopes. But suppressed spot prices plus suppressed contract prices add up to shortages, because potential supply investors would have no way to recover sufficient market revenues to support the level of supply investment needed to meet regional reliability (reserve margins) requirements. And APPA doesn't have a solution to that problem except to assume the RTO will fix it. But what options would the RTO have to ensure adequate supply?

A major goal of RTO market critics, especially APPA, has been to eliminate the RTO capacity markets and the associated requirement that LSEs make capacity payments to generators. Such payments provide revenue to generators to cover their investment costs and supplement energy market prices. If market prices are suppressed, and capacity payments are eliminated, then we have an investment problem, which will eventually become a shortage or reliability problem. On this point, APPA is sticking its head in the sand. The APPA reforms would not achieve the stated APPA objectives, much less achieve an improvement in RTO market design.

Most RTOs use capacity markets and payments because (among other reasons) their energy market rules prevent spot prices from reflecting scarcity costs when resources are short of desired reserve levels, as may happen a few hours each year. The problem is not that spot energy prices are too high, as APPA contends, but that they are too low when the system is short of resources. The "missing money" must be recovered in some fashion in order to provide the total revenues needed to support the desired investment level. That is the function of capacity markets in RTO regions. The same need to recover full investment revenue requirements would also apply in a fully regulated, cost-of-service regime.

APPA's Unworkable Framework: The Wrong Path Again

Altogether, APPA's proposals would create an unworkable framework and impossible dilemma for RTOs and the regions they serve. Spot market prices would be suppressed, reducing incentives for generators and demand-side response to be available when most needed. Spot prices would be even further below levels needed to support investments. Contract prices would also tend to be suppressed, but if not, the RTO would somehow force prices to levels acceptable to APPA members, while capacity payments were eliminated. Yet despite suppressed market prices, somehow investors could be persuaded to build enough capacity to meet the regional reliability standards.

The math doesn't add up, and the formula would lead eventually to shortages and necessary discriminatory rules. Once this became apparent, we would need to return to better spot pricing to improve incentives and encourage contracts, and probably some form of capacity payment to achieve the desired investment levels. But APPA proposals work against the direction of improving spot markets and providing improved incentives for real-time availability, long-run investments and energy efficiency that RTOs need and are developing.

The APPA analysis is internally inconsistent, and its proposals disconnected from the real requirements of operating electricity systems. As the accompanying paper demonstrates in detail, the APPA proposals point down the wrong path, again.

John D. Chandley and William W. Hogan[1]

I. Introduction

Beginning with the Energy Policy Act of 1992, the United States undertook an intense period of experimentation and regulatory innovation to restructure electricity systems.[1] Repeated policy reviews and supporting legislation have reinforced this process.[2] As part of this restructuring, the US developed widespread organized electricity markets coordinated by Regional Transmission Organizations (RTOs).[3]

RTOs now reliably serve over two thirds of US electricity consumers. But they are under attack by critics who argue that RTO markets are causing or enabling higher electricity prices and that the way to force lower prices is to substantially reduce RTO market functions. The American Public Power Association (APPA) and other early market supporters have called for major reforms of RTO-organized markets. Here we examine their key proposals.

The critics write that RTO markets are not working as anticipated and do not produce prices consistent with "just and reasonable rates." These concerns have intensified in recent years as electricity prices rose, against unrealistic expectations that markets alone would lead to lower prices even if the industry's underlying cost structure was rising. But during this period, prices were rising nationwide, in RTO and non-RTO regions alike, because of rising costs of the coal and natural gas that power a majority of US generating capacity. In many US regions, coal and gas-fired resources are often the marginal units that determine market-clearing wholesale prices. Costs for basic construction materials and equipment also increased markedly.[4]

APPA's main criticism focuses on RTO bid-based auctions for buying and selling power. The RTOs operate hourly spot markets for energy and ancillary services (such as operating reserves) which select the lowest-cost plants to keep the lights on, and forward capacity markets that help pay for adequate resources. The claim is that spot prices are too high, capacity payments are unwarranted, and suppliers are exercising market power.

A. The APPA's Search for RTO Reforms Has Been on the Wrong Path

The critics' reforms wrongly target RTO markets without acknowledging what these markets do or why they are needed. It is beyond dispute that electricity systems require central coordination to keep the lights on. RTOs use bid-based markets to select the lowest-cost resources to perform these coordina-

---

1 William W. Hogan, "Electricity Market Restructuring: Reforms of Reforms," (hereafter "Reforms of Reforms"), Journal of Regulatory Economics, Kluwer Academic Publishers, Vol. 21, No. 1, 2002, pp. 103-132.

2 Joseph T. Kelliher, "Statement of Chairman Joseph T. Kelliher," Federal Energy Regulatory Commission, Conference on Competition on Wholesale Power Markets AD07-7-000. February 27, 2007.

3 Prior to 1996, the US grid was operated by publicly and privately owned electric utilities and federal power marketing agencies. Closely interconnected utilities sometimes "pooled" their operations, allowing a central regional power pool to operate the interconnected systems as one system. For example, the PJM power pool was created in 1927 to operate the combined grids and dispatch generation for member utilities in Pennsylvania, New Jersey, Maryland (hence "PJM") as well as Delaware and the District of Columbia. PJM became an "Independent System Operator" (ISO) in 1997 and began coordinated market operations then. After 2000, the Federal Energy Regulatory Commission (FERC) redesignated several ISOs—PJM, ISO New England (ISO-NE), Midwest ISO (MISO)—as RTOs. Essentially similar organizations now include the New York ISO (NYISO), California ISO, the Southwest Power Pool (now an RTO), and ERCOT (an ISO covering most of Texas but not subject to FERC jurisdiction). Each coordinates organized markets with many of the features of a regional power pool. For simplicity, we refer to all such markets as being administered by an RTO, even though some of the markets are actually administered by ISOs rather than RTOs.

4 In the mid-1990s, state restructuring decisions froze retail rates for a transition period. When the transition periods expired in recent years, retail rates had to be adjusted to reflect current (higher) wholesale market prices facing utilities and other load-serving entities (LSEs).

tion functions. Faced with the need to provide both this coordination and non-discriminatory access to the electricity grid, RTOs must operate under a market design that is internally consistent and compatible with the special technical features of the electricity system. Eliminating the RTO coordinated markets would make the necessary coordination function more difficult and expensive, while creating a need to invent solutions to solve the coordination problems.

Predictably, the critics' redesign proposals vary widely, but because the new solutions are often unworkable, they keep changing, suggesting an evolving, incomplete appreciation of how RTOs must function. The result has been a series of *ad hoc* proposals that often resurrect flawed approaches that have already been considered and rejected, or tried and failed.

APPA's latest *Competitive Market Plan* proposal[5] reflects this continuing *ad hoc* search for a workable redesign. Beginning with the publication, *Consumers in Peril* [6] and continuing with an article in the Energy Law Journa,[7] APPA officials argue for a redesign of RTO-coordinated electricity markets, dismantling some markets, restricting others, and forcing parties into contract arrangements they claim the RTO markets do not support. With its *Competitive Market Plan*, a hopeful evolution is apparent, but it is not complete; a coherent framework that recognizes what RTOs must do and why is still missing. Also missing is the realization that its proposed reforms would not achieve APPA's expressed goals.

In *Consumers in Peril*, APPA acknowledged important transmission benefits provided by RTOs but argued these benefits do not exceed RTO costs. We examine these claims in Appendix A and show how they misread cost/benefit studies and ignore quantitative and qualitative empirical evidence of benefits provided by RTOs.

The *Consumers* paper called for the elimination of current bid-based RTO spot markets. These would be replaced by restricted access (no more than 5 percent of load) to a limited, residual balancing market. But the discussion left unclear how the reliability-related functions of these spot markets would be performed. Indeed, it is doubtful that the vaguely defined framework APPA proposed in *Consumers in Peril* would have allowed RTOs to perform the most basic functions of keeping the lights on, let alone supporting a viable electricity market.

APPA promised further details, and the November 2009 *Kelly/Caplan* article included proposals to implement APPA's intentions, offering a "hybrid" design the authors dubbed "Day 1.5 RTO."[8] (Current RTO designs are called "Day 2 RTO.") However, the Day 1.5 RTO model actually described a pre-RTO market design familiar to those in non-RTO regions and included features that were tried and failed in the 1996-2001 period in California and elsewhere. That design, which we call a "contract-scheduling" model, would probably have required the dissolution of RTOs and regional grid operators (power pools), forcing a reversion to utility-by-utility dispatch to sustain reliable operations.

It is important to understand how APPA's earlier proposed contract-scheduling model would have functioned, so we devote Chapter IV and related Appendices to explaining its features, costs and disadvantages. While APPA has now moved to yet another proposal, there are still remnants of this flawed approach in the most recent *Competitive Market Plan*. Returning to that framework, once advocated by Enron and partly implemented at great cost in California (see Appendix C), would be a serious policy mistake.

---

5 APPA, *Competitive Market Plan*, February 2009.

6 APPA, *Consumers in Peril*, February 2008; available at http://www.appanet.org/pressroom/index.cfm?ItemNumber=18029&sn.ItemNumber=16668. Similar critiques come from Electricity Consumers Resource Council (ELCON) available at www.elcon.org.

7 Susan Kelly and Elise Caplan, *Time for a Day 1.5 Market: A Proposal to Reform RTO-Centralized Wholesale Electricity Markets*, 29 *Energy Law Journal* 491 (2008). Kelly is APPA's Vice President of Policy Analysis and General Counsel and Caplan is the Coordinator of APPA's Electric Market Reform Initiative, but the authors state at 491: "All statements in the article, however, are the authors' alone and should not be attributed to the APPA." This article is hereafter referred to as "*Kelly/Caplan*" and the revised proposal is the "Day 1.5" proposal. In *Consumers in Peril*, APPA promised further details, but *Kelly/Caplan* has now been superseded by yet another proposal.

8 *Kelly/Caplan* at 533.

In the contract-scheduling model, parties would be required to secure owned generation and/or forward contracts to cover their entire load and demonstrate the sufficiency of those plans in advance to the RTO.[9] Parties would then gain access to the transmission grid by obtaining rights to schedule their own generation and contract deliveries with the RTO, but without the flexibility provided by the RTO spot markets in facilitating those schedules.[10]

Forcing such schedules and limiting planned access to the balancing market implies some form of physical transmission rights, instead of the financial rights used by RTOs today. While *Kelly/Caplan* alluded to the transitional problems,[11] the preferred approach was a transmission access scheme that would revert to a system of physical transmission rights in which parties would reserve physical capacity on specific transmission lines to match their desired schedules. RTOs largely abandoned the physical rights approach years ago, because it does not account for actual power flows and thus requires subsequent curtailments to maintain reliability. Under the APPA framework, the RTO's current system of "financial transmission rights," which do not require physical reservations but do account for actual flows, would be phased out.[12]

A contract scheduling approach with physical transmission rights would be a costly step backwards. As we illustrate in the Appendices, just to perform the essential dispatch function, the design would have required at least some utilities to reacquire — at an estimated cost of $130 billion — generating capacity they divested a decade ago. In addition, the contract-scheduling model would have reduced inter-area trading, making transactions less likely or more costly, thus increasing costs to consumers by another $13.6 billion over a decade.

B. APPA's Evolving *Competitive Market Plan* Is Still Seriously Flawed

In its latest proposal, *Competitive Market Plan*, APPA has abandoned (for now) portions of the earlier approach, particularly its formal reliance on physical transmission rights. Financial transmission rights (FTRs) would be preserved. Other elements of the contract-schedule model are retained, such as the requirement that LSEs secure advance RTO approval for each LSE's plans to serve its load through owned generation and/or bilateral contracts.[13] The new *Plan* does not explain what would occur if all requested generation were not simultaneously deliverable to all loads given current transmission limits. But these are the details that matter and that help determine the current RTO design.

Once again, planned reliance on the spot market is forbidden, even though the spot market would be available in real time for unplanned imbalances. Inevitably, limiting access to the spot market would lead to resurrecting the associated contract scheduling requirements.

A consistent APPA goal has been "to deemphasize the role of RTO-run centralized power supply markets and provide support for a stronger bilateral power supply contracting regime."[14] These goals then translate into forcing load serving entities (LSEs) and power suppliers to rely almost exclusively on bilateral contracts while restricting their option to use the RTO auction-based spot markets to buy and sell power. There would be a limited "balancing market," (which *Competitive Market Plan* calls an "optimization market") while today's day-ahead and real-time bid-based spot markets would be phased out.[15]

APPA's "optimization market" would perform system-wide central dispatch, arrange and pay for ancillary services (*e.g.* operating reserves), and provide balancing for parties' schedules. Importantly, this optimization market retains certain core functions found in the RTO real-time spot market: security-

9 APPA retains this feature in *Competitive Market Plan* at 27.
10 *Kelly/Caplan* at 539.
11 *Consumers in Peril* at 27. *Kelly/Caplan* at 534, footnote 204.
12 *Kelly/Caplan* at 534, footnote 204, 535.
13 *Competitive Market Plan* at 4.
14 *Kelly/Caplan* at 491.
15 *Kelly/Caplan* at 535, 539.

constrained economic dispatch, clearing prices ("for the near future") based on locational marginal pricing (LMP)[16] and FTRs to help offset congestion costs. These are essential features for a real-time spot market, but then APPA compromises them without explanation. Thus, while there is a partial evolution away from the flawed contract scheduling design, there are still troubling departures from how today's markets function; if implemented, these departures would prove to be unworkable and even exacerbate the problems APPA claims to be solving.

For example, APPA's *Plan* would force suppliers to reveal and the RTO to use a simplified statement of each generator's short-run marginal costs (SRMC). Without explaining how, these SRMCs would be individually verified and continuously updated by the RTO's Market Monitors.[17] Mandatory participation would be enforced by a must-offer requirement.[18] The RTO would then be required to use these SRMC estimates as the basis for dispatch and associated optimization market pricing.

With some exceptions, all generators would be required to participate in the RTO dispatch/optimization market. This is a form of "mandatory pool," a model that hasn't been mentioned since the early California debates on market design. When combined with the forced use of SRMC, the approach is analogous to how vertically integrated utilities conduct a dispatch when they own all of the generation and dispatch is based on internal company information. It is not an exaggeration, therefore, to describe this approach as akin to detailed less-than-cost-of-service regulation.

Another serious concern is APPA's proposal to prohibit generators (and the RTO) from considering a generator's opportunity costs — *e.g.*, what a supplier could receive from selling into a neighboring market — as a basis for dispatch and spot market participation.[19] Recognition of opportunity costs is standard, textbook economics, and prohibiting any supplier from using opportunity costs would result in distorted incentives and suppressed market prices. These features would encourage withholding or discourage generators and demand-side responses from being available when most needed.

The suppressed prices would also undermine investment. Indeed, by suppressing spot prices, the prohibition would exacerbate the "missing money" problem that currently serves to justify capacity markets, which APPA also seeks to eliminate.

Yet another problem with APPA's *ad hoc* "optimization market" design is that it would not actually optimize the RTO's decisions about which generators should be dispatched for energy and which held as operating reserves. Getting that right requires that each set be paid clearing prices for each service provided, so that the prices both minimize the RTO's total costs and maximize value to each provider. Each provider then has the incentive to follow the RTO's instructions, and no generator regrets being told to provide reserves instead of energy (or the reverse). The APPA mistake is proposing to pay generators at cost to provide reserves, rather than a clearing price optimized between energy and operating reserves.[20] This design error would distort incentives to follow dispatch instructions and encourage reserve shortages.

---

16 APPA obscures the fact that its *Plan* retains Locational Marginal Pricing (LMP), which APPA disparaged in *Consumers in Peril* and *Kelly/Caplan*. In *Competitive Market Plan*, the term "LMP" is never mentioned, but the *Plan* states: "RTOs would continue to provide transmission service under open access transmission tariffs (OATTs), dispatch generating units in merit (lowest cost) order subject to system constraints, *determine price differentials arising from congestion*, and assist LSEs in hedging congestion." (emphasis added) *Competitive Market Plan* at 29. In today's RTOs, LMP spot prices reflect price differentials arising from the dispatch to deal with congestion. Similarly, APPA's "optimization market" is in fact a real-time spot market, which APPA previously sought to eliminate. To improve the dialogue, it would be helpful if APPA would use the terms everyone else uses, and simply acknowledge that RTO elements it once criticized are in fact essential and must be retained.

17 APPA concedes that it would be difficult for the RTO to maintain an accurate, up-to-date analysis of every generator's SRMC, as fuel and other cost components varied daily. APPA merely assumes the RTO could solve this without explaining how. Yet this is a principal reason why RTO pricing rules create strong incentives for the generators themselves to determine, and bid, their marginal costs. Today, RTO Market Monitors set SRMC-based limits on supply offers only in those situations in which market power might be expected. This more manageable approach gives the RTO reasonable assurance that offers will tend to track actual SRMC and/or opportunity costs, without requiring the RTO to track and verify every possible component and change in every generator's cost structure.

18 *Competitive Market Plan* at 25, 27.

19 *Competitive Market Plan* at 25.

20 *Competitive Market Plan* at 28.

Another misguided reform relates to FTRs. While retaining FTRs (and giving up on a return to physical rights for now), APPA would complicate the task of allocating FTRs to grid users. APPA would eliminate the RTO's monthly and annual FTR auctions by which short- and long-term FTRs are currently allocated and traded. Instead, the RTO would simply allocate FTRs to LSEs annually, giving a preference (without any apparent justification) to those with long-term contracts.[21] This would seem to favor APPA members who do not own transmission, but the details are missing and the intent is not clear.

Eliminating periodic FTR auctions is consistent with APPA's general opposition to "bid-based markets," but just as APPA now concedes the RTO must have a real-time spot market (renamed "optimization market") it would eventually discover the RTOs also need periodic FTR auctions. No RTO began with such auctions, but every RTO and their members eventually concluded such auctions were worthwhile.

In the meantime, eliminating the current FTR auctions would leave unanswered how the RTO would solve the problems these auctions address. For one thing, the auctions are a simple, proven way for parties to acquire FTRs, to exchange those they have for ones they'd prefer, and to exchange them with other parties. That is why many commodity markets create central exchanges.

The auctions also solve the difficult problem of deciding how many FTRs to allocate, and which ones. The grid cannot support an unlimited number of FTRs, nor can it support a condition in which all the requested FTRs are between the most preferred grid locations. While noting that the RTO would have to confirm the simultaneous feasibility of any FTR allocation requests,[22] APPA does not explain how the RTO would solve the problem if, as frequently occurs, LSE requests for FTRs were not simultaneously feasible.

Today's RTOs use periodic FTR auctions to solve the feasibility problem. Each auction allocates that period's FTRs to those who value them the most, up to the limits of simultaneous feasibility, but no further. The auction winners receive a set of FTRs that are simultaneously feasible. Auction revenues then revert to those who pay the grid's embedded costs, a solution most parties agree is fair and workable. Without explanation, APPA would eliminate these useful auctions without offering any alternative means to implement their functions.

