Revesion 03/18/04

Resource Adequacy and Market Power Mitigation via Option Contracts

Hung-po Chao and Robert Wilson
Electric Power Research Institute and Stanford University

Resource Adequacy and Market Power Mitigation via Option Contracts

Hung-po Chao and Robert Wilson
Electric Power Research Institute and Stanford University

Abstract

This paper is one of a pair in which we propose a unified scheme for utilities' risk management in wholesale and retail markets. This first paper focuses on wholesale markets and issues that arise in current efforts by state PUCs to specify implementations of resource adequacy obligations. It demonstrates the advantages of a portfolio of long-term contracts, each written as an option with a specified "strike price" for a specified quantity callable at that price. The second paper shows that, from its costs of purchasing the spectrum of options in its portfolio, a utility can infer the appropriate retail price for nearly any retail service plan for core and non-core customers – including those that provide compensation when the customer is required to curtail its load.

All long-term contracts take advantage of the common interests of utilities and suppliers to mutually insure each other against subsequent price variation in spot markets. They can also fulfill the ISO's objective of ensuring adequate supplies of energy and reserve capacity, and the PUC's aim of reducing utilities' cost of capital by reducing their exposure to financial risks. An important effect of long-term contracting is that it transfers transactions from spot markets to forward markets where the elasticity of supply is greater because the longer time frame enables investments in capacity. That is, forward markets are more contestable and thus dissuade an incumbent supplier from inflating its bids above its competitors' long-run incremental cost of capacity expansion or the costs of entry by new suppliers.

Compared to a contract that specifies a fixed quantity, an option has the advantage of reducing quantity risks by enabling a utility to purchase power at the strike price only when it is needed and the spot price exceeds the strike price. We especially emphasize the further advantage of options that they directly suppress price volatility and mitigate market power in spot markets. Indeed, the spectrum of strike prices can be designed to

implement a load-contingent cap on the spot price; that is, a price cap that is higher when the load is greater. A portfolio of options does this because, at each spot price, the demand net of the options callable at lower strike prices has greater price elasticity than aggregate demand. Even if aggregate demand is completely inelastic, the elasticity of net demand can be made arbitrarily large by specifying an appropriate portfolio of option contracts. To illustrate, we provide a simple model that demonstrates that the typical PUC objective of maximizing consumers' surplus subject to non-negative profits for suppliers implies that utilities should obtain a positive fraction of their energy purchases via options.

The paper addresses the practical aspects of implementing this approach to resource adequacy. In particular, we propose an annual auction of a specified quantity of multi-year option contracts at each strike price in a specified range. Each contract is an option on physical capacity since it requires the supplier to back the contract with available capacity, to submit a standing bid at the ISO for the contracted quantity at a price no higher than the strike price, and to be dispatchable for either energy or reserve capacity. Thus, even though the option contracts might be tradable in secondary markets, they are not solely financial instruments.

Resource Adequacy and Market Power Mitigation via Option Contracts

Hung-po Chao and Robert Wilson
Electric Power Research Institute and Stanford University

1. Introduction

Our purpose is to propose a unified scheme for utilities' risk management strategies in wholesale and retail markets. This first paper focuses on wholesale markets and issues that arise in specifying resource adequacy obligations. It demonstrates the advantages of a portfolio of long-term contracts, each written as an option with a specified "strike price" for the quantity that is callable at that price. Our companion paper shows that the costs of options imply corresponding retail prices for service plans.

All long-term contracts take advantage of the mutual interests of utilities and suppliers in insuring each other against subsequent price variations. They can also ensure adequate supplies of energy and reserve capacity. Long-term contracting transfers utilities' purchases from spot markets to forward markets where the elasticity of supply is greater. Because the longer time frame enables investments in capacity, forward markets are more contestable and thus prevent incumbents from inflating their bids above the long-run incremental cost of capacity expansion.

Compared to a contract that specifies a fixed quantity, an option reduces quantity risks by enabling a utility to purchase additional power only when the load is high and the spot price exceeds the strike prices of some of its options. A further advantage of options is that they directly suppress spot price volatility and mitigate market power. We show that the spectrum of strike prices can be designed to implement a variable price cap; i.e., one that allows a higher price when the load is greater. This possibility derives from the fact that net demand – i.e., aggregate demand net of the quantity callable via options with strike prices below the spot price – has greater price elasticity than aggregate demand. Even if aggregate demand is completely inelastic, the elasticity of net demand can be made arbitrarily large by specifying an appropriate portfolio of option contracts. Using a stylized model we show that there is a range of cost parameters in which maximization of consumers' surplus (subject to non-negative profits for suppliers) implies that a positive fraction of energy purchases should be obtained from options.

Section 2 reviews relevant background. Then Section 3 provides an overview of the basic concepts and Section 4 addresses practical aspects of implementation.[1] Section 5 illustrates the basic ideas with a simple theoretical model.

2. Background

At the state level, the argument for restructuring wholesale electricity markets was to strengthen incentives for efficient investments and facility management. This objective stemmed from dissatisfaction in some states with the burdens of cost recovery for nuclear plants, and with the prices specified in long-term contracts with "qualified facilities" [QFs] under the terms of the federal Public Utility Regulatory Policy Act enacted in 1978. Before California initiated restructuring in 1994 the issue was symbolized by the disparity between average retail prices of about 12 cents/kWh and wholesale prices below 3 cents/kWh at the Oregon border. Proponents conjectured that retail prices could be reduced to 6 or 7 cents/kWh if utilities procured their supplies in wholesale markets – after paying off "stranded costs" from previous investments. The federal Energy Policy Act of 1992 had strengthened prospects for competitive wholesale markets; in particular, the Federal Energy Regulatory Commission's [FERC] subsequent Order 888 established requirements for open access to transmission and nondiscriminatory pricing of transmission services. The existence of wholesale markets in Alberta, Australia, Scandinavia, and England added encouragement. Moreover, FERC's Order and subsequent directives allowed that a transmission system can be managed by a nonprofit

[1] The scheme described here builds on previous work by Robert Entriken and Stephen Wan, and takes account of comments from Shmuel Oren and Alex Papalexopoulos. Portions of their work are described in Report 1007755, "Pushing Capacity Payments Forward: An Agent-Based Simulation of Available Capacity Markets," Electric Power Research Institute, December 2003. This report is available at the URL http://www.epriweb.com/public/000000000001007755.pdf

System Operator [SO]. In addition to engineering management of the grid, the SO can conduct markets for energy, reserves and congestion management, but it must be independent of participants in the energy markets.

Subsequent experience with wholesale markets differed markedly among the states that restructured. The SOs in the northeast (New England, New York, Pennsylvania-New Jersey-Maryland [PJM]) inherited the organization and procedures of pre-existing power pools that maintained tight control of dispatch, including optimization of unit commitment and scheduling for energy and reserves; also, NY and PJM used local marginal pricing [LMP] at the "nodal" or bus level to account for the cost of redispatch to alleviate congestion. In contrast, the SO in California relied on highly decentralized markets (including a separate power exchange, and a separate market for procuring reserves), self-scheduling of units, and zonal pricing that reflected congestion only across major interfaces. Soon after operations commenced in 1998, California encountered high prices in its reserve markets because procurement requirements were defined inflexibly. A year later, further problems occurred when some suppliers exploited the "dec game" in which, in effect, they were paid for alleviating congestion that they themselves caused by scheduling excessive transfers over intra-zonal transmission lines. California also had continuing problems ensuring provision of must-run generation for local reliability under long-term contracts. These adverse experiences illustrated that wholesale markets must be carefully designed when the SO has weak control of dispatch and the utilities have divested most of their dispatchable generation. Deficiencies of the California market design were most evident in the California crisis of 2000-2001. The crisis was precipitated by reduced imports from hydro sources in the northwest due to prolonged drought, later by scarcity of fuel supplies and emission permits for gas-fueled generators, and finally by collapse of the utilities' financial resources. Nevertheless, the crisis was exacerbated by the thermal generators' ability to raise spot prices further. The financial consequences for utilities were severe because they were fully exposed to the high spot prices. California's restructuring allowed a utility to purchase supplies without *ex post* prudency reviews only if they were bought in the power exchange's day-ahead market or the SO's real-time market – and late in the crisis, in the power exchange's block-forward market, but the utilities did not use this option to any large degree.

Although there were many factors contributing to the California crisis, we focus here on one salient lesson that became a major ingredient of subsequent regulatory policies at the federal and state levels. This lesson has three components that work together to sustain a viable wholesale energy market:

(1) ample supply-side generation,

(2) substantial use of long-term contracts for energy purchases, and

(3) stringent measures to mitigate market power in spot markets.

These components are addressed in FERC's Standard Market Design [SMD] via procedures for strengthening resource adequacy, a dominant role for bilateral contracting, and measures for market-power mitigation. The SMD includes some specific proposals but ultimately FERC allows each SO to design its own markets for installed or available capacity [ICAP or ACAP] and its procedures for automatic market-power mitigation [AMP]. The dominant role of state regulators in choosing how to implement ACAP stems from the states' exclusive jurisdiction to authorize cost recovery via charges added to retail electricity rates.

Our purpose here is to describe a single scheme that addresses all three components simultaneously.

3. Using Option Contracts for ACAP and AMP

The scheme uses long-term contracts in the form of options on capacity. Each option has a strike price at which a utility or other load-serving entity [LSE] can exercise the option, and thus purchase energy supplies at the strike price. In the long-term forward market, a generation company that sells options to utilities can use the funds to finance new capacity, or as security to obtain debt financing. In the day-ahead and real-time spot markets, the portfolio of option contracts ensures that LSEs can purchase energy at prices that do not exceed the strike prices of the options. Option contracts therefore address all three aspects simultaneously: long-term contracting, stimulus for investments in new capacity, and mitigation of suppliers' market power in spot markets. We explain these effects in detail later.

