NATIONAL ECONOMIC
RESEARCH ASSOCIATES
1166 AVENUE OF THE AMERICAS
NEW YORK, NEW YORK 10036
TEL: *212.345.3000* FAX: *212.345.4650*
INTERNET: *http://www.nera.com*

n/e/r/a
Consulting Economists

The Allocation of Locational Rents

By Jonathan Falk

Presented in Connection with the Federal Energy Regulatory Commission Technical Conference in Docket No. PL04-2 -000 et al.

The treatment of "local market power" is but one of the many remaining contentious issues in the restructuring of the US electric power industry. This paper represents an attempt to sort out local market power issues and create a framework for the analysis of local market power mitigation proposals. I believe that much of the controversy over this issue can at least be put on a common footing by adopting an explicit view of the problem as one of rent allocation.

I have two central points. The first is that an undue focus on local market power can undermine the benefits we expect competition to provide. What is "undue" is to some extent in the eye of the beholder, but it is easily demonstrated that consumers are often (indeed, routinely) better served by markets in which there is some amount of market power than they are in markets in which market power is eliminated by fiat. The sooner we can think of electric markets as "normal" markets with "normal" amounts of local market power, the sooner we will reap the benefits of competition.

The second point is that, even if we had a generic framework for the diagnosis of "undue" market power (and we don't), administrative procedures to mitigate market power should be avoided whenever feasible. Rent-seeking behavior by participants in those proceedings will chew up even more resources, further reducing social welfare by converting social costs to private gain. Auction procedures to mitigate local market power problems may not be perfect, but they are often preferable to administrative devices, even those that attempt to impose a second-best alternative of mimicking what a competitive market would do. In general, it is better to replace a dysfunctional market with a workably competitive market than attempt to simulate market outcomes.

I think it is clearly in the best interest of consumers and generators that some of what look like local market power rents be allowed to flow to generators as a matter of standard practice. In particular, the PJM proposal to impose offer caps set at a slight premium to variable costs fails on several counts. It is excessively fearful of market power; it risks causing disruptive situations in which markets would actually work quite well if only allowed to do so; and it invites second-guessing by both generators and loads, depending on who feels wronged in a particular circumstance. More light-handed application promises substantial benefits, albeit at the price of some market power.

In many cases, market interference of any sort is unnecessary as competitive price signals will quickly resolve any local market power concerns. Nonetheless, there will be some instances in which chronic local market power persists. These instances are, I believe, far less common than the situations to which the PJM proposed procedures would be applied. For those remaining "hard cases," an auction proposal is a laudable substitution of a functional market for a dysfunctional one.

I. WHAT ARE RENTS?

If all productive assets were the same then there is no reason to assume that they wouldn't all earn the same level of profit. Indeed, if this were not the case, we could argue on that ground alone that they weren't "the same." **Something** has to be different between productive assets to allow them to earn different rates of expected[1] profit.

Rents are nothing more than differences in profit measured relative to the normal level of economic profit. Whatever it is that causes the difference in profit is the source of those rents. Any accurate representation of the asset must include any source of rent as a factor in the productive quality of the asset. Thus, all rents are merely the extra value of some factor of production, no different than the return to any other factor of production, except in one respect: scarcity. If a given factor of production were not scarce, everyone would use it and it would no

[1] It is conceivable that identical firms would expect the same profits but earn different actual profits because of luck. Holders of lottery tickets have the same expected profits before the drawing but very different actual profits afterwards. Chance variation does not affect the problems I'm discussing here, however, so we can use either expected profit or profit interchangeably.

longer be a source of rents. The theory of rent is the theory of the returns to a scarce factor of production.

Since rents represent a difference in profit, the cost of the factor of production matters. If coals of different heat content sell in the market place for differences that exactly represent their relative ability to produce electricity per ton, then a generator with access to superior coal will earn no incremental rent. The greater efficiency in electricity production per ton of fuel is exactly offset by the extra price of the fuel. If there is an incremental rent to be earned in this system, it must be further back in the productive chain. Thus, the owner of the mine might be earning an incremental rent (the additional profit per ton). Or, the rent might have been captured by a previous owner of the mine when he sold the mine for a premium.

Thought of in this fashion, almost any factor of production we can think of in electricity production can earn rents. Cheaper fuel can earn rents. Lower heat rates can earn rents. Smarter traders can earn rents. Improved O&M procedures can earn rents. And, the location of a particular generating unit can earn rents.

