Ben Hobbs *is Professor of Geography and Environmental Engineering, and of Mathematical Sciences (Joint) at The Johns Hopkins University, Baltimore, Maryland, and a consultant to the Federal Energy Regulatory Commission. His Ph.D. in environmental systems engineering is from Cornell University. He has been on the research staffs of Brookhaven and Oak Ridge National Laboratories, and from 1984 to 1995 he was a member of the systems engineering faculty at Case Western Reserve University. He is a member of the* TEJ *Editorial Advisory Board.*

Javier Iñón *is a Ph.D. student in the Department of Geography and Environmental Engineering at Johns Hopkins. He has also been a consultant with Cepri S.A. of Santiago, Chile, and PA Consulting. He has a degree in physical chemistry from the University of Buenos Aires.*

Steve Stoft *is the author of* Power System Economics: Designing Markets for Electricity, *to be published by Wiley–IEEE Press later this year. He is also a Senior Research Associate at the University of California Energy Institute and a Senior Advisor to The Brattle Group, Cambridge, Massachusetts. He previously served as an economist with the Office of Economic Policy at FERC. He holds a B.S. in engineering mathematics and Ph.D. in economics from the University of California at Berkeley.*

The authors thank Matt Kahal, Diane Brown, Sandi Patty, Udi Helman, Dick O'Neill, Eric Hirst, Pete Lalor, Mike Rothkopf, Kevin Kelly, Shmuel Oren, and Harry Singh for stimulating discussions. This work was funded by the Maryland Power Plant Research Program and the National Science Foundation, Grant ECS0080577. The authors alone, however, are responsible for the opinions and any errors contained in this article.

Installed Capacity Requirements and Price Caps: Oil on the Water, or Fuel on the Fire?

ICAP markets do not inherently inflate generator revenues, distort prices in a harmful way, or subsidize existing capacity that is otherwise uneconomic. Rather, they can be used with energy price caps to provide effective incentives for system adequacy.

Benjamin F. Hobbs, Javier Iñón, and Steven E. Stoft

President Bush and Vice President Cheney, past and current chairmen of the Federal Energy Regulatory Commission (FERC), and many others insist that price caps are a terrible idea for electricity markets even in the most dire circumstances. This opinion is based on inaccurate comparisons with the price controls of the 1970s. At that time, hamburger was price-controlled at $0.80/lb and sold at this price in almost every store at all times. The Eastern market price caps that are now questioned by FERC limit the price of electricity to the equivalent of $240/lb for hamburger in 1970. Had President Nixon set such a price cap on hamburger it would have caused no problems.

Such enormously high caps can make a difference in electricity only because the electricity is not easily stored and the market is flawed. A well-functioning spot market for electric energy should motivate the optimal amount of investment in generating capacity. After all, dairy farmers don't peddle "udder capacity" separately from milk, and consumers are not required to buy refinery capacity to gain the right to purchase gasoline. Sadly, however, electricity spot markets lack the consumer responsiveness that Saint Fred (Schweppe) envisioned as their

  PII S1040-6190(01)00212-3

linchpin in his epistle "Power Systems 2000."[1] Such responsiveness would discipline generator market power and give signals for the optimal amount, location, and type of capacity. Spot markets would then equate the marginal cost of providing reliable power with the consumer's willingness to pay (WTP) for it.

For many technical and institutional reasons, WTP for reliability is not reflected in price spikes that occur during times of capacity shortage. Metering limitations along with consumer and supplier skepticism about the value of real-time pricing mean that the vast majority of consumers neither face nor respond to spot prices. Despite FERC pushing independent system operators (ISOs) to encourage price-responsive demand and retail load participation in wholesale markets, this situation will not greatly change soon. Indeed, restructuring in some places has actually led to the cancellation of extant real-time pricing programs. As a result, the height of price spikes today reflects the unwillingness of ISOs or load-serving entities (LSEs) to stomach the political fallout of curtailment, rather than the WTP of the marginal power user. The present market is simply failing to provide information on the value of reliability.

Although this market flaw is recognized, many still worry that even very high caps could reduce generation investment. Perhaps, in spite of the flaw, all caps should be avoided because of this danger. We investigate here whether price caps could hinder investment, and conclude that when done right, they do not. In particular, we model markets with three styles of price capping with caps that range from $1,000/MWh and below, up to $12,000/MWh, and find that with each style of capping, too little, too much, or just the right amount of generation capacity can be summoned by the market. The three styles are pure price spike markets, installed capacity (ICAP) markets, and operating reserves markets.[2] In fact, under the ideal conditions we consider, not only can each approach induce the desired total capacity, each can also induce the optimal mix of base-load, shoulder, and peaking plants.

Willingness-to-pay is not reflected in price spikes that occur during times of capacity shortage.

FERC Chairman James Hoecker's final opinion on the California situation, issued just six days before leaving his post, called for California to adopt a capacity market to avoid "periodic reliability crises with energy price booms followed by price busts."[3] While capacity markets have some shortcomings, they do allow much lower price caps, which then control price spikes and thereby reduce the fluctuations in profits to generators. This, together with capacity revenues, should help stabilize the stream of investment and reduce the chance of another market meltdown. The good news is that the stability of lower price caps need not be accompanied by any attenuation of investment incentive.