We use the term "option" because it conveys the main idea that, like a financial option that is a derivative of an underlying stock, an option on power supply obligates the seller to reimburse the buyer for any positive difference between the subsequent market price for power and the strike price. A somewhat similar instrument used in commodity markets is a one-sided contract-for-differences [CFD]; as with an option, the key feature is that the seller reimburses the buyer for market prices above the strike price of the contract. Even so, options and CFD differ in terms of the physical obligations. An option imposes no obligation on the seller to deliver power at prices below the strike price, whereas a CFD does. Thus a CFD provides a fixed quantity at a variable price that is capped by the strike price. In contrast, an option provides a somewhat variable quantity (either the full amount or none, at the discretion of the buyer) at a fixed price.

A key feature is that the total portfolio of options can include a schedule of increasing strike prices. The different strike prices reflect the higher marginal cost of supplies sufficient to meet higher loads; and on the demand side, the LSEs' need for financial hedges against the high prices that usually accompany high loads. For LSEs the crucial advantage of options over CFDs is that they hedge against both price and quantity risks, which in power markets are highly correlated. With a portfolio of options having many different strike prices, an LSE can call more options the higher is the spot price, thus obtaining more supplies when the spot price is higher, which usually occurs precisely when its load is greater.

The scheme also includes procedures enabling the SO to exercise options as needed to maintain reliability. In particular, a firm selling an option is obligated to submit corresponding bids in the spot markets for the optioned quantity at a price no higher than the strike price, thereby fulfilling the must-bid requirements mandated by FERC. We emphasize this specification rather than alternative in which options could be purely financial; that is, the supplier who sells an option is financially liable to the LSE that buys the option for the difference, if positive, between the spot price and the strike price. An LSE could also buy an option from a retail customer with a dispatchable load; that is, sale of the option makes the customer liable for the excess of the spot price over the strike price and thus provides an incentive for the customer to curtail demand when the spot price exceeds the strike price.

In the following subsections we describe reasons for using option contracts in this way, rather than relying on a separate ACAP market for capacity contracts, and separate AMP procedures for market-power mitigation. We begin with the role of options in spot markets, and then follow with their role in forward markets.

A. Options in Spot Markets: The Basic Mechanism

We begin with a schematic illustration of how spot prices and quantities are determined when suppliers and LSEs have previously traded option contracts in the forward market.

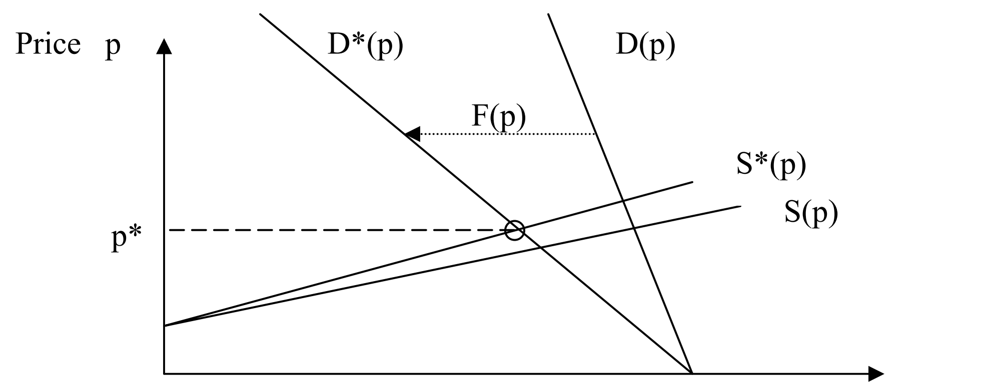

Figure 1

For the illustration, assume that the spot market is perfectly competitive; that is, because there are many small suppliers, each supplier offers each MWH at its marginal cost. Figure 1 depicts an aggregate demand function D(p), representing the total load when the price is p. For simplicity, D(p) is shown as a linear function of the price p. Assume that LSEs can call options to obtain a quantity F(p) at strike prices less than p, and that F(p) is also a linear function of the price. Then the residual load served by spot purchases is D*(p) = D(p) – F(p). Similarly on the supply side, Figure 1 depicts an aggregate supply function S(p) that specifies the total quantity that can be produced at marginal costs less than p. Again, S(p) is shown as a linear function of the price. To construct the net supply function S*(p), we make an assumption that is explained in more detail below. The assumption is that at each price p the options with strike prices less than p are fulfilled by a representative sample of those supplies that can be produced at marginal costs less than p. With this assumption, $S^*(p) = S(p) - F(p)$. The spot market price p^* equates net demand and net supply, namely $D^*(p^*) = S^*(p^*)$, as shown in Figure 1.

In the situation represented in Figure 1, the spot price p^* when options exist is no different than the price that equates aggregate demand and supply in the absence of options; that is, $D(p^*) = S(p^*)$. Thus in general there is no necessary implication that the presence of options alters the spot price. This coincidence results from a combination of the two key assumptions that were invoked. The implications of these assumptions can be sketched as follows.

1. We assumed that the spot market is perfectly competitive, so suppliers have no incentive to inflate their offers above marginal costs. Therefore, the aggregate and net supply functions reflect only the distribution of marginal costs among suppliers.
2. We assumed that at each price p the options with strike prices less than p are filled by a representative sample of those supplies that can be produced at marginal costs less than p. Actually, this is less an assumption that a prediction that selection effects will largely determine which suppliers provide the options sold at each strike price. For each price p, those suppliers with marginal costs below p can sell options with lower strike prices for less than can those suppliers with marginal costs above p – since the latter incur some operating losses when such options are called.

These two assumptions are relaxed below as we develop in more detail the role of options in forward and spot markets. However, it is worth noting now that in practice the first assumption is significant but the second is usually not. The reason for this is that aggregate demand in the spot market is typically inelastic whereas aggregate supply is moderately elastic. As illustrated in Figure 1, this configuration implies that the presence of options can have a large effect on the elasticity of the net demand function $D^*(p)$, whereas the effect on the net supply function $S^*(p)$ is relatively small. When the first assumption is relaxed, the effect of suppliers' market power becomes important. Equally, the ability of options to alter the net demand elasticity becomes important, because the effect of market power declines as the elasticity of net demand increases.

B. Options in Spot Markets: Market-Power Mitigation

The intent of hard bid caps and AMP procedures is to limit how far the prices in the SO's spot markets can deviate from the competitive norm that is the marginal cost of the last MWh dispatched. The variable bid cap aims more directly at this objective, because it makes a direct connection between the heat rate for a specific unit's marginal MWh and its fuel cost. At the level of the entire market, the appropriate generalization of variable bid caps is a bound on the aggregate supply function obtained from all the submitted bids. This bound can be represented by a function, P(q), that represents the maximum price ($/MWh) payable when the aggregate load is q (MWh). As with variable bid caps, the bounding function P is a matter of regulatory policy. The SO knows the heat-rate curves of the generation units in its control area, and the average prices of fuels. Therefore, it can calculate the aggregate supply function, C(q), based on the competitive norm of marginal cost; indeed, every SO's market monitor routinely calculates this function because it is the standard against which market performance is measured. Because suppliers face additional costs of capital, maintenance, startup, and no-load conditions, however, variable caps typically allow latitude for the desired bound P(q) to exceed the aggregate marginal cost function C(q). A possible policy objective is therefore to keep spot prices under the bounding function P(q). When long-term contracts are included, the relevant policy objective is to keep average prices under the bounding function, where the average accounts for the quantities obtained from both contracts and purchases in spot markets.

The fact that the bounding function P(q) increases as the aggregate load q increases is fundamental to electricity markets. A generation unit's marginal cost varies significantly with its operating rate, and escalates substantially as the operating rate nears the unit's capacity; further, more costly units are committed and dispatched as the load increases. The aggregate load varies depending on season, day-of-week, and hour-of-day, and also it varies substantially with weather and other contingencies. It is therefore intrinsic that spot prices, even according to the competitive norm, are volatile. Realistically, regulatory policy regarding a bounding function like P(q) must account for the escalation in marginal costs as the load increases.

The policy of keeping either average or spot prices under a bounding function P(q) is arguably within the jurisdiction of regulators at the state level. A state PUC has sole responsibility for regulating LSEs and the retail rates they charge, and since retail rates are materially affected by wholesale prices, a state PUC has an evident interest in limiting the wholesale spot prices paid by the LSEs under its jurisdiction. However, a state PUC faces substantial impediments if it relies solely on the provisions of the FERC-approved tariff of the local SO that conducts wholesale spot markets. In principle, FERC's jurisdiction encompasses interstate trade, but in practice it focuses mainly on transmission. Although the Federal Power Act gives FERC authority to ensure "just and reasonable" prices in wholesale markets, judicial interpretations have repeatedly limited FERC's authority to intervene in bilateral contracts. In contrast, a state PUC is vitally concerned to regulate the terms of the contracts undertaken by the utilities it regulates. Practical means of limiting average and spot prices must therefore address specifically how the design of these contracts can indirectly affect prices in the spot markets that are FERC-regulated and therefore outside the immediate jurisdiction of state regulatory agencies.

Option contracts provide a specific mechanism for limiting average and spot payments by LSEs. To illustrate we focus initially only on spot prices and suppose that state regulations ensure that, for each load level q, the options purchased by LSEs are sufficient to call at least the quantity Q(q) of energy from suppliers at strike prices no more than P(q). In terms of the notation used in subsection 2.A, such regulations can be phrased alternatively as specifying that at a price p the LSEs can call options on a quantity F(p) at strike prices less than p, where P(F(p)) = p. Thus, whenever the SO's spot price exceeds p the LSEs can obtain at least a portion F(p) at strike prices below p. Figure 2 shows how a bounding function P(q) can be implemented using options. For simplicity we consider just two load levels, q and q', that parameterize two corresponding demand functions D(p) and D'(p), and we assume that the net supply function S*(p) is the same for both.

