GETTING THE PRICES RIGHT IN PJM:
Analysis and Summary: April 1998 through March 1999
The First Anniversary of Full Locational Pricing[1]

April 2, 1999

William W. Hogan[2]

The connection between prices and operating decisions often receives cursory treatment in the electricity restructuring process. Market participants want flexibility and choice, but object to consistent pricing as too complex. This is a mistake, and produces only an illusion of simplicity. If customers have flexibility in the choice of generation, spot purchases, bilateral transactions, and so on--then prices matter and competitive prices should reflect marginal costs. In large part, control of operating decisions is moving from engineers motivated by principles of technical efficiency, to market participants motivated by prices and profits. This is a major purpose of electricity restructuring--to change the locus of such key decisions. If we want the market to be guided by prices, and we expect and intend for people to take these prices seriously, it becomes important to follow the usual advice to "get the prices right." The experience in the first year with a consistent market pricing system in PJM underscores the point and provides empirical heat to help dispel the fog of confusion covering one of the central problems in electricity markets: pricing to allocate use of scarce transmission capacity.

INTRODUCTION

In the United States, the move to a competitive electricity market with a consistent pricing system for allocating scarce transmission capacity entered a new phase beginning in April 1998 with the introduction of spot market locational pricing in the Pennsylvania-New Jersey-Maryland Interconnection (PJM). The new system includes a spot market coordinated by the Independent System Operator (ISO). The ISO accepts both bilateral schedules and voluntary bids of the market participants. Using these schedules and bids, the ISO finds an economic, security-constrained dispatch for power flows and the associated locational marginal cost prices. Even without transmission constraints, this coordinated and transparent spot market provides significant benefits. When the transmission system is constrained, the spot prices can differ substantially

---

[1] This note provides an update of an analysis with the same title for previous months beginning in April 1998. See also the Electricity Journal, September 1998.

[2] Lucius N. Littauer Professor of Public Policy and Administration, John F. Kennedy School of Government, Harvard University and Principal of the Law and Economics Consulting Group. 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 American National Power, British National Grid Company, GPU Inc. (and the Supporting Companies of PJM), GPU PowerNet Pty Ltd, Duquesne Light Company, Electricity Corporation of New Zealand, National Independent Energy Producers, New York Power Pool, New York Utilities Collaborative, New England Power Company, Niagara Mohawk Corporation, PJM Office of Interconnection, Putnam, Hayes & Bartlett, Inc., San Diego Gas & Electric Corporation, Transpower of New Zealand, Westbrook Power, Williams Energy Group, and Wisconsin Electric Power Company. The views presented here are not necessarily attributable to any of those mentioned, and any remaining errors are solely the responsibility of the author. (http://ksgwww.harvard.edu/people/whogan).

across locations. Sales through the spot market are at the locational prices. The transmission usage charge for bilateral transactions is the difference in the locational prices between origin and destination. An accompanying system of Fixed Transmission Rights (FTR) provides financial hedges between locations. These FTRs are the equivalent of perfectly tradeable firm transmission rights.[3]

IN MARKETS WITH CHOICES, PRICES MATTER

The new PJM locational pricing system was embraced after an experiment during 1997 with an alternative zonal pricing approach that proved to be fundamentally inconsistent with a competitive market and user flexibility.[4] The experiment made the point in a dramatic way, as discussed further in the appendix. The important issue is not the total cost of transmission congestion, which may be small on average if the system is used efficiently, and when the cost is often mistakenly dismissed as irrelevant. Rather, the point is the incentives at the margin when the system is constrained. In designing the rules for transmission access and pricing for a competitive market, it matters little how the rules perform when the system is unconstrained. The important question is how the rules deal with the market and participant choices when the system is constrained. The earlier zonal pricing system allowed market participants the flexibility to choose between bilateral transactions and spot purchases, but did not simultaneously present them with the costs of their choices. The circumstances created a false and artificial impression that savings of $10 per MWh or more could be achieved simply by converting a spot transaction into a bilateral schedule. Faced with this perverse pricing incentive, market participants responded naturally by scheduling more bilateral transactions than the transmission system could accommodate. In effect, using the wrong prices induced behavior which greatly increased the cost of congestion. Inevitably, in June 1997 the ISO had to intervene by restricting the market and constraining choice to preserve reliability. The PJM ISO was fully aware of the perverse incentives of zonal congestion pricing and the problems they created. But without the authority to change the pricing rules, the ISO had no alternative but to restrict the market.

Even if the total cost of congestion might be modest over a year, a gap of $10 per MWh between the true costs of transmission usage and what participants pay is more than sufficient to get the attention of market participants at the time when it matters most, when the system is constrained. Given the margins in this business, market participants will change their behavior for $1. And the changes in behavior can substantially affect system operations; in fact, the whole point of electricity restructuring is that changes in behavior can affect system operations and lead to different patterns of electricity use and investment.

---

[3] An FTR is the same as a transmission congestion contract (TCC). For further details, see William W. Hogan, "Independent System Operator: Pricing and Flexibility in a Competitive Electricity Market," Center for Business and Government, Harvard University, February 1998, available on the author's web site.

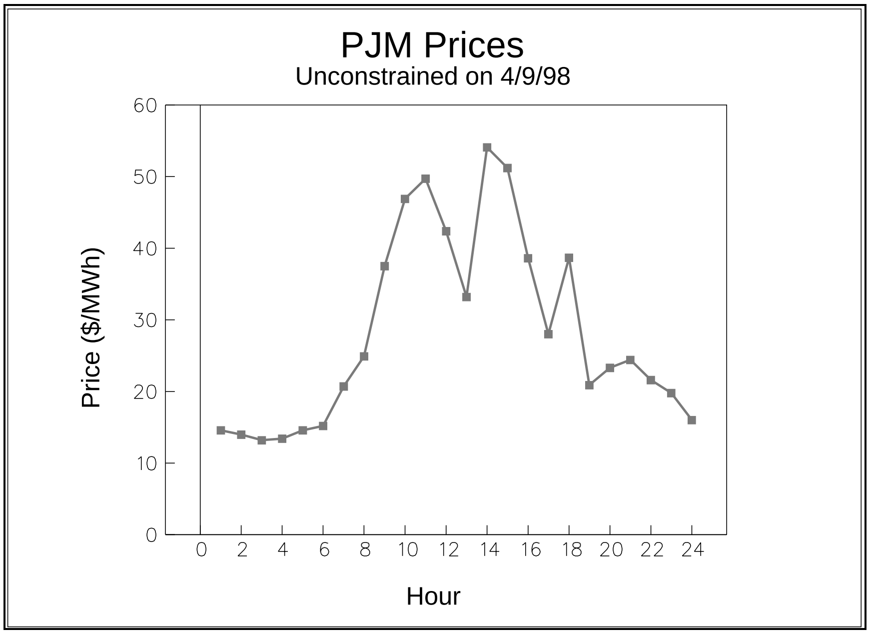
[4] Here the issue is pricing for transmission congestion. The recovery of embedded costs of transmission investment through access charges is a separate matter that is amenable to a zonal approach. Locational pricing has long been available in other markets, such as New Zealand. The PJM case is of interest because of its size, the sharp contrast of the debates, the experiment of trying two different pricing systems, and the availability of the data.

By contrast, the locational pricing system avoids this perverse incentive. By construction, the locational prices equal system marginal costs. Every generator would be producing at its short-run profit maximizing output, given the prices. The market equilibrium would support the necessary dispatch in the presence of the transmission constraints. Spot-market transactions and bilateral schedules would be compatible. Flexibility would be allowed and reliability maintained consistent with the choices of the market participants.

Faced with this reality, the Federal Energy Regulatory Commission (FERC) acted to approve the locational pricing system that became operational in PJM at the beginning of April of 1998. The developing experience with this full locational pricing of the use of scarce transmission capacity deserves close study by the Commission and all system operators.

