# Electricity Scarcity Pricing Through Operating Reserves

William W. Hogan[i]

Draft April 25, 2013

Suppressed prices in real-time markets provide inadequate incentives for both generation investment and active participation by demand bidding. An operating reserve demand curve developed from first principles would improve reliability, support adequate scarcity pricing, and be straightforward to implement within the framework of economic dispatch. This approach would be fully compatible with other market-oriented policies. Better scarcity pricing would also contribute to long-term resource adequacy.

## Introduction

Electricity resource adequacy programs often target the "missing money" problem.[1] The missing money problem arises when occasional market price increases are limited by administrative actions such as offer caps, out-of-market calls, and other unpriced actions. By preventing prices from reaching high levels during times of relative scarcity, these administrative actions reduce the payments that could be applied towards the fixed operating costs of existing generation plants and the investment costs of new plants. The resulting missing money reduces the incentives to maintain plant or build new generation facilities. In the presence of a significant missing-money problem, alternative means appear necessary to complement the market and provide the payments deemed necessary to support an appropriate level of resource adequacy.

In the United States experience, resource adequacy programs designed to compensate for the missing money create in turn a new set of problems in market design. The resource adequacy approaches become increasingly detailed and increasingly prescriptive to the point of severing the connections between major investment decisions and energy market incentives. Consideration of these unintended consequences reinforces interest in seeking ways to operate an electricity market with little or no money missing. (Newell, 2012)

Inadequate scarcity pricing is important but it is not the only problem that contributes to the missing money and concerns with resource adequacy. There is a related gap between traditional reliability planning standards and the standards that would follow from conventional cost-benefit calculations. (Telson, 1975) However, better scarcity pricing would both narrow the gap and simplify the policy discussion about the required level of resource adequacy.

Forward markets focus on long-term incentives. But the long-term is a succession of short-term markets. Inadequate scarcity pricing in the short-term design makes everything harder, now and in the future. Better scarcity pricing incentives would reinforce reliable operations now and in the expected future. (Hogan, 2005)

1 The characterization as "missing money" comes from Roy Shanker. For example, see (Shanker, 2003).

Wholesale electricity markets have demonstrated the feasibility of including operating reserve demand curves as part of economic dispatch. While demonstrating the practical possibilities, these other market implementations have not gone far enough incorporating first principles of economics and reliability to design the prices and related parameters of an operating reserve demand curve. A description of the benefits of better scarcity pricing, and an outline of a model for pricing of operating reserves, points to an opportunity to improve efficient electricity markets.

## Scarcity Pricing and Electricity Market Design

Efficient electricity market design follows the principles of bid-based, security-constrained, economic dispatch with locational prices. (Hogan, 2002) The prices for energy and ancillary services reflect the underlying requirements of the electricity system. These prices can vary substantially across locations, reflecting congestion in the transmission grid. Prices also vary over time due to the large differences in opportunity costs of different electric load and generation alternatives. These efficient prices can be highly volatile. Efficient electricity market design incorporates a variety of forward contracting opportunities and financial transmission rights to share the risks through market operations.

In principle, efficient electricity prices provide good incentives for both short-run operations and long-run investments. In the short run, prices reward generators who make their plants available when needed and in response to the changing dispatch conditions. The same prices provide incentives for loads to moderate demand during the most expensive hours and manage load to shift requirements to lower priced hours. In the long run, the expected value of future short-run payments for energy and ancillary services provides revenue for investment in new generating facilities or energy conservation.

In an idealized setting, this efficient design and associated electricity prices should be sufficient to support new investment when it is needed. In practice, as is now well known, actual electricity markets often produce results where energy and ancillary services prices are not sufficient to support new investment. (Newell et al., 2012) There are many practices that contribute to the aggregate pressure to keep prices too low to provide adequate incentives for load and generation, but the primary explanation is that prices do not adequately reflect the value of capacity scarcity. (Joskow, 2008) This focus on scarcity pricing provides a useful analytical framework and can incorporate many of the other practices that suppress electricity prices.

When there is excess available capacity, competitive pressure should drive the electricity market-clearing energy price to the variable opportunity cost of the most expensive generator running. Simultaneous consideration in the economic dispatch should produce compatible prices for ancillary services, with little or no value for additional capacity. This is the commonplace rule that animates most discussion about normal pricing conditions. However, when generating capacity becomes scarce it should become valuable. The price for operating reserve capacity should rise to reflect the scarcity conditions. The corresponding price for energy should increase to reflect this opportunity cost of reserve scarcity. This scarcity pricing could and should produce a large increase in prices under scarcity conditions, providing better incentives at just the right time when and where capacity would be especially needed.