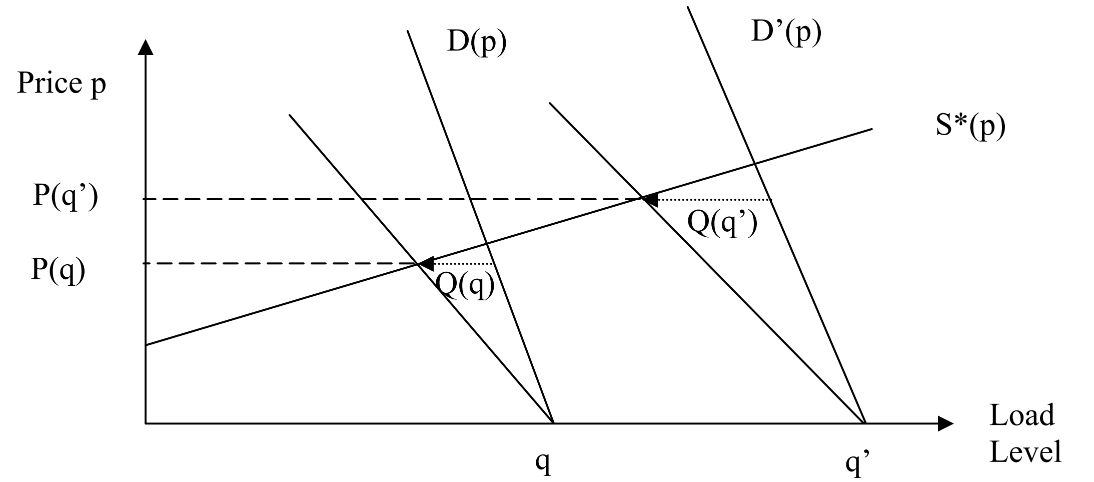

Figure 2

The construction in Figure 2 illustrates the principle that the bounding function P(q) is implemented by ensuring that the quantity Q(q) options that can be called at strike prices below P(q) is sufficiently large that when the load level is q the net demand function intersects the net supply function at the price P(q).

Of course, implementation can be done in various ways. The PUC could mandate the quantity of options purchased by LSEs at each strike price. Alternatively, if P(q) is the maximum spot price that the regulator will accept as prudent when the load level is q, then each LSE has an incentive to purchase sufficient options in forward markets to hedge against prices exceeding this bound. Due to a free-rider problem, however, this alternative is not fail-safe: any one LSE could forego purchasing options in hopes that other LSEs will purchase enough to keep the spot price within the bounding function.

There are additional beneficial effects from the presence of the option contracts. As mentioned previously, those suppliers who sold options are strongly motivated to submit bids at or below their strike prices. This strengthens bid sufficiency and prevents any supplier from being so "pivotal" that it can raise prices arbitrarily. It has further incentive effects on those suppliers who did not sell options, and also those suppliers who did sell options but have additional capacity that they can bid into the spot markets. These incentives work in two ways. First, each such supplier faces at each price a residual demand function (net demand minus the sum of other suppliers' offers at that price or lower) whose elasticity is at least as high as the elasticity of the net demand function D*(p). Suppose for example that aggregate demand is unresponsive to price, and that the aggregate marginal cost function C(q) is such that a 10% increase in the load raises the marginal cost by 10%. If the bound P(q) allows a 20% margin so that P(q)=1.2µC(q) then also the bound allows at most a 10% increase in price for a 10% increase in load. Then a supplier faces a residual demand whose elasticity is such that the supplier can get at most a 10% increase in price if it curtails its supplied quantity by an amount that is 10% of aggregate load. Faced with this tradeoff, such a supplier prefers to offer a bid price that is about 10% above its marginal cost. Thus, even those suppliers who do not sell options are significantly affected by the existence of the options sold by others. The potential market power of those suppliers who sold options is diminished in a second way: besides facing a more elastic residual demand, they have less un-contracted capacity that can benefit from driving up the market price.

The actual impact of option contracts in spot markets is subtler than the above scenario suggests. Suppose for example that if no options are called then the market price is \$100, but in fact 60% of the load is hedged by options at strike prices less than \$100. Then supplies obtained by calling these options serve 60% of the load. The net result is that in the spot market both demand and supply are reduced by 60%, leaving a residual market for the other 40% in which the spot price might still be \$100.[2] Thus, as in the example in Figure 1, the prevalence of option contracts need not alter spot prices. To obtain a full analysis of the effects of option contracts, therefore, one must consider the total payments made by LSEs, including both the cost of option contracts negotiated in forward markets and the cost of the quantities purchased in spot markets net of the amounts called under the previously purchased options. As argued below, the ultimate

[2] During the California crisis of 2000-2001, this effect was evident in the market in the Northwest. Mid-Columbia spot prices often exceeded those in California. But most LSEs in the Northwest had long-term contracts for as much as 95% of their loads and therefore they paid average prices that differed little from previous years. In contrast, the three major utilities in California were effectively excluded from long-term contracting and therefore paid spot prices for all their purchases.

advantage of options is that they transfer a major portion of energy trading from spot markets to forward markets.

C. Options in Forward Markets: Resource Adequacy

As seen above, spot markets are substantially affected by the existence of a portfolio of options held by LSEs. A seller of an option is liable for the excess of the spot price over the strike price, and due to the option, the seller has less uncommitted capacity that can benefit from raising the spot price. In addition, each supplier faces a residual demand whose greater elasticity diminishes market power. However, these spot-market effects of options cannot eliminate market power; instead, market power is transferred to the forward market where LSEs negotiate option contracts. We now elaborate why it is that moving negotiations to the forward market is advantageous from a regulatory perspective and beneficial for LSEs and their retail customers.

Attention is usually focused on the inelasticity of demand in spot markets; that is, on the lack of price-responsive loads. The motive for bid caps in spot markets is to counter the ability of pivotal suppliers to extract high prices when the SO desperately needs more energy to meet peak loads. An alternative view, however, is that the short-run inelasticity of supply is equally important. It is true that spot market supply is somewhat elastic, as measured by marginal cost, since marginal cost increases moderately as the load increases, provided capacity is not nearing exhaustion. But the key impediment to a competitive spot market for power is that entry is limited; i.e., in the economic jargon, the market is not contestable. For the most efficient units, capital costs are substantial and construction takes years, and new investments are affected further by difficulties obtaining siting and transmission connections. Even peaking units such as CTs cannot be installed within the time frame of a summer heat wave or a winter cold snap when spot prices surge and supplies are scarce.

In forward markets, supply elasticity is substantially greater, and correspondingly the market power of incumbent suppliers is less. The most important factor is that capacity expansion and new entry are feasible on a time frame some months or years before the spot markets occur. At a price that recovers the cost of a CT, the supply elasticity is essentially infinite, since any higher price attracts substantial entry.[3] To stimulate investment, a price must be sustainable on average over the life of a plant, which is not assured by volatile spot prices, but can be assured by long-term contracts. Further, long-term contracts can attract equity investors, provide financing directly, or secure lenders who provide debt financing. Related aspects also enhance supply elasticity in forward markets; e.g., a longer horizon enables long-term contracting for fuel supplies and emission permits.

In substantial part, long-term contracting provides risk sharing between the parties. As in any market with volatile prices, by contracting long term at a fixed price or an option's strike price, a seller and a buyer mutually insure each other against price variations over the term of the contract. As mentioned previously, there is the secondary effect that the greater is the supply contracted long term, the stronger is the incentive for suppliers in the spot market to offer bids closer to marginal cost. These advantages of long-term contracting are well known and familiar.

What then are the advantages of a contract written as an option with a strike price? It is important to address this question because an option contract sacrifices some of the advantages mentioned above. For the purpose of financing investments in new plants, options provide lenders with less security than contracts for fixed-price fixed-quantity contracts. The mutual insurance obtained from a contract with a fixed price for a fixed quantity is replaced by one-sided insurance in which the supplier ensures the LSE against prices above the strike price, but the LSE provides no insurance for the option seller. Both of these deficiencies reflect the fact that the seller of an option retains substantial risk: the seller remains exposed to price declines and yet cannot benefit from prices above the strike price, and there is an accompanying quantity risk because no supply will be called under the terms of the option when the market price is less than the strike price.

A practical argument for option contracts is based on the regulatory environment. Recall that the terms of long-term contracts with QFs were an important motive for the restructuring initiated by state agencies and supported by utilities burdened with prices

[3] This assertion must be qualified in states that limit emission permits.

substantially above the wholesale market prices then prevailing. State regulators' continuing aversion to fixed-price fixed-quantity contracts is strengthened by California's recent experience. At the height of the crisis the state signed long-term contracts that specify fixed prices and quantities, but within months the market prices were half as high and the quantities were excessive after imports resumed, new generation plants were completed, and higher retail prices curtailed loads. This adverse experience makes regulators reluctant to again (in effect) insure to suppliers via rigid contracts, even though they remain interested in supplier-provided insurance for the LSEs they regulate. Perhaps a minor factor is that options with strike prices substantially above expected spot prices are inexpensive and require less scrutiny in reviews of prudency.

But our view is that the main argument for option contracts is that they address precisely the lessons derived from the California crisis. Option contracts are negotiated in forward markets that are more contestable than spot markets – because long-term contracting enables financing and construction of new generation assets. And they mitigate market power in spot markets. Options are also flexible regarding quantity in that an option need not be called until extra supplies are needed and the spot price exceeds the strike price. A well-designed portfolio includes options with a variety of strike prices tailored to the LSE's load-duration curve and tolerance for risk. A portfolio can be designed to moderate an LSE's financial risks in spot markets to any desired degree. Equally, sellers of options can tailor their strike prices to the characteristics of their facilities in relation to the system's load-duration curve. Due to the selection effects described in subsection 2.A, base-, shoulder-, and peak-load plants have comparative advantages in selling options at low, medium, and high strike prices. The lower is the strike price, the more assured the seller is that its options will be called; thus, an entrant with financial constraints imposed by lenders may prefer lower strike prices than one better able to bear the risks of low prices.

It is important to recognize, however, that option contracts do not provide the perfect solution. They do not exploit fully the potential for mutual insurance between buyer and seller, they do not provide perfect security for lenders, and they cannot eliminate market power completely. Regulators must expect that LSEs will pay for these deficiencies. The contestability of forward markets will remain imperfect to the extent that lenders view an option contract as less than perfect security for a loan, or equity investors demand higher returns to compensate. The cost of an option will presumably include a "risk premium" for the seller to compensate for the residual price and quantity risks it bears. But regulators and utilities may prefer to pay this premium to obtain the resulting hedge against spot prices above the strike price, while retaining the access to lower prices and the quantity flexibility of being able to call the option only when extra supplies are needed to meet high loads.