TRANSMISSION CONGESTION AND LOCATIONAL PRICES

To put the problem in context, note that market clearing prices can vary substantially, even without transmission constraints.[5] The accompanying figure shows the prices in PJM over the day of April 9, 1998. This was not the most volatile day, and there were no transmission constraints during the day. However, the market clearing price varied from a low of $13 MWh to a high of $54. During June, the variation in unconstrained prices on the most expensive day increased by almost an order of magnitude to a high of $300, on June 26, 1998. As the summer continued, higher prices appeared even without internal transmission constraints in PJM. For example, on August 24 at 1400 and 1500 hours, the price reached the regulated maximum of $999. Clearly market participants must deal with substantial changes in prices, even without transmission congestion.[6]



---

5 The data used here were taken from the PJM web site at www.pjm.com.

6 Prices outside of PJM reached even higher levels, reportedly as high as $7,000 per MWh in the Midwest; Wall Street Journal, June 29, 1998, p. C1. The apparent market disequilibrium between PJM and the Midwest is an important issue, but that is another story. The emphasis here is on the price implications for allocation of scarce transmission capacity within PJM.

Although these were the early days, the new locational pricing mechanism worked as anticipated by the ISO and the supporters of the approach, but apparently not as anticipated by many who dismissed the importance of this issue. April and May are not typically highly constrained periods in PJM, and it would not have been impossible for the first days of locational pricing to have been boring. With no constraints, the locational prices, ignoring losses, would be identical at all locations. The cost of transmission between points--the difference in the locational prices--would have been zero. Nothing much might have happened until we approached the summer, when congestion would be more likely, as for the previous year in June.

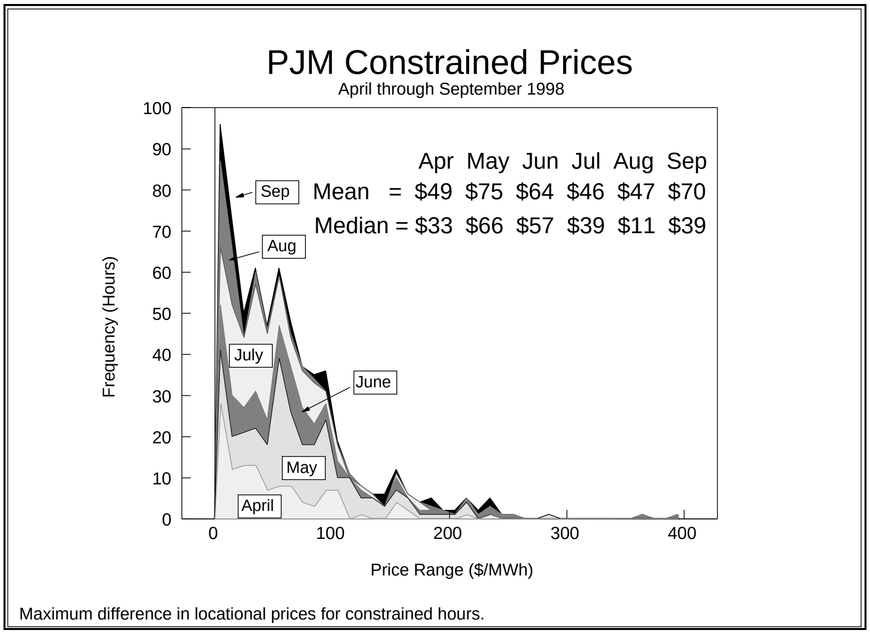
In the event, the results were not boring. The system experienced transmission constraints, locational prices separated, and the opportunity cost of transmission was quite large. The lowest locational prices were sometimes negative, reflecting the value of counterflow in the system where it would be cheaper to pay participants to take power at some locations and so relieve transmission constraints. The highest locational prices were larger than the marginal cost of the most expensive plant running, reflecting the need to simultaneously increase output from expensive plants and decrease output from cheap plants, just to meet an increment of load at a constrained location. Over all hours in April 1998, for example, the low price was -$45 at 1500 hours on April 18 at "JACK PS," and the highest price was $232 at 1100 hours on April 16 at "SADDLEBR," both locations being in the Public Service Gas & Electric territory. Over the first year with the locational pricing system, the maximum difference between the lowest and highest contemporaneous prices was $412, at 1100 hours on November 19, reflecting the difference between $322 at "SADDLEBR" and -$91 at "BELLVIL." The second highest difference was $399, at 2000 hours on August 26, reflecting the difference between $437 at "ESAYRE" and $38 at "NYPP-W." This maximum price separation reached the same level as in the relatively unconstrained month of March 1998 before the locational prices were charged, when users could ignore the cost of congestion.[7]

The contemporaneous difference in locational prices, which is the price of transmission usage, has been large quite often. It does not take much of a difference to change behavior when the reported trading margins may be as low as $1 per MWh. If we take the $1 per MWh standard as an arbitrary threshold to define a constrained period, the range of highest to lowest price across locations exceeded the threshold for 119 hours in April, or approximately 17% of the time. As shown in the accompanying figure, the frequency distribution of the price range in constrained hours is skewed, with a median hourly price range at $33 and a mean of $49 in April. When the system is constrained and the market incentives matter the most, the marginal costs of transmission can be large indeed.

The monthly data for May through September, covering the summer peak, reinforce this initial impression. In general, May saw both higher prices and more transmission congestion. The difference between the highest and lowest locational price in May exceeded the $1 threshold for 183 hours, or approximately 25% of the time. As shown in the accompanying figure, the

---

[7] On March 26, 1998, at 2200 hours, the difference between the highest to the lowest marginal cost was almost $400.

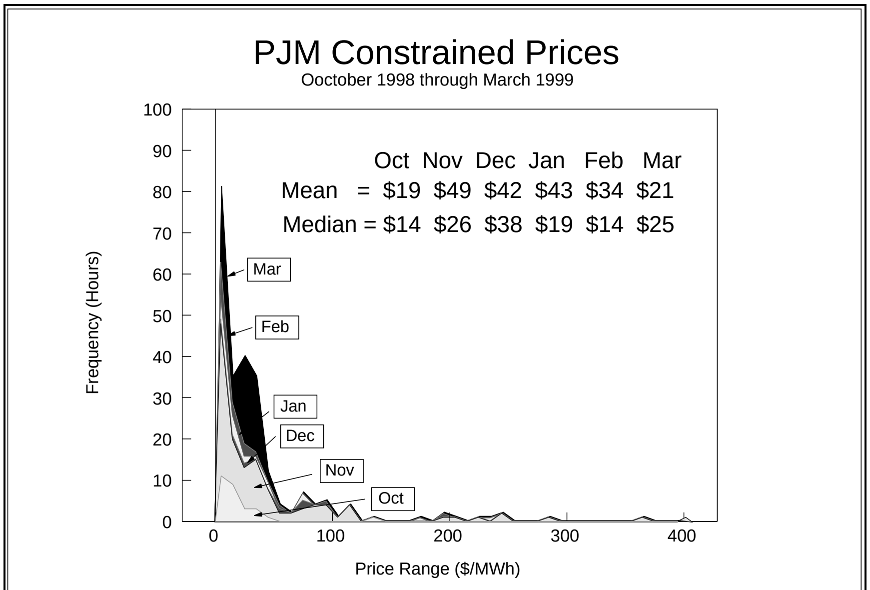


frequency distribution of this congestion price shifted to higher costs. In May, the median of the hourly price ranges doubled to $66 and the mean increased to $75. June was less constrained, exceeding the $1 threshold for 95 hours or 13% of the time. The June median of the hourly price ranges was $57 and the mean was $64. During July, constraints appeared more often, as in May, with 151 constrained hours or about 20% of the time. July saw a median hourly price range of $39 and a mean of $46. By contrast, August showed locational constraints only for 48 hours or 7% of the time. The median hourly price range in August was $11 and the mean was $47, reflecting a few hours when the difference between the lowest to the highest price reached almost $400. September was like August, with 46 constrained hours or 6% of the time. However, the average price of congestion was high in September, with a mean of $70 and a median of $39.