In a fully developed wholesale electricity market, demand bidding would interact with scarcity prices in a natural way. Loads would specify the schedule of maximum willingness to pay, i.e., the energy demand curve, and when prices rose above these levels the load would dispatch down. In short-run equilibrium the resulting electricity price would be set by the variable opportunity cost of the least-expensive load being served. The electricity price could be different from the variable energy cost of the most expensive generator running, with the difference being the short-run scarcity price. In the long-run equilibrium the expected future scarcity prices would be just high enough to justify the capacity costs of new investment in generation and load management.

In the actual wholesale electricity markets we have, this idealized version of an "energy-only" electricity market does not exist. In particular, a missing part of the picture is the active participation of demand response bidding in the short-run market. (Faruqui, Hledik, & Sergici, 2010) This combines with the practices that suppress energy and ancillary services prices to create a type of vicious circle. With prices suppressed, there is not much incentive to participate in demand bidding or make the investments needed for active load management. The absence of demand bidding keeps demand up and prompts system operators to intervene in short-run operations in ways that suppress energy prices.

The result in wholesale electricity markets has been the "missing money" problem. There is money missing in the wholesale electricity market in the sense that average prices are not high enough to sustain new investment. Looking forward, the lack of investment raises the specter of reliability problems. This produces a variety of approaches to address the problem, provide the missing money, and meet the reliability needs of the system. The two most prominent approaches are to create forward capacity markets or to raise generator offer caps. Both methods present their own challenges.

## Forward Capacity Markets

Forward capacity requirements were a staple of electricity systems operated under cost-of-service regulation. In the absence of markets and without the incentives of market prices, investment decisions were made on the basis if planning requirements rather than market opportunities. Capacity obligations were part of the mechanism to avoid problems of under investment that would be rife if utilities could have leaned on their neighbors at regulated prices. The view that there should be some type of capacity obligation did not disappear with electricity restructuring, but the problem changed character when transported into the framework of electricity markets.

The basic idea of forward capacity markets is to arrange additional payments to those who offer capacity up to some estimated level of total capacity needed to meet projected reliability requirements. The total forward installed capacity requirement is either fixed as under the Independent System Operator New England (ISONE) or set according to a demand curve as in PJM Regional Transmission Organization (RTO) for the Mid-Atlantic states. Generators and demand resources then compete in a forward auction by making offers to supply capacity. The auction determines a market-clearing price for capacity that is paid to all clearing resources.

The capacity payment is intended to cover the missing money. The putative product is installed capacity and not energy, and the capacity payment is generally separate from the energy market payments. This approach, therefore, requires a regulatory definition of a "capacity product" that is unlike energy in that there is no simple way to measure and observe delivery. This forward capacity product definition is distinct from the definition and provision of short-run operating reserve capacity where the uncertainties are reduced from looking ahead years to looking ahead minutes or hours.

The many challenges of defining and implementing forward capacity markets have been under active discussion in studies and proceedings. It is difficult to properly define the capacity product, determine the amount and location of capacity needed many years ahead, and integrate diverse products that blend capacity and energy in a variety of configurations. Experience has shown that forward capacity markets, with their preset procurements, are subject to manipulation by generators and loads. (Harvey, Hogan, & Pope, 2013) For example, in PJM the independent market monitor regularly finds that aggregate energy markets are workably competitive and capacity market structures are not competitive. (Monitoring Analytics, 2012) This leads to requirements for capacity market regulations on offers and performance, bid mitigation, and other complications. The problems are fundamental. It is not easy to build a good forward capacity market model based on first principles.

Importantly, whatever the choice of whether to have a capacity market and what design to choose, the focus on the forward market produces at best weak connections with real time operations. The socialization of capacity payments does not send the right scarcity signals to generators or loads in real-time operations. Capacity markets may provide additional capacity that could be available in real-time. But capacity markets themselves do not create the correct incentives to operate capacity or change load in response to short-run scarcity conditions. Something more is needed. (Hogan, 2006) Capacity markets may help with scarcity, but not with scarcity pricing.

