MARGINAL COST PRICING HAS LITTLE PLACE IN ELECTRICITY REGULATION

Jeffrey M. Jakubiak, J.D. †

TROUTMAN SANDERS LLP
ATTORNEYS AT LAW
www.troutmansanders.com

401 9th Street NW, Suite 1000, Washington, D.C. 20004
(t) 202 274 2892 – (f) 202 654 5613 – (e) jeffrey.jakubiak@troutmansanders.com

1. Introduction.

In economic theory, the perfect competition model postulates that producers should be willing to sell a unit of a given product at or above the marginal cost of producing the unit. The Federal Energy Regulatory Commission (FERC) – charged with ensuring that rates for wholesale electricity are "just and reasonable" – seized on this postulate in investigating the California "energy crisis" of 2000-2001, and adopted the position that a generator should be willing to sell electricity at the generating unit's marginal cost of production. The FERC defined this cost as equal to the marginal heat rate of the generating unit multiplied by the cost of fuel, plus an adder for operations and maintenance. On this basis, the FERC ordered that electricity supply prices in California's markets be reset retroactively to reflect the marginal cost of the least efficient generating unit dispatched during a given period of time.

As the litigation subsequent to FERC's orders have shown, however, marginal cost pricing cannot be applied easily to electricity markets. The reality is that complexities inherent in electricity production – particularly in bid-based markets such as California's – make it virtually impossible to determine the "true" marginal cost of a unit of electricity. That is, the extensive complexities of electricity production in the real world preclude any simple formulaic approximation of marginal cost from yielding a reliable guide to pricing. One obvious complexity is the fact that not only do generating units incur large start-up costs, but unit operators in a market like California's rarely know the period of time over which their units will run.

As a result of the disconnect between simple formulaic approaches and the complexities of electricity production, the regulated suppression of market prices to alleged marginal cost levels – even if applied only to the marginal generating unit – could have adverse consequences for the market and consumers. For example, if received prices are forced below long-term average cost, supply will exit the market and place upward pressure on prices. If supply is unable to meet demand fully, blackouts will be unavoidable.

2. The FERC's Remedy for the California Energy Crisis

Although the start of California's energy crisis is difficult to pinpoint, it is generally accepted that the crisis began in May 2000 and continued through the spring of 2001. Shortly after prices settled down in California's wholesale markets in mid-2001, the FERC directed sellers in California's centrally-operated wholesale electricity markets – those operated by the California Independent System Operator (Cal-ISO) and the California Power Exchange (Cal-PX) – to refund a sizable portion of the revenues they received during the crisis period.[1] The "refund formula" – as it has come to be called – was based on a principle espoused by the FERC and others that, as would be appropriate in a perfectly competitive market, the price for electricity in California should be the marginal production cost of the least efficient unit dispatched by the Cal-ISO during a given period to serve load.[2]

Specifically, the FERC directed that the "clearing prices" in the Cal-ISO and Cal-PX markets be reduced to the estimated marginal production cost of the least efficient unit dispatched by the Cal-ISO during a given 10 minute period. Further, this so-called "mitigated market clearing price" (MMCP) was to be calculated by the following formula:

MMCP = ([Heat Rate] x [Gas Price]) + $6.00,

Where:

> Heat Rate = the MMBtu required by the least efficient generating unit dispatched by the Cal-ISO in order to produce an additional MW of electricity;
>
> Gas Price = the daily spot price for natural gas in California; and
>
> $6.00 was an "adder" for operations and maintenance (O&M) costs.[3]

While most nuances of this formula are not relevant for this paper, one does bear particular attention – the Heat Rate variable was to be the "incremental" heat rate of the generating unit, not the "average" heat rate. It is important to note that, first, the difference between these two rates can be significant. Whereas the average heat rate includes a portion of the minimum load fuel requirement for the unit (allocated across all MW being produced), the incremental heat rate does not. For most generating units in California operating during the crisis, the incremental heat rate was below the average heat rate for any operating level.[4] Second, neither curve includes the costs of the fuel required to start up a generating unit and bring it to its minimum operating level.