The experience of higher unconstrained prices and fewer constrained hours in June, August and September reminds us that the period of peak system load is not necessarily the time of greatest transmission congestion. Transmission congestion reflects an imbalance in the location of load and generation. At peak load, more generation comes on line and may relieve system congestion. In addition, the particular flow of power into the Midwest, reversing the usual direction, tended to unload the transmission constraints during the summer of 1998.

This record of continuing constraints was reinforced by the events in the following months from October 1998 to March 1999. After the heavier loading of the PJM summer, the winter months would be less constrained but the constraints did not disappear. As shown in the accompanying figure, the frequency diagram of price ranges showed that some significant constraints applied. November alone accounted for 105 of the 242 constrained hours over the period. The median price range for the constrained hours in November was $26 and the mean was $49. The corresponding median and mean price ranges of the other months for the hours that the system was constrained appear in the figure.[8]

[8] The numbers of constrained hours for October 1998 through March 1999 were 28, 105, 6, 20, 18, and 65, respectively.



Maximum difference in locational prices for constrained hours.

The evidence shows many things. For example, calculating and reporting the locational prices for each point on the grid are not especially complex tasks, at least for the system operator who has the necessary information available. The prices can be available every five minutes on the Internet. Faced with these prices, the market participants adjust their behavior, just as intended. The transition was not painless, especially for those who ignored many warnings and entered into "seller's choice" contracts that gave the seller the maximum theoretical financial advantage for relieving congestion. Presumably, this form of contract will disappear, or be properly priced in the future, and market participants will become more attuned to the use of fixed transmission rights to hedge much of the cost of congestion. But market participants who rely on the spot market, and are not prepared to pay for congestion hedges that fix the cost of transmission in advance, will see price signals that align their incentives with the reality of system operations.

Full locational pricing is fully compatible with a trading system built on a hub-and-spoke framework. The hub becomes a common trading point, and the cost of moving to and from the hub, along the spokes, is just the difference in the locational prices.[9] If the nodal prices are available from the ISO, market participants can define their own hubs. In the PJM case, however, market participants asked the ISO to handle the accounting to create several hubs, of which the western hub has so far developed as the preferred trading point.

The full market response to all these changes is not known because the data are not all in the public domain. However, one information source is a sampling of trader activity reported in the Wall Street Journal.[10] According to these data, the immediate response of the market was to reduce reported spot trading in April of 1998. However, by mid-May of 1998 reported transactions had returned to volumes comparable to those seen just before the new locational pricing system went into effect. Subsequently, and reversing its earlier objection that the nodal pricing market would not be sufficiently liquid, in March of 1999 the New York

---

9 See the appendix for a further discussion.

10 For example, see "DJ Electricity Price Indexes," Wall Street Journal, June 3, 1998, p. C19.

Mercantile Exchange launched a new futures contract to capitalize on the highly liquid trading market that had developed at the PJM western hub. Apparently inadequate liquidity was not a problem. Even further, the spot and forward markets at the western hub were reported to be so liquid that the futures contract might not be able to compete.[11] Although market liquidity is often vaguely defined and seen only in the eye of the beholder, as reported by these sources the market appears to have adjusted to the new environment within a framework that supports transactions with consistent prices.

The operational problems experienced by the ISO in the year before full locational pricing, where profit driven market participants undermined reliability, did not appear in the year after adoption of full locational pricing. Locational pricing presents profit driven market participants with the right incentives consistent with the true opportunity costs. This same pricing system was applied by PJM for managing inter-regional transmission loading relief. With full locational pricing, the prices reinforce reliability. In addition, the anecdotal evidence suggests that investments in new generation and transmission were being considered with careful attention to the effects of system congestion, just as intended.

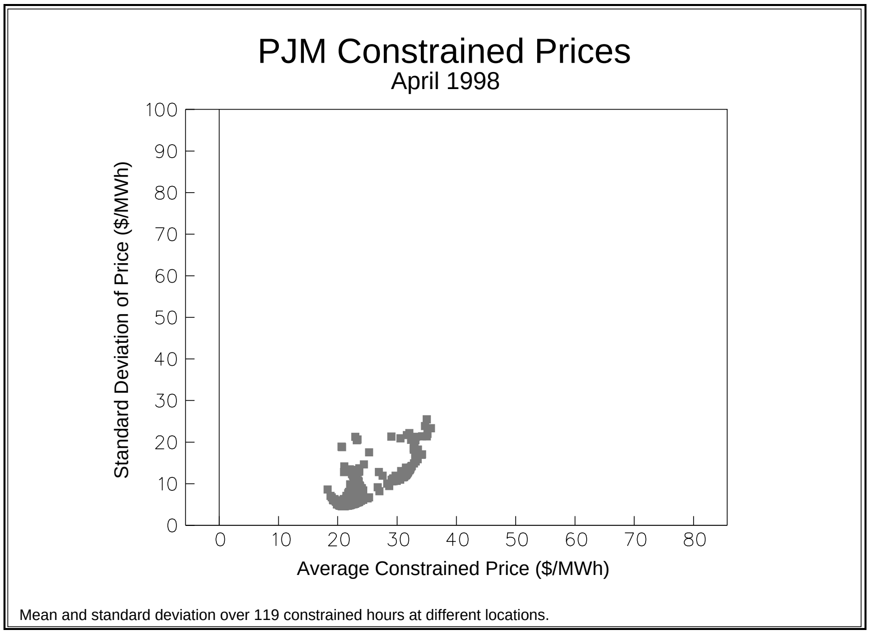
In the first year, generators within PJM's boundaries faced bidding constraints intended to impose competitive behavior. Generators outside of PJM selling into the market had no bidding restrictions, and often set the market-clearing locational prices. Based on a conclusion that there is no significant market power within PJM, at the start of the second year after application of full locational pricing, generators inside PJM also received FERC authorization to go to full market-based bidding. The experience with the first year and the success of PJM locational pricing will provide one useful benchmark for comparing the performance of the market without bidding constraints. As discussed in the appendix, if there is market power in PJM, full locational pricing should help mitigate that power.

FULL LOCATIONAL PRICING IS THE TRULY SIMPLE APPROACH

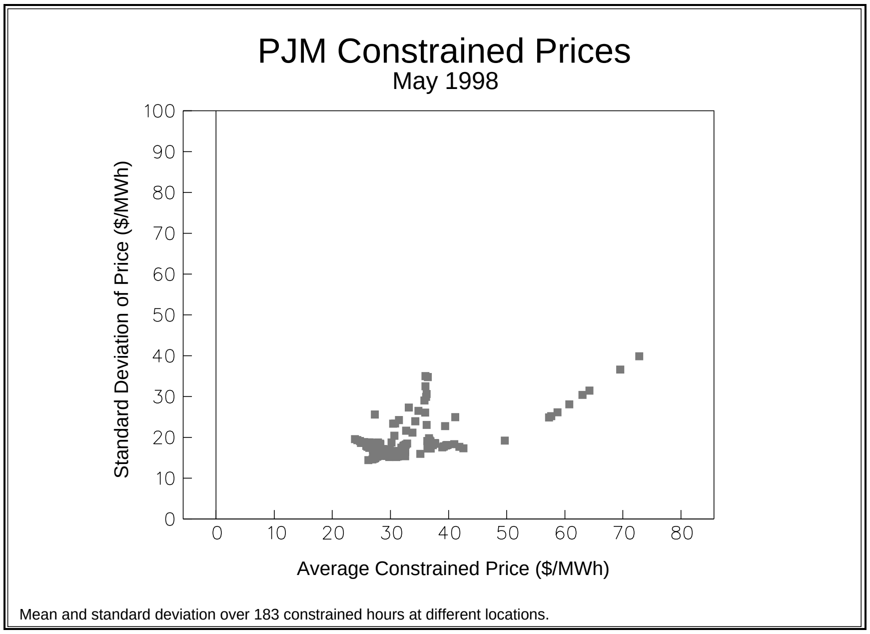
What about aggregating PJM into a few zones, if not just one? The PJM ISO is providing prices for approximately 2000 locations. This is a convenient way to represent the information, because it is how the data are organized for actual system operations. However, some of these locations are really just multiple units at the same point on the grid, and would necessarily have the same prices in most circumstances. For other points on the grid, the zonal argument is that the locational differences would be minor, and could be represented by a relatively few zones. This view has been subjected to a test over the first year of operation.