## Generator Offer Caps and Scarcity Pricing

In principle it would seem that scarcity pricing would arise naturally in the absence of offer caps on generators. An offer cap is one of the mechanisms for suppressing real-time prices. If there is no offer cap, or if the offer cap is very high, then generators could increase the offer prices during periods of scarcity and market-clearing prices would increase accordingly. For example, in 2012 the Public Utility Commission of Texas (PUCT) confirmed plans to increase generator offer caps.[2] At the same time, the PUCT was reviewing recommendations for possible modifications of the wholesale electricity market in the Electric Reliability Council of Texas (ERCOT). A common feature of these policies is a concern with the conundrum created by inadequate scarcity pricing in the short run and the possibility that resources will not be adequate to meet reliability requirements in the near future. (Newell et al., 2012)

Likewise, the model for scarcity pricing in PJM is based on increasing under offer caps under certain conditions indicating real-time capacity scarcity. (FERC, 2012a) A similar approach is found in decisions on the Southwest Power Pool (SPP) market design order from the Federal Energy Regulatory Commission (FERC). (FERC, 2012b)

---

[2] On October 25, 2012, the PUCT confirmed that "[t]he maximum wholesale rate will rise from $4,500 a megawatt hour now to $5,000 in June 2013, $7,000 in June 2014 and $9,000 in June 2015." http://www.star-telegram.com/2012/10/25/4365061/texas-regulators-vote-to-double.html.

A high offer cap may help in addressing one problem that leads to suppressed prices, but it does not deal with the treatment of operating reserves or the real-time reliability problems that arise under shortage conditions. Furthermore, to the extent that system operators turn to other out-of-market interventions to address reliability issues, the scarcity conditions need not translate into higher prices despite the high offer cap. (Joskow, 2008)

A problem with increasing offer caps arises in the tradeoff for mitigating market power. A principal purpose of generator offer caps is to mitigate the exercise of market power through economic withholding. The concern with an exercise of market power is especially acute during shortage conditions. However, a high offer-cap policy is built on the expectation that generators will withhold supply under scarcity conditions! This presents a series of problems for generators, other market participants, and regulators. For instance, generators may misjudge and make their offers too high, and their supply might not be taken in a high price market. Similarly, shortage conditions will give rise to high prices defined by generator offers. The observation of high scarcity prices would be difficult or impossible to distinguish from the exercise of market power. It will be difficult for regulators to maintain a hands-off policy that defends a high offer cap when scarcity conditions arise. And the expectation that regulators of the future may not have the ability to preserve the policy will inevitably dampen incentives for investment today in anticipation of this future.

A high offer cap would be consistent with a reasonable market for addressing scarcity market conditions, but it is not likely to be sufficient to ensure the appropriate market response. Furthermore, a high offer cap is not necessary to provide the proper incentives under scarcity conditions. An alternative approach would return to first principles and the role of operating reserves.

## Operating Reserve Demand Curves

Operating reserves for responsive spinning and quick start capacity are a regular feature of all electricity markets. These reserves are distinct from the installed capacity that is the focus of forward capacity markets. Operating reserves are a subset of the installed capacity that is both available and standing by to produce energy on short notice. In any given real-time dispatch interval, reserves are maintained to deal with unpredictable events such as a sudden surge in demand, loss of a generator, or loss of a transmission line. Balancing generation needs to ramp up very rapidly to meet the immediate emergency and to give the system operator time to reconfigure the energy dispatch.

Although it is difficult to forecast requirements for installed capacity many years ahead, it is a comparatively easier and more familiar task to forecast operating reserve requirements and availability for the next instant or parts of an hour. Supply, demand and transmission conditions are known. Weather forecasts are on hand. System operators have experience and procedures for defining and evaluating standby capabilities.

The most immediate requirement is for the operating reserves needed to meet security contingency conditions. The flows of electric power respond much faster than operators to sudden events such as the loss of a generator. In order to avoid cascading failures that could blackout most or all of the system, operators must maintain a minimum level of contingency reserves. From an economic perspective, a way to interpret and define these contingency reserves would be as the level at which the system operator would take administrative action such as involuntary controlled curtailment on selected loads in anticipation of the contingency in order to maintain minimum adequate capacity that could provide additional energy but must be kept in reserve.

System reliability would be improved if more operating reserves than the minimum were available in terms of response to increase generation or quickly decrease load. Over the next few minutes or parts of an hour, events may arise that deplete operating reserves and bring the system below the minimum contingency requirement, in which case the operator will have to impose involuntary load curtailments to restore the minimum contingency protection.