Finally, we emphasize that resource adequacy requirements can be more effective when formulated in terms of options. ICAP markets like those in the northeastern SOs have no impact on current spot prices for energy, nor on market power – and when capacity is scarce they provide an additional forum for incumbents to exercise market power. An ICAP requirement's potential stimulus to long-term investments and new entry is diminished because monthly prices for capacity claims are highly volatile and often nearly zero, which precludes them from serving as security for loans. The fundamental deficiency of monthly ICAP and ACAP markets is that they do not transfer negotiations to forward markets for long-term contracts where the effects on resource adequacy are greater because of the greater elasticity of supply from capacity expansion and new entry.

D. Reliability Considerations

We conclude this section by mentioning briefly the subsidiary effect of options on the grid management conducted by the SO.

From the SO's perspective, every bid submitted in the spot market is an option. That is, the supplier's bid enables the SO to exercise the option at any market price above the bid price. However, the recurrent problems that afflict spot markets for electricity are that either there are insufficient bids to meet demand (called bid insufficiency), or that the bids are offered at exorbitant prices that the SO cannot reject without jeopardizing reliability or curtailing loads.

The two measures adopted by many SOs and endorsed by FERC are mandatory bidding requirements and bid caps. A bidding requirement obligates each supplier to offer bids in the day-ahead and real-time spot markets for all its available capacity, i.e., operable capacity that is not committed to bilateral contracts and not scheduled in advance for maintenance. In the prevalent AMP procedures, there is a hard bid cap (e.g., no bid can offer a price above $1000/MWh) and also a variable cap that allows a supplier to be paid a price above a multiple of a thermal generation unit's heat rate times its fuel cost only if the combined effect of all bids above the variable caps does not fail an "impact screen" (e.g., would increase the market clearing price by more than 100%). These measures are *ad hoc* attempts to ensure sufficient supplies at moderate prices. In effect, they are contractual obligations included in the SO's FERC-approved tariff that the supplier must accept if it is to participate in the SO's markets. In the following we will be describing ways that portions of these obligations can be supplanted by option contracts regulated by state agencies.

For the SO's routine operations, the key requirement is bid sufficiency. The SO needs a rich stack of bids that it can draw upon to ensure voltage support and stability, alleviate congestion, establish reserves for contingencies, and follow the load in real time. To the extent that the terms of bilateral contracts between LSEs and suppliers ensure sufficient quantities of bids to meet these needs, the SO can conduct its operations without heavy reliance on bidding requirements imposed via the terms of its tariff. A purpose of the scheme that uses option contracts is therefore to reduce the SO's reliance on punitive sanctions in its FERC-approved tariff. Thus we propose that option contracts include provisions that strengthen bid sufficiency via strong incentives for sellers of options to bid all available capacity into the SO's spot markets. The chief incentive is that a seller who does not bid into the SO's market runs the financial risk of paying for the energy called by the LSE at the strike price if the SO's market price exceeds the strike price. Thus, the seller of an option at a strike price above its marginal cost prefers to offer a bid, at a price no higher than the strike price, since otherwise it runs the risk of paying the difference between the market price and the strike price. This incentive for option sellers to offer bids voluntarily is potentially more efficient in its effects, and less costly to administer, than punitive sanctions in the tariff.

4. Practical Aspects of Implementation

We divide the discussion of implementation into regulatory and technical aspects. The technical aspects include such matters as whether the SO has discretion to call options directly and the role of options in reserve procurement.

Our focus initially is on the formulation of the regulations that mandate coverage of an LSE's load by option contracts sufficient to enforce a bound on average price paid, depending on the magnitude of aggregate demand. The problem is not trivial; in particular, the regulator or the SO cannot simply require each LSE to cover its load with sufficient options, since that need not result in a pattern of strike prices that fits the intended bound. Other complications abound:

- There is a significant free-rider problem. Each LSE prefers to buy options with high strike prices because they are cheaper, and hope that other LSEs will buy the ones with lower strike prices. Conceivably this difficulty could be overcome by having one agency be the sole buyer of options. The options themselves, or the cost of purchasing them, would then be distributed among LSEs in proportion to their loads, and in a way that recognizes (via their load-duration curves) their differing load profiles. But this highly centralized procedure goes against the grain of restructuring.
- LSEs already have large portfolios of long-term contracts that for the most part cannot be renegotiated. In important cases, these are contracts with affiliates or involve must-run resources such as nuclear and hydro generators that usually are price-takers in the spot markets, and therefore not the most important ones for option contracting. The transition from existing contracts to those that would implement a bound is nontrivial.
- Inducing suppliers to sell option contracts voluntarily at prices acceptable to the state regulator could be problematic since suppliers are regulated by FERC and therefore largely exempt from state jurisdiction. The greater elasticity of supply and the greater ease of entry in the forward market encourage our view that this problem is solvable, but again it is nontrivial. At the very least one expects that entrants would bargain vigorously for full price protection via Power Purchase Agreements [PPA], not just options, so that they could obtain financing at least cost.

- The terms of option contracts require careful study. Options are invariably cheaper if the buyer pays the fuel cost and thus absorbs the risk of volatile fuel prices. Because of regulations that affect relations with affiliates, there is a material difference between an option on energy output and a "virtual capacity" contract in which the buyer has an option on capacity that the seller manages as directed by the buyer. Options are sensitive to the terms governing default by the seller on its obligations, especially in the case of bankruptcy.
- An alternative policy would allow LSEs to purchase option contracts in established markets for futures and options (such as NYMEX or CME) that are regulated by the Commodity Futures Trading Commission. These are purely financial options, not backed by physical assets, and their duration is limited to a few months or a year. The CFTC imposes margin requirements that are rather small, usually 10%. Allowing LSEs to purchase such options would be a major departure for regulators and raise risks of default.

The regulator is certain to encounter further difficulties. One is monitoring and enforcement of the regulator's mandate, possibly including sanctions or penalties for noncompliance. But there are limits that constrain the regulator, including the right of a utility as a private corporation to decline to enter into contracts. Between the carrot and stick approach we prefer the former, and therefore emphasize the benefits of options that will win the support and participation of utilities.

In the following subsections we suggest specific mechanisms that address some of these matters.

A. Design of Option Contracts

A fundamental difficulty posed by an option contract is the risk of default by the supplier who sells it. Some measures to reduce the consequences of default are obvious. For example, the LSE should make installment payments rather than an upfront payment for the option. The supplier should be required to assign its revenues from sales through the SO's markets to an escrow account, where the funds are held until the supplier compensates the LSE for market prices exceeding the strike price. In the event that the supplier's unit is inoperable, or the supplier does not submit bids to the SO, or its bid is rejected by the SO, the payment due the LSE should be secured by a lien at least on the supplier's liquid assets, and preferably also on its physical generation assets. These are fairly standard ways of dealing with default risk in situations where credit-worthiness is not fully assured.

However, long-term options on wholesale electricity are afflicted with a default risk that is inherently more severe. If market prices decline below the strike price over a long period then an option becomes worthless and default by the supplier is inconsequential. On the other hand, if market prices remain substantially above the strike price for a long period then the supplier's incentive to renege on the contract can become an important factor. Default risks are more severe if the supplier uses a separately incorporated firm as the entity that sells the option contract, and thereby protects its other assets from liens and seizures. A particular risk is that the supplier's separately incorporated entity could seek bankruptcy protection under the provisions of Chapter 11. In the reorganization of the entity the LSE's option need not have status as a senior creditor, and the bankruptcy judge would have discretion to decide how much of the LSE's claim would be honored when the entity emerges from bankruptcy. These considerations imply that an option contract acceptable to a state regulator must include stringent provisions ensuring that financial and physical assets back the option by providing security against nonperformance. Although the design of these provisions will be a major task, it is one that is encountered also in every long-term contract between an LSE and a generator.

A focus on default risk might seem to imply that it is immaterial whether an option is a physical or financial instrument. Any single LSE is made whole if the seller reimburses the LSE for spot prices above the strike price, regardless of whether the seller actually generates the energy or purchases it from the SO's market. From a regulatory viewpoint, however, it is important that the option is backed by physical capacity so that it can serve the larger purpose addressed by ACAP and AMP requirements. The beneficial effects of options on spot prices and market power, and in forward markets their stimulation of new investment and entry, depend substantially on the requirement that they entail physical obligations. The obligation might be for delivered energy, but alternatively the contract could be for virtual capacity in which the buyer has a residual right (if the seller defaults) to dispatch the seller's generation units, provided it pays all operating and fuel costs. The simplest scheme requires the seller to maintain a default bid at the SO, equal in price and quantity to the terms of the option.

There are other details that contracts must address. For instance, the contract might allow exceptions for periodic scheduled maintenance. The buyer might be required to reimburse the cost of start-up and/or the ramping necessary to reach the generation rate required by the option. Because the SO's market sets nodal prices, the contract must define the point of injection if the buyer is liable for congestion charges and losses, or point of withdrawal if the seller is liable.

B. Auctions of Option Contracts

The free-rider problem mentioned above has several possible solutions. For instance, the regulator could require each LSE to purchase option contracts in proportion to their historical loads, and also take account of their load-duration profiles and vulnerability to congestion charges. These contracts and their purchase costs would then be assigned to a pool in which the LSEs share proportionally based on actual loads and load-duration curves. The deficiency of such a scheme is that no single LSE has a strong incentive to minimize the cost of the options it purchases, because this cost is shared among all LSEs. Thus such decentralization requires the regulator to monitor each LSE's performance in acquiring its assigned portfolio of option contracts. This deficiency suggests that a central auction of option contracts would be superior. State regulators have experience with similar procurement auctions conducted under the requirements of PURPA in the years before 1993 when they contracted for supplies from QFs, and similar auctions have been conducted recently in Alberta and France.

An auction of option contracts involves so-called package bidding for diverse collections of contracts. Auctions of this type have become common in the telecommunications industry (for spectrum licenses) and in the electricity industry (e.g., for firm transmission rights [FTRs] in California, for Power Purchase Agreements in Alberta and France, and for subsidies of low-emission generators in the UK and California). Here we describe briefly the features of such an auction of option contracts, omitting many of the more detailed aspects that can be found in the large literature on this subject, and embodied in software available for conducting such auctions.