APPA's proposals appear to be still evolving. So these and other spot market design errors in *Competitive Market Plan* might be cured as APPA continues to work through the reasons why RTOs do what they do. But there are more fundamental problems with APPA's approach to forward contracting and APPA's unrealistic expectations about eliminating the need for capacity payments (or some other solution) to solve the "missing money" problem. We examine these issues next.

21 *Competitive Market Plan* at 29-31.
22 *Competitive Market Plan* at 31.

II. Effective Contracts Require A Consistent Market Design

APPA's *Competitive Market Plan* continues its goal of forcing virtually all trading into (preferably long-term) bilateral contracts and self-owned generation.[23] There are numerous problems with this coerced approach, starting with the point that there is nothing prohibiting suppliers and LSEs from contracting today, for whatever period they choose and at whatever price they agree. Hence, the problem is not contracting *per se*.

Long-term contracts can be and are an important part of the electricity market. However, effective contracts depend on a market design that is both internally consistent and reflects the essential features of the electricity system. The APPA discussion of the role of contracts and their connections to the remainder of the market illustrates the critical missing chapter in its critiques. There is no coherent diagnosis of how the electricity system works or how the pieces fit together.

APPA's complaint is not that bilateral contracts aren't possible and fully accommodated by today's RTOs — they are used extensively in today's RTOs, as APPA concedes[24] — but rather that suppliers won't agree to terms APPA's members prefer. APPA claims this is because suppliers can always sell into RTO spot markets, and spot prices are inflated by excessive supplier offers setting the clearing prices. But RTO Market Monitors have periodically evaluated and rejected these claims,[25] and FERC has agreed.

APPA further claims that suppliers are unreasonably demanding that contract prices reflect expectations of future spot market prices and related risks.[26] But this connection between contracts and expected spot prices over the life of the contract is an expected feature of properly functioning markets. Rather than being evidence of failure of the market design, the connection between contracts and spot markets is a sign that the RTO markets are functioning as designed.

APPA rejects this well-understood logic. It insists on breaking the logical link between expected spot and forward contract prices, without considering the poor incentive effects this would have on parties' contracting or dispatch behavior.

A persistent priority of RTO market critics has been to constrain or discontinue RTO "centralized bid-based locational capacity markets." In the evolving APPA proposals, these capacity markets would be replaced by a vaguely defined capacity resource planning and acquisition scheme involving LSEs, state regulators and the RTO. Regional planning is desirable, and RTOs currently coordinate it, so there is little new here. At the end of its process, APPA claims, LSEs would ultimately build their own capacity or acquire it under contract (just as they can do today), but there would be no RTO-coordinated auctions for buying and selling capacity.[27]

It is here that the APPA bilateral contracting structure starts to break down, with APPA making several

---

23 *Competitive Market Plan* at 19-21.

24 APPA notes in *Competitive Market Plan*, at 21, that the vast amount of trading in RTOs already occurs through bilateral contracts. In PJM, it's 96 percent and varies in other markets. Note that the RTOs do not (because they have little need to) track all bilateral contracting; most contracting occurs between the parties without the RTO's knowledge, even though both may be buying and selling through the RTO spot markets, while settling net differences between themselves. Thus, rather than limiting bilateral contracting, the RTO spot markets appear to be fostering and accommodating massive bilateral trading.

25 See, *e.g.*, PJM's *2007 State of the Market Report*, available at www.pjm.com One such finding states: "The overall results support the conclusion that prices in PJM are set, on average, by generating units operating at or close to their marginal costs. This outcome is strong evidence of competitive behavior." PJM press release, March 11, 2008, accompanying release of the 2007 Report. More recently, the *2008 State of the Market Report*, at 2, summarizes the Market Monitor's findings: "The MMU concludes that in 2008: *The Energy Market results were competititive; *The Capacity Market results were competitive; *The Regulation market results cannot be determined to have been competitive or to have been noncompetitive; *The Synchronized Reserve Markets' results were competitive; *The Day Ahead Scheduling Reserve Market results were competitive; and *The FTR Auction Market results were competitive."

26 *Competitive Market Plan* at 19.

27 *Kelly/Caplan* at 535.

simultaneous but incompatible assumptions. APPA seems to assume that the RTO can force generators to accept suppressed prices in the spot market and lower contract prices, while *at the same time* supply investors agree, independent of these prices, to build the amount of capacity the RTO sets as the resource adequacy goal. And investors will do this even though the structure does not provide capacity payments to make up for the suppressed spot and contract revenues and even though the market prices fail to provide price signals about the need for investments. There is no workable business or economic model to match this set of inconsistent expectations, unless it is premised on confiscation.

So how could it all work? APPA leaves that important question unanswered, essentially assigning to the RTO the task of "implementing" the proposed resource adequacy framework but without the ability to set spot prices sufficient to support the required level of investments or to use capacity payments to make up the difference.[28]

Today's RTOs could solve the problem either by allowing spot prices to reflect scarcity — prices are capped below that level today — or by supplying the "missing money" through capacity payments, or more likely, some combination balancing the two. But APPA leaves RTOs with the unresolved investment problem while prohibiting the logical solutions, which means the entire framework is unworkable.

The flaws and inconsistencies in the APPA analysis appear in related ways such as confusing the distinction between prices and costs, breaking the logical link between spot and forward markets, and failing to confront the real challenges of long-run resource adequacy and its costs.

A. Confusing Prices and Costs

The APPA analysis discusses several different measures of price and relates these to its view of costs under traditional regulation. In some discussions, "prices" refer to spot prices for energy: "The prices for electric power in these centralized markets are set at specified intervals (every hour or a given time interval within an hour) based on the offers to sell power submitted by generation owners, operators and marketers to the RTO." [29] In other contexts the reference is to retail prices which include the payments for energy and ancillary services as well as the capacity payments required by resource adequacy programs: "Restructured wholesale markets are producing both higher prices and higher profits than one would expect in a competitive market. Resulting retail prices exceed those prevailing in regions that have not restructured, but that instead retained traditional retail cost-of-service regulation and eschewed the formation of RTOs." [30]

Although prices were higher to begin with in regions that restructured, the argument is that prices are too high and would have been lower under cost-of-service regulation. However, that conclusion is based on posing the wrong question. And the conclusion for the correct question is more nuanced when we look at the performance of organized wholesale markets. In particular, and contrary to the critics' argument, the evidence indicates that a principal problem with RTOs is that spot market *energy* prices have been too low to support needed investment.

The critics' characterization of the theory for determining spot energy prices is only partly correct. "The RTO takes all power supply offers for a particular upcoming time interval in ascending price order, stopping with the last offer needed to meet the power needs of loads during that time interval." [31] Ignoring the effects of congestion, this is true during periods of excess capacity. Unfortunately, this rule does not establish the efficient price during periods of limited capacity and associated scarcity of generation offers.[32] During periods of scarcity, the market-clearing price should reflect the scarcity

28 *Competitive Market Plan* at 3.
29 *Kelly/Caplan* at 496 (footnote omitted).
30 *Kelly/Caplan* at 494.
31 *Kelly/Caplan* at 496.
32 William W. Hogan, "On An 'Energy Only' Electricity Market Design For Resource Adequacy," Harvard University, September 23, 2005, (www.whogan.com).

costs and clear at a sometimes (much) higher level. The higher energy revenues would be an important part of the contribution to recovery of the fixed costs of generation assets.

Compounding the difficulties, most RTO designs include features that preclude these sometimes higher prices. The lower prices result in the so-called "missing money" problem.[33] For a variety of reasons that include price caps, operating procedures and conceptual mistakes in translating theory into practice, energy prices in the electricity market have not been high enough to support investment in new generating plants.

For example, over the nine years from 1999 through 2007, the market monitor for PJM estimates that average energy market revenue under economic dispatch for a combustion turbine peaking unit was $16,401 per MW-year compared to an average fixed cost charge of a new unit of $75,158 per MW-year. The difference of $57,757 per MW-year is the missing money.

Estimates of expected net revenues going forward should be the proper benchmark, but this retrospective look at the actual revenues achieved net of variable costs is sobering and suggests a real problem in the underlying market design. There is inadequate attention to scarcity in spot energy prices, and spot market revenues are too low. The average net revenues were approximately 22 percent, 45 percent, and 63 percent of the levels needed to justify investment in a new combustion turbine, gas fired combined cycle or coal plant, respectively.[34]

The policy response has been to address the underinvestment and missing money problems by developing forward capacity markets. PJM's approach for providing these capacity payments, the Reliability Pricing Mechanism (RPM), receives much criticism from APPA and others.

In comparing market and cost-of-service paradigms, it is important to formulate the right question for evaluating prices. A comparison with cost-of-service rates presents the wrong question because it does not hold constant the allocation of risks. Implicit in the traditional regulated model is the assumption that customers bear the risk that the long-run cost of providing energy and other services will exceed the value of those services. Implicit in the restructured electricity model is a different allocation of risks, with the generators assuming the risk that the prices they receive for providing energy and other services will not be sufficient to cover the long-run costs of providing those services, unless they have entered into long-term contracts with buyers. Without controlling for these different risk allocations, and looking across the distribution of uncertain outcomes, there is not much of interest in the observation that under one set of conditions there is a price difference.

In evaluating the performance of RTOs and organized wholesale markets, the more relevant question is how the prices observed compare with the competitive outcome. Here the natural benchmark, particularly with growing demand, would be in the cost of new entry. If sustained prices were higher than needed to support the cost of new entry, there would be a cause for concern and we would be looking for the policy design flaws in RTOs that were preventing otherwise profitable entry. But in the present case the facts are reversed. Spot energy prices by themselves are too low to support entry, partially because the current RTO market designs give too little attention to the theoretical and practical requirements for better scarcity pricing.[35]

The resulting creation of capacity markets was intended to address the missing money problem and make up the net of the expected costs of entry to support new investment. The RTO critics seem well

---

33 The characterization as "missing money" comes from Roy Shanker. For example, see Roy J. Shanker, "Comments on Standard Market Design: Resource Adequacy Requirement," Federal Energy Regulatory Commission, Docket RM01-12-000, January 10, 2003.

34 PJM Market Monitoring Unit, *2007 State of the Market Report*, Volume 1: Introduction, Volume 2: Detailed Analysis, March 11, 2008, Tables 3-7 thru 3-9 (Vol. 2) & 1-3 (Vol. 1), respectively.

35 An additional contributor is that reliability requirements, such as 15 percent or higher reserve margins, may mandate the development of more capacity than is economic. The additional reserve capacity would tend to suppress the prices we observe in the spot markets below the levels required to induce entry even if those markets properly reflected scarcity.

aware that the costs of new entry are high and increasing: "Consumers are already very likely to face increased electricity prices in the coming years, given increasing infrastructure requirements, rising fuel and construction costs, and the need to comply with future carbon regulation." [36] However, the critics do not connect this reality to their analysis of the costs of capacity payments, the impacts on existing retail rates, or the evaluation of the performance of RTOs.

The implications of clarifying the question are straightforward and important. Spot energy prices should include scarcity costs and not be determined solely by the variable cost of the most expensive plant running. To the extent that spot energy prices are too low, there is an increase in the net capacity payment required to support new entry, and this higher capacity payment plays a greater role in determining retail prices.

B. Breaking the Logical Link Between Spot and Forward Markets

The varying APPA approaches target new rules for requiring forward bilateral contracts as a key ingredient in the program to change the operations of RTO markets. It is true that forward markets are important, and much of the RTO design is motivated by the intent to facilitate use of forward contracts struck between willing buyers and willing sellers.

The critique of the existing RTO markets implies that there is something wrong with the existing opportunities for forward contracting. Notably, the critics do not claim it is impossible to obtain forward contracts in the current markets. "Buyers and sellers in Day-Two markets can minimize purchases and sales of energy and capacity in the RTO-run markets by entering into individual power supply contracts (called 'bilateral contracts')."[37] Apparently, the problem is not the existence of contracts or contract counterparties; the problem is the price available.

> "But, the forward prices for energy sold under those contracts are substantially influenced by the prices the sellers can obtain for their power in the RTOs' centralized markets."[38]

> "A recent study which the APPA commissioned examining the relationship between RTO-run spot markets and bilateral contracting in RTO regions found that power supply transactions in the organized markets are dominated by the spot markets, even when much of the energy used to serve load is not directly procured through the RTO's spot markets."[39]

Given its critique of existing spot markets, APPA's implication is that there is a failure in forward contracting under reasonable terms and conditions.

It is for buyers and sellers to decide how much to transact under contract, at the price each is willing to accept. However, the RTO critics' analysis is not really a critique of the forward contracting opportunities under the RTO design. The analysis says little more than that customers would prefer to have contracts at lower prices. There is nothing in the analysis that translates into evidence of a failure of the RTO model to support forward contracts.

The observation that forward contracts are driven by expectations about future spot prices is fully in keeping with the economic theory underlying the RTO market design. It would be surprising if anything else were true in a market where buyers and sellers have a choice to contract or to rely on the spot market. In equilibrium, the natural forces of arbitrage should be enough to eliminate any risk-adjusted difference between new forward contracts and expected spot prices. This would be true for a competitive market, or for many other possible market structures that are not competitive.

36 *Kelly/Caplan* at 491.
37 *Kelly/Caplan* at 502.
38 *Kelly/Caplan* at 502.
39 *Kelly/Caplan* at 503.

However, the APPA approach anticipates breaking this logical connection between spot and forward markets. "This is consistent with the theory that sellers should be recovering their fixed costs (including return) through long-term bilateral contract arrangements, and not relying on short-term RTO market sales to recover such costs." [40] The main thrust of the contract-scheduling framework is to restrict choices, limit participation in spot markets and somehow mandate long-term contracts in order to create such a disconnect.

The forward and spot markets are both important. However, the observation that there is a linkage of prices between the two markets should not be construed as a failure of the RTO model. The linkage is part of the design, and a conclusion that the linkage is working should be interpreted as consistent with the design of the RTO markets. The critics' proposed structure would break something that isn't broken and that should be preserved.

40 *Kelly/Caplan* at 537.

III. APPA's Framework Would Not Reduce Long-Run Resource Adequacy Costs

Critics have voiced strong objections to RTO capacity markets, especially PJM's Reliability Pricing Mechanism (RPM). PJM market rules require each LSE to provide or purchase sufficient generating capacity to meet its share of the region's planning reserve requirement. PJM's RPM is one means to meet that requirement. But the requirement exists independent of RTOs and their capacity markets.

In the *Kelly/Caplan* article, and again in *Competitive Market Plan*, APPA proposes a different resource planning and acquisition framework for meeting the planning reserve requirement. That framework would do nothing to lower total electricity prices.

Under any framework, utilities and other LSEs must meet reliability standards set by their states and/or the regional entities responsible for setting those standards. In the PJM region, the standards currently require that each LSE own or purchase enough capacity to cover that LSE's expected peak loads plus a planning reserve margin of 15 percent or more.[41]

In PJM and other eastern RTOs, parties meet their reserve capacity obligations by building their own generation, purchasing capacity under bilateral contracts, or purchasing capacity in the RTO capacity markets. In PJM's capacity market (RPM), all capacity is accounted for in the annual capacity auctions; LSEs with capacity receive credit for the capacity they offer or self-supply; they are then paid the capacity market price for any surplus they offer beyond their own requirements, while paying the market price for any additional capacity they must purchase.

PJM's RPM forward auction framework defines the forward market value of capacity in each planning period, given the reserve target, the available resources and the offers/bids from auction participants. It also incorporates demand-side resources and transmission upgrades. Parties can use PJM's RPM auctions to meet their capacity obligations and to buy/sell capacity through means other than bilateral arrangements if they choose to do so.

LSEs are required to meet the mandatory capacity reserve requirements whether they function within PJM, whether PJM uses RPM or doesn't, or whether LSEs function as vertically integrated utilities outside an RTO framework. This means that for the same capacity reserve target, such as 15 percent reserve margins, the total costs of meeting the standards would likely be about the same under any structure.

To illustrate this point, Appendix F examines an alternative scenario that assumes PJM's RPM capacity market did not exist. Instead, each utility would be required to build or purchase sufficient generating capacity to meet its *pro rata* share of the mandatory capacity requirements. The comparison set forth in Appendix F illustrates that loads would have to pay essentially the same total costs in either case over time to acquire the same amount of capacity, because no matter what, the total costs associated with developing, maintaining and operating that capacity must still be paid. The assumption that eliminating PJM's RPM would reduce long-run resource adequacy costs is illusory.

The RTO critics call for an alternative process involving LSEs, generators, the RTO, state regulators and interested parties to determine the amount of capacity that should be developed. Each LSE would

41 The reliability standard is typically expressed to require sufficient capacity to ensure that the system will run short of capacity no more than one day (or one event) in 10 years. The one-day-in 10-year standard is then translated into an equivalent reserve margin for each system, which can vary depending on the reliability of transmission and generation available to that system. Systems with plants that suffer more frequent outages or more limited transmission import capacity must meet a higher installed reserve requirement; those with less frequent outages or greater import capability meet a lower installed reserve requirement.

then acquire the necessary capacity, except there would be no auction market operated by PJM where those with surplus capacity could sell it and those who needed capacity could buy it. At the end of this process, every LSE or utility would essentially self-supply its own or acquired capacity. PJM would play an expanded role in this planning process, with state regulators playing supporting roles. But even if states found this arrangement attractive, there are three observations worth noting.

First, if parties in the PJM region were enamored with APPA's approach, its results could be accommodated today within the PJM capacity market structure. In other words, states or utilities could undertake extensive integrated planning exercises, select the capacity resources they preferred and direct LSEs under their jurisdiction to build the resources or acquire the capacity through contracts. The LSEs could then offer that "self-supplied" capacity into PJM's RPM auction and receive credit towards meeting their PJM-regional capacity obligations. PJM's RPM construct accommodates self-supply.

Second, because LSEs/utilities can and do use this self-provision approach today to minimize net purchases through the PJM capacity markets, it is not correct to assume parties would save money over time by replacing the current PJM approach while retaining the same reserve requirement. All this would achieve would be to eliminate parties' ability to use RPM's periodic auctions to sell their excess capacity and purchase deficiencies to meet their respective requirements.

Third, this analysis implicitly assumes that it is possible to forecast each LSE's share of the capacity requirement far enough in the future to allocate each LSE's responsibility to develop additional capacity. But this ignores the existence of retail access and competitive LSEs. In states with retail access, loads may choose which LSE supplies their electricty. Since LSEs generally do not lock up their customers far in advance, LSEs do not know their shares of the capacity requirement that far in advance.[42]

An underlying misconception, running through much of the criticisms of RTO energy markets, is the assumption that prices are artificially higher in bid-based RTO markets, and that if we could just foreclose or limit use of the RTO markets, we could force suppliers to sell energy, ancillary services and capacity through contract arrangements for less than buyers/consumers pay now.