In this article, we focus upon the combination of low price caps and ICAP markets. In the Pennsylvania-New Jersey-Maryland (PJM), New England, and New York markets, ISOs have imposed ICAP requirements upon LSEs, together with relatively modest ($1,000/MWh) price caps for energy. Either by providing its own capacity or load management programs, or buying credits from generation owners, each LSE must have sufficient ICAP credits to cover its peak demand plus a required reserve. To its advocates, ICAP has provided dependable revenue for power developers along with assurance to regulators and the public that needed capacity will "be there." Like oil on water, it is argued that ICAP has calmed the stormy seas of power markets buffeted by changing political and economic winds.

To its critics, however, ICAP and its accompanying price caps are an unnecessary relic of regulation that provides nothing of value. ICAP is instead seen as adding fuel to the fire of market instability by defining artificial markets that give generators additional opportunities to manipulate prices

  PII S1040-6190(01)00212-3

and extract undeserved revenues. In addition, ICAP systems are also accused of discouraging innovation and increasing pollution by keeping uneconomic existing capacity on life support.[4]

This article poses several questions concerning ICAP markets and energy price caps, and offers some answers based in part upon a market modeling exercise. These questions concern whether ICAP can be an efficient means of ensuring adequacy, or whether it instead leads to market distortions that justify its modification or elimination. After introducing our modeling approach, we address each of these questions in turn.[5]

1. Does ICAP distort energy prices and capacity additions?
2. Does ICAP distort capacity retirement decisions?
3. Is ICAP subject to migration, so that it cannot ensure adequacy?
4. Is ICAP a fictional product with no inherent value?
5. Does ICAP magnify market power problems?

We believe that ICAP markets together with energy price caps can provide effective incentives for system adequacy. Such systems do not inherently inflate generator revenues, distort prices in a harmful way (unless load faces real-time prices or maintenance decisions are affected), or subsidize existing capacity that is otherwise uneconomic. As actually implemented, however, ICAP markets have several problems, most of which can be corrected by altering market rules and penalties. In the case of the PJM market, the most important correctable problem is that of capacity "delisting," or migration to neighboring pools when energy prices are higher elsewhere. A problem that is more difficult to correct within a market that has both ICAP and an energy price cap is that energy prices will not reflect system conditions during periods of tight capacity. This could discourage cost-efficient measures to reduce loads or increase generating unit availability at those times. Correcting that problem would require more fundamental reforms. The alternative of substituting operating reserves markets or payments for ICAP is one widely discussed reform. A less well-known but nevertheless promising proposal is to require LSEs to obtain call options to cover their forecast peaks.

ICAP markets, together with energy price caps, can provide effective incentives for system adequacy.

I. A Simple Model for Assessing Capacity Incentive Proposals

The model used here, like most models of power markets, defines a price and investment equilibrium among generators in the market. The equilibrium sought by the model is a "long run" equilibrium in which no generating firm in the market is losing money, while no generator considering entering the market would earn a positive profit. Profits for a generating firm equal its revenue from power sales minus its costs. Revenues are influenced by the presence or absence of energy price caps, operating reserves markets, and ICAP credit markets; in this way, market incentives for providing capacity can affect entry and exit decisions by generators.

Our market model is a simple one that highlights the fundamental differences between different systems for providing capacity incentives. Some of the simplifying assumptions include: there are just two types of generating plants; the probability of a given generator being on forced outage is statistically independent of the availability of other generators and of the load; a generating company is risk neutral, and will build a plant if its revenues exceed its annualized capital and expected operating costs; maintenance outages are not considered; and generators behave competitively ("price takers") and therefore do not manipulate prices by withholding capacity. Capital and operating costs, capacities, and forced outage rates for baseload and peaking plants are based on data for new coal and combustion turbine facilities, respectively.[6] Meanwhile, load distributions are based on recent load history in PJM, with a peak load of 48,524 MW and average load of 27,694

MW; demand is assumed to be unresponsive to price. Our particular assumptions about load distributions and generator cost and reliability do not affect the general conclusions of this article and are only presented for illustration purposes.[7]

Expected profits for a generator are calculated by systematically considering all possible combinations of load levels and generator outages using a method called probabilistic production costing, a standard approach to calculating generator outputs and costs.[8] For instance, if there are 30 baseload plants (@1,200 MW) and 100 peakers (@150 MW) installed, one possible combination of load and outages that would be considered would be a load of 42,000 MW together with 28 of 30 baseload plants and 90 of 100 peakers available (or 47,100 MW total). As a result, peakers will be the marginal power source, and their marginal cost would be expected to set the market price, assuming no market power. If load was instead below 33,600 MW (the available baseload capacity), price would instead be determined by the marginal cost of baseload capacity. On the other extreme, if load equaled 48,000 MW, there would be a capacity shortage, and the energy price would spike upward to a price cap.[9]