3. Electricity Markets are Not Perfectly Competitive

The FERC's intent with its MMCP formula was to approximate what a competitive market price would have been in the Cal-ISO and Cal-PX markets during the California energy crisis. The formula, however, is likely to yield a significantly lower market "clearing price" than what one could reasonably expect in a real-world competitive electricity market. As a result, the formula, when used as the basis for calculating refunds demanded of electricity sellers, will overstate refunds, and require that sellers pay back more revenues than would be intended under the FERC's stated objective of approximating what would have occurred in a real-world competitive market.[5]

The FERC's belief that its MMCP formula should seek to approximate marginal cost pricing in the Cal-ISO and Cal-PX markets is rooted in the principle that sellers should be willing to sell a unit of a product at any price above the marginal cost of production of that unit. The shortcoming of this approach is that it is only in a perfectly competitive market will sellers facing equal production costs be forced through competitive forces to reduce their offering prices to marginal cost levels. In other less perfect markets, sellers will recognize that they possess some modicum of market power and will price their product accordingly. The greater the market power, the greater will be the seller's margin above its marginal costs of production.

The FERC and market scholars have come a long way in recognizing that electricity production is not a natural monopoly as the crisscross of electric wires across the United States permits competition among independent generators to serve electric load. However, this fact does not permit the FERC and others to conclude that the nation's electricity markets have the potential to be perfectly competitive. The reality is that, given present technology, these markets at present do not approximate the perfect competition model, but rather are natural oligopolies.[6] In large part, it is the capital-intensive nature of electricity production that makes generation a natural oligopoly – and it will remain such until capital requirements fall markedly.

Electric generation is a natural oligopoly because of significant economies of scale, resulting in high entry costs. For example, a new 1,100 MW highly efficient gas-fired steam generating facility may cost upwards of $660 million.[7] With these characteristics, it is clear that electricity supply markets do not meet the perfect competition model requirements that are the basis for marginal cost pricing. Perfect competition is a case where there are an infinite number of identical competitors making a uniform product and where entry and exit costs are zero or close to zero.[8] Neither characteristic applies in the case of electric generation.

The high entry costs in generation derive from the natural economies of scale associated with electric generation – at least insofar as capital construction is concerned.[9] Thus, although technologies suitable for distributed generation are developing at a rapid rate, at present, the per MW capital costs for electricity generation are at their lowest with facility capacities in the hundreds of MW and not with facilities with small capacities. In sum, until there are fundamental changes in the technology of electricity generation, the industry will remain highly capital intensive and thus will be characterized by *natural* barriers to entry. Such barriers will keep electricity generation a natural oligopoly, and one in which prices cannot and should not be expected to fall to the level of the marginal costs of production.

4. What is the Marginal Cost?

In the simple "widget-based" world of perfect competition, a supplier produces widgets by expanding production along a smoothly continuous marginal cost curve. In some real world industries, such costs of production may be relatively easy to determine. In electricity production, however, determining marginal costs are not so simple. The FERC itself appears to recognize these facts, referring in an investigative order to "full incremental costs".[10] Still, the FERC did not answer the key question: what is included in these full incremental costs?

Certain characteristics of the electric industry differentiate it from most any other industry. First, since electricity cannot be stored, instantaneous supply must equal instantaneous demand (plus transmission or similar losses).[11] As a result, a MW of electricity produced now is not really the same as a MW produced later. In a very real sense, the two are different products facing two different sets of market conditions. Second, most generating units cannot produce only 0.1 MW (or some similarly marginal amount above zero). Rather, most units have some non-zero minimum operating limit. Third, generating units often may not be able to smoothly increase production levels through their maximum limits. In other words, their supply curve may be "lumpy" or discontinuous. Fourth, most generating units cannot be turned on or off in an instant, but instead have non-zero ramp rates. Thus, units may take anywhere from several minutes to an entire day (in the case of large coal-fired units) to move from being shutoff to their minimum operating levels. Fifth, once shut down, most generating units need some off-time before they can be restarted.

One practical result of these limitations is that the output of a generating unit cannot easily be changed from hour to hour across a wide spectrum of output levels. Rather, if a generating unit is on, the latitude for changing output levels from hour to hour is limited by unit's ramping capabilities – both its ramp rate and any "lumpiness" in its supply function. If a generating unit is off, it may require well over an hour to ramp up to its minimum operating limit – and then only if it has been off for a sufficient period of time. Such operating limitations make it exceedingly difficult to determine what is the marginal cost of a generating unit.

The marginal cost of production of any good is the cost of producing one more unit of that good. Applying this concept to electric generating units, if a generating unit is off, the marginal production cost includes the cost of fuel needed to bring the unit to its minimum operating level. If the unit also has a multi-hour minimum operating time – which is a common occurrence – the marginal production cost of starting up the unit also includes all operating costs during this minimum running time period.