The benchmark rate of profit for the calculation of rents is the economist's hypothetical point of zero economic profits. In theory, *ex ante,* at the margin, the expected profit on a unit of capital that can be invested in any project in the world must be the same as on all projects that receive any investment. If some investment had higher expected risk-adjusted returns, capital would flow into that project and the extra returns on that investment would disappear. This *ex ante* expected return is what we mean by zero economic profits. In practice, it can be thought of as the "normal" level of return on invested capital.

Like the factors of production from which they derive, rents have a temporal element. Rents that are embodied as a technical innovation erode over time as more and more goods come to contain this innovation. What is actually happening in this case is the market increasingly insists upon the level of value offered by the innovation, such that the innovation no longer supports charging a premium price. Other rents, for example, those flowing from a superior CEO will erode over time as the CEO ages and eventually leaves the firm unless he is able to pass this human capital along intact to his successors. Thus, no rent is really permanent.

While I have defined rents in terms of profits, final consumers earn a type of rent as well, which is called **consumer's surplus.** It is defined as the difference in the value a consumer places on something he has purchased and the price he paid for it. If there are two consumers, one of whom values a kWh of electricity at $5 per kWh and another who values it at $4 per kWh, and both pay the same price of 10 cents per kWh, the consumer for whom electricity is more valuable will have made a higher "profit." More importantly, a reduction in the price of electricity, while reducing any rents earned by the producers of electricity (since it directly lowers profits, holding everything else constant) will (again, holding everything else constant) increase the aggregate consumer's surplus by exactly the same amount as it reduced the producer's profit. We can think of this case in which we somehow lower the price consumers pay (and producers receive) as a transfer of rents from producers to consumers, even though strictly speaking rents are only earned by producers.

II. WHO GETS RENTS?

Whenever one asset is more productive than other, someone gets rents even if it is not the owner of the productive asset. If generating unit A has a lower heat rate than generating unit B, and are otherwise the same, rents are produced somewhere. Who captures those rents is not clear, however. If the owner of unit A purchased his unit from someone, some of the rents might have been captured through a higher purchase price. If the lower heat rate flows from the cleverness of some engineer at the plant, the rent might be captured by the engineer in his next (or previous) salary request. If regulation permits the units to charge no more than their costs, including the return of and on invested capital (and assuming the efficient unit was in fact no more expensive) the lower costs of the efficient unit will be flowed through to consumers in the form of increased consumers' surplus. The critical fact is that the difference in productivity must go somewhere. The act of using the more efficient asset to produce electricity created the surplus. Someone somewhere will get it. The only question is who[2].

[2] There is an exception to this proposition from a societal standpoint, however. There is a large economic literature on what is known as **rent-seeking activity.** This is the practice of spending resources not to actually perform an activity more productively, but to capture the rents that the productive activity generates. For example, if a lottery has as its prize an asset that has rents associated with it, the submission of millions of entries to increase the chance of winning has no productive effect, but may be rational to pursue from a private standpoint. The limit to rent-seeking activity is the size of the rent pursued. Thus, it makes no sense to spend $1 million to augment one's

(continued...)

III. DOES IT MATTER WHO GETS THEM?

There are at least four reasons why we might care who gets rents, other than as an exercise in intellectual curiosity.

A. Rent Allocations As A Spur To Productive Efficiency

Rents are usually not the residue of luck, but of skill and effort. Someone had to find the better coal, invest in the better location, or hire the better CEO to capture the rents. That person's incentive to do so may be critically determined by their ability to capture some of the rents. Without those incentives, potentially productive investments will never be made at all. A premise of the drive to deregulate generation is precisely the notion that the introduction of the ability to capture these rents will spur innovation and efficiency.

A clear example of this process comes in patents. The profits earned on a valuable patent are rents earned by the innovation. The whole point of the patent system is to allow these rents so that the innovation is created to capture them. If the innovator cannot capture these rents, copying activity becomes more lucrative than innovative activity. Innovation will cease, as will copying as there are no future innovations to copy.

Where an asset could be placed in different locations, it is more likely to be sited at the location of its most productive use if a siting incentive in the form of incremental rents is available. The use that generates the highest rents is generally the most productive use. If we want assets to go to their most productive uses, we should make sure that those siting them have some stake in the outcome.