The importance of operating reserves has always been known, but the requirements for operating reserves were given only a simplified consideration in wholesale electricity market design. (Hogan, 2005) The assumption was that the operating reserve requirement at any moment and location could be represented by a fixed requirement, and that economic dispatch would produce simultaneous optimization that would incorporate the dispatch of energy and reserves. Pricing, especially during shortage conditions, would be provided by demand bidding to voluntarily reduce load at high prices, and the value of operating reserves would be determined by the implied scarcity prices. While this was a workable approximation in theory, it failed in practice when the associated demand bidding did not materialize.

One approach to the problem includes a welfare maximization framework with a mix of price sensitive and price-insensitive loads. (Joskow & Tirole, 2007) In this analysis, both load responses and reserve levels are set to be different depending on the outcome of uncertain events. By contrast, system operators set operating reserve levels ex ante and use the reserves established to respond to uncertain events through administrative actions. In this case, "[d]ecentralization through an operating reserves market together with a mandatory reserve ratio is delicate, as the price of reserves is extremely sensitive to small mistakes or discretionary actions by the system operator." (Joskow & Tirole, 2007, p. 82) There is an implicit value for incremental investment in reserves, but this value may not be fully recognized in the dispatch. There is a difference between the value of marginal investment in capacity and the marginal value of reserves in the dispatch.

One solution to this problem is to revisit the treatment of operating reserves within the framework of current economic dispatch models. In effect, the administrative requirement for a fixed level of operating reserves is equivalent to a vertical demand curve. As outlined above, this cannot be correct from first principles. The error and its impact could have been small with vigorous demand bidding in the dispatch, but the chicken-and-egg problem of inadequate scarcity pricing inhibiting demand bidding makes the error much more important and calls for a better representation of an operating reserve demand curve (ORDC).

To be sure, an operating reserve demand curve would be an administrative intervention in the market. But this is already true of the administrative requirement for operating reserves. In the presence of a necessary and inevitable operating reserve requirement, it is clear that the superior administrative rule would be a better model of the demand for operating reserves that goes beyond the fixed quantity requirement. (Hogan, 2005) The problem is not administrative rules per se, the problem arises with unpriced administrative actions.

The basic outline of an ORDC follows from the description above. The key connection is with the value of lost load (VOLL) and the probability that the load will be curtailed or similar emergency actions taken. Whenever there is involuntary load curtailment and the system has just the minimum of contingency operating reserves, then any increment of reserves would correspondingly reduce the load curtailment. Hence the price of operating reserves should be set at the value of lost load net of any energy savings.

At any other level of operating reserves, set to protect the system for events in the immediate future, the value of an increment of operating reserves would be the same VOLL multiplied by the probability that net load would increase enough in the coming interval to reduce reserves to the minimum level where emergency action would be taken to restore contingency reserves. Hence the incremental value of operating reserves would be the analogous to the product of the loss of load probability (LOLP) and VOLL, or LOLP*VOLL. This is similar in spirit to the capacity payment built into the original UK electricity pool market design, with the important distinction that the implementation is intended for real-time dispatch along with any forward markets rather than just for the day-ahead schedule. (Newbery, 1995)

The clearest example of the application of this logic is from the implementation by the Midwest Independent System Operator (MISO).