Apart from start-up considerations, one might think that if a unit is already running and producing electricity, determination of the marginal production cost is easier. But this case is still fraught with some complexities. First, as noted above, the production function can be lumpy. In other words, a unit may not be able to go from X MW of output to X+1 MW. Rather, X+5 MW may be the next feasible output level. The marginal cost curve, calculated as the first order derivative of the total production cost curve, would not reflect this fact. Second, a further issue arises from a somewhat more theoretical problem: what is the true "margin" of electricity production? Is it the next MW of output within a given hour, or is it the next hour of production at the same MWh operating level?

Taken together, these issues strongly underscore the difficulty of providing answers to the question: what is the marginal cost of electric generation? Clearly, there are no easy answers. Particular note needs to be given to the fact that the characteristics of each generating unit are unique and, for each unit, its marginal cost will depend on each of a multitude of operating characteristics of the unit, as well as the market in which that unit operates (discussed further below).

In contrast to these operating limitations of most generating units, centrally-operated energy markets – such as those controlled by the Cal-ISO – tend to operate on an hourly basis. In this context, the system operator generally dispatches generating units based on bids submitted on an hourly basis – or less in some cases. Unit operators seeking to participate in these markets generally have two opportunities to submit hour-specific bids to supply energy – once during the morning of the day prior to energy delivery (the "day-ahead" market) and once during the hour prior to delivery (the "hour-ahead" or "real time" market). In both markets, the operator must anticipate – up to a day or more in advance – what its production function will be in order to submit bids that can be realistically honored if accepted in the markets.

For a combustion turbine with only nominal start-up and minimum down times, submitting and honoring bids will be relatively easy. Assume the unit's operator offers to sell – through the day-ahead market – 25 MW of energy in each peak hour (*e.g.*, 7 a.m. to 10 p.m.) of the next day. If the unit is selected in only some of these hours – such as during only the "super peak" period of 1 p.m. to 5 p.m. – it can easily start-up at slightly before 1 p.m. and shut down slightly after 5 p.m. Even if the unit is selected to run from 1 p.m. to 2 p.m. and then again from 3 p.m. to 5 p.m., it can do so, as it has only nominal start-up and minimum down times.

In sharp contrast to the combustion turbine, a large, fossil-fueled steam unit lacks such flexibility to start-up and shut down and, as a result, its operator must structure its bids so as to take into account the likely scenarios for bid acceptance. Assume a unit has a 12-hour start-up time, three hour minimum run time, and a 12-hour minimum down time. The unit would thus not be able to "cycle" like the combustion turbine – on one hour, off the next, and on in the following hour. The steam unit's operator must structure the bid such that, if one hour is accepted, the next two hours will be accepted as well owing to the unit's three hour minimum run time.

The marginal cost seen by this operator is – at the very least – the cost associated with starting up the unit and operating for three hours. If the unit operates for only three hours, the start-up costs can be allocated over the three hours. If the unit operates for longer, the start-up costs can be allocated over this longer period, reducing the per-MWh marginal cost. Thus, when bids are to be submitted from such a fossil-fueled unit, accurate determination of what will be the unit's marginal cost of operation requires perfect foresight as to how long the unit will be operating.

Recognizing such disconnects between hourly markets and multi-hour operating characteristics of most generating units, some markets now guarantee operators recovery of their start-up costs if they are selected to run – thus providing a blunt solution to the problem of perfect foresight. Some markets also allow operators to submit tripartite bids that specify start-up costs, "no load" costs, and energy output costs.[12] These market systems come much closer than the one-part California system to recognizing the true cost function of a generating unit. Still, they are sub-optimal solutions given the complexities of generating unit operations.

Some critics of market participant behavior in California have alleged that bids exceeding the marginal cost of a generating unit evidenced an attempt to exert market power.[13] Putting aside issues of fixed cost recovery needs (addressed in the next section), there may be some validity to these critics' arguments. As noted above, however, determining the marginal cost of a generating unit is not nearly as simple as looking at the marginal cost curve of a unit when it is up and operating. Thus, to simply compare the bid to a place on the operating curve will often result in the over-identification of "excessive" bids.

Similarly, returning to the FERC's refund solution for the California energy crisis, it is incorrect to assume that the prices in a properly functioning market will reflect the marginal costs of production of the least-efficient unit if one defines marginal cost as marginal heat rate multiplied by fuel costs plus an O&M adder. An accurate marginal cost calculation is far more complex than that produced by the FERC's formula.

