Generation Adequacy via Call Options Obligations: Safe Passage to the Promised Land

In contrast to the controversial LICAP, the author's plan relies on standard hedging instruments that a mature energy-only market can support without regulatory intervention. Unlike payments for an artificial capacity product, for which there is no natural demand, energy call options provide intrinsic value to customers, since the generators who are paid for such options must pay back any windfall profits. This amounts to a risk trading arrangement where the consumers assume some of the investment risk, in exchange for reducing their price risk.

Shmuel S. Oren

Shmuel S. Oren *is Professor of the Industrial Engineering and Operations Research department at the University of California at Berkeley. For the past 10 years, he has served as Berkeley site director of Power Systems Engineering Research Center (PSERC), sponsored by the National Science Foundation and industry members. He has been a consultant to private and government organizations including the Brazilian Electricity Regulatory Agency (ANEEL), the Alberta Energy Utility Board, the Polish system operator (PSE) and the Public Utility Commission of Texas (PUCT). He holds a Ph.D. in Engineering Economic Systems from Stanford University. He is a Fellow of the IEEE and of INFORMS.*
This work has been partially supported by NSF ECS/EPNES Grant 0224779 and by a Department of Energy grant administered by the Consortium for Electric Reliability Technology Solutions (CERTS).

I. Introduction

In the Promised Land of electricity deregulation it was envisioned that generation companies would be paid for energy and operating reserves and bear the investment risk in generation capacity, consumers would bear the price risk for the energy they consume, and the system operator would take care of reliability through competitive procurement of resources. Competitive forces, consumer choice, speculative arbitrage, and bilateral contracting are the vehicles through which the system was supposed to achieve efficiency in operation, investment, and risk allocation, which are the pillars of an ideal competitive energy market. In the ideal competitive energy market paradigm, generators are price

 1040-6190/$–see front matter doi:/10.1016/j.tej.2005.10.003

takers and offer their product at marginal cost and are paid the market clearing price at each location, which is determined by the locational marginal cost resulting from least-cost dispatch (subject to network and reliability constraints). In theory, scarcity rents during rare capacity shortfalls covered by demand response, and the inframarginal profits due to the difference between the locational clearing prices and generation cost, should provide the exact income to generators for recovering their amortized fixed cost of capacity when the total generation capacity and technology mix are at their optimal long-run equilibrium levels. Such a long-run equilibrium (assuming it exists) is achieved through a process of entry and exit by producers in response to excess profits or profit shortfall. Surprisingly enough, this idealistic process relying on competitive market forces appears to be quite robust and has worked in most industries with minimal regulatory intervention, stimulating innovation and investment.

Unfortunately, a decade of experience with electricity deregulation, which includes the California energy crisis, taught us that reliance on competitive market forces to provide the ''correct'' level of generation capacity and technology mix could be risky. Among the obstacles to reaching the Promised Land of an energy-only market are market imperfections – including technological and political barriers to demand response as well as local market power due to network constraints. Attempts to overcome such market imperfections and to keep prices and system reliability under control have led to market designs that are encumbered by layers of Band-aids. On the market side, price caps and automatic mitigation protocols that constrain idealized free trading and market signals have been implemented in order to curb market power abuse and politically unacceptable price spikes. On the reliability side, while the system has been relatively reliable, such reliability has been achieved at high cost and through a fundamental disconnect between market and system operation. Operational reliability is achieved largely through ''out of market'' actions by the system operators who recovers the cost of such actions through uplifts. Such out-of-market actions distort spot prices, requiring further Band-aids to the settlement system.

Public concern for system reliability and supply adequacy has led to a proliferation of capacity mechanisms designed to supplement generators' income in order to incent new investment.

Some electricity systems like the one in Australia have relied on ''energy-only markets,'' bilateral contracting, and demand response to support investment in generation capacity with significant success. That success is often attributed to the scarcity pricing policy that allows electricity prices in Australia to rise to US$8,000/MWh, to the absence of market mitigation and to market rules that allegedly promote liquidity. However, there are also unique aspects such as ownership structure, technology mix, absence of market power, and a set of not-well-defined ''good faith'' constraints on bidding behavior, that have contributed to that success. More recently, the public utility commissions of Alberta and Texas have adopted decisions that support energy-only markets as the mechanisms of choice for incenting generation investment but the specific implementation details are still under development.

With the few exceptions mentioned above, public concern for system reliability and supply adequacy has led to a proliferation of capacity mechanisms designed to supplement generators' income in order to incent new investment or prevent early retirement of unprofitable generation capacity. Furthermore, all the designs (with or without capacity mechanisms) in the U.S. (including the Federal Energy Regulatory Commission's now-dead Standard Market Design proposal) and worldwide, abandoned the principle that ''the *market* should determine the desirable level of investment'' and are relying on engineering-based determinations of generation adequacy requirements. This is done through explicit specification of global or locational installed capacity targets (e.g., the Northeastern ISOs) or through publication of a long-term statement of investment opportunity, as in Australia.

The need for regulatory determination of generation adequacy criteria is often justified by the view that supply reliability of electricity is a public good. The various capacity mechanisms vary, however, with respect to how explicit is the regulator in prescribing the level of generation capacity as opposed to providing financial incentives and relying on market forces to provide the desired level.