This central assumption is false; it ignores repeated findings by RTO market monitors that the total revenues generators receive from all energy and ancillary services markets is typically less than the generators' total fixed and variable costs, when calculated using cost-of-service methods. If that is true, then what critics are implicitly advocating is a set of rules that would discriminate among suppliers of capacity and allow buyers to force sellers to accept prices below market levels, while still expecting suppliers to build sufficient capacity to meet regional reserve requirements.

There is no theoretical basis to support these assumptions. Whether buyers and sellers rely on RTO-administered auctions or on self-supply or bilateral contracts, suppliers must receive total revenues that cover the fixed and variable costs of developing new resources. There is no magical, non-discriminatory set of rules that will allow the region to meet its planning reserve requirements without paying those costs.[43]

Finally, PJM's RPM construct does not add a capacity payment on top of the generator's revenue requirements. Instead, capacity payments under RPM reflect the difference between the margins that generators are expected to realize in energy and ancillary services markets and the total cost of developing and maintaining new capacity. Thus, if expected energy market profits went up, capacity payments under RPM would go down.

---

42 If forced to designate their loads before they are locked in, non-utility LSEs would have an incentive to underestimate their future loads and lean on the residual utility.

43 PJM's RPM construct has the effect of spreading out total capacity costs over time, rather than having those costs imposed in lumps as each set of "needed" new capacity is added to the system. This means that in any given year, buyers could pay more or less than they might under a fully regulated cost of service regime. But over time, the expected costs should be about the same.

This is an important relationship for those seeking to reform RTO capacity markets. Reducing the level of capacity payments required by controversial capacity market mechanisms can be accomplished by improving the rules by which RTOs price energy and operating reserves. Particularly during periods when the system is short of operating reserves, current RTO pricing rules fail to reflect the higher value of energy and operating reserves. The result is a gap in revenues that must be recovered through some other means–and that gap drives much of the need for capacity payment mechanisms.

Capacity markets and payment schemes are difficult to design and are probably not the ideal way to ensure sufficient revenues to support resource adequacy objectives. While they can fill the gap in revenues from incomplete energy and reserve markets, they may not provide the best incentives to encourage strong demand-side responses or generator availability in those rare periods when short-term capacity shortages arise. Improvements in scarcity pricing would close this gap and thereby reduce the need for capacity payments and the importance of capacity market constructs. Equally important, improved spot market pricing for energy and operating reserves would improve real-time price incentives for more responsive generation and demand-side investments and actions. The combined effect would be to lower total costs and improve reliability.

IV. Contract Scheduling, The Wrong Contract Path, Again

The contract-scheduling structure APPA originally proposed in *Consumers in Peril* and *Kelly/Caplan* would take us down the wrong path, again. At its core, the proponents' analysis assumes that it would be an easy matter to support non-discriminatory transmission open access that relies on a system of physical transmission rights. If this assumption were true, then it would be possible, even natural, to restructure the electricity system around physical rights and long-term bilateral contracts that would reduce the need for RTO market coordination and avoid the need to exercise great care in designing the details of the (small) balancing system and associated spot market.

This is an old, discredited idea. For example, in the early days restructuring the electricity market in the United Kingdom, more than twenty years ago, there was an extensive effort to develop a contract-scheduling scheme that was famously abandoned in favor of the spot market coordinated by the independent system operator.

The idea has a superficial appeal, and like a bad penny it keeps cropping up in various guises. Most prominently, in the mid-1990s, Enron championed essentially the same contract-scheduling model for electricity markets built around physical transmission rights, forced bilateral transactions, and suppressed spot markets. (We summarize California's experience with the Enron model in Appendix C.) But the model violates basic principles of economics and physics, and was eventually abandoned. The history of failed attempts makes clear the danger of repeating experiments that have been tried and rejected, often at substantial cost.

A. Reduced Grid Access and Trading

Just as Enron argued in the mid-1990s, more recent proponents of a contract-scheduling framework believe it is wrong to give parties the choice of relying on RTO-coordinated spot markets even when it is economic for them to do so. Therefore, market rules should somehow limit access to spot markets to force parties into bilateral contracts or self-supply arrangements.[44] For example, APPA argued in Consumers that "[p]rices for power sold under bilateral contracts (individual contracts between a buyer and a seller) have been substantially influenced by the high prices sellers can obtain in the RTOs' centralized markets."[45] In *Competitive Market Plan*, APPA's intent remains to reduce the use and influence of these spot markets.

The contract-scheduling framework would force parties to arrange and submit to the RTO balanced (supply matching demand) bilateral (or self-supply) schedules, months, weeks, days and hours ahead.[46] This would be counterproductive. Restricting the scheduling parties' access to the spot market would undercut the stated goal of promoting bilateral trading.

A contract-scheduling framework suffers from a fundamental flaw: it fails to recognize that scheduling bilateral contracts on a finite transmission grid must be facilitated by the dispatch coordination provided by the central system operator (the RTO). The central operator uses the spot market offers and bids to arrange its generation and load dispatch to accommodate the many parties' schedules so that

---

44 Although APPA has not explained how it would restrict access to the spot market, such restrictions would be necessary to implement the APPA model. As described in Appendix A, in the failed early model in California, with its separate Power Exchange and Independent System Operator, the goal was to limit the role of the spot market operating through the balancing function. To achieve this goal, the designers found it necessary to create explicit restrictions on the spot market to prevent economic dispatch with efficient balancing and to require balanced bilateral schedules. These restrictions contributed to the California crisis in 2000-2001, which led to the subsequent abandonment of the Enron-type model and demise of the separate Power Exchange.

45 *Consumers in Peril*, at vi.

46 *Kelly/Caplan* at 539: "LSEs could be required to submit anticipated loads at specified intervals (*e.g.*, month ahead, week ahead, day ahead, hour ahead), and the schedule of generation resources they have the right to call upon to serve those loads (including both generation and demand response resources)."

the total flows across each transmission element do not exceed safe operating limits. To ensure reliability, the operator must also use the offers and bids to balance the dispatch so that the total injections (supplies) match total withdrawals (demand plus losses) every moment, while respecting the limits imposed by the transmission system.

Without the essential coordination performed through the spot markets and associated dispatch, a bilateral contract and scheduling regime cannot function without limiting access to the grid and discriminating among the parties. Limiting access to the spot markets and associated dispatch would force restricted access to the grid, which would in turn limit and complicate economic trading, thus raising the costs of serving load.

The transmission grid must be able to accommodate the simultaneous flows associated with parties' schedules and ensure delivery, but effectively meeting this condition requires the RTO's central dispatch coordination. Together, the pattern of schedules and the operator's dispatch determine flows across limited transmission elements. In short, dispatch and feasible scheduling cannot be separated.

B. Unworkable Physical Transmission Rights

Under the contract-scheduling framework, the grid could not be scheduled in advance with assured delivery without first rationing grid access. Without rationing, parties could submit advance schedules, but some schedules would be infeasible (exceed grid limits) and have to be rejected in advance or curtailed in real time. To solve the problem it created, this framework would force the RTO to ration use of the grid, forcing parties to purchase a limited number of physical transmission reservations in advance and submit to possible physical curtailments after the fact.

There has been a great deal of experimentation and analysis devoted to this problem in the past.[47] As the experience in physical rights systems has shown, prior rationing and later curtailments are needed to "unschedule" the grid when, as often happens, scheduled flows exceed limits and cause congestion. Historically, physical rights rationing rules permit fewer accepted schedules and hence fewer trades, leaving the grid underutilized. These results would be the opposite of the stated goals of promoting forward bilateral contracts and scheduling. [48]

The contract-scheduling approach thus has the solution exactly backwards. Restricting access to the dispatch/spot market undermines contracting, while open access to an RTO's spot market facilitates contracting. Open access is thus not a design flaw; it is an essential feature benefiting all parties.

The spot market defines and prices the RTO's dispatch, and the dispatch provides the coordination needed to support bilateral trading, providing balancing for schedules and dispatch adjustments to avoid transmission congestion that would otherwise force schedule curtailments. Open access to the dispatch/spot market allows robust forward markets, which then help support investments in generation, transmission and demand response.

Furthermore, the existence of the RTO market provides the framework to create workable Financial Transmission Rights that substitute for the unworkable physical transmission rights. Importantly, *Kelly/Caplan* asserts that "[t]he Day 1.5 market design proposal presented in this article could potentially work with either a physical rights or financial rights transmission service regime."[49] This is a critical claim that is never explained; it glosses over the intrinsic requirement that financial transmission rights be integrated with an open spot market with locational prices.

As the RTO adjusts dispatch to manage congestion, it produces differences in locational prices.

---

47 Hogan, "Reforms of Reforms."
48 *Kelly/Caplan*, at 491.
49 *Kelly/Caplan* at 534, footnote 204.

Schedule imbalances are settled at these prices. Furthermore, parties with transmission schedules are charged the difference in the locational prices between each schedule's source and destination. Hence, transmission usage is equivalent to the identical physical transaction of selling at the source and buying at the destination. This fundamental equivalence means that the spot market outcome, transmission schedules, and economic dispatch are internally consistent. Therefore, there is no incentive for market participants to distort decisions because of different treatment of physically identical transactions. This critical consistency would not be available under any other system, and breaking the connections creates incentives to game the schedules or dispatch.

By design, FTRs are entitlements to payments (or debits) from the spot market, based on the differences in locational spot prices between each FTR's points of injection and withdrawal. Thus, FTRs are defined by and consistent with the spot market prices. Without the open spot market with locational prices, the financial transmission rights would no longer be consistent with actual opportunity costs of using the grid. This disconnect would create incentives for parties to game their schedules and the dispatch, and to submit schedules that make congestion worse.[50]

On this point, there is a real danger in offering the contract-scheduling framework while overlooking a fundamental contradiction. It is not possible to have it both ways, with restricted spot markets and consistent financial transmission rights. It is the open spot market with efficient pricing that creates the possibility to offer a consistent system of financial transmission rights in lieu of the unworkable physical rights.

Limiting access to the spot market would limit the ability to utilize FTRs and re-create the previously unsolved problem of defining a workable system of physical transmission rights. The restrictions are inherent in the earlier Enron contract-scheduling framework, which APPA and other RTO critics seek to recycle.[51]

In *Consumers in Peril*, APPA proposed to limit parties' ability to rely on the RTO spot markets to "no more than 5 percent of load." [52] With only a limited balancing market, LSEs would be forced to meet their loads using self-supplied generation and energy purchased through bilateral contracts; they would need additional arrangements to cover large imbalances if they needed to replace or supplement their contracted supplies. However, APPA's arbitrary 5 percent limit on RTO balancing would result in LSEs facing greater risks and higher costs in reliably serving their loads. And as experience shows, specifying the details of how to achieve this limited access would expose the problems that have arisen when such limitations have been tried.

Under the "Day 1.5" proposals, virtually all trading would still be restricted to bilateral contracts between suppliers and LSEs, just as proposed in *Consumers in Peril*. To be clear, bilateral trading *per se* is not a problem; it is used extensively in RTOs today by parties choosing that option; RTO dispatch fully accommodates the resulting schedules. Many of the details of RTO market design were created to facilitate the ability to use bilateral contracts in the face of the complex multilateral interactions across the transmission grid. But facilitating bilateral contracts is a quite different thing from prohibiting alternatives and compelling reliance on an unworkable system. The difference is that parties in today's RTOs also have another choice: trading through the auction-based RTO dispatch/spot markets when it is more economic to do so, an option APPA still seeks to limit or foreclose.

It makes sense that APPA has moved beyond the contract-scheduling framework. That framework falls short of ensuring open, non-discriminatory access to transmission and supporting a robust forward contracting market. As critics continue to reinvent electricity market design with all the mechanisms that must be in place both to ensure reliable operations and to allow their members and other parties

50 There is ample experience from California and elsewhere with parties gaming bilateral schedules and dispatch offers to exploit inconsistencies between the spot market and the opportunity costs of using the grid. See Appendix A.

51 *Consumers in Peril*, 27-28.

52 *Consumers in Peril*, 27.

open, non-discriminatory access to the transmission grid,[53] they will rediscover virtually every feature of the current RTO Day 2 Markets.

Any model that prohibits access to the spot market or depends on undermining economic dispatch should bear a strong burden of proof, and its proponents should not be allowed to ignore this central dilemma. Without the restrictions on the spot market, the distinction between physical and financial transmission rights disappears, and the Enron-like model for transmission scheduling and dispatch would quickly evolve into the existing RTO design. [54] The costly restrictions are an essential part of the RTO critics' preferred approach.

53 Federal statutes and FERC rules require that transmission owners/operators provide parties access to the transmission grid on terms that are "not unduly discriminatory." APPA claims to support these goals, though its proposals would make their full achievement unlikely.

54 PJM and most other RTOs operate two short-run auction-based exchanges: a day-ahead financial market, in which parties can buy and sell energy for the next day and lock in day-ahead transmission charges, and a real-time market that determines the RTO's actual real-time physical dispatch. Parties may choose to participate in the voluntary day-ahead market, but every party using the grid must settle its schedule imbalances and additional redispatch costs in the real-time market.

V. The APPA Reforms Could Frustrate Federal and State Policies

There is an inherent tension between the APPA market framework and the electricity policies being pursued in many RTO states. At its core, the APPA approach focuses on limiting or discouraging access to RTO spot markets, while forcing utilities/LSEs to rely almost exclusively on bilateral contracts or their own generation. However, many state electricity policies depend on open access to the RTO spot markets.

In recent years, many states have supported wholesale and retail competition, alternative/renewable energy development and enhanced demand-side response. Various RTO rules, including open access to spot markets, are expressly designed to accommodate these policies while also accommodating states that choose not to pursue one or more of them.

For example, PJM facilitates more efficient demand response by basing its spot prices on LMP and making them transparent (through settlements and publication on PJM's web site). Since the LMPs represent the marginal cost of serving load at each time and location, they provide the correct price signals for the value of consumer demand reductions at each time and place, indicating when and where demand-side efforts make economic sense. Utilities/LSEs facing these LMPs then have efficient incentives to implement demand-side efforts. In addition, several states now use the LMPs as the default price for at least those larger customers eligible for retail choice, encouraging the customers to implement their own demand reductions at times when prices are highest.

In the PJM region, several states have adopted policies to promote both wholesale and retail competition. Some states supported those policies by requiring or encouraging their regulated utilities to divest generation; others did not. The PJM region thus includes many utilities and non-utility LSEs that do not own generation to serve their loads. They purchase power from the wholesale markets, either through bilateral contracts or through purchases from the PJM spot energy and capacity markets. The ability to use the PJM markets provides an important option for these LSEs, allowing them to cover any part of their load that is not covered by bilateral contracts. The LSEs can choose that option in advance, or they can attempt to arrange and schedule sufficient contracts to meet their loads, while using the PJM spot markets to cover day-ahead or real-time imbalances associated with those schedules. The choice is left to the LSEs, but APPA would take away that choice.

PJM also accommodates dozens of non-utility third-party generators, including independent developers of renewable and alternative resources. These third-party generators can sell energy through PJM's spot markets and capacity in PJM's RPM capacity market, whether or not they have contracts with utilities or LSEs. If such a generator sells energy through a bilateral contract, it can use the PJM spot markets to cover any imbalances. Intermittent generation resources, like wind turbines, can also sell their power directly to the spot market and receive the spot price for the power they inject. This open access to the PJM spot market reduces entry barriers, allowing these generators to secure financing and compete more effectively with utility-owned generation.

The ability of LSEs and third-party generators to access PJM's energy and capacity markets is essential to their viability and hence the success of state policies favoring their development. Indeed, it is difficult to imagine how third-party generators, divested utilities and independent LSEs could function successfully without unrestricted access to the PJM markets. For example, how could the capacity structure work with load switching (in retail choice states) without detailed rules requiring capacity to follow loads? And how could parties contract long term with load switching?

Utilities in states that do not have these same policies also benefit from access to RTO markets, but they could conceivably tolerate more restrictive access to those markets, as shown by the approach taken in the Southwest Power Pool (SPP). In SPP's market region, states have generally retained vertically integrated utilities, avoided divesture and not implemented retail choice; there are no non-utility LSEs competing for retail load and fewer independently owned generators. The utility members of SPP own most of the generation they need to meet their own loads and purchase the remainder through bilateral contracts. Development of renewable resources or other alternative generation technologies occurs primarily, if not exclusively, under contracts with vertically integrated utilities.

Given this structure, SPP employs some features common to other RTOs but retains other features left over from its pre-RTO days. For example, SPP requires each utility member either to arrange schedules in advance to serve its own loads or to offer sufficient generation to SPP's central dispatch to meet those loads. The utilities must also reserve physical transmission rights to support moving this generation to load, an approach that tends to reduce grid access and leave the grid underutilized. Such "balanced" schedules and physical rights reservations are not required in other RTOs. On the other hand, once these schedules are submitted, SPP arranges a bid-based security-constrained economic dispatch to balance the system and adjust the dispatch to avoid congestion, just as PJM does. Moreover, SPP then prices the dispatch using "locational imbalance pricing" (LIP) which is essentially the same as LMP.

SPP is likely to evolve by developing the remaining functions now provided by PJM and other RTOs. For now, however, SPP can function with this still-developing hybrid because it does not have to provide non-discriminatory access to non-utility LSEs nor support significant third-party generation that is not contracted to utilities.

VI. The APPA Proposals Would Overburden Market Monitors But Not Reduce Prices

APPA seeks something closer to cost-of-service regulation for all generators, including those not owned by utilities. Generators would be required to offer their power to the RTO for dispatch, but every dispatch offer would be fixed by the RTO market monitors at each generator's individual short-run marginal cost (SRMC), but ignoring opportunity costs (recall the discussion in Chapter I of how this would suppress spot prices).

The task APPA has in mind for the market monitors is more detailed and intensive than anything currently performed by state utility regulators and public utility boards. PJM's region has literally hundreds of generators, each with a different cost structure. For each supplier, fuel prices can vary daily, and countless other factors shape each generator's marginal costs. APPA does not explain how the RTO market monitors could perform this level of cost-of-service regulation at the required level of detail.

In addition to assuming this burden, market monitors would be required to examine all contracts between generators and LSEs.[55] This task would not be easy, since APPA states that each contract would be tailored to match not only the needs of each LSE but the operating costs and investment requirements of each generator.[56] Since generators could not expect to recover their fixed investment costs from the suppressed spot market prices, and APPA would eliminate capacity markets and payments, generators would be need to recover those investment costs in contracts overseen and approved by the market monitors.

The proposed monitoring framework is a substantial departure from how RTOs operate today. In today's RTOs, price incentives encourage generators to bid their marginal costs, thus minimizing the need to monitor every generator's bids. Market monitors can then selectively focus on specific generators bidding into the RTO spot markets.