If it turns out that one type of generating plant or the other makes a positive profit (net of a normal return to capital), then additional capacity of that type is added until expected profits fall to zero. In this manner, the number of plants of each type is adjusted until the equilibrium condition is met (nonnegative profits for generators in the market, nonpositive profits for generators outside the market).[10]

We have simulated the operation of three different approaches to capacity incentives. The first is a pure "price spike" approach in which the price cap is high enough so that energy revenues by themselves are enough to motivate capacity additions. Ancillary services, such as spinning reserves or ICAP, are assumed to provide negligible revenues in this case. In our system, the model shows that a price cap of $12,000/MWh would suffice to achieve a loss-of-load probability (LOLP) of 1 day in 10 years. Price hits this level a few hours per year. The exact value of the required cap depends, of course, upon the particular generator and load data used.[11]

We have simulated the operation of three different approaches to capacity incentives.

The second approach is an idealized ICAP market, in which ICAP payments are made to generators in proportion to their capacity, derated by their forced outage rates. We assume that no leakage of ICAP occurs to neighboring systems via delisting during peak periods. We also presume that use of ICAP payments rather than energy price spikes to incent capacity does not alter maintenance policies or forced outage rates during peak periods. We assume that the ICAP system is accompanied by a relatively tight cap on energy prices ($1,000/MWh, as in PJM). Under this cap and our generator and load assumptions, ICAP prices will reach $61,000/MW/year (or $7/MW/hour) in equilibrium if the amount of ICAP required in the market is sufficient to meet a LOLP of one day in 10 years. In general, the lower the price cap, the higher the equilibrium ICAP price must be.[12]

The third and final capacity incentive system we considered is an operating reserves payment system in which the market operator pays a fixed price per MW for operating reserves if such reserves fall below a given target. Here, with a target of 7.5 percent operating reserves, a payment of $1,815/MW/hour for operating reserves during times of reserves shortages, together with an energy price cap of $1,860/MWh, will be sufficient to ensure a LOLP of one day per decade. (Of course, in equilibrium, the price of energy at such times must exceed the price of operating reserves by the cost of running some marginal generating unit who will be indifferent between providing reserves and energy—here, the peaker.) Reserves and energy prices hit this level in 0.35 percent of all hours in a year, six times as often as energy prices alone hit the price cap in the pure price spike and ICAP systems. Unsurprisingly, the cap in the operating reserves system is therefore about one-sixth the value needed in the price spike system to motivate sufficient capacity additions.[13]

Clearly, the particular approach selected from the three will affect the distribution of energy prices and revenues to generators. What are the implications of these differences for the efficiency of the market? Will the right amounts and types of capacity be added in each case? We now turn to the five questions we posed earlier.

II. Does ICAP Distort Energy Prices and Capacity Entry Decisions?

Let us imagine for the moment that the ISO can somehow identify the socially efficient level of LOLP and capacity reserves. Then an ICAP system together with a price cap can be designed to achieve at least that level (subject to the caveats below about capacity migration and the effects of peak period prices on loads and maintenance). Similarly, given that target LOLP, caps for the pure price spike and operating reserves systems can also be identified to reach that reliability.

Compared to a price spike market, a combination of an ICAP system and energy price cap moderates price spikes. Generators receive less revenue from the energy spot market, but more from the ICAP market. Thus, the reduction of energy spot prices does not eliminate the signal that capacity additions are needed, which is instead given by the ICAP market prices. But does this change in the distribution of prices affect incentives to construct capacity, and alter the resulting mix and total cost?

Given our assumptions, the following conclusion results from the modeling exercise. *If* the ICAP system is designed to yield the same level of system reliability as the pure energy price spike system, then total revenues to each generator, the equilibrium capacity mix, and the total cost of supply are the same under the ICAP and pure price spike market systems. These results are also the same for the operating reserves payments approach. Under the particular generation and load assumptions we used, each of the three approaches results in 30,640 MW of baseload capacity and 24,870 MW of peaking capacity, which is the least-cost way of meeting energy and capacity needs subject to the LOLP of one day in 10 years. All generators break even in each case; revenue lost because of a tighter energy price cap under ICAP or the operating reserves approaches is made up by ICAP or operating reserves revenues, respectively.

The particular approach will affect the distribution of energy prices and revenues to generators.

Therefore, if the ISO identifies a target level of LOLP and capacity reserves, then an ICAP system can be designed to efficiently achieve that level under our assumptions. The changes in price distributions over the year do not distort the mix of capacity additions in our model relative to other systems that attain the same reliability. In other words, when it comes to capacity, it's "pay me now or pay me later." However, dampening of price volatility by the ICAP and operating reserves systems might be appreciated by risk-averse producers, their capital providers, and their consumers.

On the other hand, if the different approaches yield different levels of reliability—for instance, if ICAP requirements were set so that a LOLP of 0.5 days/10 years resulted, or if energy prices were tightly capped but no ICAP or operating reserve revenues were provided in compensation—then different revenues and capacity mixes would indeed occur. The baseload capacity would be the same in all cases, but differing amounts of peaking capacity would be added. We think that there is a distinct danger that if markets are designed to rely only on energy prices to motivate capacity additions, then too little capacity will be added because of political pressures to impose tight price caps well below consumer WTP to avoid outages. Furthermore, even if the right level of capacity is added on average in price spike markets, fluctuations in capacity levels are likely to be much larger than in ICAP markets, resulting in years of very low prices followed by years of very high prices, just as has happened in California.