The two principal capacity payment mechanisms employed to date are capacity payments and capacity obligations. Both approaches are based on the notion that electricity supply is a bundle of two distinct products, capacity (which serves as a proxy for supply reliability) and energy. Hence, customers consuming electricity must pay for the bundle. In the case of capacity payments, generators are paid by the system operator for their available capacity in addition to what they get for selling the energy they produce. Capacity payments can be fixed or can vary according to offered quantity, as was the case in the U.K. system before New Electricity Trading Arrangements (NETA) or in the so-called ''demand function'' approach implemented by the New York ISO [1] and proposed as part of LICAP [2] at the New England ISO. Such payments can also vary by location as in the LICAP proposal.

The basic motivation underlying capacity payments is to provide extra income to incumbent generators who cannot recover their fixed costs through inframarginal energy profits due to suppression of energy prices by regulatory intervention or distortion of prices due to reliability-motivated actions by the system operator. The capacity payments are intended to keep such generators from going out of business. It is also conjectured by advocates of this approach that direct payment for capacity will induce more investment in capacity when such payments exceed the amortized cost of the investment. Considering all the factors and uncertainties involved, calibrating the capacity payment so as to attract the targeted level of reserves is a challenging task which is subject to political manipulation by stakeholders.

The need for regulatory determination of adequacy criteria is often justified by the view that supply reliability of electricity is a public good.

The capacity obligation approach is more direct in the sense that, once the target quantities of generation capacity are determined by engineering considerations, that quantity is allocated as a prorated obligation to the load-serving entities (LSEs) based on their peak load. The LSEs must then procure their obligation share from generators (or curtailable load) bilaterally or through an installed capacity market (ICAP) where prices vary according to supply and demand. The two fundamental problems plaguing ICAP markets have been deliverability and the time step over which the capacity product is defined. Because capacity has been defined as a system-wide product, the procured reserves were not where needed from a reliability perspective. As to the time step, since the product was defined for short durations with a short lead time, the supply has been totally inelastic so the ICAP price exhibited ''bipolar'' behavior, being either at zero or close to the penalty level imposed on LSEs that fall short in meeting their obligation. The LICAP proposal made good strides in addressing the deliverability and bipolar price problems but has the other shortcomings of capacity payments which we discuss below.

Overall, both capacity obligations and capacity payment approaches rely on introducing an artificial product into the market for which there is no natural demand and which customers do not value.

 1040-6190/$–see front matter doi:/10.1016/j.tej.2005.10.003

Consequently, it becomes necessary to create artificial or administrative demand for the capacity product, which distorts the market for energy. Generators receiving capacity payments can be more aggressive in pricing the energy they produce. This in turn may suppress energy prices, making it impossible for generators to recover their capacity costs from inframarginal profits on energy, thus perpetuating the need for the capacity revenues.

From the consumer's point of view, however, paying for capacity provides no explicit value, since the deliverability requirement on the generators is a must-offer obligation which is equivalent to selling an ''at the market'' call option. Such options have zero value and therefore provide no intrinsic value for customers. Furthermore, since selling a capacity product entails no financial liability in the form of liquidation damage for failure to deliver energy, nonperformance penalties are needed which have no relation to the economic damage associated with the nonperformance. In summary, introducing capacity products distorts the market for the energy commodity and is self-perpetuating, making the prospects of ever reaching the Promised Land of deregulation an impossibility. Furthermore, while the capacity product shifts risk away from the generators to consumers, there is no reciprocity, since consumers are still exposed to energy prices that can rise to the price cap.

In the remainder of this article we will outline an approach to assuring generation adequacy which is based on the concepts of risk management and risk-sharing arrangements. Instead of introducing artificial products such as capacity, which are viewed as an anomaly of electricity markets, we will employ market instruments that participants in a mature and functional energy-only market would employ naturally to reduce their exposures and transfer risk between investors and consumers. To account for current market imperfections that stand in the way of the Promised Land and to ensure supply adequacy during this transition period, we propose a temporary backstop hedging obligation, which will function as ''training wheels'' for the market and will eventually become obsolete.

We propose a temporary backstop hedging mechanism, which will function as ''training wheels'' for the market and will eventually become obsolete.

The idea of employing call options as a capacity mechanism has been discussed in several papers and proposals such as [3–6] but it is yet to be implemented. Here we present a refinement of earlier proposals which attempts to address some of the criticisms raised, and outline a workable scheme that deals with market realities in the U.S. restructured electricity industry. We also contrast some of the main features of this proposal with the controversial LICAP approach proposed by the New England ISO and endorsed by FERC.[1]

We will first describe the mechanics of the proposed approach and its key features and then highlight the differences and similarities between our proposal and LICAP.

II. The Backstop Call Options Mechanism

We take the position that generation capacity is not a distinct commodity but should rather be viewed as *insurance* against energy supply shortfalls that could result in involuntary outages or in energy price spikes. To rationalize some of the choices we make in the proposed approach we first outline what we consider to be desirable properties of such an insurance mechanism.