The regulator's specification of a bound on the average price P(Q) depending on the load Q translates into a specification of a quota q(p) of options desired at each strike price p, such that Q is the sum of the quantities of options procured at strike prices p § P(Q). In addition, the regulator will likely specify one bundle of procurements for the ensuing season or year, and possibly another for years 2-3, and another for years 3-6, with the intention that those in later years enable entrants to compete for contracts that can be used to finance investments in new generation units.

As part of its prudency procedures, the regulator might establish a maximum price (i.e., its reserve price) that it is willing to pay for options at each strike price: if so, then the existing theory of so-called "real options" can be applied to derive this reserve price from an analysis of the probability distribution of predicted spot prices over the duration of the option contract. Imposition of reserve prices necessarily creates a risk that the auction attracts insufficient bids. One must anticipate also that the set of allowed bidders might be restricted; e.g., must-run resources might be excluded.

In the auction each bidder submits a series of bids, continuing until no lower bids are received and the auction closes. Each bid specifies a total payment asked for a specified collection of option contracts. For example, a bid can specify a quantity q_i (measured in MWh) of contracts offered for each of several strike prices p_i , and a total payment requested for this collection if the bid is accepted in its entirety. In each round of the auction, optimization software is used to calculate the set of bids that could be accepted at the least total cost, in terms of payments to bidders, subject to the requirement that the specified quota q(p) of options at each strike price p is not exceeded. Each round ends with *tentative* acceptance of this set of cost-minimizing bids. In the next round, bidders can submit new bids, provided they improve on the previous bids according to an "activity rule" that assures that the auction progresses steadily and that the total procurement cost declines from one round to the next. The auction ends when there are no new bids that would reduce the total procurement cost. At the close of the auction, those bids tentatively accepted in the last round are actually accepted.

Because the theory and practice of auctions with package bidding for diverse collections of items has progressed greatly since their introduction a decade ago, we consider the implementation of such an auction for option contracts to now be a routine matter. The substantive issue is whether such an auction will be vigorously competitive, and thereby achieve the intended quota q(p) at each strike price p at a low total cost to the LSEs. We describe below the considerations that indicate favorable prospects for a successful auction.

1. For contracts with long durations and strike prices above the marginal costs of peaking units, it seems evident that the supply of bids is perfectly elastic. Essentially, an entrepreneur can submit a bid in which the requested payment for the option is the capital and maintenance cost of installing a CT, and if the bid is accepted, thereby assure a profitable net present value for the investment.

2. For low strike prices, one expects weak competition from those few resources, such as hydro and nuclear, with low marginal costs; and besides these must-run resources might be excluded from bidding. However, there are two factors that mitigate. One is that, for the purposes of implementing ACAP and AMP, the regulator need not specify a positive quota at low strike prices, such as those below the marginal cost of the most efficient gas-fired combined-cycle units. Second is the observation that an option at a low strike price, say zero, is indistinguishable from a standard long-term contract with a delivery price equal to the strike price. Therefore, bids for those option contracts with low strike prices must compete with bids from those suppliers who desire such contracts because they provide complete price insurance, for both the LSE and the supplier. Thus, one expects that at low strike prices the cost of an option is bounded above by the cost of an ordinary two-sided long-term contract at a delivery price equal to the strike price.

3. For a strike price in the intermediate range, a potential bidder must essentially choose between offering a bid to sell an option with that strike price, or instead competing directly in the subsequent spot markets. The expected profit from opting for direct participation in the spot markets implies a lower bound on

what it is willing to bid in the auction, but since this is true for all such potential bidders, the salient conclusion is that the lowest bid offered in the auction reflects at most the second-lowest of these lower bounds. That is, each potential bidder has an incentive to request a payment no more than its expectation of the second-lowest among the payments demanded by other bidders to compensate for what they could otherwise earn by direct participation in the spot markets.

4. Competition among the bidders is enhanced further by the fact that the portfolio of option contracts tends to dampen prices in the spot markets, and thus leads bidders to seek their profit opportunities in the forward market, here represented in large part by the auction. As described at the end of section 2.B, the fact that options are exercised whenever the spot price exceeds their strike prices has the secondary effect that each supplier bidding into the spot markets perceives a residual demand that is substantially more elastic than would otherwise be the case, and hence, the percentage profit margin of the market clearing price over marginal cost is substantially reduced.

To this list of favorable prospects one must add the unfavorable ones.

1. To the extent that LSEs and suppliers are already encumbered by existing long-term contracts, participation in the auction is reduced. Our view is that this consideration is not fundamental. To whatever extent that LSEs already have price insurance via long-term fixed-price contracts, the regulator can specify the smaller quotas of option procurements that are appropriate for the residual demand and supply that are not fully contracted, and importantly, the quotas can still include the desired amount of new investment to ensure resource adequacy.

2. A significant unknown is the degree to which suppliers are so risk averse that they strictly prefer to contract forward at fixed prices rather than selling option contracts, or alternatively, so tolerant of risk that they prefer the volatility of spot markets. If these are significant factors then participation in the auction could be weak. Again our view is that this is not fundamental as regards existing capacity, since the regulator would presumably reduce the quotas

proportionately to account for the extent of pre-existing price insurance already obtained by suppliers and LSEs via fixed-price contracts, and a prediction about the future prevalence of such bilateral contracts. The matter is more problematic for new capacity: if investors and lenders regard option contracts as insufficient security for new investments then this could dampen participation. In theory this need not be so since a financial intermediary could "complete" the security by selling the investors a put option at the strike price, and thus arbitrage the difference. But in practice there is ample evidence that in wholesale electricity market there are great risks in purely financial arbitrage that is not backed by physical resources, and thus few financial institutions other than the major investment banks have been eager to sell such put options.

C. Technical Considerations

We now address the technical aspects of implementation that affect the SO. We have already mentioned a settlement procedure that involves the SO, namely, retaining funds due a seller until the buyer is repaid the excess of the spot price over the strike price of the option, or even arranging that the SO withholds the difference from the amount due the seller and charges the buyer only the strike price. But from an engineering viewpoint the SO is less concerned with settlements than with the implications for grid operations.

An option functions somewhat like a claim on reserve capacity. The SO might prefer that it can itself call an option to obtain additional supply when the spot price in the real-time balancing market exceeds the strike price. If this were allowed then optioned capacity could be included in the merit order at the strike price of the option. An extreme form of this policy would even allow the SO to dispatch units covered by options.

As we emphasized previously, in the day-ahead market each seller of an option should be required to offer a bid at the strike price if the unit is available. This is necessary to ensure that the nodal prices reflect the existence of options. Because the SO's day-ahead optimization is comprehensive – including unit commitment, scheduling of units, and assignments to reserve status – calling additional supply afterwards is too late to have the intended effects on spot prices described in section 2.B.

There are bound to be further complications. For instance, if the SO assigns optioned capacity to reserve status then ordinarily it must be considered unavailable for generation to fulfill the option if it is called. In this case the seller is presumably liable for purchasing replacement energy at a real-time price that exceeds its own lower cost of generation were the capacity not assigned to reserve. The resolution of such tensions between seller and buyer could be negotiated between them, but if an auction is used then all such operational details must be spelled out explicitly in a single standard contract form that applies uniformly to all parties.

5. A Simple Model of Option and Spot Markets

In this section we present a mathematical model that illustrates the strategic interactions between the forward market for option contracts and the subsequent spot market for energy. This model is too simple to be realistic; e.g., aggregate demand and supply are assumed to be linear functions of price, there is no explicit distinction among generating technologies and their differing roles in serving the load-duration curve, the only entry barriers are a fixed cost and capacity costs, no allowance is made for reserves nor for congestion management, etc.[4] However, the model indicates some of the main ingredients that would be included in a general mathematical development. Moreover, it enables explicit illustrations of several effects that might otherwise be speculative:

- Utilities' use of forward markets to purchase long-term contracts that supplement their purchases in spot markets is a demand-side strategy akin to product differentiation by sellers. Although suppliers retain market power in the spot market, its role is reduced by demanders' earlier purchases of options that substitute for spot purchases when the load and the spot price are high.
- When utilities cover less than 100% of their requirements with options obtained in the forward market, the spot market allows an equilibrium among suppliers' bid functions that exceed their respective marginal costs.

[4] For numerical simulations that use more realistic demand conditions and supply technologies see Robert Entriken and Stephen Wan, "Pushing Capacity Payments Forward: An Agent-Based Simulation of Available Capacity Markets," Electric Power Research Institute, December 2003, Report 1007755.

- Equilibrium in the two markets jointly requires absence of opportunities for suppliers to arbitrage expected price differences between them. Even though we introduce no countervailing factors such as risk aversion, the no-arbitrage condition depends on the variance of the resulting probability distribution of spot prices. This distribution results from the interaction of the forward market for options, the supply-function equilibrium in the spot market, and the probability distribution of the load in the spot market.
- Option markets have a negative feedback effect on contestability. The more that options suppress suppliers' profits, the weaker is the incentive for entry and capacity expansion that disciplines incumbent suppliers in the forward market and drives bids closer to marginal costs in the spot market.
- For some cost parameters, consumers' surplus is maximized by requiring a moderate fraction of utilities' requirements to be covered by options. (In extreme cases, however, contestability of the forward market is so effective that it suffices to rely 100% on forward purchases via options purchased from a single monopoly supplier.) Optimizing the fraction covered by options is an important part of implementation.

Some of these features will be evident as we set forth the model and others we mention in context.

In the model there are just two stages, corresponding to the forward market and the later spot market.

(1) In the first stage, the demand for energy that will appear at the second stage is still uncertain. Many firms decide whether to enter the market as suppliers. Entry incurs an irreversible fixed cost h. Entry is costly but otherwise unconstrained, so a smaller entry cost implies greater contestability of the market. The model allows that firms are heterogeneous, each characterized by a parameter of its cost function for energy production. The number n of firms that enter is endogenous; that is, the number of entrants increases until the expected profit of the next entrant is nil or negative, net of the entry cost. The n firms that enter become the suppliers in the forward market for options and then later in the spot market for energy. The model does not include any forward contracts other than options. The suppliers initially compete in an auction at which they

offer options at a variety of strike prices. These options are purchased by LSEs to meet their resource-adequacy obligations imposed by the PUC.