These conclusions are robust with respect to particular cost and reliability assumptions for generators. But some of the structural assumptions of the model are strong ones, and we need to discuss their implications. Our assumptions concerning the absence of market power and the inability to "delist" ICAP and sell it outside the market are discussed later in this article. Another assumption was that generators only care about expected net revenue; if, instead, generators are risk-averse, then mixes and costs of generation might change because ICAP together with price caps will dampen price volatility and decrease risk-premiums.

Two other key assumptions behind these conclusions are (1) no consumer response to real-time prices and (2) the tying of ICAP definitions to actual ability to help meet load during peak periods. Because the prices of energy and operating reserves better reflect actual operating conditions under a price spike or operating reserves payment system, responsive loads and measures aimed at improving plant availability during system peaks will be more attractive under the latter approaches. For example, if some customers pay real-time prices, an operating reserves system—with its attendant price spikes due to reserves shortages—will motivate customers to cut back just at those times when the system is at risk. This incentive is removed if only long-term installed capacity is traded, and its cost buried in an

"uplift" on energy prices. Rolling ICAP charges into demand charges could be at least a partial fix, at least to the extent that demand peaks for individual customers correspond to system peaks. Another possible fix is to separate price to load from price to generation, allowing the former to exceed the generation price cap when reserves go to zero, and refunding the excess revenues as a reverse price uplift.[14]

The final structural assumption concerned generators' ability to lower outage rates during peak periods. Yet paying for installed capacity only, irrespective of its ramp rates and other operational characteristics, does not reward efforts to improve plant availability and flexibility during times of system peak. In contrast, operating reserves payments can provide such an incentive. Efforts to base ICAP credits on generation capacity availability during peak periods, as has been proposed in PJM, could lessen this problem. The success of this measure would depend on the ability of the ISO to predict peak forced outage rates for different types of units; fans of operating reserves and price spike approaches argue that it would be better to reward actual performance during peak rather than "iron in the ground" multiplied by an administratively-determined derating factor.

In sum, whether the energy price changes caused by ICAP represent economically significant "distortions" depends on the extent to which (1) load is responsive to real-time prices and (2) the calculated ICAP differs from on-peak availability. Additional analysis is needed to assess the possible significance of these effects, and proposed reforms.[15] The reforms just mentioned will necessarily be incomplete, because the inherent nature of a market combining ICAP and a price cap is to spread revenues over time that would otherwise be concentrated during price spikes. A more thorough reform summarized later in this article would involve mandatory call options and lifting of the energy price cap; it could provide incentives that appropriately reflect operating conditions, while

  PII S1040-6190(01)00212-3

at the same time providing ICAP-like price stability for LSEs and incentives for capacity construction for generators.

III. Does ICAP Distort Capacity Retirement Decisions?

A change in the amount and timing of revenues from energy, capacity, ancillary service, and other markets can alter incentives for life extension or retirement of existing plants. In particular, some of ICAP's critics have argued that ICAP markets prolong the life of old, inefficient plants that would otherwise be retired because ICAP provides a stream of revenue that those facilities would not receive from just the energy market.[16] These plants also tend to have higher emission rates, so extension of their operating lives could worsen air quality problems.

However, under our modeling assumptions, the analysis shows that *there is no such bias toward existing plants* in competitive ICAP markets. Plants that lose money in a pure price spike market or operating reserves payment scheme will also lose money in an ICAP market that results in the same system reliability. Existing plants that are profitable under one capacity incentive approach are profitable under all of them. The revenues and operating costs are the same for any particular plant, just distributed differently between energy and capacity payments.

For instance, given our data assumptions, an existing plant with a marginal cost of \$40/MWh and an avoidable fixed cost of more than \$78,770 per MW per year would lose money on average under either ICAP, a pure price spike system, or an operating reserves payment system. If the avoidable fixed cost is less than that, it would expect to be profitable in all cases.

On the other side of the market,

exit/entry is also an issue for LSEs. In particular, complaints have been made about the fact that a large fraction of the cost of power for new LSEs has consisted of capacity payments, and that these payments have been predominantly made to owners of existing plants. Our modeling analysis indicates that under our assumptions, the total cost of power to LSEs (including both energy and capacity payments) should not be affected by the presence of an ICAP market. It would be useful to test this conclusion empirically by examining prices in markets with and without ICAP.

IV. Is ICAP Subject to Migration, So That It Cannot Ensure Adequacy?

If two connected markets are experiencing capacity shortages, and one of those markets has a lower price cap than the other, generators in the former market will be tempted to instead sell their output to the latter. But if the generators in the former market are under ICAP contract, they (theoretically) should not be able to divert capacity in that manner.