A. Capacity mechanism desiderata

- The mechanism should replicate investment incentives in a functional energy-only market with forward contracting and hedging;

• The mechanism should facilitate risk sharing between consumers and producers;

• The mechanism should provide intrinsic value to consumers in exchange for risk sharing (not a subsidy to generators);

• The implicit value of capacity should reflect energy market consequences;

• The mechanism should incent new investment and enable direct participation by new entrants;

• The mechanism should provide stable income to generators to reduce cost of capital in exchange for relinquishing windfall profit potential;

• The mechanism should incent performance and have meaningful penalties for nonperformance;

• Implicit capacity value should reflect locational difference in investment needs due to constraints;

• Generators should be able to opt out and be rewarded for risk taking by higher spot price income potential;

• Load should be able to opt out (self-insure) by avoiding capacity payments in exchange for taking spot price or curtailment risk;

• The mechanism should not interfere or displace private risk management practices;

• Market participants should be allowed to self-provide their insurance needs through bilateral contracts;

• The mechanism should attempt to address credit problems associated with insurance needs on the demand side and investment financing on the supply side;

• Since the load is the ultimate beneficiaries of supply adequacy assurance, any obligations imposed on LSEs should reflect customer base (follow the load); and

• Regulatory interventions to support an insurance mechanism should have an automatic sunset or easy phase-out once the market provides sufficient insurance through contracting and load response (the Promised Land).

The mechanism should provide stable income to generators to reduce the cost of capital in exchange for relinquishing windfall profit potential.

B. Energy call options basics

An energy call option is a financial instrument that specifies quantity, delivery time, delivery location and a strike price for energy and it gives its holder the right but not the obligation to obtain the specified energy at the specified strike price. A call option can be exercised physically or financially. Physical exercise entails delivery of the contracted energy by the counterparty (by scheduling it through the ISO as a bilateral contract). Financial exercise entails a financial settlement where the counterparty pays to the holder the difference between the locational spot price and the strike price for the contracted amount. As a practical matter, an option can be exercised physically by imposing a ''must offer'' obligation (on the counterparty) into the ISO spot market, purchase of energy by the option holder from the spot market, and a financial settlement of the call option between the two parties. To insure the possibility of physical exercise, a call option must have physical cover, meaning that it is tied to a specific present or future physical resource which will be able to deliver the contracted energy. In normal risk management practices, bilateral forward contracts and options do not have physical cover, which allows the volume of risk hedging activities to exceed the actual volume of physically delivered energy. This capability improves market liquidity and contributes to the efficiency of the energy market. For the purpose of ensuring supply adequacy, however, we will restrict ourselves to the use of call options with physical cover which can be provided either through physical generation capacity that will be available on line at the delivery time and location or through verifiable load curtailments.

For a given strike price, the settlement value of a call

1040-6190/$–see front matter doi:/10.1016/j.tej.2005.10.003

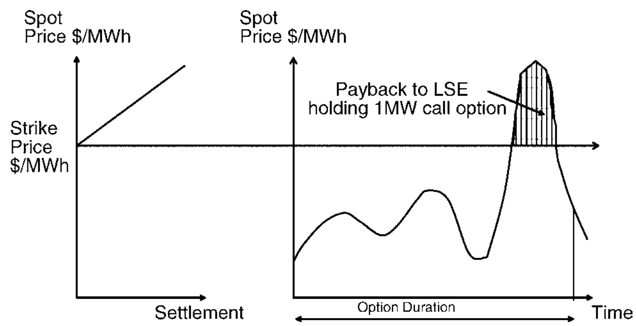

Figure 1: Payoff of an Energy Call Option Strip

option is given by the larger of zero or the difference between the spot price and the strike price. In the context of electricity markets, call options are traded as strips which cover continuous delivery (rather than a single time slot) at a specified delivery rate. For the purpose of this discussion, we will use the term monthly or annual call option to mean a call option strip covering the corresponding period. **Figure 1** illustrates the settlement value of a call option strip for one MW. An LSE holding such a call option issued by a generation company can purchase energy at the spot market price and collect the settlement from the generation firm. The call option acts as price insurance, guaranteeing that the LSE will not pay more than the strike price for the energy it insured. To obtain such insurance the LSE paid a fixed premium to the generation firm. The generation firm, on the other hand, which received a fixed upfront payment for the call option it sold, must surrender the windfall profits it made (above the strike price) when the spot price spiked. This arrangement serves as a risk-sharing mechanism which enables consumers (or the LSE representing them) to reduce their exposure to spot price risk by assuming from the generators some of the fixed-cost recovery risk.

Load-serving entities which have to serve core customers at fixed or regulated prices while purchasing power in a competitive wholesale market are typically exposed to price and quantity risk. The inherent correlation between load and prices (typically wholesale prices rise when load increases) amplifies the revenue exposure of an LSE that covers its expected load through forward contracts while relying on the spot market to adjust its positions according to its realized load. When load is at a low level, an LSE is most likely over-contracted so while its revenues are low due to low sales it may also find itself selling its excess energy at spot prices below its contract prices. On the other hand, when demand is high so are spot prices. Thus, the LSE would need to cover its unhedged portion of the realized load at high cost, which most likely exceeds the fixed rates it can charge its customers. In order to hedge such revenue risk, an LSE could utilize a portfolio of forward and option contracts as described in [7,8].[2] Thus, in a mature and competitive energy market where prices are unencumbered by caps and market power mitigation schemes, energy call options are natural means for LSEs to use to insure themselves against spikes in load and prices.