(...continued)

chances to win something that, even if won, gains one only $100,000. In the limit, with free competition to capture rents through rent-seeking activity, rents are **dissipated,** *i.e.,* the spending by others to capture the rents exhausts them. Thus, the rents are created through wasteful social expenditures that, while privately valuable, eliminate the social gain created by the rent. In saying this, we should separate two types of rent-seeking activity. Some rent-seeking activity simply shifts rent around. If I bribe someone to sell me a rent-producing asset at a below-market price, the rents are partly captured by the bribe-taker. In this case, rent-seeking activity does not diminish the aggregate rents. In the case where I produce millions of lottery entries, however, the cost of the paper (assuming the paper producers do not themselves earn rents on the extra paper sold) are simply a social waste that, in the limit, might actually exhaust the rents earned.

In traditional rate-of-return regulation, we try to ensure consumers capture all rents.[3] The main criticism of traditional rate-of-return regulation is that this scheme blunts incentives to make the utility more productive. All of the innovations to rate-of-return regulation, including but not limited to performance-based regulation, price caps, rate-of-return dead bands, institutionalized regulatory lag, and competitive generation markets have as their primary focus the goal of putting at least some rents back into the incentives provided to utility stockholders.

B. Rent Allocations In Allocative Efficiency

Unregulated monopolies can usually earn higher profits than non-monopolies. By our definition above, these are rents and must therefore have some source. The answer is that a monopoly must be sustained by some barrier to entry; it is this barrier to entry which is the source of monopoly rents. Patents confer temporary monopolies on the patented process. Copyrights confer near-perpetual monopolies on the copyrighted work. The time to build new electrical generating units combined with transmission limitations might confer a monopoly in some time periods to some generators. Up to now, my treatment of rents has been sympathetic, but there are few who will stick up for monopoly rents.[4] What's different?

The answer is that up to now I have focused on productive efficiency, *i.e.* the ability to produce a given quantity of output at the lowest possible cost. Monopolies generate rents not through cost advantages (though, like any other firm, they may have those as well) but through the ability to affect price. If I succeed in increasing the price consumers pay for my product (without losing market share) I will earn profits, *i.e.* rents.[5]

---

[3] There is a school of economic thought called Austrian economics in which the existence of rents is the driving force to virtually all entrepreneurial activity. Denying the capture of rents to innovators is identical to stifling the ability of capital to earn any return at all.

[4] With the prominent recent exception of the United States Supreme Court: "The mere possession of monopoly power, and the concomitant charging of monopoly prices, is not only not unlawful; it is an important element of the free-market system. The opportunity to charge monopoly prices - at least for a short period - is what attracts "business acumen" in the first place; it induces risk taking that produces innovation and economic growth." – Verizon Communications v. Law Offices of Curtis V. Trinko, LLP.

[5] Even where we do not have monopoly, most firms have some degree of control over price. The average level of control over prices is included in the normal benchmark level of profits in a typical workably competitive industry. We are discussing here only extraordinary levels of control over price which is not competed away due to some sort of barrier to entry.

We condemn price-setting (as opposed to price-taking) behavior on two grounds, one of which will be mentioned here and the second will be covered in the next section. The traditional economist's objection is that monopoly price-setting is allocatively inefficient. There is a deadweight loss associated with monopoly price-setting because there was incremental output that the monopolist could have produced at a cost that was less than some consumer's incremental willingness to pay, but that potentially mutually beneficial exchange was thwarted by the erosion of profit on the other sales made by the monopolist.

This deadweight loss is critically dependent on the sensitivity of the customer to price. A rough estimate of the deadweight loss is $0.5 \times \Delta P \times \Delta Q$, where ΔP is the change in price that the monopolist manages to cause and ΔQ is the reduction in quantity purchased by consumers. Thus, if there is very little diminution of consumption, there is very little deadweight loss, even if there is a substantial increase in price. In the limit, if the increase in price causes no change in consumption at all, there is no deadweight loss, since the increase in price did not reduce production of worthwhile goods. Since the elasticity of demand for electricity is fairly low (at least around currently observed average prices) pricing of electricity at something other than marginal cost does not result in a substantial loss of allocative efficiency.