5. The Potential Harm of Forced Marginal Cost Pricing

Forced marginal cost pricing in an environment in which marginal costs are not calculated properly will have adverse market consequences. In the short-run, generators near the market margin – *i.e.*, those that are often or occasionally the least-efficient unit – will cease to bid into the market as a result of the risk that they may not recover their true marginal costs over their dispatch period. The consequences of this supply shrinkage will be less market liquidity and during peak demand periods (such as the summer) even greater upward pressure on market prices.

In the long-run, these generators will exit the market altogether. While economic theory would suggest that these exiting generators would be quickly replaced by new, more efficient ones, such may not be the case in electricity markets. New, efficient fossil-fired steam generating units often take upwards of five years to construct – and that is if one can find a suitable site in a world resistant to environmental degradation. Existing generating facilities can however be retired quickly. Some fixed costs are naturally sunk, such as in-ground capital. But some fixed costs (such as labor or capital repair) are recurring. Thus it may be reasonable for a generator earning only its marginal cost – even properly-calculated marginal cost – to retire a unit rather than pay such recurring fixed costs. This lack of parallelism between the time required to construct a new unit and the decision-period for retiring an existing unit has the strong potential to result in a shrinkage of supply to the market.

A possible regulatory response to this problem might be to provide generators with some incentive to stay operational until new supply sources can enter the market. Such a response, however, would be unworkable in reality since determining the incentive level raises the same quantification problems faced in determining a generating unit's marginal costs. Short of evaluating each separate expenditure, there is no accurate means for determining how much generators should be paid in order to keep them on-line and no more – which moreover is the apparent goal of regulators in seeking to drive market prices down to marginal operating cost levels. Further, such financial incentives are little different in the end than cost-based regulated pricing in which utilities earn a regulated rate of return on investment, after recovery of total, prudently-incurred operating costs. In all, one would only be replacing one severe regulatory scheme with another – not with deregulation.

The core problem here is that regulators seeking marginal cost pricing effectively want the benefits of market competition without the adverse consequences of price spikes. But it is these spikes that discipline the market and create the incentive for new generating units to enter the market. Regulators should not expect a half-deregulated system to work effectively, or to yield the efficient market results that a truly free and competitive markets could deliver. Not the least of the problems here is the lack of effective retail competition, or even more importantly the lack of effective price-responsive demand. In theory, demand should fall if prices rise. But even when rates are not artificially capped, retail consumers rarely see the price consequences of their high demand during peak periods until a month or two later when they get their bill.[14] Then, it is too late for price responsive demand to be effective, so they turn to regulators for a response – re-regulation.

Fully free competition in electricity markets has the strong potential to deliver greater market efficiency in the long run. But these benefits can only be accomplished if (1) prices are permitted to fluctuate freely, and (2) there is a direct and immediate pass-through of prices to end-use customers. If regulators wish nonetheless to suppress market prices to marginal cost levels, they must find some way of determining true marginal costs, rather than the "rough justice" used by the FERC for the California energy crisis. Continued suppression of prices to inaccurately determined marginal cost levels will surely result in another energy crisis as it will engender too much demand fed by too little supply.

[†] The views expressed in this paper are solely those of the author and do not necessarily represent the views of Troutman Sanders LLP or its clients. The author would like to thank Troutman Sanders LLP for its generous support of this paper and Kimber Shoop for his helpful assistance.

[1] Due to the limits of the FERC's authority under the Federal Power Act, it was only able to order refunds for the period beginning October 2, 2000. *San Diego Gas & Elec. Co.*, 96 FERC ¶ 61,120, 61,504 (2001) ("July 25 Order").

---

[2] *See San Diego Gas & Elec. Co*., 95 FERC ¶ 61,418, 62,560 (2001); *San Diego Gas & Elec. Co.*; 102 FERC ¶ 61,317, PP 11, 13, 56-63 (2003) ("March 26 Order").

[3] There are numerous nuances and limited exceptions to the MMCP formula established by the FERC. Most notably, the MMCPs calculated beginning January 6, 2001 were elevated by 10 percent to reflect a "creditworthiness adder" owing to the eroding financial situations of the state's largest utilities. Also, the gas price input to the formula originally contained two different prices – one if the generating unit was in the North of the state, the other if it was in the South. Later, the FERC substituted a proxy gas price for all generating units after finding that gas prices in the state may have been manipulated or, at the very least, did not reflect competitive pricing. *See* July 25 Order at 61,516-19; March 26 Order at PP 56-63.