However, a flaw in the PJM ICAP market is that, depending on market conditions in PJM and neighboring markets, it can be advantageous for PJM ICAP providers to commit to meet loads outside PJM even during times of capacity shortages within PJM. Capacity can be diverted in this manner by delisting an ICAP resource with a two-day notice. If neighboring systems such as the East Central Area Reliability Coordination Agreement (ECAR) have a higher energy price cap than PJM or essentially no cap at all, there will be a temptation to divert capacity when there are large price differences. The effect of this can be that PJM generators will be paid twice for their capacity—once through ICAP payments during most of the year, and a second time from price spikes in neighboring systems. The PJM consumers who pay for ICAP will not receive the capacity precisely when it is most needed.

Generators who delist from PJM are subject to penalties, including an ICAP deficiency charge and upward adjustments to the outage rate assumed in calculating their ICAP payment. These incentives are widely recognized to be insufficient, and changes in the deficiency charge have been proposed to try to prevent the delisting problem.[17] However, any change in the penalty will be *ad hoc* to some extent, as there is no widely accepted basis for determining the "right" penalty. More fundamental changes may be required and should be considered, such as increasing the length of notice required, harmonizing price caps within a region (as FERC ordered PJM, New York, and New England to do on July 26, 2000), or switching to an operating reserves system or the mandated call option system discussed in the next section.

V. Is ICAP a Fictional Product With No Inherent Value?

If an ICAP requirement improves system adequacy, it gives something of value to consumers. ICAP provides this value if otherwise energy and operating reserves prices by themselves would fail to motivate sufficient construction. Nonetheless, ICAP credits are indeed an artificial commodity whose demand is created by ISO rules. Owning an ICAP credit does not confer the right to any cash flows independent of the ICAP market itself; its value to the owner derives solely from demonstrating compliance with ISO regulations and avoiding penalties for noncompliance. Because credits have no inherent value, ICAP credit markets would not spontaneously spring up in the absence of ISO rules.[18]

Although the value of ICAP credits to purchasers is related to noncompliance penalties imposed by the ISO, the cost of ICAP credits to suppliers is more ambiguous and depends on market rules. In the case of New England, it has been correctly pointed out that the true short-run cost of ICAP is no more than the modest incremental

expense of maintaining the plant so that it can produce power.[19] This is because all power plants in New England could be called to operate when ISO-New England (ISO-NE) has insufficient generation, even if they do not submit bids to the ICAP market or commit the capacity in the bilateral market. As a result, there is no opportunity cost stemming from foregone sales of capacity outside ISO-NE, because such commitments are not allowed. Thus, when capacity is in excess, as it is now in that region, ICAP prices should be at or close to zero (assuming competitive behavior!).

In PJM, however, supplying ICAP does have an opportunity cost. During times when energy prices are higher elsewhere and PJM recalls its ICAP, PJM generators have the option to delist from ICAP and not be subject to recall, in which case they are then free to sell the energy in another market. Thus, PJM generators will require that ICAP revenues compensate them for foregone sales opportunities in more profitable markets, net of any penalties. This opportunity cost will be reflected in ICAP bids even if there is excess capacity in PJM.

In contrast to ICAP, an example of a product with inherent value is a call option on energy: It entitles the owner to certain cash flows or energy when energy prices exceed the strike price. Proposals have been made to eliminate the ICAP system and instead require LSEs to obtain call options (perhaps backed by capacity) sufficient to cover their loads during peak periods.[20] If the seller of such an option cannot provide the energy (perhaps because of an outage), then it must pay the right holder the spot energy price, which would be subject to a much higher cap than at present. This proposal would substitute a product with a value independent of ISO rules (call options) for a product with no such value (ICAP). The proposal would provide ICAP-like incentives for capacity construction along with protection against price spikes. In addition, as mentioned earlier, the mandatory call option proposal would prevent the delisting problem, while providing

  PII S1040-6190(01)00212-3

appropriate incentives for improving availability on peak. Therefore, we recommend that the merits of the mandatory call option proposal be considered by markets now using or contemplating ICAP.

VI. Does ICAP Magnify Market Power Problems?

An important objection to an ICAP requirement is that capacity markets can be gamed. Yet the existence of a market for installed capacity does not inherently lead to more market power than already exists. Generally, if power markets are concentrated, strategic manipulation of price will be attempted in all markets in which larger firms participate, including energy, capacity, transmission, and ancillary services. The presence of ICAP together with an energy price cap may simply move market power around, from the spot energy market to the capacity market.

But poorly designed rules for ICAP markets can indeed magnify market power. This is indicated by the experience in the ISOs that have ICAP markets. Three different pools with different ICAP rules have shown different results, depending on the design of the market. New England has seen market power exercised in the monthly ICAP. Very high bids had been submitted for supplying capacity in hopes that those bids would set the price. This occurred even though ICAP was in excess supply and had no opportunity cost, as it cannot be delisted and diverted outside the region. What invited this gaming was that the demand for ICAP was automatically entered into the auction as a perfectly inelastic demand curve. Design of ICAP markets to allow submission of price-sensitive demand bids (as is permitted in PJM) could help eliminate this gaming.