(2) In the second stage, the suppliers and the LSEs participate in the spot market. This is an auction in which the spot price is determined by the market-clearing condition that at that price the realized demand equals the sum of (a) those quantities callable via options with lower strike prices, and (b) those quantities offered by suppliers at lower prices. Those transactions of type (a) are settled at their respective strike prices, and those of type (b) are settled at the spot price.

Equilibrium in the spot market is a "supply-function equilibrium" [SFE]. That is, each supplier offers an entire supply function. The actual quantity it sells depends on the competing offers from the other suppliers as well as resolution of the uncertainty about the actual demand. Anticipating this equilibrium in the second stage, in the first stage those firms that choose to incur the entry cost submit offers in the auction market for options. This auction is a simultaneous auction of many different option contracts, each characterized by its strike price. For each strike price, the quantity of options demanded by LSEs is fixed by the PUC's resource-adequacy requirement. All options with the same strike price are settled at the market-clearing price for this type of contract.

A. Supply and Demand Conditions

On the supply side, the key features are that (a) entry incurs a fixed cost, and (b) each firm's marginal cost increases linearly. We ignore investment decisions other than entry; in particular, the fact that marginal cost increases plays the role of a capacity constraint in a more general model.

Assume that *n* firms will enter. These are the suppliers who participate in the forward market for options and subsequently bid in the spot market. Their cost functions are assumed to be quadratic of the form $C_i(x, y) = h_i + k_i y_i + c_i x_i^2 / (2 y_i)$, where h_i is the entry cost for firm *i*, and k_i is the unit capacity cost for firm *i*. Note that when the entry cost is zero, the cost function is homogeneous of degree one, implying constant returns to scale.

On the demand side, the aggregate demand function is linear of the form, $D(\overline{p}, p, \varepsilon) = a - b\overline{p} - \beta(p - \overline{p}) + \varepsilon$, where p is the spot price, $\overline{p}$ is the expected spot price, and ε is an independent random variable with mean zero and variance σ^2. We assume that the aggregate demand is more elastic in the forward market than in the spot market. We allow the realistic possibility that $\beta = 0$, in which case the aggregate demand is completely inelastic in the spot market.

B. The Forward Market for Option Contracts

In the forward market, the firms sell option contracts, and the LSEs purchase these contracts to meet their resource-adequacy obligations. The aggregate demand for options at each strike price is determined by the obligations of the LSEs to satisfy their resource-adequacy obligations specified by the PUC. These requirements are assumed to require the LSEs to acquire in total a quantity $F(p) = \theta p$ of option contracts at strike prices below each possible spot price *p*. We adopt this linear formula for *F* to simplify calculations. The PUC might choose the parameter θ to maximize either consumers' surplus or the total surplus of consumers and producers, but initially we make no specific assumption about how θ is determined.

The total quantity of production capacity of firm *i* that is optioned at strike price *p* is assumed (again, for simplicity) to be a linear function of the form $R_i(p) = r_i p$. Since the aggregate demand for option contracts is given by $F(p) = \theta p$, the forward market clears at a price that ensures that $\theta = \sum_{i=1}^{n} r_i$. The forward price of the standard bundle of option contracts corresponding to $r = 1$ is denoted by λ. Such a bundle provides an option on a quantity *p* at strike prices less than *p*.

C. Equilibrium in the Spot Market

The spot market is an auction in which all trades are settled at the market-clearing price. The firms announce their bid functions simultaneously. The bid functions are assumed to be linear such that the total quantity of energy offered by firm *i* at price *p* is $S_i(p) = s_i p$. A supplier's bid function need not be the same as its marginal-cost function, since in the spot market the *n* suppliers have market power and demand is imperfectly elastic. At the market-clearing price *p*, each firm *i* produces a total amount, $S_i(p) + R_i(p) = (s_i + r_i)p = x_i p = X_i(p)$, which includes both its spot sales and the amount called from the options it sold in the forward market. The spot price $\tilde{p}(\varepsilon)$ is determined by the following market-clearing condition:

$$\sum_{i=1}^{n} X_i(p) = D(p,\varepsilon) .$$

This formula reflects that key fact that suppliers' previous sales of options in the forward market reduce the aggregate demand in the spot market to the net demand $\sum_{i=1}^{n} S_i(p) = D^*(p,\varepsilon) = D(p,\varepsilon) - \sum_{i=1}^{n} R_i(p)$. Solving this market-clearing condition, one obtains

$$\tilde{p}(\varepsilon) = \frac{\alpha+\varepsilon}{\beta + \sum_{i=1}^{n} x_i}$$

as the formula for the spot price, where $\alpha = a-(b-\beta)\overline{p}$. Given the resource adequacy requirement, the spot price will not exceed $(\alpha+\varepsilon)/(\beta+\theta)$. To complete the formulation, we include bilateral contracts under the terms of contract-for-differences with a contract price $\hat{p}_i$ and a contract quantity Q_i. To analyze the suppliers' behaviors in the spot market, we first construct the realized profit for firm *i* in the spot market as follows:

$$\begin{aligned}\Pi_i^0\left(x_i,r_i,y_i\right) &= \tilde{p}S_i(\tilde{p}) + \int_0^{\tilde{p}} z dR_i(z) - C(S_i(\tilde{p})+R_i(\tilde{p})) - k_i y_i + (\hat{p}_i - \tilde{p})Q_i \\ &= \tilde{p}X_i(\tilde{p}) - \int_0^{\tilde{p}} R_i(z)dz - C(X_i(\tilde{p})) - k_i y_i + (\hat{p}_i - \tilde{p})Q_i \\ &= \left[x_i - \frac{r_i}{2} - \frac{c_i x_i^{\,2}}{2y_i} \right]\tilde{p}^2 - k_i y_i + (\hat{p}_i - \tilde{p})Q_i .\end{aligned}$$

As mentioned earlier, the fact that the square of the spot price enters this formula implies that suppliers' expected profits depend on the variance of the spot price, which formula for the spot price shows to stem from the variance of the load that occurs.

Given the limited number of suppliers, a supply-function equilibrium arises naturally in the spot market. Each supplier chooses the bid function that maximizes its expected profit under the assumption that the bid functions of the *n-1* other firms are fixed at their respective optima. The condition for a supply-function equilibrium in the spot market is found to be:

$$x_i^* = \left(1 - \frac{c_i x_i^*}{y_i}\right)\left[\beta + \sum_{j \neq i} x_j^*\right] + r_i + \frac{\gamma Q_i}{\overline{p}} \qquad \text{for all } i,$$

where $\gamma \equiv \alpha^2 / \left(\alpha^2 + \sigma^2\right)$, the asterisk indicates the value of the optimal parameter for the firm's total supply function.

We shall focus on the option contracts starting with the assumption that $Q_i = 0$, and will revisit the bilateral contract issues in a simple example.

Theorem 1. An increase in a firm's supply of option contracts in the forward market increases the firm's equilibrium supply in the spot market and reduces the equilibrium spot price.

Proof: Differentiating the previous formula yields,

$$\frac{\partial x_i^*}{\partial r_i} = \frac{y_i}{y_i + c_i(\beta + \sum_{j \neq i} x_j^*)} \geq 0 .$$

Intuitively, an increased sale of option contracts reduces a firm's later incentive to withhold supply in the spot market, since it has less residual capacity that can benefit from raising the price. The effect of an increase in x_i^* on the other firm's supply can be derived similarly as

$$\frac{\partial x_j^*}{\partial x_i} = \frac{y_j - c_j x_j}{y_j + c_j(\beta + \sum_{\ell \neq j} x_\ell^*)} \geq 0 .$$

Therefore, we conclude that the equilibrium market supply $\sum_{i=1}^{n} x_i^*$ increases. Clearly the increase in supply reduces the equilibrium spot price. Q.E.D.

D. Equilibrium in the Forward Market

In the forward market, additional entry occurs whenever a positive expected profit net of the entry cost could be obtained after a single new entry. The expected net profit for a supplier in the forward market includes the revenue from selling option contracts; i.e.,

$$\Pi_i(s_i, r_i, y_i) = E\left\{\Pi_i^0(s_i, r_i)\right\} + \lambda r_i - ky_i - h_i + (\hat{p}_i - \overline{p})Q_i$$

$$= \left[x_i - \frac{r_i}{2} - \frac{c_i x_i^2}{2y_i}\right] E\left\{\tilde{p}^2\right\} + \lambda r_i - ky_i - h_i + (\hat{p}_i - \overline{p})Q_i .$$

We assume that there is no profitable arbitrage between the spot and forward markets. From this condition, one derives the market clearing price in the market for options:

$$\lambda^* = E\left\{\frac{\tilde{p}^2}{2}\right\} = \frac{\alpha^2 + \sigma^2}{2\left(\beta + \sum_{i=1}^{n} x_i^*\right)^2} .$$

That is, the price of a standard bundle of option contracts is the opportunity cost of limiting expected revenues in the spot market. This does not say that the option contracts have no effect, however, since the fact that other suppliers have sold options reduces the spot price resulting from the supply-function equilibrium.

Theorem 2. For each supplier, the marginal value of an option contract is non-positive if there is no resource-adequacy requirement; that is, $d\Pi_i / dr_i \leq 0$, when $r_i = 0$.

Proof: This result follows from applying the Envelope Theorem to the expected net profit:

$$\left.\frac{d\Pi_i}{dr_i}\right|_{r_i=0} = \left.\frac{\partial \Pi_i}{\partial r_i}\right|_{r_i=0} = -\frac{1}{2}\frac{\partial E\left\{\tilde{p}^2\right\}}{\partial r_i} \leq 0 . \quad \text{Q.E.D.}$$

This result indicates that the market for option contracts cannot arise spontaneously. Since it is unprofitable for firms to offer such options unilaterally, a resource-adequacy obligation must be imposed.