[4] An average heat rate is determined by dividing: (i) the amount of fuel (in MMBtu) needed to produce a given amount of MWs of electricity at a steady level for an hour by (ii) the number of MW produced for that hour. By measuring fuel input volumes at different operating levels, one can develop both a total fuel input curve, calculating the fuel input at each MW level of output. The average heat rate curve is the total fuel input curve divided by the number of MWs. The incremental heat rate curve is the first order derivative of the total heat curve.

[5] In reality, many sellers have not yet seen many of those revenues as a result of, among other things, Pacific Gas & Electric failing to pay the Cal-ISO and Cal-PX for a significant amount of energy purchased in early 2001. For these sellers, the refunds will be offset against what they are still owed by the Cal-ISO and Cal-PX. A FERC administrative law judge estimated that, before any refund offset, the Cal-ISO and Cal-PX owe sellers approximately $3 billion, and that refunds will total approximately $1.2 billion. *See* March 26 Order at P 4; *San Diego Gas & Elec. Co.*; 101 FERC ¶ 63,026, P 7-8 (2002). The FERC's subsequent substitution of a natural gas "proxy price" into the MMCP formula will increase the refund amounts. *See supra* n.3.

[6] *Compare* Borenstein at 3-4 (stating that there are markets "in which virtually no market power exists" that share important similar characteristics to the electric industry).

[7] *See, e.g.*, Public Utilities Comm'n of Ohio, Press Release, "Ohio Power Siting Board Approves the Construction of Three New Major Electric Generation Facilities," May 20, 2002, available at http://www.puc.state.oh.us/emplibrary/files/media/MediaReleaseArchive/2002/02041.cfm.

[8] James M. Henderson & Richard E. Quandt, *Microeconomic Theory, a Mathematical Approach*, at 136 (3rd Ed. 1980).

[9] Fox-Penner at 299.

[10] *Investigation of Terms and Conditions of Public Utility Market-Based Rate Authorizations*, 97 FERC ¶ 61,220, 61,976 (2001).

[11] While pumped storage facilities may be considered an exception to this rule, the fact remains that these facilities still consume electricity and just convert electric energy to potential energy in the form of water at a higher elevation. Even if one were to consider such facilities an

---

exception to the supply-equals-demand rule of electricity, pumped storage facilities represent such a small portion of the United States' electric system that we can effectively ignore them.

[12] *See, e.g.*, *New England Power Pool*, 95 FERC ¶ 61,123 (2001).

[13] *See, e.g.*, Anjali Sheffrin, Director, California Independent System Operator Corp., Dep't of Market Analysis, *Empirical Evidence of Strategic Bidding in California ISO Real Time Market*, March 21, 2001 <http://www.caiso.com/docs/2001/04/27/2001042710305919478.pdf>.

[14] *See* Van Doren at 12.

Selected Bibliography:

Severin Borenstein, *Understanding Competitive Pricing and Market Power in Wholesale Electric Markets* (1999).

Phillip S. Cross, "Electric Rate Request Denied," *Public Utilities Fortnightly*, Vol. 123, NO. 3, at 59 (1989) (discussing empirical problems with marginal cost determinations in the then-deregulating natural gas markets).

Peter Fox-Penner, *et al.*, "Competition in Wholesale Electric Power Markets," 23 *Energy Law J.*, Vol. 23 at 281 (2002).

Michael J. Gergen, *et al.*, "Market-Based Ratemaking and the Western Energy Crisis of 200 and 2001," *Energy Law J.*, Vol. 24 at 321 (2003).

William H. Hieronymus, "Market Power Analysis of the Electricity Generation Sector," *Energy Law J.*, Vol. 23 at 1 (2002).

Peter Navarro and Michael Shames, "Electricity Deregulation: Lessons Learned from California," *Energy Law J.*, Vol. 24 at 33 (2003).

Gerald Nordlander, "May the FERC Rely on Markets to Set Electric Rates?" *Energy Law J.*, Vol. 24 at 65 (2003).

Craig R. Roach, "Measuring Market Power in the U.S. Electricity Markets," *Energy Law J.*, Vol. 23 at 51 (2002).

Peter Van Doren and Jerry Taylor, "Rethinking Restructuring," *Forthnightly*, Vol. 142, No. 2, at 12 (2004).