The period April through at least early June could have been relatively unconstrained, presenting a low hurdle for the zonal approach. The choice of the appropriate zones would not be an easy matter, and there is some ambiguity in clustering criteria. However, one simple way

[11] "The New York Mercantile Exchange will launch an electricity futures contract March 19 at the PJM western hub, one of the most liquid markets in the Eastern grid. ... The PJM hub already features an active and growing over-the-counter forwards market. A liquid hub can have a downside [for the futures contract] given that players are content trading in the OTC, said one Northeast broker." Power Markets Week, February 8, 1999, p. 14.



Mean and standard deviation over 119 constrained hours at different locations.



Mean and standard deviation over 183 constrained hours at different locations.

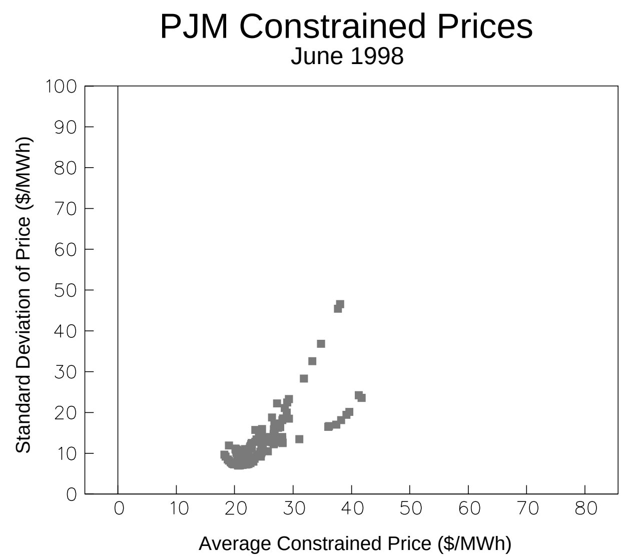
to summarize the data would be to examine both the average and the variation of prices at different locations during constrained hours. If two locations always have the same prices, then the two averages of prices over the period would be the same and the two standard deviations of the prices would be the same. These conditions would be necessary, but not sufficient, for the prices to be the same at the two locations. Hence, this straightforward calculation gives a lower bound on the number of different locations with sometimes unique prices.

The accompanying figure for April plots the data on average price and standard deviation of price across 119 constrained hours in April for all the locations reported by the PJM ISO. There are 2000 points in the graph, one for each location. Were it true that there were only a few zones, the graph would show a few clusters of locations where the average prices were the same and the standard deviations were the same. In fact, there is substantial dispersion.[12] After the first month of operation, there were 766 locations within PJM where the price points did not overlap and were different by this lower bound test. The corresponding data for May show a similar dispersion and higher costs of congestion. The accompanying figure plots the locational

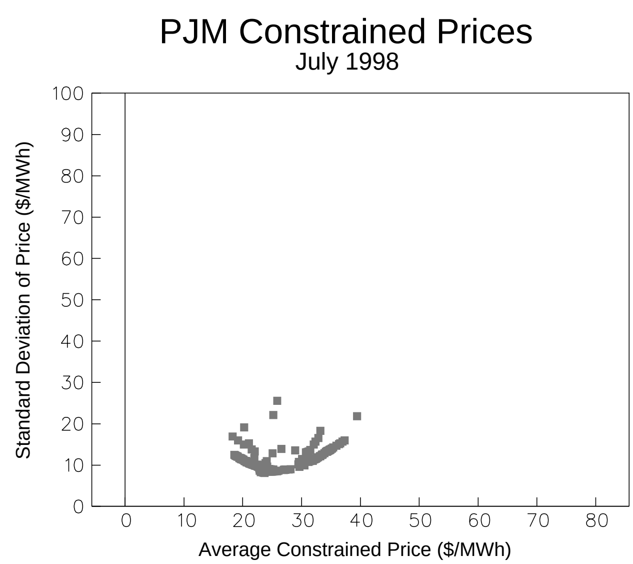
[12] The large scale in the graphic allows for consistent comparison with later months, when the price dispersion is even greater.

means and standard deviations for May, for which there were 789 different locations by this lower bound test. There were not as many limiting constraints, but even one constraint can produce a substantial range of prices.

In June there was a similar dispersion of prices. As shown in the corresponding figure for June, there were 689 locations with unique prices according to the lower bound test of having different averages or different standard deviations. The data for July in the subsequent figure show more congested hours and more locations, with a total of 785 different points, much as in May. By contrast, August revealed relatively fewer congested hours. Nonetheless, even in August the data indicate 693 different locations. September showed greater price dispersion with 825 different locations appearing in the price data.



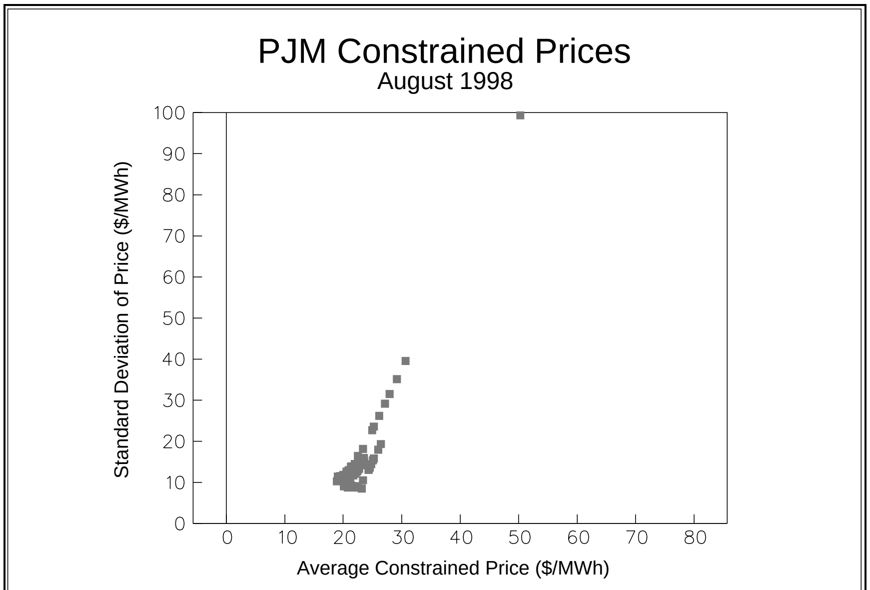
Mean and standard deviation over 95 constrained hours at different locations.



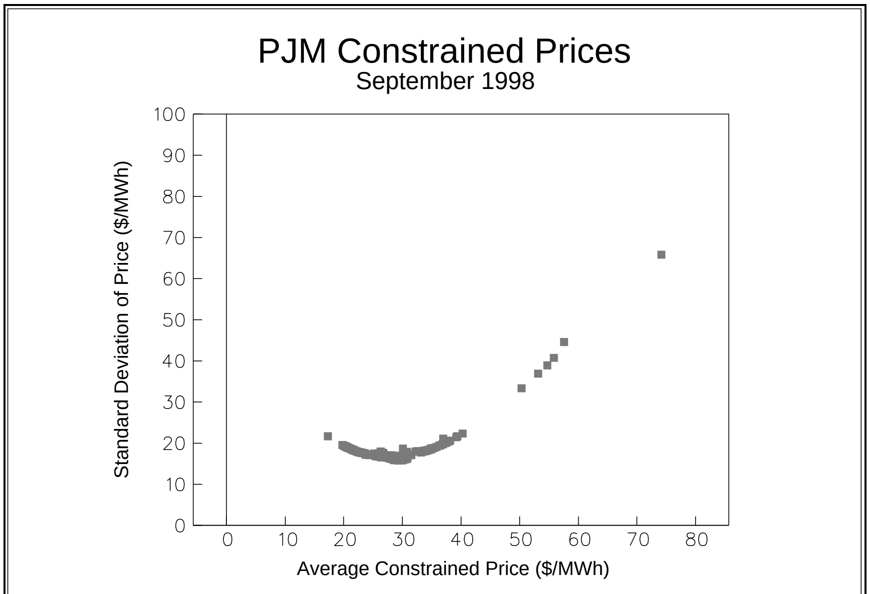
Mean and standard deviation over 151 constrained hours at different locations.