Today's market monitors evaluate market concentration in specific areas (load pockets) in which transmission limits create conditions vulnerable to market power. The monitors apply market power mitigation measures to specific plants to prevent economic or physical withholding. In such situations, the market rules require that spot prices be based on offers/ bids mitigated in advance by the market monitors to reflect marginal costs or prices acceptable to the market monitor. The results, confirmed by RTO *State of the Market Report*s, is that generation offers generally reflect marginal operating costs, while overall spot prices do not exceed competitive levels.[57] There is no need for this extensive, and probably unworkable, system of cost-of-service regulation APPA describes. The proposed reforms would work against the principles of efficient operation.

Spot prices then have an effect on forward contract prices. Since both buyers and sellers can always buy/sell power from the spot market at competitive prices, forward contract prices tend to reflect expected spot prices. Accounting for differences in risks, contract prices are not likely to be significantly higher (or lower) than expected average spot prices over the same period. Market monitoring and mitigation in the spot markets thus helps to ensure contract prices are also competitive. Under the APPA framework, spot prices would not include opportunity costs and contracts would not reflect expected spot prices. Hence every contract would differ. The APPA framework would overwhelm the workable structure market monitors have developed.

It would also likely create supply shortages. The RTO's focused mitigation efforts in spot markets help

55 *Competitive Market Plan* at 19-22.
56 *Competitive Market Plan* at 22.
57 PJM's *2007 State of the Market Report*, available at www.pjm.com

keep contract prices competitive, but by suppressing spot market prices, APPA's proposals would tend to suppress contract prices as well. Suppressed spot prices and suppressed contract prices translate to shortages, unless there are capacity payments to provide the "missing money" needed for investment. But APPA would eliminate capacity market payments.

Importantly, *Kelly/Caplan* recognizes that if spot prices were based on marginal operating costs, as PJM's Market Monitor finds they generally are,[58] then PJM's spot prices would be attractive to buyers. But that happy result is seen as a problem that requires rules to force buyers not to use the spot markets:

> Without market features requiring purchasing LSEs to maintain a portfolio of longer-term generation resources to serve their loads, the temptation for them to simply rely on the short-run marginal cost-based short-term optimization markets for a substantial portion of their power supplies could be quite high. In such an environment, the generators' claims that they were suffering "missing money" (*i.e.*, that they were not recovering their fixed costs) might well be justified. To prevent this result, we propose to impose on LSEs a "resource adequacy" requirement to obtain sufficient longer-term generation resources to serve their anticipated loads, thus preventing them from "leaning" on a short-term optimization market intended primarily as a balancing market.[59]

This is an important insight. The reason the "missing money" arises is, as *Kelly/Caplan* suggests, because prices in PJM's current energy and reserve markets do not cover the full fixed costs of building and maintaining the capacity needed to meet reserve requirements. By definition, that means that spot market prices are not "too high," because they don't fully cover the cost of sustaining existing or developing new generation. If capacity payments are not to be used to cover this missing money, then logically spot market prices must be higher if total market revenues are to cover the cost of developing generation.

Having encountered the underlying dilemma, it is important to address its remedy. Today's spot prices are insufficient when the region experiences shortages. Under current rules, spot prices are not allowed to reflect scarcity when the system is short of operating reserves. That means the underlying "missing money" problem could be mitigated by reforming energy and operating reserve pricing to include shortage-cost pricing, thus reducing the "missing money" problem that creates the need for capacity payments.

Without that preferred outcome, the RTO needs a capacity payment regime, such as RPM, to collect the "missing money" and pay it to generators. The *Kelly/Caplan* statement implicitly confirms that some type of capacity payment system is justified. But instead of paying the "missing money" through PJM's RPM, critics would require LSEs in the region to pay the missing money through higher prices in long-run forward contracts. That approach would then need some type of enforcement mechanism to force LSEs to contract for the right amount of capacity.

We have been here before. The original eastern power pools required each utility to build or contract with sufficient capacity to meet its share of reserve requirements, just as APPA proposes. When utilities fell short of their requirements, they could not be allowed to lean on others, so they were penalized. But to be effective and provide the correct incentives to acquire the right amount of capacity, the penalties must be related to the market value of capacity, and capacity auctions administered by the RTO (such as RPM) help define that value.

APPA apparently believes its forced contracting regime would provide contract prices sufficient to support the level of investment needed to meet the regional reserve margin targets. But having rejected capacity payments, its proposal includes no mechanism by which the necessary revenues would occur through contracts. Unless something intervened to force higher contract prices, suppressed spot and contract prices would produce market revenues insufficient to cover investors' fixed costs. That

58 Ibid.
59 *Kelly/Caplan* at 539 (emphasis added, footnote omitted).

means the region would fail to meet its resource adequacy (reserve margin) targets.

But suppose there were some magic way to compel LSEs to sign contracts with enough generators at prices high enough to support the required reserve margins. By definition, that means the contract prices would provide the “missing money” so that the total market revenues received by generators would be enough to cover the investment costs of the target level of reserves. But of course, that is essentially what the capacity markets do today; they provide capacity payments to generators to replace the missing money. So replacing capacity payments with sufficiently higher contract prices would not reduce total electricity prices.

Regardless of any concerns APPA may have about the details of each RTO’s capacity market structure, APPA’s basic assumption is simply wrong. It cannot reduce total electricity prices by dismantling the RTOs’ capacity payment systems (see Appendix F) while simultaneously suppressing spot and contract prices. The revenues to meet the investment requirements must come from somewhere.

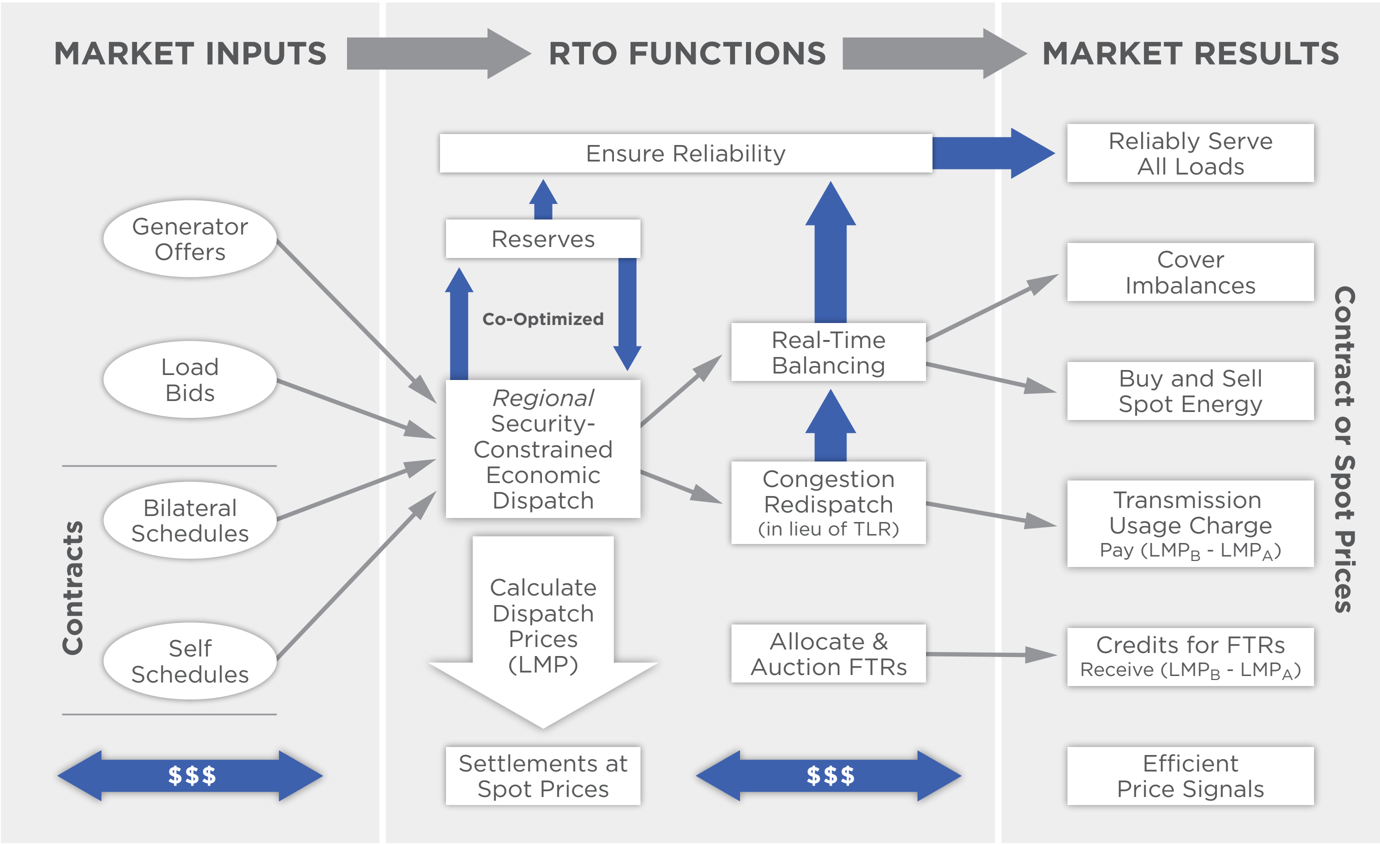
To meet the region’s reserve margin targets, the total revenues needed to support that level of supply investment must be provided, one way or another. Much of it can come from spot prices and contract prices linked to spot prices, if those prices are not artificially suppressed. Moreover, if spot markets were improved to include scarcity pricing when the RTO is short of operating reserves, then spot markets and contract prices would recover even more of the needed revenue requirement, while the need for capacity payments would diminish. But whatever is left must be recovered through some mechanism, presumably capacity payments. There is no magic formula that allows APPA to escape this equation.

VII. The RTO Approach Is A Proven, Superior Model

PJM and other RTOs operate on a reliability foundation common throughout the industry. To maintain reliability at each moment, every system operator must keep generation (supply) exactly in balance with consumption (demand) plus transmission losses. To avoid system failures and outages, it must also ensure the energy flows across any transmission line or element do not exceed the safe operating limits for that component.

Every modern system operator accomplishes these tasks through a process called "security-constrained economic dispatch," by which it determines which generators will operate and how much energy (or voltage support) each unit will produce at each location on the grid. The dispatch is "economic" in the sense that the system operator chooses the lowest-cost (or lowest bid cost) mix of plants available to dispatch to balance the system and manage flows to avoid congestion. All RTOs follow this standard industry approach. The modern RTO framework is illustrated in the following figure.

Figure 1: RTO Maintains Reliability Using Security-Constrained Economic Dispatch. The Spot Market Defines/Prices Dispatch



Voluntary offers and bids to determine the RTO's dispatch, and settlements for energy bought and sold through the dispatch, create a "spot market," which is the financial side of the essential physical dispatch. After each dispatch interval, the RTO determines the locational spot prices, based on the actual dispatch.

While many generators make their plants available for economic dispatch (it makes economic sense for them to do so), most energy is eventually priced to retail consumers through contracts or cost-of-service approaches. As Figure 1 shows, utilities may self-schedule their own plants to serve their own loads, or schedule bilateral transactions to move power from sellers/suppliers to buyers/loads. The RTO accepts and accommodates all of these schedules by arranging its security-constrained economic dispatch to accommodate any fixed schedules submitted by the parties.

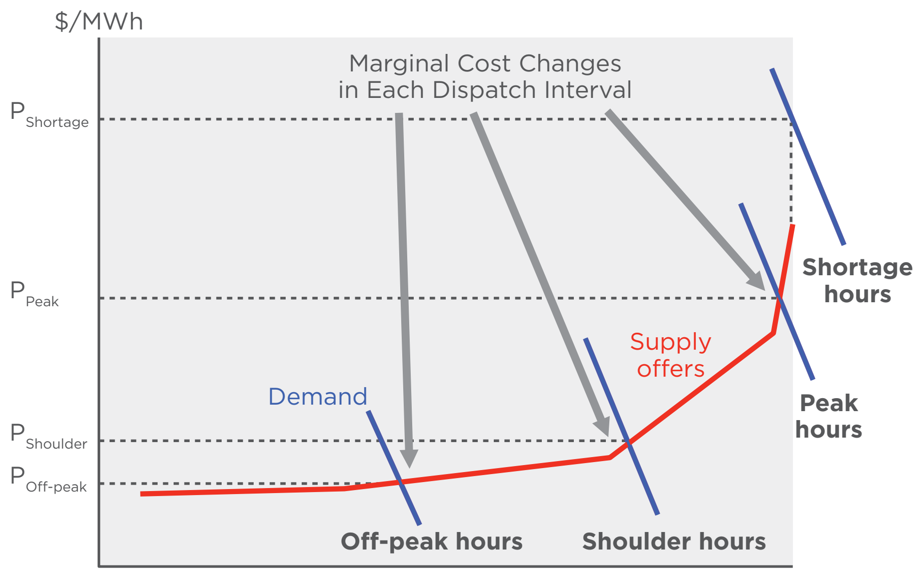
If these fixed schedules would overload any transmission line or element, the RTO routinely adjusts the dispatch (a step sometimes called "redispatch") to redirect flows and relieve the congestion. The RTO then charges each affected schedule the marginal costs of any redispatch needed to accommodate that schedule, so that no party is "leaning on" or subsidizing any other party.

If parties deviate from their set schedules, the RTO covers their imbalances by simple adjustments to the dispatch, ensuring that all load is reliably served at all times.[60] The RTO then settles the deviations at the LMP spot price, which reflects the dispatch marginal cost for covering each imbalance at the time and location where it occurred.

Thus, the principle used by all RTOs is that every grid user gets access, each user pays for whatever costs it imposes on the dispatch and is compensated for whatever benefit it provides. There are no cross subsidies, and no party "leans" on any other party. Since every party has access to the grid via open access to the dispatch (and redispatch if needed to relieve congestion), these principles mean that RTOs fully satisfy the federal mandate to provide open access transmission service to all parties without discrimination.

The LMP spot market prices reflect the marginal cost of operating the dispatch and thus the value of power at each time and location. They are used to pay parties that sell power through the dispatch and to charge parties for the power they purchase through the dispatch in each hour. The LMP spot prices routinely change each hour (or shorter dispatch interval) as a different mix of plants with different marginal costs is dispatched in each interval to match ever-changing demand.[61] See Figure 2. If there is congestion, LMP will also vary at different locations, reflecting the change in marginal costs as different power plants are instructed to raise or lower their outputs at each location so as to redirect flows and avoid congestion.[62]

Figure 2: Marginal Costs Define RTO Spot Market Clearing Prices



Clearing Price rules encourage generators to bid marginal (or opportunity) cost and to follow dispatch instructions

The principal gap in implementing the LMP design is during shortage hours. Although the shortage price defined by

60 For example, suppose a party scheduled a bilateral transaction in which it planned to inject 100 MW at location A and withdraw 100 MW at location B. In real time, however, the party only injected 95 MW, instead of the scheduled 100. In that case, the RTO would simply dispatch 5 more megawatts from other generators to cover the imbalance, thus reliably serving all load. The dispatch is thus a "balancing market." The party with the imbalance would be charged for the 5 MW supplied by the dispatch times the LMP at location A. Any imbalance, on the supply or demand side, would be automatically covered by the dispatch and the parties charged (or paid) accordingly.

61 RTOs readjust the dispatch every five minutes to follow rapidly changing demand. Thus, marginal costs of serving loads, and hence spot prices, are also changing every five minutes. In some RTOs (New York), the five-minute spot prices are used directly for settlements; in PJM, five-minute prices are aggregated into hourly prices for settlements.

62 Marginal costs will also vary due to losses, but we ignore this for the purposes of this discussion.

the demand curve is evident in Figure 2, it has been harder to include this logic in RTO market design and software. The issue of scarcity pricing is one area where RTO reform is under consideration.[63] However, the effect will be to make spot prices higher, not lower, and this well-founded reform is not addressed in the APPA critique.

Transmission scheduling problems are at the center of the challenges in providing open access and non-discrimination. But critics are often silent on how they would address these well-known problems; they assume that the contract-scheduling model with physical rights could be made to work and meet non-discrimination requirements, despite the experience.

The modern RTO design solves this scheduling/delivery problem. As Figure 1 illustrates, the RTO model gives parties expanded access to the grid by allowing them access, though the spot markets, to the system operators' security-constrained, economic dispatch (including "redispatch" service to avoid congestion). The dispatch accommodates the parties' schedules without physical rationing, requirements for obtaining physical rights in advance, or risking curtailments later. By going down the wrong path, again, restricted access to spot markets and the resulting need for contract scheduling with physical rights would discard these hard-won advantages.

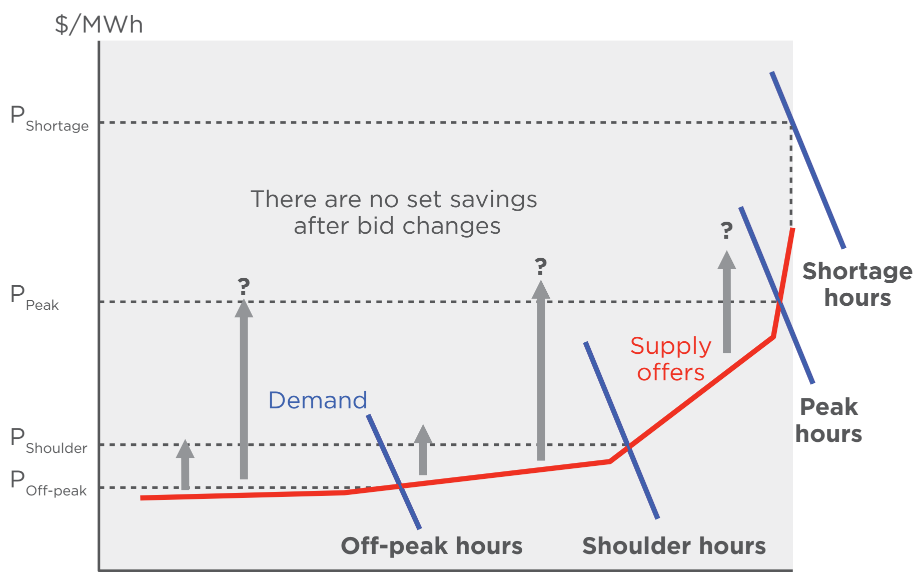
In *Consumers in Peril*, APPA proposed to replace the "clearing price" approach for the spot market, but it did not explain how any other method would work. In *Competitive Market Plan*, APPA would retain clearing prices. RTO dispatch/balancing markets use LMP to price the energy bought and sold through the dispatch. LMP is a clearing price approach, in which the price at each location represents the marginal costs of serving load during each dispatch interval. This means that lower-bid-cost units in the dispatch are paid a clearing price set by the marginal units in the dispatch.