On the other hand, it does not appear that high ICAP prices in the NYISO result from the exercise of market power. New York has

defined transmission constrained submarkets for ICAP, such as one for the New York City area. In the summer of 2000, shortages of capacity meant that ICAP demands in that area could not be met, and the result of any auction in that circumstance, even under pure competition, is an ICAP price that hits the cap. In general, however, geographically separated submarkets will have higher supplier concentrations and more opportunities for exercising market power. An important question facing market designers is the tradeoff between a more accurate representation of capacity needs (by defining submarkets based on transmission constraints), the greater opportunity that segmented markets present for strategic manipulation, and the added complexity involved in setting up inter-regional trading. Modeling tools are available to address such questions.[21]

Finally, in PJM there are claims of market power based on the illiquidity of the market, concentration of suppliers, and the high ICAP prices that have recently been experienced. The PJM Market Monitoring Unit has concluded, however, that the opportunity cost of ICAP is a plausible explanation of those high prices, as there is a correlation between PJM's ICAP prices and energy prices in neighboring systems.[22] Such a correlation is to be expected because PJM generators (unlike those in ISO-NE) can commit their capacity to other markets where they can at some times get higher energy prices. Continued monitoring is necessary, however, because manipulation may still have occurred undetected and is possible in the future.

VII. Conclusions

Under our assumptions, the analysis leads to the following conclusions about the three basic approaches to capacity incentives (no operating reserve requirements with a very high price cap during blackouts; operating reserves requirements with medium price caps; and ICAP with reserve requirements and low price caps). First, any of the

  PII S1040-6190(01)00212-3

approaches can in theory motivate the provision of a targeted level of system reliability through selection of the appropriate values of the price cap, operating reserves payment, and/or ICAP requirement. Consumers pay the same amount in each of the systems; the creation of a separate capacity market should not raise consumer costs in the long run. Significant revenues from capacity sales are not necessarily an indication of market power, nor are they necessarily "excess." The long-run cost of power is approximately the same as would occur under other systems if designed to ensure the same reliability. Second, if alternate market designs give the same total capacity, the same mix of capacity results from each. Third, regulators have a crucial responsibility that should not be ducked: to set the market parameters that determine the reliability level in each of the three approaches. These parameters include an energy price cap level (used in all three approaches), an operating reserve requirement in the second system, and ICAP parameters in the third.

But theory and experience indicate that reliance on energy prices alone to motivate sufficient capacity is economically and politically risky. Appropriate incentives are provided only if prices during times of shortage approximate the value of unserved energy. Yet, since real-time price signals to consumers are essentially absent at this time, there is little reason to believe that this is true. We could just set the energy price cap equal to some guess about the value of unserved energy, as the Australians have done. But as has been pointed out in these pages before, there is much more uncertainty about the worth of reliability than there is about the right level of installed reserves.[23]

For these reasons, we believe that pouring some ICAP oil on the troubled waters of power markets for the next few years can help

improve economic efficiency and public confidence in the restructuring enterprise. "(I)n this highly integrated business, where the system requires everyone, and not just the visionary, to be prudent or face losing service and paying high spot prices, enforced customer-side planning ahead will be a small price to pay to avoid a cycle of boom and bust."[24] If ICAP is not adopted, then a low price cap with a high operating reserve requirement would be second best. If, as we earnestly hope, a large portion of demand eventually becomes responsive in real time to market conditions, and energy prices better reflect consumer valuations of reliability, then ICAP requirements can be phased out.[25]

The potential value of an ICAP requirement could escape our grasp, however, if the system is poorly designed. To be effective, ICAP systems must minimize opportunities for generators to game ICAP auctions (as occurred in New England) or to divert ICAP resources to other markets when prices are higher elsewhere (as happened in PJM). ICAP systems should also reward load management efforts (as in PJM) to hasten the day that most loads respond to real-time prices. If the portion of the load that is subject to real-time price does increase significantly, it may eventually become desirable to switch to an operating reserves market system with a higher energy price cap so that consumers can make better consumption decisions during peak periods. Incentives to improve generator availability during peaks would also be improved under such a system. Alternatively, proposals for a mandatory call options together with higher price caps could accomplish the same goals while eliminating the incentive to divert ICAP resources. ■

Endnotes:

1. Fred C. Schweppe, *Power Systems '2000': Hierarchical Control Strategies*, IEEE Spectrum, July 1978, at 42–47. For a quantitative presentation of the role of demand responsiveness in generation investment decisions, see Michael Caramanis, *Investment Decisions and Long-Term Planning under Electricity Spot Pricing*," IEEE Trans. Power App. Sys., *PAS-101*(12), 1982, at 4,640–48. Prospects for increasing the role of the demand side in

  PII S1040-6190(01)00212-3

price determination are discussed in Eric Hirst and Brendan Kirby, *Retail-Load Participation in Competitive Wholesale Electricity Markets, Consulting in Electric-Industry Restructuring*, Edison Electric Institute, Washington, DC, Dec. 2000.