Furthermore, the differences were not the same across the months. The publicly available data for June through September cannot be merged easily with the previous months due to changes in the names and number of locations. However, using pooled data that calculates the averages and standard deviations for each location across both April and May, there were 821 different locations where either the average price or the standard deviation differed from other locations. Presumably as more experience develops and different constraints appear, this number will grow. Recall that in a sufficiently interconnected network, a single thermal limit on a transmission line could create different prices at every location, sometimes very different prices.

This is contrary to the intuition that arises from the misleading analogy to a simple radial transmission connection without any network interactions, where a single constraint results in only two prices. But the network interactions and the many different prices are quite real and no surprise to the system operators.



Mean and standard deviation over 48 constrained hours at different locations.



Mean and standard deviation over 46 constrained hours at different locations.

The criterion of no difference in prices may be too strict, and we might be willing to declare two locations as the same if they are close enough. The criterion should be set so that the maximum difference in prices across a zone would be small enough not to affect behavior. Note that this is not the same having a small contemporaneous standard deviation of the prices across the zone, which would leave many locations with the perverse incentives that create the complicated rules for market interventions.

For instance, consider the criterion proposed to select a zone such that the standard deviation of prices across the zone is less than 10% of the average prices.[13] Apparently the view

---

[13] This 10% standard deviation rule across a zone was adopted in Richard D. Tabors, "Transmission Pricing in PJM: Allowing the Economics of the Market to Work," Tabors Caramanis & Associates, February 24, 1999, p.17. Even with this relaxed rule, there are major regions in PJM where the zonal approach would fail, notably in northern New Jersey during the period April 1998 to September 1998. The analysis illustrates some of the problems with the zonal approach. For example, consider the constraint "shift factor" evaluation that suggests only a few zones

is that if this 10% criterion were met, *and market participants did not respond to the incentives*, cost shifting would be small. However, with average prices of $25 and normal distributions, this would mean that at least one third of the locations within a zone would see price differences of more than $5 per MWh, surely enough to cause the same problems that PJM experienced. The problem is that the participants *will* respond to the incentives and this will create the inevitable need for the ISO to adopt arbitrary rules to restrict the market. The goal should be avoid the perverse incentives, and not have to hope that market participants will ignore trading opportunities.

Defining the standard for "close enough" could be contentious, but it may be moot. If we accept the $1 per MWh threshold above for the maximum deviation of prices across a zone, and ask how many separate zones would be necessary to cover all the points in the figures, the answer is 94 zones in April, 83 zones in May, 75 zones in June, 57 zones in July, 52 zones in August, and 64 zones in September.[14]

Again, these are not the same zones in each month. If we pool the months of April and May and apply the consistent threshold of $0.50 per MWh average difference over two months, we find 132 different zones needed to capture the variability in locational prices, so far. This is many more than the few zones predicted, and there is no reason to believe that we are finished adding to the list.

For the months of October through March, the data are easier to combine across locations. The corresponding presentation of the average prices and their variation appears in the accompanying figure. For these six months, there are 842 different locations with distinct prices. Applying the same $1 threshold, these would aggregate into 61 zones. The number of zones that would be required in any given month, which range from a low of 38 in October to 71 in January.[15] Note that the number of zones required is not solely a function of the number of constrained hours. For example, in December, with only 6 constrained hours, the constraints were such that 64 zones would have been required to capture the variability in price. In short, a few zones wouldn't do, many would be required.

Furthermore, as is obvious, the effect of locational pricing for "natural" zones would be to produce the same price for all such locations. Hence, aggregating to zones where prices are almost the same would result in many zones and no apparent simplification with any

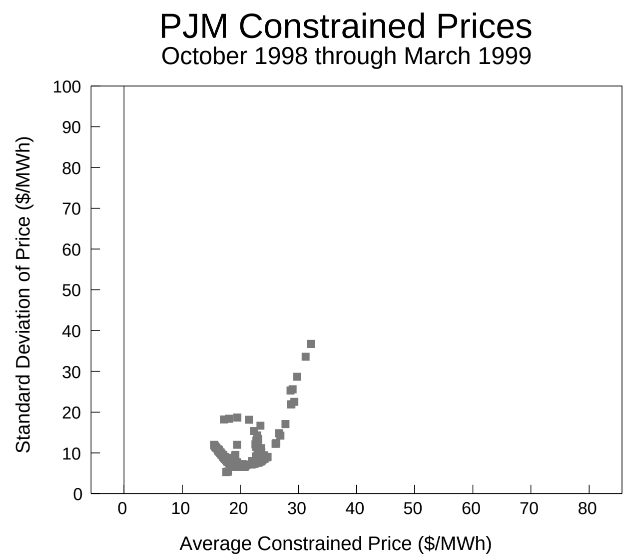
---

for PJM. The data show that the 28 constraints examined (Tabors, p. 31.) include *none* of the 43 constraints that actually limited transmission during the first six months of locational pricing, as reported on the PJM web page.

[14] Starting with separate zones for the ten service territories and sorting on the average and then the standard deviation. If we do not require the individual service territories to be separate zones, the number of distinct price zones is 60 in April. The corresponding figures for a $5 per MWh threshold would be 28 and 10 zones in April, with and without separating service territories.

[15] The numbers of zones for October 1998 through March 1999 were 38, 67, 64, 71, 62, and 65, respectively.

meaning. We might as well do the simpler thing of using the locational prices at each location. Aggregating across the many real locations to a few zones, by contrast, necessarily means combining like with unlike, thereby recreating in microcosm the perverse incentives of the failed experiment with a single zone, leading to a breakdown of the non-discriminatory market and administrative restrictions on choice.



Mean and standard deviation over 242 constrained hours at different locations.

The interaction between reliability (with its inescapable physical realities) and economics will limit the acceptable ISO access and pricing rules for allocating scarce transmission capacity. It would be desirable to offer market participants flexibility in their own decisions. A great deal of flexibility in combining a range of bilateral schedules and spot transactions would be possible. However, the more varied and flexible the options for the market participants, the more important it will be to get the prices right, meaning consistent with the marginal impacts on the system. The whole point of the turn to greater reliance on competition is that the market participants will respond to incentives. As we have seen, if prices don't provide the right incentives, consistent with the impacts on the system, the participants will respond in their own interests without concern for the system effects, and the ISO will be driven inexorably to intervene in the market and restrict choice. The locational pricing system with FTRs works in PJM, and a similar system received FERC approval for the New York ISO.[16] With locational pricing, participant incentives are aligned. Buyers and sellers can buy and sell as they choose through the spot market at the locational prices. Or they can schedule bilateral transactions and pay the difference in locational prices as the charge for transmission usage. The result is flexible, non-discriminatory, and compatible with the mandates of reliability.