The use of the call option for managing energy prices and quantity risk is further facilitated by the introduction of variants such as swing options and spark spread options. Swing options are call options that give the holder some flexibility in selecting the exercise times. An interruptible service contract that specifies the strike price and a total number MWh that can be interrupted within a month at the supplier's discretion, when the spot price exceeds the strike price, is an example of a swing option. Spark spread call options are options on the price difference between the spot price for electricity (Se) and the heat-rate-adjusted spot price for natural gas ($Hr \cdot Sg$). A spark spread call option strip with a zero strike price is the financial analog of gas turbine capacity as illustrated in **Figure 2**. It accounts for the real option available to the plant owner to turn it off when the fuel cost exceeds the price of the electricity produced. Spark spread options are useful instruments for trading investment risk and capacity scarcity risk between producers and consumers and can be used for valuation of generation assets (see [9]). Since the seller of a spark spread call option does not assume fuel cost risk,

 1040-6190/$–see front matter , doi:/10.1016/j.tej.2005.10.003

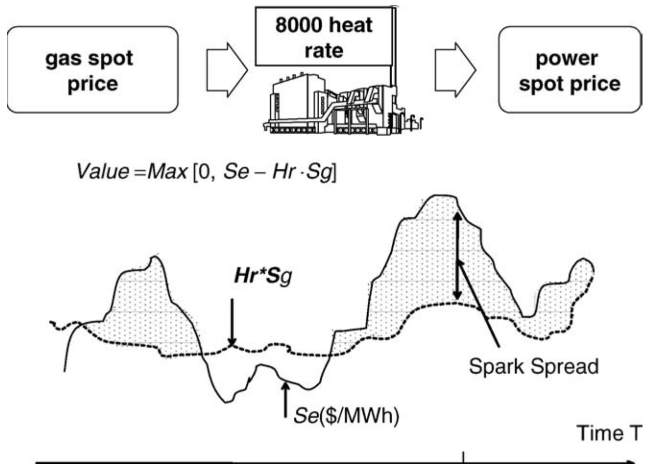

Figure 2: A Spark Spread Call Option Replicates Gas Turbine Capacity Value

such an option would be cheaper than the corresponding electricity call options.

C. The mechanics of backstop call option obligations

In theory an energy-only market should suffice to induce risk management and contracting practices that will produce investment incentives when needed. However, hedging practices are falling short due to market imperfections and the regulatory interventions detailed above. We propose to rectify that situation by employing call options, which, as described above, are a natural hedging instrument that an ideal energy only market would employ to trade risk between producers and consumers. To address the apparent market failure, we propose to impose on LSEs a backstop hedging requirement that can be viewed as mandatory insurance akin to mandatory liability insurance imposed on automobile drivers. The proposed mechanism would require LSEs to hold amounts of physically covered call option strips or forward contracts in proportion to their forecasted peak load within the delivery period.

1. *Selecting the strike prices*

In general a call option obligation can take the form of a portfolio of options with staggered strike prices. Such an approach has some advantages from a market power mitigation perspective as discussed in [7], since the call option portfolio effectively produces demand elasticity in the spot market. Indeed, participants in a mature market would employ a variety of hedging instruments with different strike prices for managing their risk. In this article, however, we propose the use of a single strike price so as to simplify the implementation of a transitional regulatory intervention. Furthermore, we envision the role of the call option obligation as a backstop hedge, rather than as a market power mitigation device that will be ''in the money'' and affect spot prices only on rare occasions. Thus, a single strike price is likeliest to achieve the set objective of a backstop hedge. The specific value of the selected strike price will be treated as a regulatory policy parameter much like the selected level of reserve requirement and the breakpoints in the demand function employed in a ''demand function'' approach. The determination of that strike price should be guided by several considerations, which we discuss below.

Contracting obligations imposed on LSEs, like any other capacity mechanism, must be designed to minimize interference in the energy market and in the voluntary risk management practices of market participants. This objective can be achieved by restricting contracting obligations to call options with a strike price that is sufficiently high (probably not lower than 50 percent of the offer caps imposed by most ISOs) to provide a backstop hedge – rather than replacing the bilateral contracts that LSEs would otherwise use for risk management. A high strike price also ensures that the options will be ''out of the money'' most of the time and hence their cost will be relatively low. Further reduction in the cost of the call options can be achieved by defining the call option in terms of the spark spread (as discussed above) so that the generator does not bear the fuel cost risk.

Because a call option provides a right but not an obligation for the LSEs to buy the contracted amount of energy, it can be used to secure reserve capacity in excess of forecasted peak demand. In particular, any wholesale customer or LSE should be required to carry call options that will cover its peak load forecast within the covered delivery period, plus adequate reserves as set by the regulator. In order to insure deliverability, the call options must be backed by existing generation capacity, or by a commitment to invest in generation capacity that will be available by delivery time, or by verifiable interruptible load contracts. Bilateral contracts held by an LSE or a wholesale customer for the covered period, provided the contract's energy price is below the mandated backstop strike price, can be used to meet the call option obligation.

On the other hand, it is important that the backstop strike price of mandatory call options be significantly below the offer cap or price cap for energy in the spot market. Maintaining such a gap serves several objectives. First, it provides a natural economic penalty for nonperformance by the generator through the financial liability entailed by the option for the price difference between the energy clearing price and the strike price for undelivered energy. For a call option covered by interruptible load, the strike price of the option sets a penalty for interruptible load that wishes to override its curtailment. In such a case the load will be liable for the difference between the spot price and the strike price for the energy it uses. Such penalties can be imposed in addition to any other nonperformance penalties such as forfeiture of the option premium.