Note that nothing in this section makes any distinction between monopolies that can raise price and those entities that exercise either unilateral or joint market power, or, for that matter, monopoly prices which are misaligned with marginal cost due to historic cost-of-service ratemaking. From an allocative standpoint, the only real question about market power concerns how effective it is at changing price and how much demand shrinks as a result. Virtually every firm possesses some ability to affect price. The allocative losses are still $0.5 \times \Delta P \times \Delta Q$; most firms simply can't affect P enough to be worth our concern, or the ameliorative efforts involve more dynamic losses through the blunting of incentives than the short term gains of correcting the allocative inefficiency.

C. Rent Allocation As Fairness

At least in electricity, the real objection to monopoly pricing is generally perceived unfairness rather than allocative inefficiency. Increases in price transfer money from consumers to producers – they monetize consumer's surplus and give it to producers. Remember that we

described consumer's surplus, the excess of value over price paid, as the analogue to profits. And so it is, with one important caveat. If something is worth $10 to me but I only have to spend $1 on it, I have enjoyed $9 of consumer's surplus, but that consumer's surplus is not directly transferable to anyone else. I can only enjoy personal satisfaction regarding the value. You can succeed in taking one dollar of the value from me only by removing a dollar from my pocket[6]. If we view this personal satisfaction as an *entitlement*, then the rent transfer involved in monopoly pricing is bad because it upsets that entitlement; in other words, it is unfair.

Fairness is not just a consideration in the case of a producer's ability to set price. Fairness can be invoked through historical considerations, neediness criteria, causality, or any of a host of grounds of merit or deservingness. The critical issue in the role of fairness in allocating rents is that the market alone cannot decide what is fair.[7]

Unlike with allocative or productive efficiency, there are no universally accepted criteria of fairness. While we can **measure** whether market rule A is allocatively or productively efficient relative to market rule B, we have nothing but opinion to rely on in determining whether the outcome of market rule A is more or less fair than market rule B. There are many solutions to the problems of allocative and productive efficiency to which the only objection is that they are unfair. Most critically, there is to my knowledge no theory under which prices based on marginal cost are "fair." It is certainly efficient, but its patterns of compensation routinely reward luck over merit and it does not discriminate between the deserving and the undeserving. While there are numerous grounds on which to criticize nonuniform prices as unfair, the uniform price that happens to equal marginal cost has no fairness properties of which I am familiar either.

D. Allocation of Rents and Regulatory Drag

Wherever there is an arbiter with the ability to allocate rents, *i.e.,* a regulator, there will be claimants for an entitlement to those rents. One of the pre-eminent failings of the traditional scheme of utility rate-of-return regulation is that the costs of adjudicating claims from more and

[6] For industrial and commercial customers, reductions in the cost of electricity reduce their overall costs; competitive conditions in their own markets will determine how much of that cost savings is embodied in the price of their products and how much their owners capture as rents.

[7] Unless there is someone to appeal to who has the ability to rectify the (perceived) unfairness, fairness will carry no weight. As John Kennedy remarked, "Life isn't fair."

more sophisticated claimants turned into a process that was highly costly and socially wasteful; with each claimant pursuing its own financial interest. Adjudicating these disputes, and the extremely difficult task of separating those claims that actually improve welfare from those that simply redistribute an existing rent, began to occupy more and more of the regulator's agenda. One hope was that the restructuring of the electric industry would, by subjecting most of these claims to impersonal market forces with no arbitrator, reduce this regulatory drag.

To some extent, restructuring has succeeded in this regard where market-based regulation of generation sales is permitted. Where we can, however, we should strive to further remove regulators from the process, not because they necessarily do a bad job, but because by offering a forum to argue over the distribution of rents, they inevitably encourage socially wasteful conduct.

IV. LOCAL MARKET POWER IN ELECTRICITY

The sections above now give us enough background to begin to look at the local market power problem in electricity. There are two different ways to think about local market power, and which of them is employed tends to color one's opinion in a way that is distinctly unhelpful to arriving at consensus, as the PJM Local Market Power Mitigation Working Group process demonstrated.

A. Local market power as a return to location

One way to look at local market power is that it is the process of capturing rents associated with one's location. Because of the physical characteristics of the electricity network, outputs from certain generators are not directly substitutable. Electricity is far from unique in this regard. Gasoline at a station in Atlantic City, New Jersey is not readily substitutable with gasoline at a station in Homer City, Pennsylvania, even though they are the same physical commodity. Indeed, one of the amazing things about electricity is that electricity generation in Homer City is in fact partially substitutable for generation in Atlantic City, although we have to take account of distribution effects, losses, and congestion costs. It is sometimes possible to turn on streetlights in Atlantic City and a generator in Homer City and, with only those two changes, have the PJM system balance. Alternatively, we could have left the generator in Homer City off

and turned on a generator in Atlantic City. Electricity is in general far more substitutable across large distances in short periods of time than other commodities.