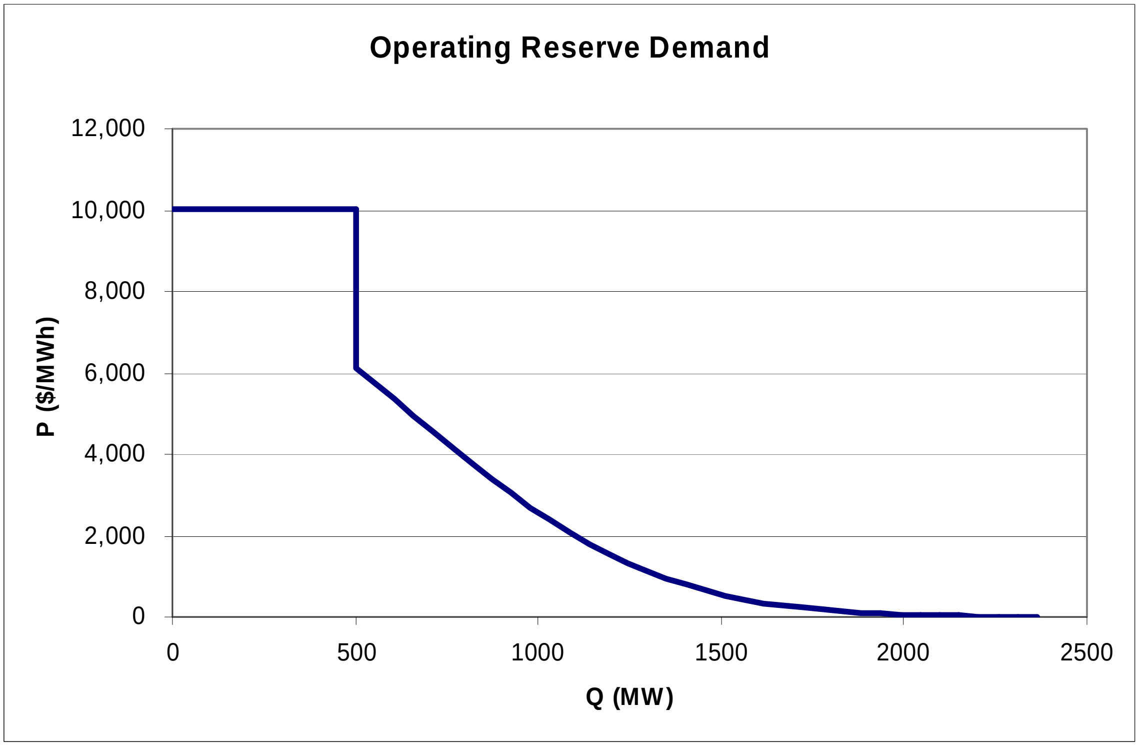
> "For each cleared Operating Reserve level less than the Market-Wide Operating Reserve Requirement, the Market-Wide Operating Reserve Demand Curve price shall be equal to the product of (i) the Value of Lost Load ("VOLL") and (ii) the estimated conditional probability of a loss of load given that a single forced Resource outage of 100 MW or greater will occur at the cleared Market-Wide Operating Reserve level for which the price is being determined. … The VOLL shall be equal to $3,500 per MWh." (MISO 2009, Schedule 28, Sheet 2226).

Combining this with the treatment of minimum contingency reserves, the resulting operating reserve demand curves would look like the hypothetical illustration in the accompanying Figure 1 drawn from a similar analysis in another region. In the figure the assumed VOLL is $10,000/MWh and the minimum contingency reserve requirement is assumed to be 500MW. The VOLL is associated with involuntary curtailments that would be incurred to preserve the minimum contingency reserves, and should be estimated as the average value for those who would be curtailed in a rolling blackout.

Note that emergency actions such as controlled load curtailments (i.e., rolling blackouts) are distinct from the event of a cascading failure that blacks out an entire system. (Joskow & Tirole, 2007) All electricity systems operate under security constraints intended to prevent cascading failures. The basic idea is to operate the system subject to so-called N-1 contingency analysis. The conservative constraints serve to limit the dispatch so that it could survive the any single contingency. The list of monitored contingencies is long, and can include multiple facilities failing at the same time. Part of the design of a workable ORDC is to integrate it with this contingency security constraint framework.

The demand curve discontinuity at 500MW occurs because of the probability that load will reduce over the interval more than the expected generation losses, in which case there is no need for load curtailment. Above the minimum reserve level, the shape of the demand curve follows the LOLP distribution. Importantly, a general property of an operating reserve demand curve derived from first principles is that the demand is not vertical and price does not drop to zero. Scarcity pricing would arise to some degree for all hours.

**Figure 1**




Depending on the needs in the regional system, the same principles could be generalized to include zonal requirements for operating reserves that interact with energy and economic dispatch, incorporate local interface constraints, and provide compatible short-term prices for operating reserves and interface capacity. With simultaneous optimization in the economic dispatch, the scarcity prices attributed to operating reserves apply as well to energy whenever there is a tradeoff between energy dispatch and operating reserve capacity. Hence, the scarcity prices would contribute to resolving the missing money problem for all generators actually providing energy or reserves.

## Features of Operating Reserve Demand Curves

The essential features of operating reserve demand curves include various properties that complement the basic electricity market design.

### *Reliability*

Market price incentives for energy and reserves would be better aligned with reliability requirements. By design, the scarcity prices would reflect the immediate reliability conditions, and both generators and load would see the benefits of responding to the market reliability needs. The focus on short-term operations would provide incentives that are difficult or impossible to capture in forward markets. Short-term changes in fuel availability, plant outages, demand conditions, load management practices and so on, would be reflected in current prices; emergency actions would be compatible with rather than in conflict with market incentives.

### *Consistent Design*

Since the operating reserve demand curve is indicated by first principles, it is inherently compatible with either an "energy only" market design or the various forward-market constructs. There is a possibility that an operating reserve demand curve by itself would provide sufficient incentives to support resource adequacy without further developing forward capacity markets. However, the benefits do not depend on resolving this larger question. There is no need to choose between the operating reserve demand curve and the other elements of electricity market design. Better scarcity pricing would help in all cases. Fixing the fundamental scarcity pricing related incentives should be the first order of business.