CONCLUSION

The experience of the first year in PJM illustrates the importance of using, or at least reporting the real locational marginal costs. The network effects can be surprising for virtually everyone other than experienced system operators. Our intuition about these impacts and their market implications is poorly informed and often wrong. Much of the policy argument on this point is misinformed. However, only the ISO would have the information needed to calculate

[16] Federal Energy Regulatory Commission, New York ISO Ruling, Docket Nos. ER97-1523-000, OA97-470-000 and ER97-4234-000, January 27, 1999.

and post real locational prices, as in PJM. The computations are easy for a given dispatch, but only the ISO has all the information about the dispatch. Given the striking gap between the previous claims that congestion is insignificant and the observed reality of full locational pricing in the first real implementation in the United States, the Federal Energy Regulatory Commission should have a strong interest in prescribing that the real locational marginal costs--considering the real network interactions, and not just simplified zonal aggregations--be made available on a regular basis.

APPENDIX

THE NODAL-ZONAL DEBATE

Full locational pricing at every node in the network is a natural consequence of the basic economics of a competitive electricity market. However, it has been common around the world to assert, usually without apparent need for much further justification, that nodal pricing would be too complicated and aggregation into zones with socialization of the attendant costs would be simpler and solve all manner of problems. On first impression, the argument appears correct. On closer examination, however, we find the opposite to be true, once we consider the incentives created by aggregation combined with the flexibility allowed by market choices. But the debate continues.

For example, the original one-zone congestion pricing system proposed for the New England independent system operator (ISO) created inefficient incentives for locating new generation.[17] To counter these price incentives, the proposal imposed limiting conditions on new generation construction. Following the FERC rejection of the resulting barriers to entry for new generation in New England, there developed a debate over the preferred model for managing and pricing transmission congestion.[18] One zone was not enough, but perhaps a few would do? Or should New England go all the way to a nodal pricing system as in PJM?

> **Fact:** A single transmission constraint in an electric network can produce different prices at every node. Simply put, the different nodal prices arise because every location has a different effect on the constraint. This feature of electric networks is caused by the physics of parallel flows. Unfortunately, if you are not an electrical engineer, you probably have very bad intuition about the implications of this fact. You are not alone.
>
> **Fiction:** We could avoid the complications of dealing directly with nodal pricing by aggregating nodes with similar prices into a few zones. The result would provide a foundation for a simpler competitive market structure.
>
> There are many flaws in this seductive simplification argument.[19] In reality, the truly

---

[17] The use of zones for collecting transmission fixed charges is not the issue here. The focus is on managing transmission congestion. For a critique of the previously proposed one-zone congestion pricing system, see Peter Cramton and Robert Wilson, "A Review of ISO New England's Proposed Market Rules," Market Design, Inc., September 9, 1998.

[18] Federal Energy Regulatory Commission, New England Power Pool Ruling, Docket No. ER98-3853-000, October 29, 1998.

[19] William W. Hogan, "Nodes and Zones in Electricity Markets: Seeking Simplified Congestion Pricing," in Hung-po Chao and Hilliard G. Huntington (eds.), Defining Competitive Electricity Markets, Kluwer Academic Publishers, 1998, pp. 33-62. Steve Stoft, "Transmission Pricing in Zones: Simple or Complex?", The Electricity

simple system turns out to be a market that uses nodal pricing in conjunction with a bid-based, security-constrained, economic dispatch administered by an independent system operator. Purchases and sales in the balancing spot market would be at the nodal prices. Bilateral transactions would be charged for transmission congestion at the difference in the nodal prices at source and destination. Transmission congestion contracts would provide price certainty for those who pay in advance for these financial "firm" transmission rights up to the capacity of the grid. The system would be efficient and internally consistent.

The debate over transmission usage and the earlier experiment with a zonal transmission pricing approach in PJM provide a stark illustration of the difficulty and the challenge.[20] In March of 1997, the FERC approved an interim transmission access and pricing system to operate in conjunction with a real-time spot market coordinated through the PJM ISO. Faced with opposition to a full locational pricing and congestion charging mechanism for actual use of the system, the FERC endorsed the locational approach in principle but adopted temporarily an alternative model proposed by Philadelphia Electric Company (PECO) and others. The PECO approach minimized the importance of transmission congestion and rejected the locational pricing model as too complicated and unnecessary. Instead, the PECO model would treat the entire PJM system as a single zone.

In essence, the PECO model priced all transactions through the spot-market at the "unconstrained" price, based on a hypothetical dispatch. To the extent that the actual dispatch encountered transmission constraints, the PECO model would pay the more expensive generators to run and average these congestion costs over all users.

The model included two other notable features. First, in the face of transmission congestion, the generators that were constrained not to run would be paid nothing, even though they had bids below the "unconstrained" price. There was objection to adopting any system that depended on paying generators not to run, with the attendant discrimination and perverse incentive effects. Second, market participants had the option to schedule bilateral transactions separate from the bid-based economic dispatch of the ISO, with a separate payment for their share of the total congestion cost. This flexibility to use bilateral transactions or to participate in the coordinated spot market was a major design objective not to be abandoned.

This pricing system is representative of a zonal approach, and has much in common with zonal systems adopted elsewhere in the world.[21] However, should the system become constrained, the two exceptional features noted above would create a powerful and perverse

---

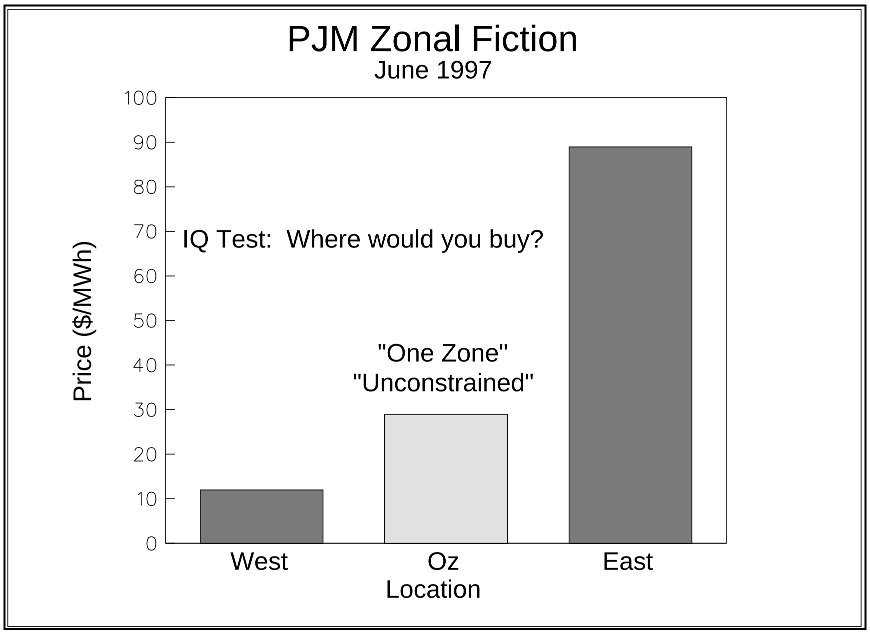
Journal, Vol. 10, No. 1, January/February 1997, pp. 24-31.

[20] For details, see William W. Hogan, "FERC Policy on Independent System Operators: Supplemental Comments," Federal Energy Regulatory Commission, Docket No. PL98-5-000, Washington DC, May 1, 1998.

[21] Here the issue is pricing for transmission congestion. The recovery of embedded costs of transmission investment through access charges is a separate matter that is amenable to zonal approach.

incentive. If there were no transmission constraints, there would be no transmission congestion and everything would work as with the locational pricing system. But when congestion appeared, everything would be different. The supporters of the zonal approach argued that the total cost of congestion would be small, summed over the year, and therefore any inefficiencies could be safely ignored.

Ignoring a difference between prices and marginal costs is a safe practice in a regulated world without flexibility and choice. The incentives don't matter and the small costs get lost in the larger system. It can work inside the closed black box. But the cost of ignoring a gap between prices and marginal costs in the world of choice can be large indeed. Witness the events when the PJM system became constrained, starting in June of 1997.