As a practical matter, generation capacity providing physical cover to a call option would have a must-offer obligation at any price at or below the price cap. If it is called and the spot market clears above the strike price, generation firms and the LSE whose call option obligations they covered, will settle the differences financially. Thus, generators that received a capacity payment in the form of a call option premium must refund to the customers (through the LSE) any inframarginal profit resulting from spot prices that exceed the strike price. If the LSE's call option coverage exceeds its load, such refunds may produce a net speculative profit that could defray the cost of meeting the hedging obligation. On the other hand, if the LSE is not sufficiently covered, i.e., is short in meeting its obligation, it will be exposed to the difference between the spot price and the backstop strike price in addition to other penalties that may be imposed on it for failing to meet its obligation. Specifically, if the strike price is set at \$500/MWh and a generator that sold the option is not available when the spot price rises to \$1,000/MWh, the nonperforming generator will be financially liable for \$500 for every MWh of undelivered energy.

2. *Opt-out provision for generators*

The gap between the backstop strike price and the energy price cap also enables differentiation between generators that sold call options on their capacity and those generators that did not. Generators that did not sell call options should be allowed to set energy prices up to the energy price cap and sell their energy above the backstop strike price if their energy is needed either due to unforeseen high demand or nonperformance of call option sellers. These situations are particularly likely in regions with high hydro concentration that are prone to occasional dry years. Hydro generators that collect capacity payments through some capacity mechanism must have an incentive to reinsure their delivery obligation through contracts with thermal plants. Such incentives are provided by allowing thermal plants who did not receive capacity payments to set the price above the backstop strike price, and by holding the hydro generators who received capacity payment liable for the price difference between the spot price and the backstop strike price.

Setting the strike price significantly below the price cap makes sense when the price cap is sufficiently high to provide headroom for such a two-tier approach. When electricity prices are artificially suppressed by a low cap as in California (currently $250/MWh) the call option value is also depressed due to limited price volatility. Hence, selling call options may not generate sufficient income to support investment in generation capacity. It would make sense in such an environment to set the strike price to the level of the current cap and raise the offer cap on generators that do not sell call options. Such an approach would effectively maintain the original low cap on energy prices as long as the aggregate backstop call option volume (corresponding to peak load plus reserves) is sufficient to meet the load and operating reserve requirements. In the rare instances where additional generation capacity is needed beyond the planning reserves either due to extreme load conditions or outages, energy procured from uncontracted capacity (typically out-of-state imports) will be subject to the new high cap and not to the low cap, which is now used as a strike price on contracted capacity. This method rationalizes the *de facto* emergency procedure of the CAISO during the California crisis and afterwards, that allowed out-of-market energy procurement at high prices (above the price cap) to avoid involuntary curtailment.

3. Demand response and opt out provision for load

The proposed mechanism provides an easy way for load participation in supporting supply adequacy or to opt out from paying for supply adequacy insurance. To do so, firm load can offer a call option at the backstop strike price or below, which should be counted against reserve requirements. Such a call option is effectively an interruptible service contract allowing the system operator to curtail the load when the spot price exceeds the strike price. By subscribing to such interruptible service, a customer forgoes the insurance coverage provided by generators selling the call options and its share of the insurance cost is offset by the income from selling a call option which is covered by its curtailable load.

4. The quantity of backstop call options obligations

The differential energy caps (the global cap versus the backstop strike price) between contracted capacity and uncontracted capacity eliminates the need to increase the quantity of call option obligation beyond the target capacity level (as in the LICAP approach). Under the proposed scheme, any excess capacity beyond the target level will either be contracted on a voluntary basis or be allowed to recoup the option value by selling on the spot market at the market clearing price rather than being limited by the strike price. In either case, such excess capacity will lower the price-duration curve and reduce the intrinsic value and market price of the mandatory call options. Thus, the downward-sloping effect provided in NYISO and the NEISO LICAP proposal via an artificial "demand curve" for capacity is provided in our scheme through the profit opportunity given to generators who do not sell options and instead take their chances on inframarginal profits from selling energy above the strike price. This aspect will be further elaborated below in the subsection discussing the valuation of backstop call options.

5. Contract duration and lead time

Contract duration is a key element in developing a capacity mechanism based on contract obligations. Proponents of short-term capacity products like ICAP argue that a stable income stream to incumbent generators for their installed capacity will provide the right price signals to new entrants to invest in capacity. According to this paradigm, potential entrants play a passive role in the capacity market. Alternatively, capacity products and call options with a long lead time enable direct participation by investors who may sell such products against investment commitments. The current wisdom advocates a two- to three-year lead time for capacity products – long enough to enable new investors to participate in the market by offering ICAP products or contracts covered by generators in the planning stage. Enabling active participation by entrants will attract capital and mitigate market power in the capacity market.

The main obstacle and source of opposition for long-term capacity products comes from LSEs and retail energy providers, who argue that long-term contractual obligations are inconsistent with a competitive retail environment where customers are allowed to switch providers. In principle this should not be a problem, since the call option obligations can be based on a three-year forecasted peak but the obligation can be adjusted monthly based on current load. A secondary market for the three-year call options would allow their holders to trade them so as to adjust their holding according to their load. The prices of the call options in the secondary market would fluctuate in the same way as the prices of 30-year Treasury bonds fluctuate on a daily basis to reflect supply and demand. An alternative central procurement approach, which addresses the credit problem associated with holding long-term call options, will be discussed later.