Sometimes, however, only one or a small number of generators can serve load or preserve the reliability of the electric system. No short-run substitution is possible– either those units are turned on, load must be shed, or we run the risk of a blackout. The value of these particular units in this situation is quite clear. Because the average value of electricity to consumers is likely more than 50 times the price, a temporary locational rent has been created. It is temporary because it will only persist as long as the underlying conditions that caused nonsubstitutability of other generation persist. When the conditions go away, the value of the unit is now no longer its value vis-à-vis lost load, *i.e.* fifty or more times "normal" price, but the value of the next substitute.

Transient or not, the rent has been created. We now need to figure out who is going to receive it. If price is not allowed to change, consumers will receive the rent in the form of consumers' surplus. The only way to allocate it away from consumers is to permit higher prices. If higher prices are charged, the resulting fraction of rent, now monetized, could conceptually flow to almost any party.

Those who have adopted this view see locational rents like any other sort of rent, *e.g.* rents from lower costs or superior foresight. Since there is no serious effort to stop generators from capturing any of these other rents, they believe the burden is on those seeking to reallocate these rents to demonstrate why disparate treatment of locational rents is justified.

B. Local market power as price gouging

The second way that one could look at local market power is a noncompetitive gouging of the consumer. In this view, the existence of local market power is simply the happenstance result of the particular way the electric system looked as we embarked on competition. Proponents of this view believe local market power represents a historical accident. There is, in this view, really no way for new locational rents to be created, and thus no particular reason to reward those who can create them. The old locational rents, to the extent they are given to the current owners of generating assets, have no redeeming value.

There may undoubtedly be circumstances where the latter is true. This view is the impetus for proposals like those in California to simply designate certain units as reliability must-run and remove them from the competitive system altogether. The problem with this result is two-fold. First, it risks classifying as permanently noncompetitive situations in which, at small expense, competition could be reintroduced. Second, it removes units from competition in all hours even when locational market power might only arise in a few hours. It is essentially an abdication of the promise of generation competition.

Thus, for advocates of this view, locational rents are simply one of the many rents that, in the traditional rate-of-return era, belonged to ratepayers. Barring good reason to transfer some of these rents to generators, there would seem to be no legitimacy to altering that allocation.

V. RENT ALLOCATION PRINCIPLES FOR LOCAL MARKET POWER

Using the principles we have outlined above, we can look at where the rent **should** go.

A. Productive efficiency

Must-run units are usually higher cost than their substitutes. This is an empirical fact, not a law of nature, but it follows from the way the electricity system was constructed. Generation close to loads is generally constrained-on. In general, it is more costly to build generation closer to loads – land costs are higher, the infrastructure necessary to fuel the plant is more costly to construct, the size of plants must be smaller and, hence, less efficient, and competing demands on that infrastructure are higher.

Productive efficiency usually implies that we should allow sufficient quantities of rents to flow to those who can eliminate the condition. If we could reduce the number of times in which these plants are called on, we can increase productive efficiency by substituting cheaper power for local expensive power.

Why would anyone bother to relieve these constraints without being allowed to capture at least some of the rents? Low cost distant resources can earn no profit on these increased sales, since by definition this generation is marginal and consequently would earn little if were substituted for elsewhere. Without allowing access to at least some of the rents, there is no incentive to relieve the constraints. In addition to allowing those who could reconfigure the

transmission system to eliminate the constraint to share in the rents, we should allow some of the rents to go to units in the constrained-on area. In doing so, we induce others, where feasible, to enter the constrained area to capture at least some of those rents. There is of course a limit to this process, for as more and more entry occurs, the output of the units in the constrained area will be substitutable for one another, and thereby reduce the locational value created by any one of them. However, failure to allow the recovery of at least some of the rents caused by locational advantage threatens both short- and long-run productive efficiency – it wastes society's resources.