2. For fuller discussions of alternative approaches, see, e.g., Eric Hirst and Stan Hadley, *Generation Adequacy: Who Decides?* Elec. J., Oct. 1999, at 11–21; Adam B. Jaffe and Frank A. Felder, *Should Electricity Markets Have a Capacity Requirement? If So, How Should It Be Priced?* Elec. J., Dec. 1996, at 52–60; N. Rau, *The Need for Capacity in the Deregulated Electrical Industry: A Review,* Proc. IEEE Power Engineering Society 1999 Winter Meeting, Piscataway, NJ; and Harry Singh and Jonathan Jacobs, *Capacity Products and ISO Markets*, PG&E Energy Services, 2000. Other approaches not discussed here include direct payments for capacity (as were once made in the United Kingdom and Argentina), and more sophisticated ICAP and operating reserves proposals in which an ISO would pay a price for capacity or reserves that would increase with the degree of shortage (Steven Stoft, *Power System Economics: Designing Markets for Electricity,* draft, http://www. stoft.com [July 6, 2001]). The New England ISO has proposed an operating reserves market of the latter sort.

3. From Chairman Hoecker's concurring opinion issued Jan. 4, 2001, on San Diego Gas & Electric Company, Complainant, Docket Nos. EL00-95-000, 002, 003, http://www.ferc.fed.us/Electric/bulkpower/4chair.final.pdf (July 6, 2001).

4. For criticisms of ICAP, see Peter Cramton and Jeffrey Lien, *Eliminating the Flaws in New England's Reserve Markets*, working paper, University of Maryland, March 2000, http://www.cramton.umd.edu/auction-papers.htm (July 6, 2001); John Hanger, John Rohrbach, and Peter Adels, *Time for PJM to Dump ICAP Too*, Penn Future, June 30, 2000; and ISO-New England Inc. Filing to FERC ICAP Termination effective June 1, 2000, http://www.iso-ne.com/FERC_filings/market_ rules/terminate_ICAP_5-8-00.pdf (June 25, 2001).

5. For details on the analysis, see Benjamin F. Hobbs, Javier Iñón, and Matt Kahal, *A Review of Issues Concerning Electric Power Capacity Markets*, report to the Maryland Power Plant Research Program, June 2001.

6. Coal plants are assumed to be 1,200 MW in size, have a forced outage rate of 0.12, variable operating costs of $25/MWh, and annualized capital charges of $140,000 per MW per year. For turbines, these assumptions are instead 150 MW, 0.05, $45/MWh, and $63,000/MW/year, respectively.

7. In particular, it can be proven that assumptions about the number of generation types and their costs and reliability do not affect our conclusions concerning the ability of different capacity incentive systems to achieve the least-cost mix and amount of capacity, given an LOLP target. The only exception is for the operating reserves system under very restrictive conditions. The conditions are that there are at least two distinct peaking technologies, one with significantly higher operating costs and lower capital costs than the other, such that both are in the optimal generation mix, and there is a significant probability that the one with the lower operating costs will be the marginal source of power when reserves are less than the targeted level (in our analysis, 7.5 percent). In that case, the mix of the two peaking technologies could be different under the operating reserves approach than in the ICAP or the price spike approaches.

8. A nice review of the use of probabilistic production costing for planning and market analysis is presented by Edward H. Kahn, *Regulation by Simulation: The Role of Production Cost Models in Electricity Planning and Pricing*, 43 Oper. Res., 1995, at 388–98. Our version of the probabilistic production costing method develops probability distributions for the amount of available capacity of each type and then convolves those distributions with the assumed distribution of loads. Given independent forced outages and identical plants, the distribution of available capacity of a given type is described by a binomial distribution, which can be adequately approximated by a normal distribution with the same mean and standard deviation. The distribution of load for PJM is modeled as a mix of normals distribution (see George Gross, Nancy V. Garapic, and Bruce McNutt, *The Mixture of Normals Approximation Technique for Equivalent Load Duration Curves*, IEEE Trans. Power Sys., *PWRS-3*(2), 368–74, 1988). The use of normal distributions means that the distributions of capacity margins (capacity − load) are also mixes of normals, which are easily manipulated. This probabilistic approach to analyzing capacity markets is a generalization of the deterministic analysis in Stoft, *supra* note 2. The ORCED model of Hirst and Hadley, OakRidge Competitive Electricity Dispatch (ORCED) *supra* note 2, is a broadly similar approach that they used to analyze price spike versus capacity payment markets that also considered demand responsive to real-time prices. The advantage of our approach is that it explicitly models the three capacity incentive systems while considering dispatch, unserved energy, and operating reserves in a consistent probabilistic framework.

9. In the simulations, it is assumed that the energy price goes up to the price cap whenever the available capacity exceeds load by 1,000 MW or less. This is a rough approximation to the ubiquitous phenomenon that price spikes will occur before a load exceeds capacity. The pre-

PII S1040-6190(01)00212-3

cise value of this MW threshold does not affect our general conclusions.

10. A generalized reduced gradient nonlinear programming method (Excel Solver®) was used to adjust capacity amounts in this manner.