From the necessary condition for optimization of each supplier's capacity expansion, the optimal scale parameter for capacity can be found to be

$$y_i^* = \sqrt{\frac{c_i E\left\{\tilde{p}^2\right\}}{2k_i}} x_i^* .$$

Therefore, the expected total payoff of firm *i* can be simplified to

$$\hat{\Pi}_i(\theta,n) = \Pi_i(s_i,r_i,y_i) = \left(2E\left\{\tilde{p}^2\right\} - \sqrt{c_i k_i E\left\{\tilde{p}^2\right\}}\right)x_i^* - h_i .$$

Equilibrium in forward market is obtained when entry is marginally profitable; i.e., $\hat{\Pi}_n(\theta,n) > 0 \geq \hat{\Pi}_{n+1}(\theta,n+1)$. Without significant loss of generality, we set the price of an option by the condition that $\Pi_{n+1}(\theta,n+1) = 0$.

The expected consumers' surplus and producers' surplus are given by:

$$CS(\theta,n) = E\left\{\frac{\left(\alpha-\beta\tilde{p}+\varepsilon\right)^2}{2\beta}\right\} = \frac{\lambda}{\beta}\left(\sum_{i=1}^{n} x_i^*\right)^2$$

$$PS(\theta,n) = \sum_{i=1}^{n}\hat{\Pi}_i(\theta,n) .$$

Theorem 3. When the number of firms is fixed, a resource-adequacy requirement can improve consumers' surplus; that is,

$$\left.\frac{\partial CS(\theta,n)}{\partial\theta}\right|_{\theta=0} \geq 0 .$$

Proof: From the formula for consumers' surplus one obtains

$$\left.\frac{\partial CS(\theta,n)}{\partial\theta}\right|_{\theta=0} = \lambda^*\left[\frac{\frac{\partial}{\partial\theta}\left(\sum_{i=1}^{n} x_i^*\right)^2}{\beta + \sum_{i=1}^{n} x_i^*}\right] \geq 0 ,$$

and from Theorem 1, the resource-adequacy requirement increases the equilibrium total supply $\sum_{i=1}^{n} x_i^*$ in the spot market. Q.E.D.

E. An Example with Symmetric Firms

We now consider the case in which *n* existing firms have the same production cost function, $c_1 = c_2 = ... = c_n$ and $k_1 = k_2 = ... = k_n$, but different entry costs. The entry cost of the marginal firm is denoted by *h*. In this case, we obtain a symmetric solution in which $r^* = \theta / n$ and $s^* = x^* - r^*$. Thus, all suppliers share equally in the provision of options at each strike price. From the previous formulas one obtains the slope of each firm's total supply function as

$$x^* = \frac{-\phi + \sqrt{\phi^2 + 4\rho\nu}}{2\rho}$$

where

$$\mu = \sqrt{\frac{\alpha^2 + \sigma^2}{2kc}}$$

$$\rho = n(n-1)$$

$$\phi = \beta(2n-1) - \mu(n-2) - \frac{n\mu\gamma Q}{\alpha} \cdot$$

$$\nu = \mu(\beta + r) - \beta^2 + \frac{\beta\mu\gamma Q}{\alpha}.$$

The equilibrium price of the standard bundle of options is then

$$\lambda^* = \frac{\alpha^2 + \sigma^2}{2(\beta + nx^*)^2}.$$

Define $y^* = \frac{\mu x^*}{\beta + nx^*}$. Then the formulas for producers' and consumers' surplus are

$$PS(\theta, n) = \lambda^* \left[2x^* - \frac{c(x^*)^2}{y^*} \right] - ky^* - h$$

$$CS(\theta, n) = \frac{\lambda^*(nx^*)^2}{\beta}.$$

Since the equilibrium condition in the forward market determines the number n of firms, the regulatory problem can be stated as one of setting the parameter θ of the resource-adequacy obligation to maximize the consumers' surplus $CS(\theta, n)$ subject to the resource feasibility constraint that $s^* \geq 0$ and the contestability constraint that $PS(\theta, n+1) = 0$.

To illustrate some numerical implications of these results we use an example with the following parameters:

α	1.0000
β	0.00001
c	1.0000
σ	1.0000
k	0.1000
h	0.0100

Figure 3 shows how the optimal choice of the parameter θ of the resource adequacy obligation varies when the number *n* of firms is assumed to be fixed. Figure 4 shows the consumers' surplus that results from this optimal choice of θ. Then Figure 5 shows how consumers' surplus varies with θ when the number of firms is endogenous.

In the following table, we illustrate with a numerical comparison of alternative market power mitigation mechanisms when there are two producers.

	Case 1 No contract	Case 2 Forward contract	Case 3 Option portfolio
Expected level of spot demand	0.998742	0.999864	0.999997
Expected level of demand with option coverage	0	0	0.969344
Quantity of forward contracts	0	0.046047	0
Expected spot price	125.82	13.58	0.33
Producers' profit	251	0	0
Consumers' surplus	749	973	999
Total surplus	999.19	972.85	999.35

Case 1 serves as the reference case, which involves no mitigation actions. Case 2 incorporates a form of forward contract with fixed quantity. Case 3 considers the option portfolio approach. These two cases support forward contracts as an effective means to reduce the producers' market power. Among the common observations, the producers' profits are reduced to virtually zero (after allowing recovery of entry costs); the expected spot price is lowered measurably; and the consumers' surplus increases accordingly.

However, only the option portfolio approach provides a consistent improvement in both consumers and the total social welfare. This corroborates the general insight that the option portfolio may be a superior approach and deserves further investigation.

5. Concluding Remarks

As emphasized in Sections 1 and 2, we view three of the basic problems of market design as aspects of the larger problem of risk management. The three problems now addressed via *ad hoc* measures are:

- Resource Adequacy: devising investment incentives for suppliers, and ICAP/ACAP requirements for LSEs, to ensure adequate resources to meet peak demands.
- Market-Power Mitigation: imposing bid caps and AMP procedures to suppress exercise of market power by suppliers.
- Long-Term Contracting: developing forward markets for bilateral contracts that mitigate the risks of price volatility in spot markets.

Viewed in a larger context, these problems are facets of the single problem of risk management. Long-term bilateral contracting enables the two parties to insure each other against volatile spot prices, but in fact this volatility is affected greatly by measures taken to attract entrants, stimulate investments, and suppress the market power of incumbents.

Our theme in this paper is that a special kind of long-term contract – a portfolio of options at various strike prices – has considerable potential to address all three problems simultaneously. Although options provide less complete insurance than fixed-price contracts, they share to some degree the advantage that they facilitate entry and investments by providing security for lenders. Most important, in our view, is their effect on the elasticity of net demand in the spot market. This effect diminishes market-power directly, and thereby reduces reliance on administrative procedures such as AMP.

The efforts of state PUCs to devise mechanisms that will ensure resource adequacy currently focus almost entirely on ICAP/ACAP requirements and on fixed-price bilateral contracts. We see a portfolio of options at various strike prices as a viable alternative. Options are less expensive and they address the precisely the PUCs' main concern about high prices. Indeed, few if any PUCs perceive insurance for suppliers against low prices as having the same importance as insurance for LSEs against high prices.

An important advantage of a portfolio of options at various strike prices is that it reduces spot-price volatility and diminishes market power by increasing the elasticity of net demand – even if aggregate demand is completely inelastic! As demonstrated in section 3.C, by requiring LSEs to purchase the right portfolio of options, a PUC can implement bounds on spot prices under various contingencies and ensure substantial elasticity of net demand in spot markets.