No one proposes to compensate constrained-on generation at below its variable costs. So, at a minimum, these generators are compensated at something resembling those costs[8]. Consideration of only variable costs is too narrow a perspective however. Another proposal would pay these plants both variable and some portion of the rents by permitting recovery of fixed costs for those times when they are put into service. Inclusion of these costs in the allowed bid of a constrained-on generator considers long-run productive efficiency, not just short-run efficiency. Again, to the extent that this value exceeds the price that would obtain were other plants substitutable (and this is generally the case) we could increase the productive efficiency of the electricity sector by reducing the number of times high-cost resources must be employed over low-cost alternatives.

In any given situation, the minimum efficient scale of generating units may render this solution impractical. If the constrained area requires no more than 100 MW of local generation and there is already a 500 MW generator in the area, new entrants may not see the opportunity to earn rents even if they are conceptually available. The incumbent generator could ensure that the new entrant never sees any of the rents simply by agreeing to turn on his own unit at a price that undercuts the variable plus fixed cost of the new entrant. But note that this threat, if enforced through an auction, automatically limits the rents that can be earned by the incumbent to a lower level. I will discuss the PJM auction proposal in greater detail below.

[8] Indeed, this is the PJM proposal,

B. Allocative Efficiency

Allocative efficiency requires price to reflect marginal cost. In practice, of course, fixed costs must be recovered as well, but this requires a reduction in allocative efficiency, in the absence of some way of recovering fixed costs that is independent of consumption. Recovering this on the basis of historic costs, as traditional cost-of-service practice does, may actually be worse. As noted above, across the range of prices typically charged and in the absence of significant demand response, price increases above marginal cost result in only small losses of allocative efficiency.

It should be noted that holding prices at levels as low as short-run variable cost can create another problem in allocative efficiency. Low prices signal to loads that there is no problem worth addressing. In choosing where to locate, neither loads nor generators will take heed of the fact that demand outside of the load pocket is in fact cheaper to supply than demand inside the load pocket. Thus, over time, low prices within load pockets actually exacerbate the problem.

C. Fairness

There are no generally accepted principles of fairness that tell us where locational rents should go. Obviously, loads in the affected areas will protest that higher prices should not be imposed on them. The appeal of this argument will no doubt depend on a host of historical circumstances, including but certainly not limited to:

- How have historic prices in the constrained area compared with prices in surrounding areas?
- Are prices lower than they would have been absent competition in generation?
- How have transmission congestion costs been allocated?
- Were there initial bargains at the outset of competition, *e.g.* guaranteed fixed rate reductions for a number of years?
- What options are available to hedge risks of significant nontransient price increases?

All of these arguments must be weighed by the regulator in determining the fairness of any particular allocation of rents. In particular, we should be very careful before asserting that any particular group of customers is in fact receiving the locational rents. Many customers have

already hedged their price risk, either through regulation or in the explicit purchase of hedges. Industrial customers may actually profit from high electricity prices within a constrained zone, making more money by shutting down and selling power back to the market than they would have made in output.

D. Regulatory Drag

As explained above, whatever the optimal allocation of rents as a theoretical matter, the fact that that allocation will be subject to adjudication between interested parties should, optimally, remove some of those rents from the adjudication process in order to account for the dissipative effect of rent-seeking behavior. Unlike academia, in the real world, the intensity of discussion declines as the stakes weaken. Wherever possible, the regulator should have as little to arbitrate as is consistent with the rules cited above.

FERC clearly recognizes this goal. For instance, FERC has consistently failed to intervene to abrogate contracts reasoning quite correctly that virtually all long term contracts create *ex post* regret for either the buyer or seller. This is simply a statement about rents as either the profits from the contract left a rent to the seller, or the buyer gains consumer surplus vis-à-vis the spot market. A world in which the fairness of those contracts can be re-examined *ex post* by FERC is a world in which long term contracts have no meaning. Even if there are certain rare circumstances in which discrete amelioration of patent unfairness can be achieved, a general willingness to reconsider these cases could bury this wheat among the chaff (and threshing expense) of unmeritorious claims. Better to bear occasional unfairness than to drown in the search for fairness.

VI. THE SPECIFIC PJM PROPOSAL AND THE PPL ALTERNATIVE

As mentioned above, PJM proposes a two-part allocation of locational rents. The basic premise is that all locational rents belong to the consumer. The proof of this proposition is that offer caps are set at variable costs[9]. Even when the price paid rises above this level, the resource

[9] Actually, they are set at variable costs plus ten percent. I ignore the ten percent here because I take it as a proxy for real, but not easily measured, "other" variable costs.

has received none of its rents from location; the higher price simply represents a rent from lower costs.