11. For instance, if we assume instead that prices do not hit price cap until an actual shortage occurs (as opposed to being within 1,000 MW of a shortage), then price spikes of $12,000 will only be sufficient to obtain a LOLP of 2.3 days in 10 years. Price caps would then have to approach $27,500 to reduce LOLP to our assumed target of one day per decade. Meanwhile, changes in plant characteristics or the load distribution would *not* affect the size of the price spike required to achieve a given LOLP because the payment needed to cover capacity costs turns out to be equal to the product of LOLP and (PriceCap − MC), where MC is the marginal operating cost of the peaking plant. Note that our definition of LOLP is literally 24 hours/87,600 hours, or 0.027 percent, as opposed to a frequency-based definition (number of outage events per decade).

12. For instance, a $200/MWh price cap would increase the required ICAP price to $65,500/MW/yr. Other sensitivity analyses show that a lower target LOLP would decrease the equilibrium ICAP payment, as would a decrease in the assumed capital cost of combustion turbines or their forced outage rates.

13. We have also analyzed the New England-style operating reserves market mentioned in note 2, *supra*. If our ISO paid a linearly declining price for operating reserves, descending from $5,005/MW when reserves are zero to $0/MW when a 7.5 percent operating reserves level was achieved, together with an energy price $45/MWh greater than that level, then the one day in 10 year target would be achieved. Further, the number and mix of generating plants would be the same as in the other systems.

14. For instance, the demand price could be allowed to go up to $2,000 during shortages; the difference between that amount and the generation price cap could be refunded on a per-MWh basis over the following year (perhaps as an offset to uplifts imposed for other reasons). This system, together with ICAP, could help mitigate market power, lower risks, and still provide sufficient incentive for investment.

15. The effect of demand elasticity with respect to real-time prices could be analyzed in our framework using the basic approach of Hirst and Hadley, *supra* note 2. They found that even a very modest elasticity of −0.05 could be very useful in moderating price spikes and changing load shapes. They and others have previously made our point that the basic ICAP proposal would dilute the incentive during peak periods for efficient load reductions and extraordinary efforts to increase plant availability.

16. ISO-New England, *supra* note 4.

17. PJM-Interconnection Market Monitoring Unit, *PJM Interconnection, State of the Market Report 1999*, June 2000.

18. Of course, there is also artificiality in the pure price spike and operating reserves systems in that the market designer interested in motivating capacity additions must select parameters (such as price caps and operating reserve requirements) in order to induce that investment. ICAP can be viewed as more direct in that its main parameter is the capacity level one wants to induce, rather than parameters whose impact on capacity must be guessed at.

19. Cramton and Lien, *supra* note 4.

20. Shmuel S. Oren, *Capacity Payments and Supply Adequacy in Competitive Electricity Markets*, VII Symposium of Specialists in Electric Operational and Expansion Planning, Curitiba, Brazil, May 21–26, 2000; Harry Singh, *Call Options for Energy: A Market-Based Alternative to ICAP*, paper presented at PJM-FAWG Committee Meeting, Oct. 2000. Columbia is seriously considering such a proposal. See Carlos Vázquez, Michel Rivier, and Ignacio J. Pérez-Arriaga, *A Market Approach to Long-Term Security of Supply*, in PROCEEDINGS OF MARKET DESIGN 2001, Elforsk, Saltjőbaden, Sweden, June 7–8, 2001, available at http://www.elforsk-marketdesign.net/archives/papers/vazquez.pdf (July 6, 2001).

21. See, for example, Carolyn A. Berry, Benjamin F. Hobbs, William A. Meroney, Richard P. O'Neill, and William R. Stewart, Jr., *Analyzing Strategic Bidding Behavior in Transmission Networks*, 8(3) UTILITIES POLICY, 1999, at 139–58. For a review of models of strategic behavior in power markets, see Edward P. Kahn, *Numerical Techniques for Analyzing Market Power in Electricity*, ELEC. J., July 1998, at 34–43.

22. PJM-Interconnection Market Monitoring Unit, June 2000 ICAP investigation, presented to the Energy Market Committee, July 5, 2000, available at http://www.pjm.com.

23. Jaffe and Felder, *supra* note 2. The Australian $10,000/MWh price cap, although intended to represent the value of unserved energy, is upon closer inspection an arbitrary regulator decision. On the other hand, as indicated by Hirst and Hadley, *supra* note 2, and EPRI's "over/under" analyses of the 1970s and 1980s (Edward G. Cazalet, Charles E. Clark, and T.W. Keelin, *Costs and Benefits of Over/Undercapacity in Electric Power System Planning*, EPRI, Palo Alto, CA, 1978), the total social cost function in the region of the optimal installed reserve is flat, and so errors of plus or minus a couple of percent on the level of ICAP should not have large economic efficiency consequences. We do not mean to understate the difficulty involved in determining a target for system adequacy; however, this task (or the no easier one of determining the "right" value of unserved energy) will be the regulator's responsibility under any system.

24. Hoecker, *supra* note 3, at 7.

25. Note, however, that the task of transitioning to (and hopefully eventually out of) an ICAP system is nontrivial. PJM's system may have been relatively easily to implement because it was a continuation of PJM's capacity obligation system under regulation, but California, for instance, does not have such a natural starting point.