Figures

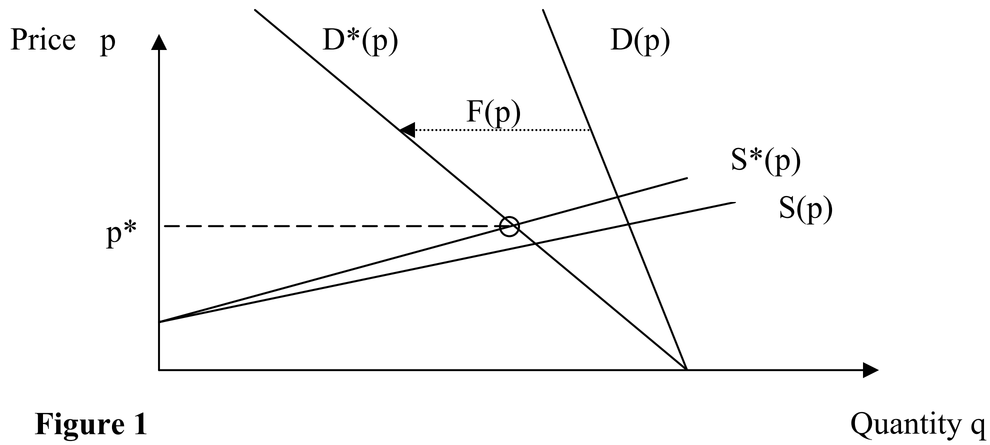

Figure 1

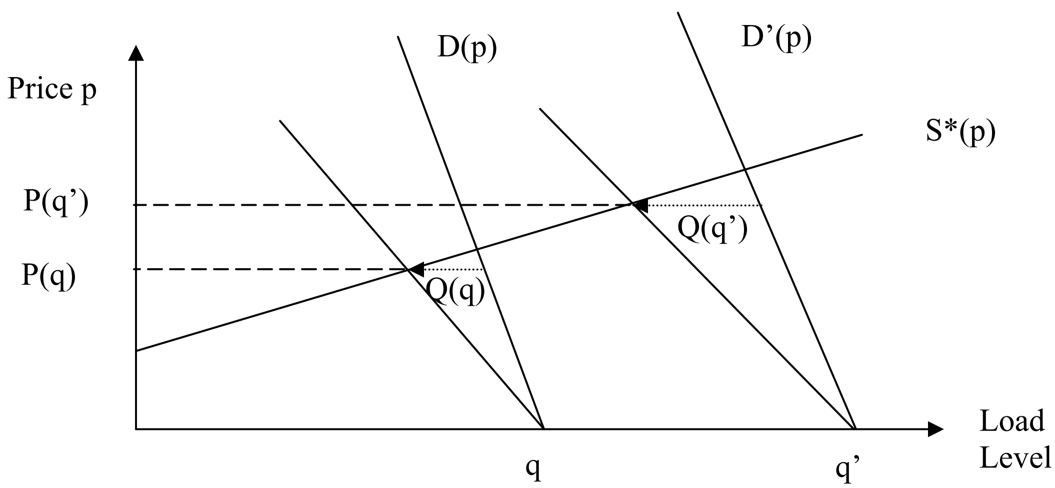

Figure 2

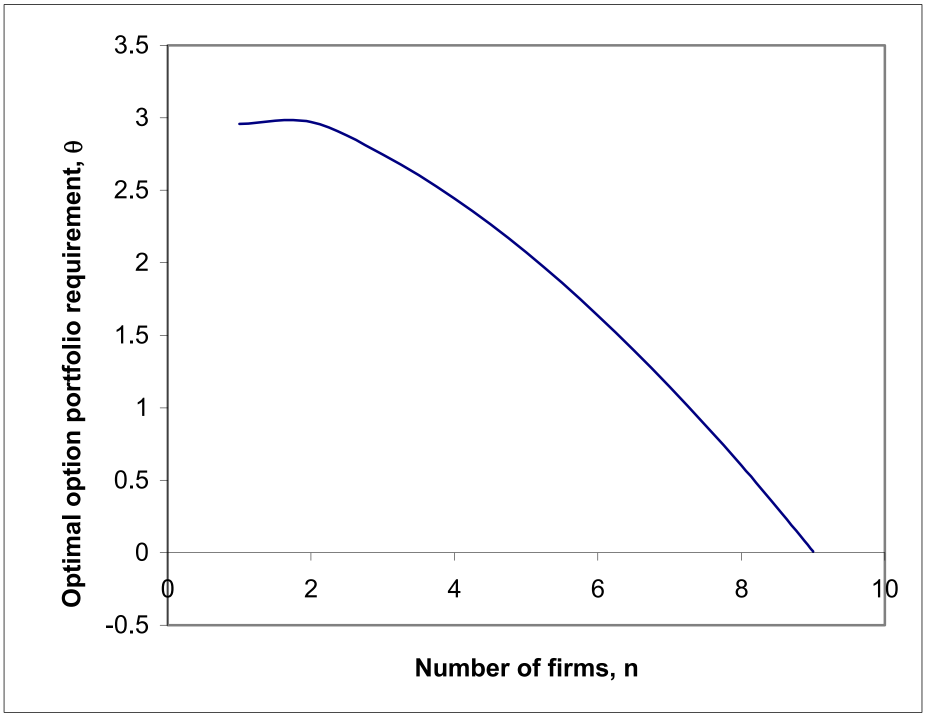

Figure 3

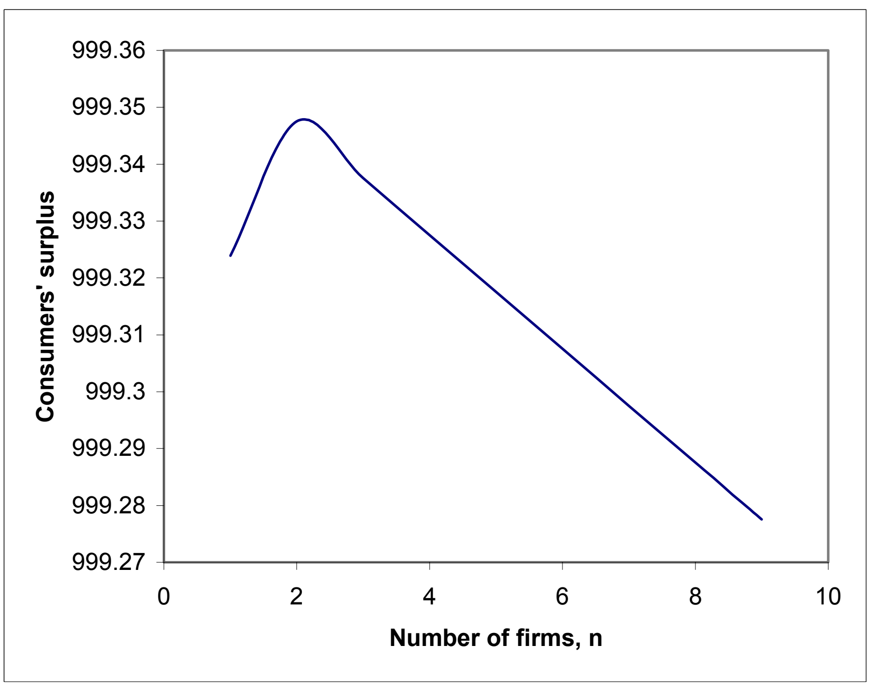

Figure 4

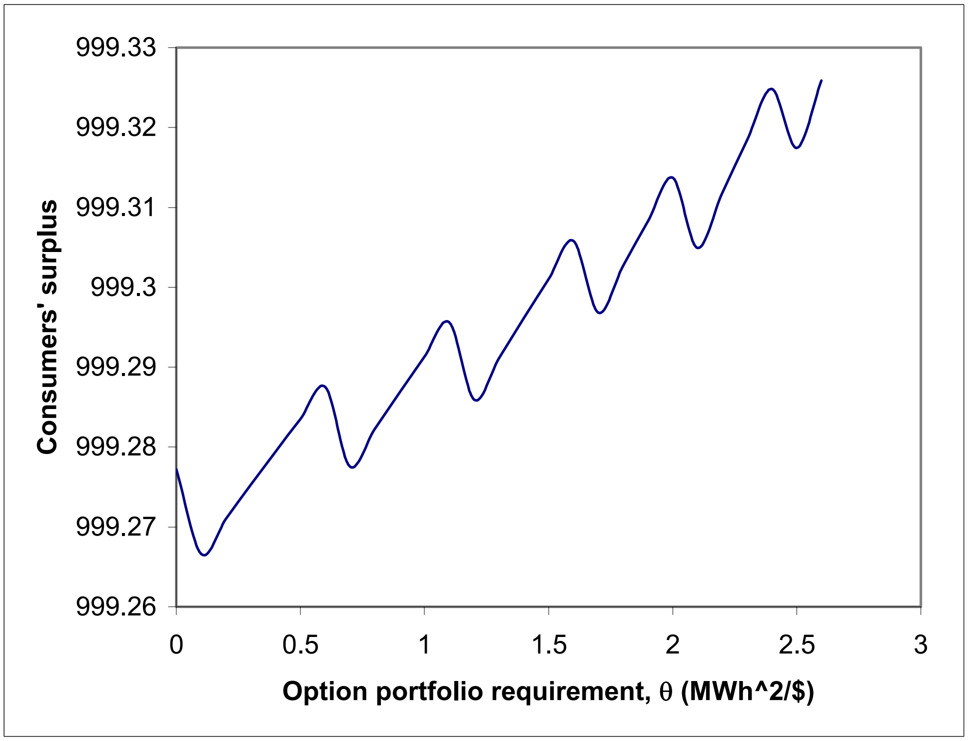
999.33
999.32
999.31
999.3
999.29
999.28
999.27
999.26
Consumers' surplus
0
0.5
1
1.5
2
2.5
3
Option portfolio requirement, θ (MWh^2/$)

Figure 5

Supplement: Derivation of key formulas for the symmetric case.

$$\left(x-r\right)=\left(1-\frac{cx}{y}\right)\left[\beta+\left(n-1\right)x\right]+\frac{\gamma Q}{\alpha}\left(\beta+nx\right)$$

$$y=\frac{cx\left[\beta+\left(n-1\right)x\right]}{\beta+\left(n-2\right)x+r+\gamma Q\left(\beta+nx\right)/\alpha}$$

$$2ky^2=cx^2\frac{\alpha^2+\sigma^2}{\left(\beta+nx\right)^2}$$

$$\left[\frac{\left[\beta+\left(n-1\right)x\right]}{\beta+\left(n-2\right)x+r+\gamma Q\left(\beta+nx\right)/\alpha}\right]^2=\frac{\alpha^2+\sigma^2}{2kc\left(\beta+nx\right)^2}$$

$$\mu\triangleq\sqrt{\frac{\alpha^2+\sigma^2}{2kc}}$$

$$n=1$$

$$\left[\frac{\beta}{\beta-x+r+\gamma Q\left(\beta+x\right)/\alpha}\right]^2=\frac{\mu^2}{\left(\beta+x\right)^2}$$

$$\frac{\beta}{\beta-x+r+\gamma Q\left(\beta+x\right)/\alpha}=\frac{\mu}{\left(\beta+x\right)}$$

$$\beta\left(\beta+x\right)=\mu\left[\beta-x+r+\gamma Q\left(\beta+x\right)/\alpha\right]$$

$$x=\frac{\mu\left(\beta+r+\gamma Q\beta/\alpha\right)-b^2}{\left(\mu+\beta-\mu\gamma Q/\alpha\right)}$$

$$n\geq 2$$

$$\frac{\left[\beta+\left(n-1\right)x\right]}{\beta+\left(n-2\right)x+r+\gamma Q\left(\beta+nx\right)/\alpha}=\frac{\mu}{\left(\beta+nx\right)}$$

$$\left(\beta+nx\right)\left[\beta+\left(n-1\right)x\right]=\mu\left[\beta+\left(n-2\right)x+r+\gamma Q\left(\beta+nx\right)/\alpha\right]$$

$$n(n-1)x^2+\left[\beta(2n-1)-\mu(n-2)-\frac{\mu\gamma Qn}{\alpha}\right]x-\left[\mu(\beta+r)-\beta^2+\frac{\mu\gamma Q\beta}{\alpha}\right]=0$$

$$x=\frac{2\left[\mu(\beta+r)-\beta^2+\frac{\mu\gamma Q\beta}{\alpha}\right]}{\left[\beta(2n-1)-\mu(n-2)-\frac{\mu\gamma Qn}{\alpha}\right]+\sqrt{\left[\beta(2n-1)-\mu(n-2)-\frac{\mu\gamma Qn}{\alpha}\right]^2+4n(n-1)\left[\mu(\beta+r)-\beta^2+\frac{\mu\gamma Q\beta}{\alpha}\right]}}$$