6. Locational differentiation

In a system where spot prices are locational (LMP), backstop call option obligations with a fixed strike price will naturally provide locational differentiation that yields higher premia for call options that cover load at locations where the expected LMP is higher. All that is needed is a deliverability requirement that makes the obligation location-specific to reflect the local load served by the LSE. The granularity of such a delivery condition can be zonal or hub-based, in which case the financial settlement of the call option would be determined by a weighted average of the differences between the LMP and the strike price over the zone or the hub. Such a settlement will reflect local scarcity due to load pockets and transmission constraints, so that the call option premia that are based on the expected settlement would also reflect such local scarcity. In other words, call options will be more expensive at locations where the expected LMP is higher due to local scarcity of capacity. Higher option prices will in turn incent more investment in generation or more load response in these locations. Note that such locational differentiation can be achieved by simply making the hedging obligation reflect local constraints, while maintaining the same backstop strike price, which simplifies the design and policy decision of determining that strike price. This should be contrasted with the LICAP approach, where the demand curve parameters vary by location and have been subject to endless negotiations among stakeholders.

7. Self-provision through normal risk management practices

One of the key features of the proposed approach is the ability of LSEs to meet all or part of their backstop call option obligation through normal risk management practices. This feature is the basic mechanism that will lead to eventual obsolescence of the regulatory intervention that enforces the call option obligation. Basically, any bilateral forward or option contract that covers the delivery period and is covered by physical resources qualifies to meet the equivalent portion of the backstop call option obligation. Given that the backstop strike price is sufficiently high, it is reasonable to assume that it will exceed the contract price or strike price used by the LSE as part of its normal hedging portfolio. The important qualification, however, is the physical cover requirement. Thus, bilateral contracts covered only by liquidation damages do not qualify towards meeting the backstop call option obligation. Furthermore, if the obligation is locational, then a bilateral contract used to meet the obligation must be deliverable to the particular location either by being covered with a resource at the same location or by being supported with the appropriate long-term transmission rights.

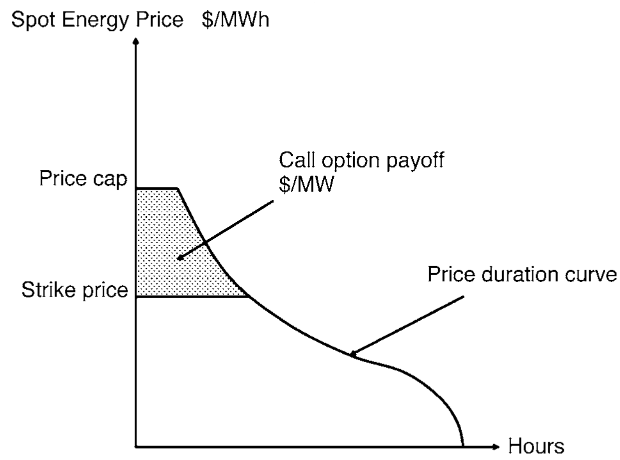

Figure 3: Expected Payoff of a Call Option Strip

III. Valuation of Backstop Call Options and the Implicit Demand Function for Capacity

The expected payoff of an electricity call option strip over a given time interval can be determined from the corresponding price duration curve describing the cumulative time over which the spot price exceeds any given level. The expected payoff of the option is represented by the area under the price duration curve and above the strike price as illustrated in Figure 3.

The price duration curve depends, however, on the available capacity and is expected to be lower as capacity increases, which in turn lowers the expected call option payoff. If we draw a family of price duration curves for different levels of available capacity, then, keeping the price cap and strike price fixed, we can derive from it the expected call option payoff, which declines with available capacity (Figure 4).

The downward sloping function determining the expected option value as a function of available capacity is supported by the opportunity cost of generators who can sell their uncommitted energy at the spot price and by the expected settlements for the call options. If the total available capacity exceeds the target capacity upon which the aggregate obligations are based, then some capacity will remain uncontracted. The availability of such excess capacity,

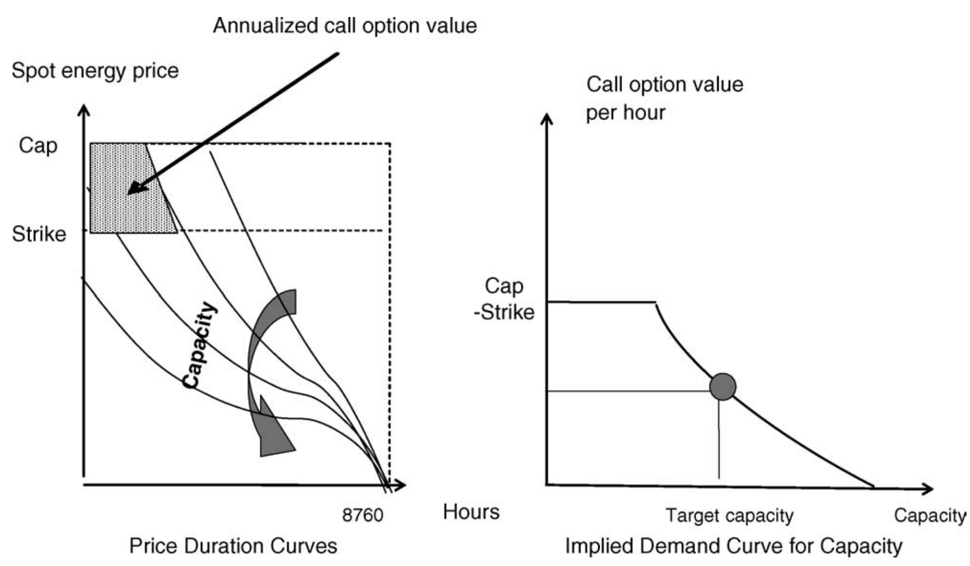

Figure 4: Expected Call Option Value as Function of Capacity

 1040-6190/$–see front matter doi:/10.1016/j.tej.2005.10.003

however, puts downward pressure on spot energy prices and consequently on the market price for the backstop call options. Thus, unlike capacity markets employing demand functions for capacity, in our proposed approach only the target capacity will receive capacity payments (as call option premiums) from the load but the payment will still reflect the total available capacity and decline with quantity.