A. The Offer Cap

The rationale for this offer cap is that it approximates bidding behavior in a competitive market. But this rationale contains a hidden supposition. It assumes that the output of this plant is a homogeneous good with respect to the output of other plants. Where this is not the case, *i.e.,* there are differentiated aspects of the output, we expect prices to reflect short-run marginal cost, but not to mirror it, because the producer of the differentiated good is capturing some of the rents from this difference. In this case, the output of the locally constrained-on plant is simply more valuable than other units, since the output of that unit is superior in substitution with other units on the system. Under these circumstances, we expect the unit to offer at a higher price, even in a competitive market. Once we recognize that the output of some units is more valuable than others, the assertion that equilibrium prices even under competition reflect marginal costs is false. The rents earned by units that have made themselves more valuable than other units is the spur to locate in areas that can yield those higher values; even where the unit's extra value is the result of historical chance, the opportunity to capture these rents forms the impetus for new entry.

Of course, in the absence of demand response, the price that might be asked in certain circumstances would be higher than fairness would warrant. The fact that new entry might take years is good reason to cap the rents that might be earned. However, the cap must be sufficient to induce entry in order not to perpetuate the historical situation.

In response to my argument that the offer caps may lead to revenues that are not compensatory to generators inside the load pocket, PJM replies that these generators are not entitled to any particular level of return. While perhaps true, the statement is misleading. In normal competitive markets, sellers who do not earn adequate returns can exit. In so doing they avoid throwing good money (and effort) after bad. If the ISO refuses to allow exit or delays, it imposes an obligation without paying for it. It is of course true that in a time of depressed market prices no units would earn a compensatory rate of return. This can clearly be the case where there is overcapacity. But plants that are not inside a load pocket are free to exit any time they wish.

Each such exit raises the revenue received by others. Eventually, equilibrium is reached with the set of units who demonstrated the greatest staying power. Units in a load pocket do not just deliver energy, or even just ISO-wide capacity. They provide another service – reliability inside the load pocket, for which, under the PJM scheme, they receive no compensation whatsoever. Worse still, they are required to continue to provide this service even where other sources of revenue are insufficient to justify doing so economically.

Loads, of course, are also capable of providing this reliability through demand response. They have no incentive to do so if they are guaranteed a bailout at the variable cost of current units. As I mentioned above, offer caps at this level do not merely perpetuate load pockets – they induce loads within load pockets to expand and send the wrong signal to new loads on the best locations for siting.

B. Imposition of the Offer cap

The PJM offer cap is not only set at the wrong level; the means of imposing it, to the extent I understand it, is unnecessarily restrictive. Too much generation is put under these restrictions, both temporally and quantitatively. As my previous discussion made clear, as described, the caps are imposed even in circumstances where a unit's output is not required, in particular, for the part of a unit's output that is not pivotal.

Very high prices for very brief periods of the year are not necessarily the sign of a dysfunctional market. Any evaluation of rents ought to examine the aggregate rent earned by a particular unit over the whole year. It is unclear why any customers should be concerned about high prices for very brief period so long as their aggregate bill is lower than it would have been under regulation.[10]

Indeed, I suspect that there are many times when the imposition of the offer cap is due more to engineering assumption than necessity. One of the more interesting outcomes of the first year of the New England PUSH bidding experiment is that plants were in fact not needed very often. We have discovered that out-of-merit units cannot actually bid "whatever they want," which

[10] Moreover, the most sophisticated retail customers actually see hourly or even daily prices. In addition, retail customers are served by suppliers that are perfectly capable of hedging short term price spikes, provided they have the incentive to do so.

is the paradigm that animates the local market power problem, but in fact are displaced by units bidding only somewhat more than the PUSH unit's short-run variable costs. This of course calls into question the notion of designating the units as "local market power" units in the first place.

C. How much rent to allocate and the use of auctions

In choosing the level at which offer caps, if deemed necessary, are set, regulators must balance fairness with productive (and to some extent allocative) efficiency. While there are no generally accepted principles for setting the an offer cap that is "fair," with respect to the allocation of rents, there are only three dynamic possibilities. The first is that enough rent will be allocated to generators that new entry will be induced which solves the problem. The second is that structural entry barriers of some sort prevent new entry even in the presence of these rents. But the truly serious problem is the one which has animated the Commission's concern – caps will be set too low, which will not only deter entry but raises the real possibility that the unit inside the load pocket has no opportunity to earn a compensatory rate of return except in the circumstance where the entire region is short of energy.