This is a straightforward application of economic marginal cost pricing. However, critics argue this approach overpays infra-marginal units. This assumes that such units would willingly and consistently sell their power at the lower prices they bid in a clearing price regime. There is no economic or logical basis for this assertion.

As various studies have explained,[64] if generators are faced with a rule that says they will be paid the prices they bid, they will change their bid behavior, raising their bids to their expectation of what the market-clearing price would be, as illustrated in Figure 3.

In the end, the "pay-as-bid" approach would be likely to raise costs, because bidders may overestimate the expected clearing prices, causing less efficient generators to be dispatched in their stead. Moreover, if suppliers could somehow be forced to sell their power for less than the market-clearing price—that is, for less

Figure 3: Pay-as-Bid Rules Result in Generators Changing Their Bids



Pay-as-bid rules encourage bidders to raise their bids to expected market-clearing levels.

63 FERC, Order 719, October 17, 2008; PJM *2008 State of the Market Report*, Recommendations, pp.6-7.
64 *E.g.*, see Alfred E. Kahn, Peter C. Cramton, Robert H. Porter, Richard D. Tabors, "Pricing in the California Power Exchange Electricity Market: Should California Switch from Uniform Pricing to Pay-as-Bid Pricing?" January 23, 2001.

than what it is worth at that moment and location—the long-run effect would be to discourage and limit investments needed to sustain the suppliers over time or replace them when they retire.[65]

The RTOs' operation of these markets is the essential ingredient that allows for competition and provides open access to the transmission grid without discrimination. Through these markets, the RTO provides coordination for competition, allowing participants to arrange bilateral contractual commitments without producing energy schedules inconsistent with safe and reliable operation of the grid. The dispatch/spot market captures the interactions among many market participants and then prices the needed dispatch services accordingly, using locational prices.

The RTO approach is not merely a workable, fair and efficient solution; it is a vast improvement over the restrictive physical rights regime that the RTO critics would force on RTO regions. The standard RTO design the critics would dismantle was recently characterized by the International Energy Agency in its review of market experiences across its member countries:

> "[L]ocational marginal pricing (LMP) is the electricity spot pricing model that serves as the benchmark for market design — the textbook ideal that should be the target for policy makers. A trading arrangement based on LMP takes all relevant generation and transmission costs appropriately into account and hence supports optimal investments."[66]

Any retreat to a failed contract-scheduling model would only raise transaction costs, restrict grid access and reduce interregional trading, while leaving the regional grids underutilized compared to today.

Electricity markets organized under RTOs are still evolving. The existing models are not perfect. But the remaining problems are in the area of scarcity pricing, transmission expansion, demand-side participation, and so on. The needed refinements call for better, not restricted, spot markets.

In short, an RTO model that uses bid-based, security-constrained economic dispatch with locational prices provides the foundation for open access and non-discrimination. It provides the framework for FTRs that solve the otherwise unsolved problem of how to make physical transmission rights compatible with open access, non-discrimination and efficient use of the grid.

Furthermore, this is the only model that meets these objectives. This is fundamental. Any suggestion to deviate from this model should bear a strong burden of proof. The critics' analysis does not fully acknowledge this critical problem, nor does it provide an analysis that meets the burden of proof in the face of overwhelming evidence against restricting access to the spot market and leaning on the contract-scheduling model.

---

65 The reason is that the difference between the clearing price and a marginal cost bid constitutes a contribution to the generator's fixed/capital costs. A properly structured clearing cost mechanism will thus cover both marginal operating costs and fixed costs. If generators are forced to forego this contribution, they will not recover their fixed costs, so future investments will be lower.

66 International Energy Agency, Tackling Investment Challenges in Power Generation in IEA Countries: Energy Market Experience, Paris, 2007, p. 116. Also see, Paul Joskow, "Challenges for Creating a Comprehensive National Electricity Policy." Speech given to the National Press Club, September 26, 2008, (http://www.hks.harvard.edu/hepg/Papers/Joskow_Natl_Energy_Policy.pdf). Joskow is a Professor at MIT, President of the Sloan Foundation and a member of the Exelon Board of Directors. He notes that the RTO model works well and is evolving in helpful directions; he recommends it be expanded and made mandatory across the country.

VIII. Conclusion

The recommendations of APPA and other critics of RTOs have undergone substantial evolution since *Consumers in Peril*. In one sense, this evolution might be viewed as progress as APPA confronts the realities of electricity systems that others have learned and embodied in the current RTO market design. However, a continuing missing chapter in the APPA analysis is any forthright description of the special characteristics of electricity systems that underpin the current RTO market structure. The several elements of bid-based auctions, economic dispatch, security constraints, locational prices, unit commitment, long-term contracts and capacity markets all work together to solve the complicated coordination problems that come hand-in-hand with an integrated transmission grid. The RTO market design elements are there for a good reason, and the lessons about missing pieces were learned at great cost.

Despite the repeated experience of failure with attempts to constrain spot markets, such proposals and the associated return to contract scheduling with physical rights would recycle the mistakes and perform radical surgery on the healthy vital organs of the working RTO markets. These recycled restructuring proposals misunderstand both the basic requirements of reliable grid operations and the prerequisites of efficient trading under a statutory requirement for open access and non-discrimination. Equally important, implementing restricted spot market access and the contract-scheduling framework would cost consumers billions of dollars. The APPA continues to sidestep the issues or give new labels to old ideas that obscure the message and ignore the lessons of the past. The APPA analysis is internally inconsistent, and its proposals disconnected from the real requirements of operating electricity systems. The APPA proposals point down the wrong path, again.

Appendix A: What Have We Learned About RTO Costs and Benefits?

The APPA-led analysis includes a number of arguments that expose critical misconceptions about how RTOs work and are working. Typically the misconceptions are implicit assumptions rather than explicit claims, but the implications of the errors are significant. The performance of RTOs is better than stated, quantitative and qualitative evidence of benefits are ignored, and critical market design connections are neglected.

Misapplying Cost and Benefit Studies

Evaluating the costs and benefits of RTOs is a challenging task. There are too many moving parts to allow for a simple comparison that might arise with a controlled laboratory experiment.[67] Some of the moving parts include the timing of various reforms within RTOs, the partly separable effects of state retail open access rulings, divestiture decisions for generation assets, transition plans for state rate restructurings, and so on. Most importantly, evaluating the costs and benefits of a component of RTO activities requires care in defining the question and developing the appropriate counterfactual for comparison.

The supporting analyses assembled by APPA do not address these details. Furthermore, in summarizing the attempts to address costs and benefits, APPA does not even consider all the costs or all the benefits, focusing on consumer impacts and finding only an absence of evidence of significant net benefits rather than contrary evidence:

> "Much time, energy and expense has been expended by all sides producing 'dueling studies' regarding the costs and benefits of RTO-run centralized markets. In our view, informed by both the literature and the actual experience of the APPA members in RTO regions, it is difficult to conclude that consumers have benefited from the implementation of these markets." [68]

The critics present a view that RTOs are not performing as intended:

> "Restructured wholesale markets are producing both higher prices and higher profits than one would expect in a competitive market. Resulting retail prices exceed those prevailing in regions that have not restructured, but that instead retained traditional retail cost-of-service regulation and eschewed the formation of RTOs. Long-term adequacy of generation resources is also a substantial concern in RTO regions."[69]

At the same time, the argument is that there are benefits from some aspects of RTOs (that apparently exceed the associated costs). For example:

> "We hasten to add that RTOs provide real benefits to consumers. RTOs provide independent and non-discriminatory transmission service under open access transmission tariffs (OATTs), charging regional transmission rates instead of individual system-by-system pancaked transmission rates. They maintain reliable transmission service through their wide-area-[view] of moment-to-moment system operations. They lead regional collaborative transmission planning processes. Such RTO functions undoubtedly benefit consumers. Yet the FERC's policies have increasingly lost sight of these core transmission-oriented RTO functions, as implemen-

---

67 John Kwoka, "*Restructuring The U.S. Electric Power Sector: A Review Of Recent Studies*," APPA, 2006, (available at http://appanet.org/files/PDFs/RestructuringStudyKwoka1.pdf)

68 Susan Kelly and Elise Caplan, *Time for a Day 1.5 Market: A Proposal to Reform RTO-Centralized Wholesale Electricity Markets*, 29 *Energy Law Journal* 491, at 514, (2008).

69 *Kelly/Caplan* at 494.

> tation of centralized markets for energy, ancillary services, and generation capacity have taken center stage. It is the RTO-run centralized wholesale markets and their performance that are the primary focus of this article."[70]

Therefore, it is only part of the RTO design that is the subject of criticism for producing more costs than benefits. This formulation of the critique would be difficult to establish based on the available evidence, and the most relevant evidence does not support the conclusion that "centralized wholesale markets" create costs greater than their benefits. To the contrary, the weight of the evidence points to substantial net benefits from RTOs under a regime of open access and non-discrimination.[71]

A recent Government Accountability Office (GAO) report concluded that "FERC officials believe RTOs have resulted in benefits; however, FERC has not conducted an empirical analysis of RTO performance or developed a comprehensive set of publicly available, standardized measures to evaluate such performance."[72] The GAO did not identify what measures to employ, but surely the right questions would include at least the operational record, risk allocations, and investment decisions. The GAO subject was the combined effect of RTOs with restructured electricity markets. The GAO did not address the narrower APPA question about the independent effect of organized wholesale markets.

Apparently the critics agree with FERC that the evidence supports the view that RTOs provide real benefits. The FERC perspective includes the effects of organized wholesale markets in this benefit calculation, but APPA argues that the independent benefits of organized markets have not been worth the costs. Yet the available empirical evidence is inconsistent with this conclusion.

The critics' focus on retail price impacts of RTOs would require an experiment or methodology to isolate the effects of wholesale markets in RTOs from the many other activities that determine retail rates. Most of the comparisons cited in the APPA analyses suffer from the inability to isolate the independent effect of the RTOs from the separate impacts of state regulation, generation configuration and other confounding factors.

Notably, the best attempt to answer the question about the retail price impact of RTOs approaches the problem by limiting the analysis of retail rates to a comparison of municipal utilities (not subject to state regulation), for regions with similar fuel dependencies, and for regions included or excluded from an RTO. That study found a statistically significant residential rate savings of $430 million to $1.3 billion per year in PJM and the New York Independent System Operator (NYISO) from membership in an RTO.[73] The criticism of the result is principally about the small number of paired comparisons between similar RTO and non-RTO regions.[74] The small number of comparables is an inherent fact that limits the possible empirical analysis, but it does not change the conclusion.

RTO Market Coordination Reduces Curtailments of Contract Schedules

A focus on the criterion of retail rate impacts addresses part of the story, but it ignores other benefits that were intended to flow from the creation of RTOs. For example, part of the purpose of RTO design was to facilitate trading and reduce the need for administrative Transmission Loading Relief (TLR)

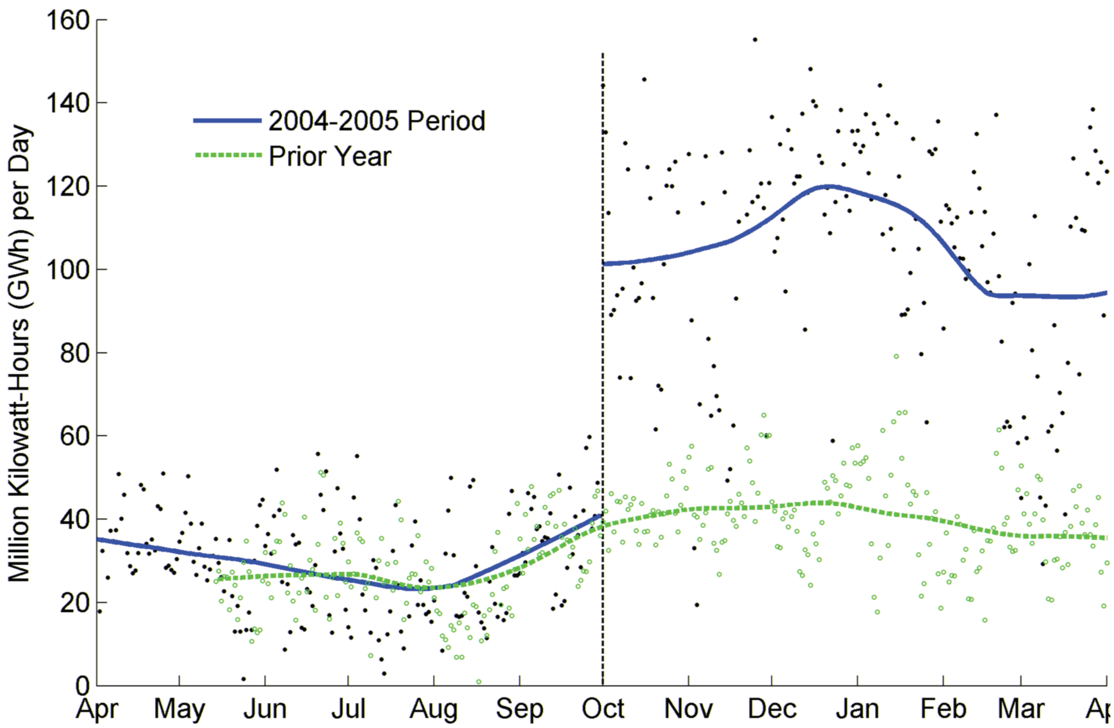
---

70 *Kelly/Caplan* at 494 (footnotes omitted).
71 Frank Huntowski, Neil Fisher, Aaron Patterson, "*Embrace Electric Competition or It's Déjà Vu All Over Again*," The Northbridge Group, October 2008, (www.nbgroup.com).
72 Government Accountability Office (GAO), "Electricity Restructuring — FERC Could Take Additional Steps To Analyze Regional Transmission Organizations' Benefits And Performance," GAO Report 08-987, at i.
73 Scott Harvey, Bruce McConihe, and Susan Pope, "Analysis of the Impact of Coordinated Electricity Markets on Consumer Electricity Charges," LECG, November 20, 2006 (revised June 18, 2007) and available at http://www.lecg.com/files/upload/AnalysisImpactCoordinatedElectricityMkts.pdf.
74 John Kwoka, "Restructuring The U.S. Electric Power Sector: A Review of the LECG Study," Northeastern University, 2007. (http://www.appanet.org/files/PDFs/KwokaLECGReview.pdf.)

orders that curtail scheduled transactions to relieve transmission congestion.[75] In the case of PJM, there was an interesting experiment in October 2004 when utilities in the Midwest, including American Electric Power (AEP), converted from relying exclusively on a *pro forma* open access regime based on contract scheduling with physical transmission rights to membership in PJM with its centralized wholesale spot market using LMP and financial transmission rights. This was a relatively clean experiment that allows a before-and-after comparison of regional trading without much need for other complicated control variables.

The result, as shown in the following graph, was dramatic and abrupt. Immediately following the expansion, the monthly average of day-ahead exports from the Midwest to PJM tripled and stayed at the new higher level.[76]

Figure A-1: Quantities Traded: Day-ahead net exports, Midwest ⟶ East



A similar result occurred after completion of the expansion of PJM's footprint in 2005. If all other things had been equal, PJM's expanded responsibility for managing additional congested transmission lines should have required more TLRs for PJM. But to the contrary, and complementing the trading statistics, PJM TLRs started to decline in 2005 and average annual PJM TLRs at level 3 or above for 2006-2008 dropped to 27% of those in 2004, despite a general overall increase in reported TLRs.[77]

The PJM experience is important because it was a clean experiment and the trading results were unambiguous. However, the PJM expansion data combine the effect of moving from a bilateral trading arrangement to an RTO and the effect of participating in a centralized wholesale market, showing an

75 Transmission Line-Loading Relief (TLR) involves a set of NERC rules under which a reliability coordinator requires certain parties to curtail their transaction schedules until the excess flows on congested transmission lines fall within safe operating limits. In RTOs, dispatchers adjust the dispatch (redispatch), which changes the flows across selected lines and thus reduces the need for TLR curtailments. In non-RTO regions without open access to redispatch service, TLR curtailments become necessary and more widespread.

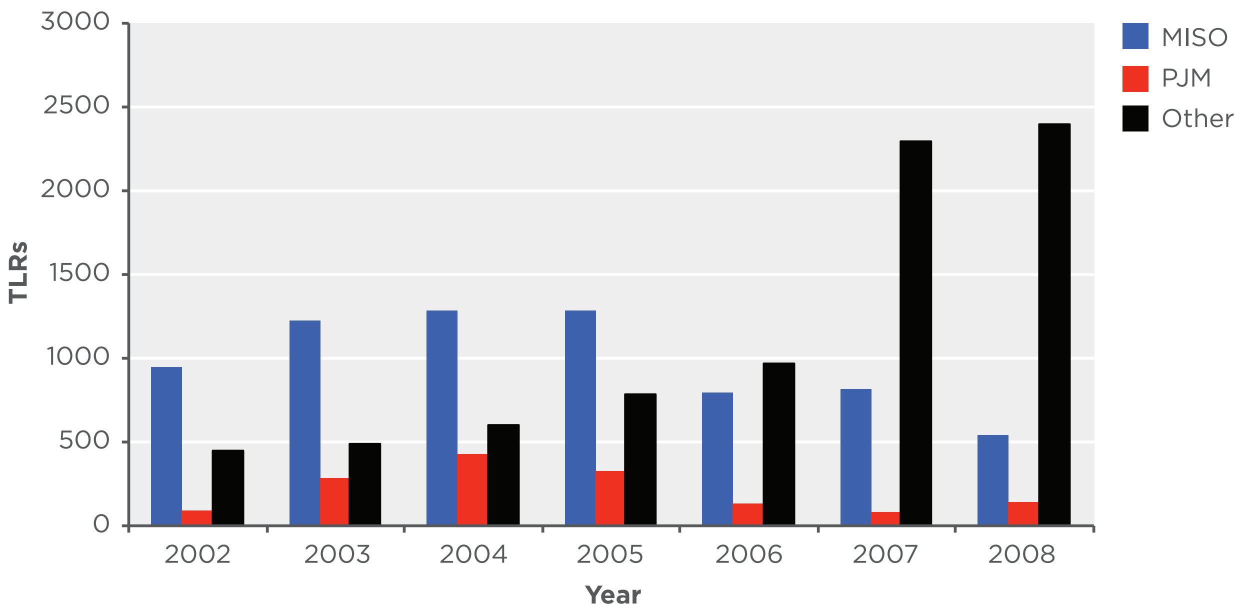
76 Erin T. Mansur and Matthew W. White, "Market Organization and Efficiency in Electricity Markets," October 2008, Figure 2, discussion draft (available at http://bpp.wharton.upenn.edu/mawhite/).

77 NERC, December 22, 2008, http://www.nerc.com/docs/oc/scs/logs/trends.htm.

increase in trading benefits. But the PJM expansion case does not isolate the effect of the organized market separate from the effect of an RTO as required under the approach advocated by APPA.