IV. Central Procurement

As indicated earlier there is a fundamental time step incompatibility between the need for long-term capacity products or call options that will serve the need of new investors on the supply side versus the short-term commitments supported by the load side. Credit problems associated with imposing a long-term contracting obligation on load-serving entities represent one of the main challenges to such an approach. This obstacle is particularly relevant in systems like ERCOT, where retail competition is thriving due to the presence of many small retail energy providers (REPs) that find the credit burden of a long-term call option obligation prohibitive. Such objections would apply to any forward-looking capacity mechanism that is designed to enable direct participation by new entrants and not just to the backstop call options approach. It should also be noted that in Australia, which has an energy-only market and where contracting is voluntary, customers are protected against risky overexposure of their retail suppliers through rigorous credit requirements imposed on LSEs. Such measures are needed in order to prevent small retailers from leaning on the volatile spot market and using the bankruptcy option as a hedge.

One way to resolve the credit problem and the time step incompatibility is through centralized procurement of backstop call options by the system operator, which can then pass the cost to the LSE on a share-of-load basis. Such an approach effectively places the burden of supply adequacy insurance and the associated credit requirements directly on the load while the system operator serves as an intermediary between the generation companies providing the insurance and the LSE representing the load. This approach is similar to the shelved central resource adequacy market (CRAM) proposal [10] developed as a blueprint of a joint capacity market for NYISO, NEISO, and PJM. However, our centralized procurement is designed around backstop call option obligations rather than an ICAP product as in the CRAM proposal. As discussed earlier, the call option obligations facilitate self-provision and demand-side participation, which will pave the way for eventual sunset of the regulatory intervention and passage to the Promised Land.

Under the proposed centralized procurement alternative, the system operator would procure in an annual centralized auction physically covered backstop call options as described earlier. The procured quantity would be based on the forecasted annual peak. The system operator will underwrite the procurement of the options and guarantee payment but the sellers will receive payments during the performance period, which will be passed through to the load (through the LSEs). The cost of procurement can be distributed over the year based on a loss-of-load probability (LOLP) calculation and allocated to the load on a *pro rata* basis. Holders of bilateral contracts can self-provide their call option obligation by offering call options into the central procurement auction against their bilateral contracts (assuming that the price, time period, and physical cover meet the backstop call option criteria). This scheme works like an ancillary service market for a reserve with a three-year lead time.

  doi:/10.1016/j.tej.2005.10.003

V. Conclusion

We close by highlighting some of the key features of the proposal articulated in this article and comparing it to capacity mechanisms that rely on an administratively set demand function, and in particular the LICAP proposal. An important aspect of our approach is reliance on standard hedging instruments that a mature energy-only market would support without regulatory intervention. Thus, the regulatory enforcement is viewed as a temporary necessary evil to correct for market imperfections. Unlike payments for an artificial capacity product, for which there is no natural demand, energy call options provide intrinsic value to customers, since the generators who are paid for such options must in return pay back windfall profits due to spot prices in excess of the strike price of the options. Thus, the payment to generators for call options represents a risk trading arrangement where the consumers assume some of the investment risk in exchange for reducing their price risk. Contrast this with the LICAP proposal, in which generators that are paid for capacity must refund their windfall profits. Because these repayments are smoothed out based on historical performance, customers will not see a direct price break on energy in return for the supply adequacy insurance premium they pay. Such smoothing distorts price signals and incentives for demand-side participation.

The obligation of generators and curtailable load supplying call options to supply or forgo profits by supplying energy at a strike price below the potential market price provides a natural penalty for nonperformance by holding the option providers financially liable for the difference between the market clearing price and the strike price. Such liability is, of course, moot if the option is not in the money, i.e., the spot price is below the strike price. The LICAP proposal attempts to achieve a similar affect by imposing nonperformance penalties for unavailability during a small number of shortage hours which roughly correspond to the hours when a call option would be in the money. In the LICAP proposal, however, the penalty is in the form of forfeiting a prorated share of the capacity payment received by the generator rather than being liable for the cost of unprovided energy. Hence, under the LICAP approach, generators do not assume downside risk in exchange for the capacity payment they receive because the nonperformance penalty they are exposed to can never exceed the payment they obtained.

The ability of load to opt out of paying for supply adequacy insurance is an important feature of our proposal, which incents demand response, an essential element of a healthy market. This is accomplished by enabling load to offer call options which are equivalent to entering into a curtailable load contract that allows the supplier to curtail the load when the spot price reaches the strike price. The payment received by load offering such call options will exactly offset their share of the LSE obligation cost.