This is the situation in which PPL has proposed that the affected unit have the ability to instigate an auction. The stated purpose of the auction proposal is to induce entry into constrained areas when deemed necessary by PJM. As I stated in my declaration attached to PPL's proposal, the PJM proposal should require an auction in another case – when the PJM offer cap results in insufficient return to the constrained generator to continue operation, but the necessity for the unit will not allow the unit to exit. In this case, the generator could petition to auction off his implied obligation in the load pocket. Whoever could relieve the constraint at least cost would win the auction and receive supplementary payments.

It should be clear from the preceding section that the cause of revenue inadequacy was simply a failure to allocate at least as much rent as a competitive market would to the generator in the load pocket. Thus, the addition of the auction simply corrects a problem that was correctable with proper capping in the first place, *i.e.,* if a cap is necessary, the auction supplies one set by the market, not the regulator.

Another factor which should make both offer capping and auctions relatively rare events are transmission infrastructure improvements. The PJM regional transmission expansion

planning protocol ("RTEPP") is designed to relieve load pockets where such investments make sense economically. There should be no reason to auction off a load pocket (or unnecessarily constrain bids there) in such situations which can be easily rectified under an already-existing transmission expansion structure.[11]

Nonetheless, there may be circumstances in which the rents paid to the incumbent generator are deemed "too high," but there is no obvious near-term path to entry; one example would be the case where the load pocket was too small to accommodate even one new efficiently-sized unit. In this case, a carefully structured auction for the obligation to serve the load pocket can find the cost-minimizing method to relieve the constraint. Since most such solutions would involve major capital decisions, the auctions would probably have to be quite long-term to induce appropriate responses.

D. Regulatory Substitution for the Market

While I accept that some regulatory intervention to allocate locational rents is inevitable, the PJM proposal, by explicitly intending to allocate all the rents to loads, ensures that continued regulatory intervention will be required. Pressures grow the farther rents are pushed from their natural levels. More important, however, is the mechanism by which allowing the rents leads to their dissipation, reducing the set of situations in which there are serious load pockets to those few cases where network topography simply does not permit a reasonable solution. In this way, the problem of local market power is self-limiting.

VII. SUMMARY

My previous arguments against the PJM local market power mitigation scheme were directed at what I still feel are the arbitrary and counterintuitive nature of its implementation. It seems to have started from a position that no locational rents should go to generators; realizing that that position leads to inadequacy of incentives to locate within a load pocket, the proposal then tacks on an auction as a last-resort method of compensating for that manifest defect. PPL's

[11] The RTEPP is not limited to long-lead-time transmission line projects. Before resorting to traditional regulated transmission projects, PJM will entertain load response, generation, and merchant transmission proposals. Moreover a transmission project could entail equipment upgrades that can be constructed in far less time than a new transmission line.

position is that if they have to remain subject to variable cost-based offer caps, at least allowing the generator to initiate the auction would allow generators to receive compensatory rates for their capital and other fixed costs within load pockets. I still believe that position to be correct as far as it goes.

My proposal, by contrast, starts with an exploration of whether the initial assumption – that there is a problem allocating locational rents to generators – has any validity. Just as we are willing, in a transition to competitive generation markets, to allow generators to reap all manner of other rents, we should keep an open mind about locational rents. In fact, there are good reasons to do so.

There are few markets in which we impose any restrictions on the ability of competitors to capture rents. As a salutary example; we do not allow landlords in New York City to capture locational rents for rental apartments constructed before 1947. By all accounts, that policy has been disastrous. At best, it simply shifted those rents (no pun intended) from landlords to tenants with no societal gain. In practice of course, it has created allocational and productive inefficiencies in New York City housing markets for the last 60 years.

The PJM plan caps too often and caps too stringently. It does this out of a mistaken notion that "market power" is a great evil out of which no benefit can come. It is not merely nomenclatural to point out that locational rents are an important part of the electrical system and that an understanding of, and optimal reaction to, locational rents are a big part of the promise of deregulated generation markets. Excessive regulation of these rents simply denies grist to the mill of competition.