An examination of the TLR experience after the startup in the Midwest Independent System Operator (MISO) market in 2005 provides another experiment that comes closer to isolating the independent effect of the wholesale market. In February 2002, the MISO inaugurated operation as an RTO with an open access transmission tariff based on contract scheduling, but MISO did not launch the RTO-wide organized centralized wholesale market with LMP and FTRs until April 2005. The MISO showed an increasing frequency of TLRs while operating as an RTO without an organized market. But the increase reversed and was followed by a sharp decline in TLRs when the organized market was put in place. Comparing average 2006-2008 TLRs for the MISO shows a similar experience as in PJM, with a reduction to 56% of the 2004 curtailments.[78] TLRs in regions without RTO Day-2 markets have been increasing.

Figure A-2: NERC TLR Orders



The benefit from reduced reliance on TLRs and associated lower costs from RTO coordinated markets is clear. The short-run system cost savings might not appear immediately in retail rates, but enhanced trading through reduced curtailments could raise profits in the short-run and reduce the cost of entry in the long run.

Hence, the quantitative experimental evidence employed must be used with care to address the proper question. As the GAO concluded, a full evaluation of the impacts of RTOs and their associated wholesale markets would go beyond the limited empirical evidence available today. However, the evidence is incomplete, not absent.

In addition to the quantitative experimental test, there are two other approaches that suggest themselves for evaluating the costs and the benefits of RTOs and organized wholesale markets. One method would be to conduct counterfactual simulations to estimate the costs and the benefits of organized wholesale markets. This is a common practice. For example, this was the approach used in Texas in 2008 to evaluate once again the costs and benefits of moving to the full locational marginal pricing (LMP) model for the wholesale spot market; the study found that benefits were substantially greater

78 NERC, December 22, 2008, http://www.nerc.com/docs/oc/scs/logs/trends.htm. Most of the increase in "other" is in SPP, which is an RTO but does not have a Day 2 market.

than costs.[79] A further example of such a simulation for PJM is provided below and in Appendix E, and this too reinforces the view that there are substantial benefits from the operation of the RTO-wide co-ordinated spot markets.

Missing the Forest for the Trees

The experimental evidence usually considered in the cost-benefit analyses focuses on the quantitative trees but often ignores an important view of the qualitative forest. In many ways, this forest is the more substantial part of the overall cost-benefit picture. The view of the forest includes the qualitative experimental evidence in the experience with RTO designs that attempted to provide open access and non-discrimination without the centralized market of the type now operated by RTOs. This experience speaks directly to the RTO critiques and the approaches APPA has recommended.

As discussed further below, there have been many attempts to develop RTO structures without the organized, centralized wholesale spot market and the associated LMP design. Given the principles of transmission open access and non-discrimination, these alternative RTO models confronted a fundamental dilemma. In short, there is no RTO design that has been shown to work and provide consistent incentives under these principles other than the basic LMP model. Every attempt to build an RTO model without the LMP framework has failed, visibly and dramatically, and either led to comprehensive reform to embrace the LMP model or compromised on the basic principles of open access and non-discrimination.[80] This dramatic evidence is "hidden in plain sight."

There are many ways to fail, but a common thread in the failed models included contract-scheduling restrictions on the spot market, inconsistent pricing models, and reliance on bilateral transactions without the support of a well-designed spot market. For example, in 1997 FERC ordered PJM to follow the recommendations of an Enron-led coalition to implement a simplified single-zone balancing model for its "small" balancing transactions, and rely primarily on bilateral schedules. This early PJM system imploded on the first hot day in 1997, and threatened to put the lights out until PJM suspended the market. As a result, PJM abandoned the failed market design and moved quickly to an LMP-based open spot market with Financial Transmission Rights (FTRs) in 1998.[81]

New England adopted its own version of a "simplified" spot market without the coordinated wholesale market based on the LMP model. Because of differences in detail from PJM, the failure mode appeared in perverse investment signals that by 1998 had been recognized as leading to the wrong new generation location decisions. The initial response was to impose discriminatory, administrative transmission cost allocations for generation investment. In the end, FERC intervened, recognizing that what was required was a more comprehensive market redesign. As a result, New England switched to the revised PJM market design, even to the point of using the same dispatch and LMP pricing software.[82]

The California case, discussed further in Appendix C, followed a parallel process with another Enron-led coalition arguing for bilateral contracts with a highly constrained balancing market and no effective spot market transactions. The resulting approach required repeated reforms until FERC concluded in 1999 that the basic design was "fundamentally flawed" and required a comprehensive market reform.[83] Although recognized and launched before the outbreak of the California crisis, the reform analysis was taken up again later and led to a new market design based on the LMP model now used by eastern RTOs.

Subsequently Texas, although innovative in the development of its retail market, embraced many of

79 CRA International, Resero Consulting , "Update on the ERCOT Nodal Market Cost-Benefit Analysis," December 18, 2008, (www.puc.state.tx.us).
80 Not all RTO regions meet the same test of open access and non-discrimination. Alberta and Ontario are examples of RTO-co-ordinated markets that do not use the LMP model and suffer the problems of restrictions on access and discriminatory pricing.
81 Hogan, "Reforms of Reforms," pp. 121-123.
82 Hogan, "Reforms of Reforms," pp. 123-124.
83 Hogan, "Reforms of Reforms," pp. 126-130.

the flawed elements of the original California wholesale market design with a zonal spot market and an emphasis on bilateral transactions. The results in Texas paralleled equivalent parts of the California experience. Since ERCOT, the Texas version of an RTO, is not subject to FERC jurisdiction, the ERCOT case demonstrates that FERC participation is not the explanation of failed market designs. And when Texas reconsidered, the final order was for implementation of an LMP-based system that moved in the opposite direction of the contract-scheduling model.[84]

The RTO-coordinated markets that from inception successfully provided open access in a non-discriminatory manner, such as MISO and NYISO, were LMP-based markets that arose from adherence to the laws of physics and basic economic principles. There was no need to enforce reliance on contract schedules and a restricted spot market, and the implosions elsewhere were avoided.

The evidence is clear that every attempt to provide transmission open access under principles of non-discrimination without using the organized, open spot market based on the LMP design has failed. These failures have been dramatic and unambiguous; the results overwhelm any simple quantitative cost-benefit assessment. Comparing hypothetical RTO models that cannot meet the objectives of open access and non-discrimination, with a proven model that does meet these objectives, is not cost-benefit analysis; it is tantamount to comparing good apples to bad oranges. The organized wholesale markets provide substantial benefits greater than the costs, and the combined weight of the evidence points not to restricting the functionality of these markets but to improving the market design in directions quite opposite of the main thrust of the APPA critique.

84 Public Utility Commission of Texas, "Rulemaking Proceeding on Wholesale Market Design Issues in the Electric Reliability Council of Texas," Project No. 26376, September 22, 2003.

Appendix B: The Contract-Scheduling Structure Would Impose Major Costs on RTO Regions

In *Consumers in Peril*, APPA criticized RTOs for using generator bids and competitive auctions to arrange the dispatch. They also faulted RTOs for using market-clearing prices based on locational marginal costs (LMP) to price dispatch services. But every RTO uses bids, auctions and market-clearing prices to arrange and sustain a reliable dispatch.

Every system operator strives to achieve an "economic" dispatch, which requires the operators to select and dispatch the lowest-cost set of dispatchable generators and loads at each moment. RTOs must use a "bid-based" auction system because they do not own generation; they need each dispatchable generator's offer prices and each dispatchable load's bid information to arrange an efficient ("least-cost") and reliable dispatch. LMP clearing prices then reflect the bid-based marginal cost of the dispatch in each dispatch interval, at each location, thus providing dispatched generators and loads with the correct incentives to follow dispatch instructions.[85]

Some entity(ies) must perform the dispatch function, and there must be some means to encourage generators and loads to follow dispatch instructions, but without these proven methods of offers, bids and market-clearing prices that reflect the marginal cost of meeting load at each location, no RTO could do so. Therefore, under the contract-scheduling approach, some other entity(ies) would have to perform this function.

Dismantling RTOs Would Incur a Multi-Billion Cost of Reacquiring Capacity

If critics were to succeed in eliminating RTO "bid-based" auctions for dispatch and locational marginal cost clearing prices, the dispatch would have to be performed by some other entity(ies) that owned or controlled sufficient generation to perform the dispatch function. Each dispatch entity would need the authority to ensure generator compliance with dispatch instructions and to overcome the perverse incentives provided by prices that do not clear markets. Vertically integrated utilities that both operate the dispatch and own their own generation fit that description.

Each entity controlling the dispatch would be required to own or have under its control sufficient generation to sustain a reliable dispatch. Each utility's dispatchers would obtain the information required to arrange and implement a dispatch from its own generators (and any other generators under its control), and would then direct those generators to follow dispatch instructions. The generators would do what they were told, since they were owned or controlled by the same entity that controlled the dispatch. In this framework, the entire region now served by PJM would need to reassemble the pieces of vertically integrated monopoly utilities to make the structure work.

APPA's original proposals eventually would have required the region to disband the regional dispatch function PJM now performs—that is, to dismantle the core function of the PJM power pool that has existed for decades—while requiring a number of large, transmission-owing utilities to perform separate dispatches in each sub-region or zone of PJM's footprint. The Day 1.5 proposals accepted the need to retain PJM's region-wide dispatch, but only for the time being and only to postpone immediate "transi-

---

85 When the transmission system faces congestion, the value of power is different at different locations, thus requiring an RTO to pay locationally different prices reflecting the marginal cost of serving load at each location. That system encourages all generators to follow dispatch instructions, a result essential for reliable operations. Failure to recognize these differences, such as by paying all generators the same price no matter where they are located, produces incentives to produce too much or too little at each location, while encouraging gaming of bid/offer prices. These problems with non-LMP systems have been widely documented in California and other systems that tried other approaches.

tion costs," not because APPA recognized the value of a regional pool-wide economic dispatch.[86]

Those who recall that a few of the RTOs (including PJM, New York and New England) evolved from power pools might ask why unraveling power pools is a logical consequence of the contract-scheduling model. The explanation is found in the national policy for transmission open access and non-discrimination. The old power pools functioned without an organized spot market, particularly without the associated locational prices and financial transmission rights. There was economic dispatch, but the benefits were shared by the member utilities through a complicated cost allocation scheme that required both closed access to third parties and discriminatory application. The old power pools cannot be reconstituted to preserve trading across vertically integrated utilities without abandoning the principles of open access and non-discrimination required under the Federal Power Act.

Several large utilities that were original members of PJM (*e.g.*, PPL, PECO, PSE&G) have long since spun off or divested all their generation. To make a utility-by-utility dispatch work again would first require those transmission-owning utilities that took over the dispatch function from PJM to reacquire generation from the current owners and thereby return to the vertically integrated utility structure prevalent in non-RTO regions of the US. It is never explained how this result could be achieved or how much it would cost.

In Appendix D, we provide a first-order estimate of the asset purchase costs utilities would face if they were forced to reacquire sufficient generation for a reliable dispatch. That estimate indicates the purchase costs for the PJM region would be more than $130 billion.

Limiting Transmission Access Would Increase Costs of Serving Load

The reemergence of separate, sub-regional dispatches by transmission-owning utilities would also have cost consequences for parties seeking to contract for inter-regional trades. With multiple dispatch zones, inter-regional trades would become more difficult and costly to arrange. Equally important, each zonal dispatch entity would be functioning under transmission access rules and physical rights regimes that would reduce parties' ability to gain access to the grid to implement their contract schedules. With reduced trading and higher costs, electricity prices would rise throughout the region.

The ability to make economic trades across dispatch boundaries is partly a function of how well trading parties discern feasible and economic opportunities in the face of unpredictable transmission congestion. With multiple dispatch zones, each subject to separate dispatch, individual traders cannot easily determine which trades are feasible or profitable. However, if security-constrained economic dispatch is applied across the entire trading region, as occurs when an RTO assumes a regional dispatch over previously separate dispatch zones, then the RTO coordination can facilitate trading that might not otherwise occur from uncoordinated bilateral trading; as Mansur and White found when PJM expanded its RTO dispatch to include Midwest utilities, the net exports from the Midwest to the Eastern parts of PJM almost tripled.[87]

> "We find that these changes enabled the organized market to direct production to the most efficient available resources, realizing significantly greater gains from trade than occurred under the bilateral trading system."[88]

If restricted access to the RTO's spot market coordination would reduce trading, what impact would that have on electricity prices? To examine the effect of limited spot market access and the associ-

86 APPA leaves the status of PJM's "power pool" uncertain. *Kelly/Caplan* at 535. In PJM, "pool-wide" or regional dispatch simply means that PJM arranges a dispatch to match supply from anywhere in the region with demand for the entire region. Participation in this pool-wide dispatch is voluntary and LSEs can meet their loads through self-supply or bilateral schedules. PJM does not require that every generator in the region submit to its dispatch; generators can choose to submit fixed schedules for their operations. PJM arranges the dispatch to accommodate these fixed schedules.

87 Mansur and White, Figure 2.

88 Mansur and White, at 2.

ated contract-scheduling proposals on prices, Ventyx performed a study for the region defined by the boundaries of PJM's RTO. Using a commonly accepted production cost model (PROMOD IV) to simulate how the system would operate (and set prices), Ventyx compared what electricity prices would be under PJM's current framework and market rules (PJM "as-is" case) and what they would be under a market/industry structure consistent with the contract-scheduling approach.

The alternative case assumes utility regions that joined PJM in recent years return to their pre-RTO status, when they functioned as vertically integrated utilities. This results in 14 different dispatch zones. Six of these zones correspond to the service areas of Commonwealth Edison, AEP, Allegheny, Dominion, Dayton and Duquesne, which are the six utilities that joined PJM since 1997 and turned over dispatch operations to PJM to create a much larger regional power pool. The other eight zones consist of the service areas of the transmission owners who created PJM decades ago.[89]

Sub-regional dispatches would be more costly than a larger regional dispatch. And acquiring transmission rights across multiple utility dispatch zones would increase transaction costs, reduce transmission usage and limit trading. Ventyx examined how much energy production costs and inter-utility trading costs would increase for this alternative scenario compared to the current PJM "as-is" case.

Taking a conservative view of the additional hurdles to least-cost operations and inter-area trading imposed by this contract-scheduling framework, Ventyx found that the costs would increase by at least $2.47 billion in energy costs alone over the next 10 years, compared to the current PJM "RTO as is" case. If the PJM electric demand consumers were paying market prices for all of their energy requirements, their energy purchase costs could increase by over $1.3 billion per year, or $13.6 billion over 10 years. These results are explained further in Appendix E.

89 We assume that this level of pool dismemberment would be sufficient to create enough utility dispatch zones. However, it is possible that further disaggregation would be necessary.

Appendix C: Experience From California to New England Shows The Risks of Restricting Access to RTO Spot Markets

Proposals for restricting the ability for market participants to transact in centrally coordinated markets are familiar: Enron made them in the mid 1990s, beginning in California, and then in New York and PJM. That history is important, because the contract-scheduling features Enron advocated led to systemic failure, yet the same arguments keep recurring at different times and in different forums.

Such proposals were first made by Enron and its allies in California when the original rules for the California ISO and Power Exchange (PX) were being debated in 1994-96. Initially, Enron opposed creating a regional pool to serve California and even opposed having an Independent System Operator (ISO) operate the pool. Enron and other advocates of an unfettered, decentralized market preferred to leave the dispatch function dispersed among the individual utilities.[90] However, during the stakeholder process, it became apparent that a regional pool could reduce costs, and that the pool would need to be independently operated to avoid discrimination.

Like today's critics of RTO spot markets, Enron then proposed that the ISO perform only a limited balancing function. As a trader and middleman, Enron preferred a system in which load-serving entities relied almost exclusively on bilateral trading to match supply and demand. The limited balancing mechanism would handle final adjustments of the generation plant dispatch and would charge scheduling parties for minor deviations from the fixed schedules associated with parties' bilateral contracts. There would be no organized day-ahead market; Enron argued that competitive traders would efficiently handle all forward trading.

If the Enron approach had adhered to the principles of open access and non-discrimination, with no cross subsidies between parties, and no limits on the parties' ability to rely on the dispatch for balancing when it was economic to do so, this approach could have led to the same outcome as the regional pooled dispatch and associated spot markets operated by RTOs today. However, this would have required that parties' access to the ISO's dispatch not be arbitrarily limited and that the spot prices from using the dispatch reflect marginal costs of the dispatch used to balance the system and manage congestion.[91]

In 1995, the California Public Utilities Commission directed its jurisdictional (investor-owned) utilities to create an ISO, which would operate a regional pool dispatch for the state's three investor-owned utilities. The ISO would manage a bid-based real-time dispatch (and associated balancing or spot market), while another new entity, the Power Exchange (PX), would coordinate a day-ahead market through which the utilities would buy and sell the energy they needed to match their expected loads. The PUC's 1995 order did not limit access to the ISO's dispatch and associated spot markets, but the PUC left important details of market design to a stakeholder process dominated by the utilities, industrial customer groups and Enron.

---

90 *Initial Comments of Enron Power Marketing, Inc. In Response to California Public Utilities Commission's Order Instituting Rulemaking and Order Instituting Investigation*, June 8, 1994, filed in Dockets R.94-04-031 and I.94-04-132; also, hearing testimony of Enron's Jeff Skilling before the California Commission on August 4, 1994 (transcript at 1136).

91 William W. Hogan, *Economic Dispatch, Transaction Accounting and the OPCO or POOLCO Model*, August 31, 1994, prepared for the California Public Utilities Commission and restructuring working groups during consideration of alternative models for an independent system operator. For an explanation of why open access to an ISO's dispatch (and resulting spot market) is essential to support bilateral contracting, see Hogan's *An Efficient Bilateral Market Needs a Pool*, testimony submitted to the California Public Utilities Commission, August 4, 1994 in Dockets R94-04-031 and I.9404-132. The principles explained in these and related papers from that era would eventually become the foundation for all ISO/RTO markets. The papers are available at: www.whogan.com.

Enron and other energy marketers vigorously opposed integrating the operations of the ISO and PX.[92] During the development of market rules, they insisted that the separate PX manage the day-ahead market independently of the ISO, because they feared most trading would occur day ahead in the PX's bid-based auctions without the need for marketers. But having a PX separate from the ISO would eventually prove unworkable, because day-ahead schedules accepted without regard to congestion would prove to be infeasible in real time. The proposal was never implemented in practice. The ISO quickly determined that the Enron approach would compromise its ability to manage congestion and keep the lights on. This is the same reason that advance scheduling of contracts without RTO spot market coordination would be problematic.