The expected value function of backstop call options displays a similar shape to the administrative capacity payment functions (better known as demand functions for installed capacity) employed by the NYISO and in the LICAP proposal. In our approach, these functions, which decline with the amount of available capacity, are market-based, resulting from the opportunity cost of generators that are not contracted to sell their power at prices in excess of the strike price. Enabling generators to opt out from selling call options and collect higher energy prices provides an important market force that regulates the prices of the backstop call options. If the market price of the options is low, generators will opt out and take their chances on recovering their fixed costs from energy sales at prices above the strike price. Allowing uncommitted capacity

 1040-6190/$–see front matter doi:/10.1016/j.tej.2005.10.003

to charge higher prices will enlist imports and speculative investment in excess of the target capacity at no upfront cost to consumers. Nevertheless availability of capacity in excess of that dictated by reliability criteria lowers the expected price of the call options as in the case of the downward sloping capacity payment, so that consumers can ''have their cake and eat it too.'' The latter is an important distinction from the LICAP proposal, where customers are required to pay for capacity procured in excess of the target capacity, whether needed or not. This discourages self-provision, since an LSE whose load is fully covered for peak energy and reserves according to the reliability target may still be charged for excess capacity procured by the system operator. By contrast, the call option obligations never exceed the target capacity and payment for excess capacity will only occur as energy payments in the unlikely event that such capacity is actually needed. Even if the owners of unsubscribed capacity exercise market power and raise prices up to the offer cap, the impact on consumer prices will be minimal, since most of the load and operating reserves are capped by the call options strike price.

Locational differentiation in the value of call options, which provides incentives to generation capacity at needed locations, is achieved automatically by enforcing locational deliverability on the call options

Ultimately, demand response and risk management practices will evolve to the point where the Promised Land can be reached.

obligation. The price differences between are driven by the locational marginal prices which determine the call option settlements.

The only design parameter in the proposed call option obligation scheme is the strike price, which determines how much of the spot price risk is assumed by generators and how much of the fixed-cost recovery risk is assumed by consumers. In a mature market, that decision is left to the market participants and can differ across parties in a similar manner in which individuals purchasing car insurance can select their level of deductible. However, for a backstop call option obligation, it is simpler to impose a uniform strike price (of course, the obligation can be met with call options having lower strike prices). The strike price level in this scheme will be a policy choice similar to the choice of parameter for the LICAP demand function. The overall outcomes of supply adequacy and cost to customers should not be very sensitive to the strike price level since it is a matter of ''pay now or pay later.''

Finally, our proposal recommends a term and lead time for the call options that will allow contestability of the call option prices by new entrants. We also describe a centralized procurement approach that will resolve credit problems associated with long-term hedging obligations and monitoring of compliance by the LSEs. These features can be implemented with capacity products such as ICAP and LICAP and not just with backstop call option obligations.

The vision supported by this article is that ultimately demand response and risk management practices will evolve to the point where the Promised Land can be reached and an energy-only market with minimal regulatory intervention will support innovation and investment. The proposed approach is intended to provide a safe passage to that Promised Land without creating self-perpetuating regulation and entitlements.■

References

[1] FERC, Docket No. ER03-647-000, New York Independent System Operator, Inc., Report on Implementation on ICAP Demand Curve, 2004.

[2] Peter Cramton, Stephen Stoft, *A Capacity Market That Makes Sense*, Elec. J., Aug./Sept. 2005, at 43–54.

[3] Shmuel S. Oren, *Capacity Payments and Supply Adequacy in Competitive Electricity Markets*, in: Proceedings of the VII Symposium of Specialists in Electric Operations and Expansion Planning (SEPOPE VII), Curitiba, Brazil, May 21–26, 2000.

[4] Carlos Vazquez, Michel River, Ignacio Perez Arriaga, *A Market Approach to Long-Term Security of Supply*, IEEE Trans. Power Syst. 17 (2) (2002) 349–357.

[5] Teknechron Energy Risk Advisors (TERA), *A Revised Framework For the Capacity Charge, Minimos Operativos, and Rationing Rules*, World Bank Colombian Electricity Project Final Report, Vol. 3, Feb. 28, 2001, Austin.

[6] Shmuel S. Oren, *Ensuring Generation Adequacy in Competitive Electricity Markets*, in: M. James Griffin, Steven L. Puller (Eds.), Electricity Deregulation: Choices and Challenges, Univ. of Chicago Press, 2005 (Chapter 10).

[7] Hung-Po Chao, Robert Wilson, *Resource Adequacy and Market Power Mitigation via Option Contracts*, in: Proceedings of POWER Ninth Annual Research Conference on Electricity Industry Restructuring, Berkeley, CA, Mar. 19, 2004.

[8] Yumi Oum, Shmuel S. Oren, Shijie Deng, *Volumetric Hedging in Electricity Procurement*, in: Proceedings of the PowerTech 2005 Conference, St. Petersburg, Russia, June 27–30, 2005.

[9] Shijie Deng, Shmuel S. Oren, *Incorporating Operational Characteristics and Startup Costs in Option-Based Valuation of Power Generation Capacity*, Probability Eng. Info. Sci. 17 (2) (2003) 155–182.

[10] NERA Economic Consulting, *Central Resource Adequacy Markets for PJM, NY-ISO and NE-ISO, Final Report*, Feb. 2003.

Endnotes:

1. Implementation of the proposal has been recently delayed by FERC due to political pressure from state governors and legislators in the NEISO states.

2. Alternative approaches such as the use of weather derivatives that are based on the correlation between load and temperature can also be used to hedge volumetric and revenue risk.

 1040-6190/$–see front matter , doi:/10.1016/j.tej.2005.10.003