The data for a representative constrained dispatch found the marginal cost in eastern PJM at about $89 per MWh, when at the same time the marginal cost in the west was $12 per Mwh. At the same time, the "unconstrained" price for the "One Zone" (Oz) was approximately $29 per MWh. The incentives were clear. A customer could buy from the spot-market dispatch at $29, or it could arrange a bilateral transaction with a constrained-off generator in the west at a price closer to $12.[22] The small average congestion cost would be the same either way, and would not affect the choice. The choice, therefore, presented a low-level IQ test.

Faced with these incentives, constrained-off generators passed the IQ test. They quickly arranged bilateral transactions and scheduled their power for delivery, thereby exceeding the transmission limits. This, in turn, required the ISO to constrain the output from some other generator, who would then follow the same direct path to a bilateral schedule rather than sit idle and collect nothing. Soon the ISO had no more controllable generating units with which to manage the transmission constraints. Unable to fix the perverse pricing incentives, the ISO resorted to administrative mechanisms to prohibit bilateral transactions or declare a "minimum"

[22] Power Markets Week, September 1, 1997, p. 13.

generation emergency during the peak generation period. In effect, while restructuring to facilitate a market, the unintended consequences of superficially simple pricing spawned administrative rules to prohibit the market from responding to the price incentives when they mattered most. Shackled with inconsistent pricing rules, the ISO had to resort to direct preemption of market choices.

The point was made in a dramatic way. The important issue is not the total cost of congestion, which may be small on average. The point is the incentives at the margin when the system is constrained. In designing the rules for transmission pricing and access for a competitive market, it matters little what the rules are for periods when the system is unconstrained. The important question is how the rules deal with the market when the system is constrained. Even if the total cost of congestion might be modest over the year, the gap between $29 and $12, or $89 and $12, is more than sufficient to get the attention of market participants. Given the margins in this business, they will change their behavior for $1. And the changes in behavior can substantially affect system operations; in fact, the whole point of electricity restructuring is that changes in behavior can affect system operations and lead to different patterns of electricity use and investment.

In the locational pricing system, the perverse incentives would not arise. Given the same facts as above, the locational prices would equal the marginal costs. Those customers purchasing power from the spot market in the east would have seen $89 as the price. True, they could have arranged a bilateral transaction with a generator in the west, paying $12 for the energy. But they would then face a transmission charge of ($77=$89-$12), making them indifferent at the margin, just as intended. Likewise, customers in the west would pay $12 and have no incentive to change. Every generator would be producing at its short-run profit maximizing output, given the prices. The market equilibrium would support the necessary dispatch in the presence of the transmission constraints. Spot-market transactions and bilateral schedules would be compatible. Flexibility would be allowed and reliability maintained consistent with the choices of the market participants.

The PJM ISO was fully aware of the perverse incentives of zonal congestion pricing and the problems they created, but without the authority to change the pricing rules it had no alternative but to restrict the market. Faced with this reality, the FERC acted to approve the locational pricing system that became operational in PJM at the beginning of April of 1998. The developing experience should be better understood to avoid the pitfalls of the complicated zonal "simplification."

The most obvious flaw in the zonal argument is in its very definition. If the nodal prices are not materially different, then there is no need to aggregate into zones. The nodal prices would already be simple to use in the market. Apparently the move to aggregate nodes into zones is really an effort to treat fundamentally different locations as though they were the same.

If market participants had no market choices, then there would not be much effect of such zonal aggregation, other than a certain amount of cost shifting. But a central objective of market restructuring is to give market participants as many choices as possible. Further, we expect that market participants will respond to profit incentives. If we don't “get the prices right,” the market actors will respond to the prices and make choices that at best would significantly raise costs and at worst would dangerously compromise reliability.

There are many ways that things can go wrong. The PJM 1997 experiment with a zonal pricing system collapsed as soon as the system became constrained. New England faced the problem of zonal incentives to build new generation in locations that would exacerbate constraints. In Australia, a zonal pricing system has complicated the ability to offer transmission rights that match the real capability of the system. The same physical laws that govern nodal pricing make it impossible to define a complete set of zonal transmission rights, or to guarantee that generators can always participate in the market. By contrast, with a nodal pricing system such as in New Zealand, point-to-point transmission contracts could be defined in a natural way that is inherently consistent with the pricing regime and the real capacity of the grid. Just such a system of transmission contracts is operational with nodal pricing in PJM, and plays a central role in the recently FERC-approved nodal pricing system for the New York ISO.[23]

Similar experiences revealing the hidden complications of zonal pricing can be found from England to California, with different ad hoc rules applied to create more and more complex structures to fight against the choices of market participants confronted with administrative prices. Often it is hard to recognize the connections among the isolated ad hoc decisions, or to see the root of the problem in offering choices without getting the prices right. New England is not alone, but it could learn from the mistakes of others.

A great advantage of a nodal pricing system is that it creates incentives that are "self-policing."[24] Competitive market generators and loads could bid into a spot market and find that the economic dispatch result created a solution that would meet the no arbitrage condition. The attendant locational prices would be such that, as shown in the figure, every generator with a bid less than the price at its location would be running, and generators who had bid more than the market clearing price would not be running. There would be no artificial incentive to deviate from the market equilibrium solution.

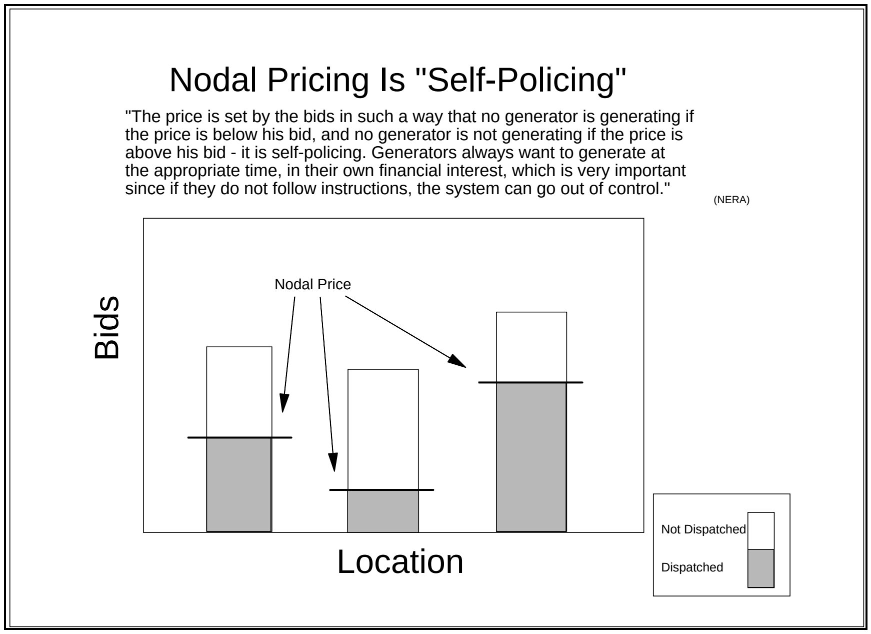
By contrast, a zonal pricing system must by definition create conflicting incentives. Set aside the complications about how to determine the zonal price. Whatever the rule, the zone will by definition have a single price. For generators who have bid less than the zonal price but who cannot run because of transmission constraints, there will be a strong incentive to leave the spot market and schedule a bilateral transaction, just as in the PJM experience. By contrast, for