Once California decided to create an ISO, Enron advocated limiting the ISO's spot market to a narrow "balancing mechanism," just as today's RTO critics urge today. While Enron argued publicly that forward contract markets could achieve more efficient results than the ISO, it was also true that limiting access to the ISO would benefit Enron. If access to the ISO's balancing mechanism could be constrained, and traders penalized for using it, traders could be forced to turn to Enron or other marketers to provide services they could not easily obtain from the ISO.

To limit the efficiency of the ISO's real-time dispatch and discourage its use, Enron also pushed for rules to prevent the ISO's central dispatch from achieving least-cost results. Astonishingly, Enron and its allies convinced enough California parties and FERC that in order to limit the ISO's balancing market, the ISO should be restrained from pursuing the lowest cost dispatch solutions to congestion and balancing, imposing a rule that *by design* raised costs and complicated reliable operations. This rule persisted through the energy crises in 2000-2001; it took years to remedy this design flaw.

Why did California regulators and FERC accept such obviously anti-consumer restrictions on the ISO? Enron argued that "the market" would function better if the ISO's market coordination was deliberately restrained and made inefficient, so as to create arbitrage profit opportunities for marketers and strategically located parties. These parties, Enron claimed, would produce lower cost results through unfettered marketer trading than the ISO could achieve through regionally coordinated least-cost dispatch. Of course, there was no evidence or theory to support Enron's claims, and simple economic logic would lead to the conclusion that a higher-cost dispatch would actually enable and shield higher-cost contract trading. Nevertheless, parts of Enron's design for California won the day in 1996.

The results were both predictable and predicted: a compromised ISO dispatch that struggled to maintain reliable operations while suffering higher costs, exacerbated by manipulation and bid gaming from savvy marketers and strategically located generators.

With some exceptions, the eastern ISOs avoided California's flawed designs, insisting instead that the ISO be required to operate a security-constrained economic dispatch. In such a dispatch, the ISO is obligated to select and dispatch the lowest-cost mix of generators to balance the system and meet all transmission safe operating limits. To be sure, Enron representatives and others made the same arguments in the East that they made in California,[93] but their proposals for a "limited balancing market" were rejected by the parties supporting the original PJM and New York ISOs. The New York ISO began operations with locational marginal pricing, over Enron's objections;[94] after an initial wrong turn down the path espoused by Enron, PJM began using LMP in 1998. Since then, other US RTOs eventually followed the New York and PJM models. Together, the improved designs have allowed the Eastern ISOs to maintain reliable operations and pursue economic dispatch solutions to the complicated physical issues that characterize electricity grids.

---

92 *Comments of Enron Capital & Trading Resources, et al on the Memorandum of Understanding Filed September 11, 1995*, filed October 2, 1995, in Dockets R.94-04-031 and I.94-04-132.

93 See, *e.g.*, Initial Comments of Enron Capital & Trade Resources on the Optimal Model for New York State's Electric Industry, submitted October 24, 1995, to the New York State Public Service Commission in PSC Case NO. 94 — E — 0952.

94 California rejected LMP and instead adopted a compromised spot pricing regime that was easily and repeatedly manipulated, with Enron inventing various gaming strategies to create artificial congestion and be paid to relieve it. Later reforms emphasized the need to implement the LMP model.

Proposals similar to Enron's have been discredited for nearly a decade, not merely because of California's experience with flawed designs, but also because no one has ever demonstrated how an ISO/RTO can facilitate forward contract scheduling and ensure open, non-discriminatory access without the core elements that RTO market opponents seek to eliminate. Ample experience has shown there are no workable solutions consistent with those goals without organized spot markets using bid-based security-constrained, economic dispatch and locational clearing prices, the core features of the RTO organized markets. And opponents have failed to describe any workable alternative that supports both market and regulated environments, while meeting the federal statutory requirement to support competition and provide non-discriminatory, open access to transmission.[95] This flexibility is a necessary requirement, because the highly interconnected eastern grid encompasses both traditionally regulated states (*e.g.*, Indiana) and "restructured" states (*e.g.*, New Jersey and Pennsylvania) with many variations in LSEs and generation ownership.

The eastern RTOs are not unique in coming to this conclusion. Every RTO in the country eventually arrived, voluntarily and through its own history, at the same conclusion. One finds the same core elements of bid-based, security-constrained economic dispatch with locational prices in PJM, the New York ISO, and the New England ISO; the same features appear in the revised rules at the California ISO, ERCOT,[96] and the rules developed by the newest RTOs: the Midwest ISO and (with some exceptions) the Southwest Power Pool. Today, more than two-thirds of electricity consumers function under this framework.

RTO market critics who seek to alter or compromise the core elements of RTO regional dispatch and associated markets have a burden to show they are at least compatible with the underlying physical requirements of grid operations and can achieve the economic benefits of efficiently priced pooled dispatch without discriminating against some users and/or creating barriers to entry.

---

95 William W. Hogan and John D. Chandley, *A Path to Preventing Undue Discrimination and Preference in Transmission Service*, comments submitted to the Federal Energy Regulatory System, August 2, 2006; this and follow-up papers on how RTOs provide open access are available at: http://ksghome.harvard.edu/~whogan/

96 Note that ERCOT, the ISO for most of Texas, arrived at the same conclusion even though ERCOT is not subject to FERC jurisdiction. California, New England and Midwest ISO independently adopted the basic PJM/New York design after watching their original alternative models undermine reliability.

Appendix D: Estimated Costs of Reacquiring Generating Capacity

In order to implement limited spot market access and the associated contract-scheduling structure, it would be necessary to purchase the fleet of generators in the PJM control area that are not already owned by public power or the regulated portions of investor-owned utilities. As we explain further in Appendix F, in order to purchase a generator, it will be necessary to pay its owner an amount sufficient to induce it to give up its rights to (1) the net energy revenues that generator would otherwise be expected to earn (defined as energy revenues net of the variable costs it incurs to produce energy), plus (2) the capacity revenues it expects to earn, minus (3) the fixed costs the owner of that generator incurs to make it available for operation, but which could be avoided if the generator were shut down.

One approach to estimating the value of these generators is to estimate the present value of each of these three cash streams for each generator in the PJM control area that would be purchased to implement the structure. However, given the large number of generators to be valued, that would be extremely difficult. Instead, we have applied a simplified approach to value these generators, which builds upon work performed by Levitan and Associates Inc. (LAI) in a study performed for the Maryland Public Service Commission. In that study, LAI stated, "The current fair market value of Maryland's power generators is at least $18 billion."[97] Based on LAI's valuation, and on an assessment of the impact that differences in generating technology, location, age, generating capacity and outage rates would be expected to have on the value of the revenue streams that each generator would be expected to realize (and hence the cost of purchasing each of those generators), we estimated the value of each generator in PJM that would have to be purchased.

Using this simplified approach, we estimated the cost of purchasing those generators at $133 billion. While a plant-by-plant evaluation of each of the cash streams described above would provide a more accurate assessment of this cost, this estimate realistically conveys a sense of the approximate cost that would be incurred to purchase this amount of generating capacity. The remainder of this appendix describes the methodology we used to calculate that estimate.

Factors That Affect Generator Value

There are five primary factors that affect the revenue streams that generator owners receive, and hence the amount that a generator owner would require in order to sell it:

- **Generating Technology.** Since baseload generators can produce energy at very low variable costs, they can produce a given MWh less expensively. They also are called upon to operate more often than other, more expensive plants. Both of these factors cause the net energy revenues that baseload generators earn, stated in terms of dollars per MW of capacity, to be higher than the net energy revenues that other generators earn. The capacity revenues that different generation technologies earn should be about the same (holding everything else equal), since capacity markets do not differentiate between different generation technologies. Therefore, the revenue stream that baseload generators realize will generally be larger than the revenue stream that other generators realize, so it will cost more to purchase a baseload generator of a given size than to purchase other generators of that size. For similar reasons, it is less expensive to purchase peaking generators, whose variable cost of generating energy is high, than it is to purchase intermediate generators, whose variable cost of generation is between the costs of baseload and peaking generators.

- **Location.** Energy prices vary from location to location within PJM, because PJM uses locational pricing. Under locational pricing, the price of energy at each location reflects the marginal cost of producing additional energy at that location. When there is transmission congestion that restricts

97 Kaye Scholer LLP, Levitan & Associates, Inc. and Semcas Consulting Associates, *State Analysis and Survey on Restructuring & Re-Regulation*, in Response to Task #2 Request for Proposals PSC #01-01-08, November 30, 2007, at 69.

the amount of inexpensive generators at one location to serve the needs of consumers at another location, thereby making it necessary to operate more costly generators to meet those consumers' needs, the price of energy in the second location will exceed the price of energy at the first location. Likewise, capacity revenues may also vary locationally, since they are based on the cost of developing capacity, which can vary from one location to another. Therefore, the net energy revenue and capacity revenue streams that a generator owner expects to realize may depend upon its location, so the cost of purchasing that generator may also depend upon its location.

- **Age.** The owner of a generator that was built long ago and is approaching obsolescence, and that will only realize energy and capacity revenues for a few more years, will be willing to sell that generator for considerably less than a newer generator that will continue to produce energy and capacity revenues for many more years.

- **Generating Capacity**. Generators with more generating capacity are able to produce more energy, so they receive more net energy revenue; they also qualify for larger capacity payments, everything else held equal. Therefore, the revenues that the owner of a larger generator expects to earn will generally be larger than the revenues a smaller generator expects to earn, so the cost of purchasing larger generators is greater than the cost of purchasing smaller generators.

- **Outage Rate.** Finally, generators that are more frequently unavailable will realize fewer energy revenues. They will also realize fewer capacity revenues, since unforced capacity, which incorporates a correction to account for unplanned outages, is the metric that is used in PJM to determine the amount of capacity a generator is permitted to provide. Consequently, all else held equal, generators that are more frequently out will sell for less than other generators.

To estimate the cost of purchasing a given generator, it is necessary to take the impact of these five factors on its value into account.

Adjusting for Differences in Generating Technology

Over the long term, there is a certain average amount of revenue that a generator owner would have to expect to earn each year, below which it would not be willing to develop new generation. That amount of revenue is called the "levelized annual cost" of building that generator. PJM's market monitoring unit (MMU) has compiled statistics on the levelized annual costs of building generators using three different technologies. These are reported in the *State of the Market Report* the MMU issues each year. The results for 2005 through 2007 are as follows:

Levelized Fixed Costs for Entrants ($/MW-yr.)

Technology	2005	2006	2007	Average
Combustion Turbine	72,207	80,315	90,656	81,059
Combined Cycle	93,549	99,230	143,600	112,126
Pulverized Coal	208,247	267,792	359,750	278,596

Source: PJM Interconnection, *2007 State of the Market Report*, Table 3-22.

As this shows, while there is some movement from year to year, the cost of building a combustion turbine generator averages about 27 percent of the cost of building a combustion pulverized coal generator, and the cost of building a combined cycle generator averages about 41 percent of the cost of building a combustion pulverized coal generator. Consequently, on average, we would expect the sum of the net energy and capacity revenues that the owner of a combustion turbine would receive would be about 27 percent of the sum of the net energy and capacity revenues that the owner of a pulverized coal plant would receive, and the sum of the net energy and capacity revenues that the owner of a combustion turbine would receive would be about 41 percent of the sum of the net energy and capacity revenues that the owner of a pulverized coal generator would receive. If it were otherwise—for example, if the owner of a combustion turbine expected revenues that were 50 percent of the revenues that the owner of a pulverized coal generator would realize, in return for only spending 27 percent as much as the pulverized coal generator developer spends—then everyone would built combustion turbines and no one would build pulverized coal generators. This would progressively reduce the revenues that combustion turbines would receive, as compared to the revenues that pulverized coal generators receive, until the point where this disparity disappears.

Therefore, on average, combustion turbines (or other peaking generators) can be expected to earn net energy and capacity revenues that are about 27 percent of the net energy and capacity revenues that otherwise identical pulverized coal generators would earn, and combined cycle generators (or other intermediate generators) can be expected to earn net energy and capacity revenues that are about 41 percent of the net energy and capacity revenues that otherwise identical pulverized coal generators would earn. So the cost of purchasing peaking capacity would be expected to be about 27 percent of the cost of purchasing baseload capacity, and the cost of purchasing intermediate capacity would be expected to be about 41 percent of the cost of purchasing baseload capacity, all other things held equal.[98]

Adjusting for Differences in Location

In the 2007 State of the Markets Report, PJM's MMU also reported the sum of net energy revenue and the capacity revenue that an entrant generator using each of the three technologies above would have earned in each of the zones within PJM. This permits us to assess the impact that location has on the total revenue stream that a generator owner using a given technology would expect to realize; for example, in 2007, a new combustion turbine in the BGE zone would have been expected to earn $94,710/MW-yr. in net energy revenue and capacity revenue, while a new combustion turbine in the MetEd zone would only have been expected to earn $46,663/MW-yr., about half as much.[99] Therefore, if these sorts of revenue differences are expected to persist, one would expect the sale price of a generator in the BGE zone to be about twice the sale price of an otherwise identical generator in the MetEd zone.

Adjusting for Differences in Age

Different generators will have different lifespans, but the value of a baseload generator in a given location with only three years remaining in its lifespan is not simply one-tenth of the value of a baseload generator with the same capacity and at the same location that is expected to remain in service for another 30 years. That is because the value of a dollar in revenues that a generator owner expects to earn 30 years from now is considerably less than the value of a dollar in revenues that a generator expects to earn this year. Consequently, the value of a generator with only three years left in its lifespan is more than one-tenth the value of an otherwise identical generator with 30 years of life remaining. In fact, using an annual discount rate of 7%, the value of a generator with only three years left in its lifespan is about 21 percent of the value of an otherwise identical generator with 30 years of life remaining.

For the purposes of this analysis, we assumed that each generator would have a useful service life of thirty years starting with its in-service date. However, many generators are more than thirty years old. Therefore, we assumed a minimum value for the remaining lifespan of three year for all units other than nuclear generators. For nuclear generators, we assumed a minimum remaining lifespan of seven years,

---

98 For the purposes of this analysis, all steam turbines, the steam portions of combined cycle units, and all hydraulic turbines other than pumped storage were classified as baseload units; peaking units included all combustion turbines, internal combustion engines, and wind turbines; and intermediate units included all single-shaft combined cycles, combined cycles not otherwise broken down, pumped storage, and units not otherwise classified. Reclassifying non-pumped storage hydraulic turbines or wind turbines had little impact on the estimated cost of purchasing the generation fleet.

99 PJM Interconnection, *2007 State of the Market Report*, Tables 3-24, 3-26 and 3-28.

since it seems unlikely that many nuclear units will shut down in the next three years. This assumption produces a value for those units that is consistent with the $4.5 billion recently paid by Electricité de France ("EDF") for a 49.99 percent interest in Constellation Energy Nuclear Group ("CENG").[100]

Adjusting for Differences in Generating Capacity and Outage Rates

Finally, the amount of unforced capacity that a generator can provide reflects its generating capacity adjusted to account for its outage rate. The capacity revenues a generator earns will be directly proportional to the amount of unforced capacity it provides, and while the energy revenues it receives are not directly proportional to this amount, they should be roughly proportional (particularly since the adjustment to account for outage rates is typically not large). Therefore, the ratio of the amount of unforced capacity that two otherwise identical generators provided approximates the ratio of the cost of purchasing those generators.

Using these Adjustments to Derive the Total Purchase Cost

Since we know the generating technology, location and age of each generator in Maryland, as well as the amount of unforced capacity it can generate, we used that information to help us determine the cost of purchasing the fleet of generators in the PJM control area that are not already owned by public power or the regulated portions of investor-owned utilities. We calculated the value of each generator in Maryland given (1) the factors above, which establish the relative values of each of those generators given differences in generating technology, location, age, capacity and outage rate, and (2) the need for the sum of the values of Maryland generators to sum to $18 billion to conform to LAI's calculation.

Illustrative Example

Before we delve into the mathematical detail of the equations that were used to perform these calculations, an example illustrating the gist of the approach is likely to be useful. Consider the value of a combined cycle generator compared to the value of a pulverized coal generator, using the latter generator as a benchmark. For the purposes of this example, assume a combined cycle generator would be worth about 40 percent of the value of an otherwise identical pulverized coal generator. (The figure that we used is actually about 41 percent.) However, suppose that the combined cycle generator is in a location where its value is 120% of the value of the combined cycle generator in the location that was assumed when calculating the annual levelized costs of such a generator while the pulverized coal generator is in a location where its value is 80% of the value of the pulverized coal in the location assumed in when calculating the annual levelized costs of developing such a generator.

In that case, everything else is not equal, so it is not accurate to state that the value of the combined cycle generator is 40 percent of the value of the pulverized coal generator. Instead, once these locational adjustments are taken into account, the ratio of the value of the combined cycle generator to the pulverized coal generator increases from 40% to 40% x (120% / 80%) = 60%.

Similarly, assume that:

- Taking age into account leads to the conclusion that the value of the combined cycle generator is twice the value of an otherwise identical pulverized coal generator.

100 The $133 billion cost of purchasing non-utility-owned generation in PJM includes $4.06 billion for the purchase of CENG's Calvert Cliffs units. Calvert Cliffs represents about 44.5 percent of CENG's capacity, the remainder of which is in New York and which therefore was not included in our study. However, if we assume that the New York capacity is just as valuable on a per-MW basis as the Calvert Cliffs capacity, then the value of all of CENG's capacity would be $4.06 billion / 44.5% = $9.12 billion, so the value of EDF's share of CENG is $4.56 billion, approximately equal to the $4.5 billion paid by EDF.

- The combined cycle generator provided 100 MW of unforced capacity, while the pulverized coal generator provided 300 MW of unforced capacity.

Adding the impact of these two factors to the locational and technology adjustments described above leads to the conclusion that the value of the combined cycle generator should be 60% Ð (1/2) Ð (1/3) = 10% of the value of the pulverized coal generator. Therefore, if the total value of the two generators was assumed to be $330 million, it would be appropriate to assign a value of $300 million to the pulverized coal generator and a value of $30 million to the combined cycle generator.

Implementation

We implemented this approach using the following three-step procedure:

1. We calculated the amount of normalized capacity that each generator provided. Normalized capacity is the amount of unforced capacity a generator provided, adjusted to account for the impact of that generator's technology, location and remaining lifespan (each as compared to a benchmark generator) on its value, as described above. Therefore, the ratio of two generators' normalized capacities should reflect the ratio of their values.[101]

2. Next, we divided the $18 billion value of the Maryland generation fleet estimated by LAI by the number of MW of normalized capacity in that fleet to determine a value per MW of normalized capacity.

3. Finally, we multiplied the amount of normalized capacity provided by each generator to be purchased and the value per MW of normalized capacity that was consistent with LAI's valuation of the Maryland fleet, and summed the result over all generators that would have to be purchased.

The number of MW of normalized capacity that each generator provided was calculated using the following equation:

$$NCAP_{g,t,z} = UCAP_g \cdot TNF_t \cdot ZNF_{t,z} \cdot LNF_g$$

where:

$NCAP_{g,t,z}$ is the normalized amount of capacity provided by a generator g of technology type t located in zone z;

$UCAP_g$ is the amount of unforced capacity provided by generator g, as reported in the 2008 PJM Load, Capacity and Transmission Report, Sch. 3, Part D;

TNF_t is the technology normalization factor for generators of technology type t;

$ZNF_{t,z}$ is the zonal normalization factor for generators of technology type t located in zone z; and

LNF_g is the lifespan normalization factor for generator g;

$TNFt$ was calculated as the ratio of the levelized fixed cost of an entrant using technology type t averaged over 2005-07, as reported in Table 3-22 of the *2007 State of the Market Report*, to the levelized fixed cost of a pulverized coal plant over that time period as reported therein;

101 In the illustrative example, using a pulverized coal generator receiving the average level of revenue in PJM as the benchmark, the combined cycle generator would have provided 100 MW x 0.4 x 1.2 x 0.5 = 24 MW of normalized capacity (with the adjustments respectively reflecting the impact of the combined cycle's technology, location and age on its value as compared to the benchmark generator), while the pulverized coal generator would have provided 300 MW x 1 x 0.8 x 1 = 240 MW of normalized capacity.

$ZNF_{t,z}$ was calculated as the ratio of the net revenue that would have been earned in 2007 by a generator in zone *z* using technology type *t*, as reported in Tables 3-24, 3-26 and 3-28 of the *2007 State of the Market Report*, to the average net revenue that would have been reported in those tables for a generator in PJM using that technology type; and

LNF_g was calculated using the following equation:

$$LNF_g = \sum_{i=1}^{RL_g} \frac{1}{(1+d)^{i-1}} \Bigg/ \sum_{i=1}^{30} \frac{1}{(1+d)^{i-1}},$$

where:

RLg, the remaining lifespan of generator g, is equal to the greater of (1) the number of years from Jan. 1, 2009 to a date 30 years after generator *g*'s in-service date, as reported in the 2008 PJM Load, Capacity and Transmission Report, Sch. 3, Part D; or (2) three years (seven years if generator *g* is a nuclear generator); and *d*, the real discount rate applicable to the cash flows resulting from generation ownership, was set at 7 percent per year.

The resulting valuation for each generator to be purchased is consistent with the each of the rules above, regarding the relative values of generators using different technologies, at different locations, of different ages, with different capacities, and with different outage rates, while also being consistent with the valuation that LAI calculated for the Maryland generation fleet. The $133 billion estimate of the cost of purchasing the non-utilty-owned portion of the PJM generation fleet corresponds to a value of $1,123 per kW of capacity purchased; by way of comparison, LAI's calculation of the cost of purchasing the Maryland generation fleet corresponded to $1,390 per kW.[102]

102 Detailed calculations are included in an Excel spreadsheet available from the authors.

Appendix E: Estimate of Increased Energy Costs within PJM

In response to the APPA proposal to reform energy markets, Ventyx has performed a *pro forma* quantitative analysis of the PJM market, to attempt to quantify the increase in energy costs that would ensue as a result of this proposed unraveling of the integrated energy market.

This analysis focuses on a view of the PJM market for the nominal 2006-2007 market year. From its latest release of the Marketvision™ database, Ventyx has extracted the data required to represent the current PJM footprint. Using PROMOD IV®, Ventyx' commercial software simulation model for electric markets, Ventyx has represented the "As Is" PJM market for the June 2006 through May 2007 market year. A second simulation was performed, based on a representation of the effective market conditions in a Revised Market, consistent with the bilateral market envisioned by APPA.

Information regarding PROMOD IV is available on the web at:

http://www1.ventyx.com/analytics/promod.asp

Basic Data Assumptions

The "As Is" market simulation represents the existing PJM under a standard assumption of coordinated unit commitment and dispatch by the centralized RTO market.

The representation of the transmission system is based on a 2008 Summer Peak MMWG case, from the 2006 MMWG series of powerflow cases. For the portion of the transmission system outside of the PJM RTO simulation footprint, PROMOD IV scales the bus generations to match the total bus loads, so as to remove any net interchange between PJM and the non-PJM powerflow areas.

Hourly demands for each of the seven PJM zones reflect the actual zonal demands as posted by PJM.

The "PJM Classic" zone was divided into eight zones, consistent with the legacy investor-owned utility structure of PJM before market restructuring. This analysis, then, represented PJM as being composed of a total of 14 zones.

In performing its security-constrained unit commitment and dispatch, PROMOD IV monitors a prescribed set of contingency constraints, or flowgates, just as the actual markets are scheduled based on a defined list of commercial flowgates. For this study, PJM staff provided a constraint set that is used in similar in-house PROMOD IV analyses.

Natural gas and oil prices, as well as SO2 and NOx prices, matched contemporary commodity prices for the 2006-2007 simulation period.

For nuclear generating units, actual generator outages lasting one week or longer (primarily refueling outages) were directly specified in the data. All other generators' scheduled maintenance outages were scheduled internally by the model, based on a reliability levelization algorithm.

In order to capture cost impacts on a zone-by-zone basis, the PJM generating resources needed to be assigned to the different zones, more or less representing a vertically integrated traditional utility. The starting point for this assignment of generator "ownerships" is a spreadsheet that PJM provides, for use by market stakeholders in planning their FTR market participation. This spreadsheet identifies the historical generating resources of each zone, prior to implementation and expansion of the PJM market. The resource assignments to the zones resulting from this historical information were adjusted so that newer resources not represented in this spreadsheet would be assigned to capacity-deficient zones so as to result in a roughly equal summer peak installed reserve margin over the zones. The overall reserve margin for the simulation footprint is approximately 15%. The resulting reserve margins for the individual zones, after assignment of newer combined cycle and CT generators, are all in the range of 14%-15%.

Spinning reserve requirements were identified from the PJM market monitor's *State of the Market Report*. For the "As Is" simulation, this results in spinning reserve requirements primarily for the ComEd and MidAtlantic zones, plus a small requirement for the Dominion zone.

Revised Market Simulation.

The "Revised Market" scenario represents a market relying on bilateral energy trades. A market coordinator would ensure that these schedules satisfied transmission constraints. Finally, some entity would operate a real-time balancing market, the intent of which is to schedule for deviations from the submitted schedules. In *Consumers in Peril*, APPA originally proposed that this balancing market would clear no more than 5% of the energy in the overall market; in *Competitive Market Plan*, APPA continues to assume most trading would be done through bilateral contracts and not the spot market.

The world envisioned by APPA would comprise energy scheduling entities ranging from traditional vertically integrated utilities to retail LSEs. It is not possible to simulate how each of these entities would arrange their bilateral schedules in a real-world unstructured market. For purposes of this study, the "Revised Market" representation of the PJM zones assumes that each of these fourteen zones would operate as the equivalent of a traditional control area, with centralized commitment and dispatch of generating resources within the zone. This is a conservative assumption with regard to energy costs of the actual scheduling entities within the zone, because it assumes that any implied market inefficiency due to the independent scheduling by the zone's members could be resolved by this coordinated dispatch within the zone.

From a modeling perspective, the diminished efficiency (higher energy production cost) of a bilateral market is due to three primary factors. First, the bilateral energy scheduling process is less efficient than the schedules derived by a centralized LMP market. This energy market inefficiency is manifested as an implied hurdle rate for scheduling economic interchange among the zones. These hurdle rates have physical components, such as the OATT through-and-out rate that must be charged to schedule a firm energy transfer, as well as out-of-pocket trade execution costs. Additionally, there is a significant non-physical component to these hurdle rates, reflecting the market inefficiencies related to a lack of price transparency and centralized market clearing.

Because scheduling firm, day-ahead transactions bears a higher cost (firm transmission charges, no centrally cleared day-ahead market), a higher hurdle rate between the zones is assumed for purposes of unit commitment than is used for the hourly non-firm interchange. Various RTO cost/benefit studies that have been performed in recent years have assumed a range of values for these hurdle rates. In 2003, Cambridge Energy Research Associates (CERA) performed a cost/benefit study for AEP's entry into the PJM market. In that study, CERA used a commitment hurdle rate of $7.25, and a dispatch hurdle rate of $4.25. Although CERA states that these hurdle rates were derived from a benchmark to historical market conditions, these rates (in particular, the commitment hurdle rate) are somewhat lower than some other calibrations. For this study, Ventyx has used a dispatch hurdle rate of $5, and a commitment hurdle rate of $8, although other calibrations suggest that a commitment hurdle rate in the range of $10-$12 might be appropriate.

The second source of inefficiency in the bilateral market is congestion management. In the LMP market, PJM schedules energy flows up to the physical flowgate limits. In a bilateral market, these flows are limited by the ATC limits posted on the OASIS sites, which reflect derations (?) for such factors as TRM. Furthermore, a security coordinator reviews actual schedules to determine their combined feasibility under current conditions, and must curtail schedules on a non-economic priority basis when infeasible flows are anticipated. APPA recognizes that this function would still need to be performed by a security coordinator in the bilateral market that it proposes.

As part of a series of cost/benefit studies performed with MISO in recent years, historical TLR records were reviewed, to determine the impact of the inefficiencies of TLR congestion management. For Level 3 TLR curtailments that had occurred over several years, the actual flows scheduled after curtailment were compared to the nominal limits that were used to apply the curtailments. For transmission in the ECAR and MAIN regions, it was found that TLR curtailments resulted in a 9% under-utilization of the transmission flowgates. Consistent with these previous studies, Ventyx has applied this 9% deration to the flowgate limits in the "Revised Market" scenario.

The third market characteristic that introduces inefficiency in the bilateral market is decentralization of ancillary services. In this study, Ventyx has assumed that the total ancillary services (spinning reserves and load-following) would remain the same. However, in the "Revised Market" scenario, the MidAtlantic spinning reserves were allocated over the Reliability First zones in proportion to their non-coincident annual peak loads.

Measure of Increased Energy Costs

For the total PJM footprint, the increase in energy supply cost due to revising the market is simply the change in energy production costs of the resources in the market. In order to identify the change in costs zone by zone, it is necessary to adopt a definition for the prices paid within the market for energy exchanges. For this and similar studies, Ventyx has assumed that a zone that is buying in an hour will pay its purchased energy times its generation-weighted zonal LMP. These revenues are then allocated over the sellers in proportion to their sold energy times their generation-weighted zonal LMP in the hour. The resulting "adjusted production cost" is used as the zonal cost measure for the analysis.

Alternative measures of the increase in cost are the change in generator revenues and the change in load payments.

The estimated annual increase in energy costs are:

	Adjusted Production Cost	Generator Revenues	Load Payments
AEC	$6,375,371	$41,949,138	$41,555,752
AEP	$5,286,571	-$397,463,001	-$162,736,564
Allghny	$30,777,824	$3,792,891	$301,899,006
BG&E	$36,588,624	-$3,650,568	$108,967,851
ComEd	-$57,700,762	-$276,402,171	-$155,260,602
Dayton	$8,759,490	$45,857,962	$41,375,117
Dominion	$52,332,676	$766,995,222	$459,796,530
DPL	$32,912,194	$135,573,670	$90,278,689
Duquesne	$7,482,096	-$10,859,380	-$16,569,218
FE	$22,646,014	$59,254,467	$172,473,389
PECO	$706,893	-$34,955,280	$28,938,435
PEPCO	$60,815,115	$218,849,693	$246,932,591
PPL	$10,371,573	-$78,825,134	$28,960,383
PSEG	$29,929,899	$244,326,418	$174,651,798
Total PJM	**$247,283,578**	**$714,443,927**	**$1,361,263,157**

This analysis indicates that, over ten years, the increase in energy costs due to reliance on a bilateral market would be $2.5B, measured as adjusted production cost, or as much as $13.6B, measured as increased load payments.

Appendix F: APPA's Proposals Would Not Save Money On Capacity Costs

In the several RTO-administered markets, including PJM, independently owned generators realize two main sources of revenue. First, they sell energy to the market. Second, they sell capacity into capacity markets, such as PJM's RPM. (They may also earn revenues from the sale of ancillary services, but those are relatively minor for most generators and will be disregarded for the balance of this discussion.)

There are costs involved in operating a generator. Some of those costs (primarily fuel costs) are directly related to the amount of energy that a generator produces, so they could be avoided if a generator were to shut down. In addition, other costs, such as the labor costs associated with staffing a generator, are not directly related to energy production, but could also be avoided if a generator were shut down. But to the extent that the energy and capacity revenues paid to generators exceed the sum of these avoidable costs, the difference accrues to the owners of those generators.

If these generators were instead paid on an embedded cost basis, end-use customers would no longer have to pay this difference. Instead, they would only need to cover the costs that are directly associated with generating energy and the other costs associated with operating a generator. Consequently, it is tempting to conclude that this difference between the market-based revenues that generator owners earn and the avoidable costs associated with operating a generator represent an amount that could be saved by returning to the vertically integrated paradigm.

But this analysis overlooks a vital element: The cost of purchasing those generators from their owners. To the extent that the energy and capacity revenues that a generator owner expects to receive exceeds the costs it expects to incur as a result of operating a generator, that generator owner will require a payment that is sufficient to compensate it for foregoing that difference. If that payment is financed over time, the cost of purchasing those generators will not be substantially different from the cost of making energy and capacity payments to the owners of those generators under the current market structure. Therefore, this potential source of savings from returning to the vertically integrated paradigm is illusory.

To illustrate, suppose for simplicity that a generator has an anticipated lifespan of three years. It expects to realize $500 million in revenues from the sale of energy in the first year, and the cost of generating that energy is expected to be $350 million. It also anticipates $100 million in capacity revenue that year, and $50 million in other operating costs, so the total operating profit it expects to realize in that first year is $500 million + $100 million — $350 million — $50 million = $200 million. All costs and revenues are expected to increase at a rate of five percent per year. Therefore, the generator expects to realize operating profits of $631 million over its lifespan, as calculated in the table below.

Anticipated Operating Profit Over Hypothetical Generator's Lifespan

	Year 1	Year 2	Year 3	Total
Energy Revenues	500	525	551	1,576
Capacity Revenues	100	105	110	315
Total Revenue	600	630	662	1,892
Cost of Generating Energy	350	368	386	1,103
Other Operating Costs	50	53	55	158
Total Operating Costs	400	420	441	1,261
Operating Profit	200	210	221	631

Suppose the restructuring rules required each utility to build or purchase sufficient generation to cover its own loads plus planning reserve requirement. The amount required to purchase this generator from its owner should be less than $631 million, as the owner prefers a dollar now to a dollar to be realized two years from now. Assume, again for the purposes of illustration, that the owner of the generator expects to be able to realize ten percent per year in returns on if it invests the revenues it receives in exchange for selling the generator. Then it would be willing to accept 90 cents in Year 1 to give up a dollar in operating profit that it expects to earn in Year 2, and 81 cents in Year 1 to give up a dollar in operating profit in Year 3, so it would be willing to sell this generator for $573 million (payable in Year 1), as shown by the table below.

Anticipated Discounted Operating Profit Over Hypothetical Generator's Lifespan

	Year 1	Year 2	Year 3	Total
Operating Profit	200	210	221	631
Discounted Operating Profit	200	191	182	573

The $573 million cost of purchasing this generator could be recouped in many different ways, over many different time periods, but the most natural assumption is to assume this cost would be recouped over the time that the generator is expected to operate, since recouping it over a shorter time period would mean that consumers in earlier years were subsidizing consumers of energy in later years, and recouping it over a longer time period would have the reverse implication. One way of collecting this cost from end-use customers over the generator's lifespan would be to collect $200 million in Year 1—which was the operating profit the generator expected to earn in that year—while financing the remaining $373 million purchase price. One year later, the amount to be paid off would have increased by 10 percent, from $373 million to $410 million. If $210 million (which was the operating profit the generator expected to earn in Year 2) is recovered from end-use customers in Year 2, that leaves $200 million to be financed. In Year 3, that amount will have grown from $200 million to $221 million, which would be recovered from end-use customers in that year. This is illustrated in the table below.

Schedule for Recovering Purchase Cost from End-Use Customers

	Year 1	Year 2	Year 3
Total Amount to Recover	573	410	221
Recovered from End-Use Customers	200	210	221
Amount to Be Financed	373	200	—

The important thing to note is that the total amounts that end-use customers pay in each year are the same, regardless of whether customers pay the energy and capacity market payments or the generator is purchased and customers pay its operating costs plus the purchase and financing costs, as the table below illustrates. If the generator is not purchased, end-use customers will pay the generator for energy and capacity. If the generator is purchased, end use customers will pay the generator's operating costs. The difference between the generator's energy and capacity revenues and its operating costs is its operating profits, so end-use customers would not have to pay the operating profits. But these savings must be offset against the cost of purchasing the generator. Since the value of the generator to its owner is the value of the operating profits it is expected to produce, the cost of purchasing it is the value of those operating profits. As a result, the cost of purchasing the generator offsets operating profits exactly.

Comparison of Payments by End-Use Customers

	Year 1	Year 2	Year 3
Energy Payments	500	525	551
Capacity Payments	100	105	110
Payments by End-Use Customers if Generator is Not Purchased	600	630	662
Cost of Generating Energy	350	368	386
Other Operating Costs	50	53	55
Cost of Purchasing Generator	200	210	221
Payments by End-Use Customers if Generator is Purchased	600	630	662

(Endnotes)

1 John D. Chandley is a Principal at LECG, LLC. 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. Preparation of this paper was supported by the COMPETE Coalition. 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, Barclays, Brazil Power Exchange Administrator (ASMAE), British National Grid Company, California Independent Energy Producers Association, California Independent System Operator, Calpine Corporation, Canadian Imperial Bank of Commerce, Centerpoint Energy, Central Maine Power Company, Chubu Electric Power Company, Citigroup, Comision Reguladora De Energia (CRE, Mexico), Commonwealth Edison Company, Conectiv, Constellation Power Source, Coral Power, Credit First Suisse Boston, DC Energy, Detroit Edison Company, Deutsche Bank, 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), Exelon, GPU PowerNet Pty Ltd., GWF Energy, Independent Energy Producers Assn, ISO New England, Luz del Sur, Maine Public Advocate, Maine Public Utilities Commission, Merrill Lynch, Midwest ISO, Mirant Corporation, JP Morgan, 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, PSEG Companies, Reliant Energy, Rhode Island Public Utilities Commission, San Diego Gas & Electric Corporation, Sempra Energy, SPP, Texas Genco, Texas Utilities Co, Tokyo Electric Power Company, Toronto Dominion Bank, 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 authors. (Related papers can be found on the web at www.whogan.com

Appendices D and F were prepared by Michael D. Cadwalader, a Principal at LECG. Appendix E was prepared by Ventyx under the direction of James Sustman.