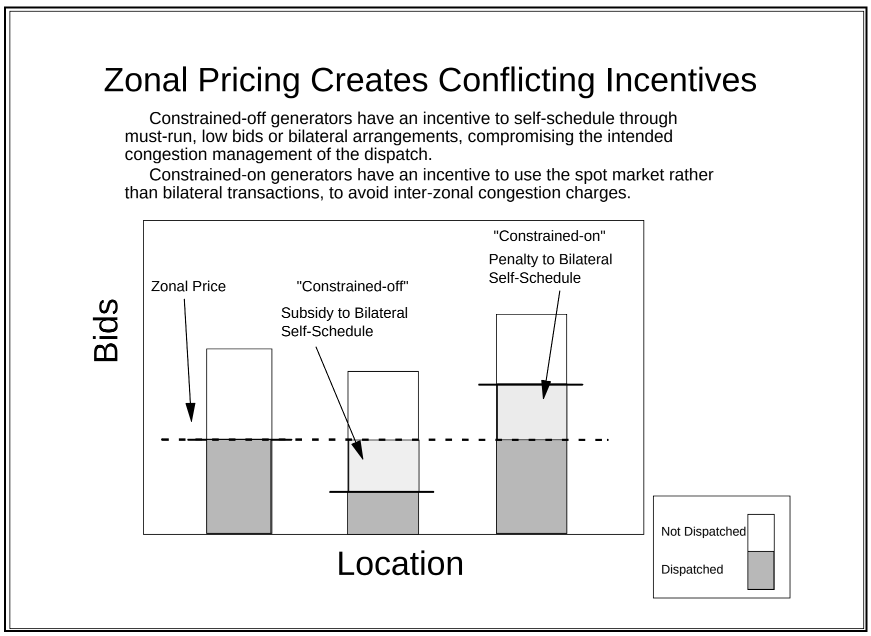
---

[23] The New York system proposed an aggregation for final loads, at least as a transition mechanism to deal with metering problems. The FERC explicitly objected to this proposed zonal aggregation for loads, calling for any metering changes needed to apply a full nodal pricing system.

[24] The term comes from National Economic Research Associates (NERA).

generators who have bid more than the zonal price but must run, again due to transmission constraints, there is a bias in favor of the spot market and these generators could not participate in the bilateral market. Furthermore, in a multi-settlement system these constrained-on generators would have a strong incentive to generate below their commitment and make up the imbalances at the lower zonal price. The details would depend on the particular rules and market setting, but the general conclusion holds that the incentives created by aggregation would greatly complicate operation of the market.





The real impact of zonal pricing is to create more administrative rules, poorer incentives for investment, demands to pay generators not to generate power, and proposals to "socialize" the higher costs by using the taxing power of the ISO. This is not the way of a market. It creates more problems than it solves.

Furthermore, there really are no major complications in implementing full locational pricing, once you look closely at what is required. The fears are misplaced, and the benefits are real and substantial.

Consider some of the arguments:

Is transmission congestion a small problem? No. On close inspection, few systems are really unconstrained. And when the constraints bind, the effects can be very important. Transmission congestion costs can easily exceed generation costs at the margin. The incentive effects have major commercial implications. Before the fact, zonal advocates argued that PJM would consist of at most only a few zones and constraints would be rare. In the event, the PJM market environment saw significant constraints in 15-20% of the hours in the first six months of operation, often under conditions that would be important in capturing commercial profits. And the number of zones needed to respect the commercially significant price differences in PJM has been far more than the promised few. Zonal pricing is not simple.

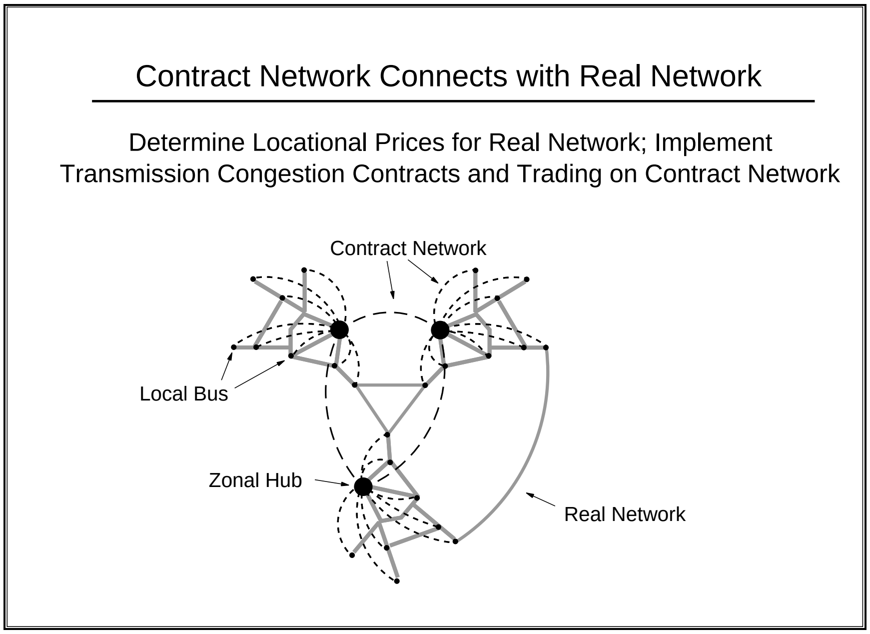
Are nodal prices produced by a black box? No. Given the dispatch, the prices are easy to calculate, explain and audit. There is ample operational experience to dispel the notion that nodal pricing is too hard. The engineers know how to do it, and have been doing it for years. The nodal prices have always been there; we just haven't used them for market transactions.

Doesn't nodal pricing preclude transmission price certainty? No. To be sure, we do not know in advance what the spot price will be, just like in any market. But those who want transmission price certainty can acquire transmission congestion contracts. And those who do not want to pay in advance for price certainty, and want to rely on the spot market, cannot socialize the costs by making others pay for the congestion they create.

The list of misconceptions about the pricing debate is longer. Given the fundamental underlying differences in marginal costs, it is not so easy to define the zonal price. It is not an easy matter to set or later change the zonal boundaries. The inherent averaging of zonal prices tends to remove incentives for energy efficiency or distributed generation. And so on. Perhaps the most oft-repeated point of confusion has to do with the impact of zonal aggregation on the ability to exercise market power.

Won't zonal aggregation mitigate market power? No. Real elimination of the physical constraints would help reduce market power where it exists. But administrative aggregation into zones simultaneously increases and obscures market power. Under the zonal approach, favored generators could take advantage of the real physical constraints, but their higher charges would be socialized and averaged over all system users, hidden from view. Market power can be a problem, but the problem is neither created by locational pricing nor resolved by zonal aggregation.

More recently, there has been the argument that the market needs zonal aggregation to support simplified trading. Won’t nodal pricing destroy market liquidity? Apparently not. The issue is more complicated in the case of electricity than other markets because the open access spot market creates its own form of liquidity that may obviate the need for vigorous trading of bilateral contracts. However, even for trading in contracts, this is largely an empirical question. The early returns from PJM suggest that the predictions of no liquidity in the market have been quite wrong at the western hub. Of course, no market has complete liquidity at every location. Typically there are trading hubs and the liquidity is found at the trading hubs, not at every location. This use of trading hubs is valuable and fully functional under nodal pricing.



Isn't a simpler system possible? Yes. The nodal pricing approach is completely consistent with a hub-and-spoke description of the market. One or more trading hubs can be established. The transmission charge for moving along the spokes, to and from the hub, is just the difference in the locational price and the hub price. This would be in contrast to a price of zero along the spoke implicit in a zonal aggregation. A hub can be selected as a single location or as a fixed portfolio of locations. This approach captures most of the intended simplification of the zonal model without embracing the hidden defects of aggregation. There is no mystery here. The hub-and-spoke approach is the system now working in PJM.

There is nothing unusual in nodal pricing. It is the natural system that falls out of an analysis of competitive market marginal-cost pricing principles in the context of the physics of the electric network. Nodal pricing does not solve all problems in electric market design, but it turns out to be important in dealing with some of the otherwise most intractable problems created by the special nature of the electric grid and the complex network interactions. Furthermore, if despite all the evidence zonal aggregation is commercially attractive, it presents a business opportunity for its advocates and need not be imposed by the ISO. But practical experience and theoretical analysis both support the conclusion that for the independent system operator, nodal pricing is the simplest system that actually works in the context of a market with choices and flexibility.

Get the prices right, and it is much easier to rely on the market.

--end--