### *Demand Response*

Better pricing implemented through the operating reserve demand curve would provide an important signal and incentive for flexible demand participation in spot markets. This would help resolve the chicken-and-egg problem. Higher scarcity prices in some hours would provide the incentive for demand bidding and voluntary demand reductions. The demand bids could become more important and more significant than the operating reserve scarcity prices determined as part of the dispatch simultaneous optimization. The incentives would be reinforcing, with voluntary demand response supplementing operating reserves.

### *Price Spikes*

A higher price in some hours would be part of the solution. Furthermore, the contribution to the "missing money" from better pricing would involve many more hours and smaller price increases. Unlike relying only on high generator offers, which may exist for a few hours, higher prices would reflect the entire range of scarcity conditions. Only when conditions are truly extreme and there is a material threat to reliability, would and should prices approach the VOLL.

### *Practical Implementation*

The technical requirements for inclusion in economic dispatch and simultaneous optimization of energy and reserves are known and demonstrated. The New York Independent System Operator (NYISO, 2012), Independent System Operator of New England (ISONE, 2012), MISO and PJM (PJM, 2013) implementations dispose of any argument that it would be impractical to employ an operating reserve demand curve. The only material issues to address are the level of the appropriate VOLL price, translation into a workable approximation of the demand curve, and the preferred model of locational reserves.

### *Operating Procedures*

Implementing an operating reserve demand curve does not require changing the dispatch practices of system operators. Reserve and energy prices would be determined simultaneously treating decisions by the operators as being consistent with the adopted operating reserve demand curve. There would be a requirement to translate other emergency actions, such as voltage reductions, into equivalent operating reserve contributions. Emergency action policies could remain as at present, but the operating reserve demand impacts would be incorporated to make sure that out of market actions resulted in higher and not lower prices.

### *Multiple Reserves*

The demand curve would include different kinds of operating reserves, from responsive spinning reserves to standby non-spinning reserves. This would be similar to the cascade models for reserves and other ancillary services found in other market designs. For example, suppose there are two types of reserve categories, responsive and half-hour standby. The same rule described for the generic operating reserve demand curve would produce two demand curves derived from the VOLL and corresponding LOLP distribution over the relevant period. Responsive spinning reserves would be able to meet both requirements. Standby reserves that could be available in thirty minutes would be able to provide only the second type of operating reserves. Hence, the price of responsive spinning reserves would never be less, and likely would be more, than the price for standby reserves.

### *Market Power*

Better reserve pricing would remove ambiguity from analyses of high prices and distinguish (inefficient) economic withholding through high offers from (efficient) scarcity pricing derived from the operating reserve demand curve. Hence, lower generator offer caps would not be inconsistent with high market clearing prices for energy and reserves. But the higher market-clearing prices would be determined by the operating reserve demand curve. Generators would not have to withhold, economically or physically, to realize high market clearing prices. And during scarcity conditions regulators would have a simple explanation pointing to the operating reserve scarcity prices rather than trying to explain high offers by generators.

### *Hedging*

Forward contracts could still hedge forward loads. The contracts would reflect expected scarcity costs, and price in the risk, but there is nothing that would prevent the market from deploying a variety of financial contracts that incorporated scarcity prices on average while retaining efficient incentives at the margin.

### *Increased Costs*

The higher average energy costs from use of an operating reserve demand curve do not automatically translate into higher total system costs. In the aggregate, there is an argument that costs would be lower. Opportunities to meet reliability requirements would expand, with the stimulus of better scarcity pricing. Investments in responsive generation and load management that would be difficult or impossible without smarter prices could appear, and this would lower average costs. Higher scarcity prices would lower the missing money challenge, and thereby reduce the impacts of problems in any forward capacity markets that might exist.

## A Model Operating Reserve Demand Curve

The qualitative outline of an ORDC provides guidance for applying the basic principles. By its nature, an ORDC is an approximation of a complex reality. Some approximation is necessary to make the power dispatch problem tractable. Furthermore, a representation of the value of operating reserves is essential for establishing prices for energy and reserves. The details will depend on the actual system constraints and operating practices. The following summarizes the major elements. The appendix outlines a model that incorporates an explicit treatment of an ORDC provides a guide for implementation.

### *Economic Dispatch and Operating Reserves*

The assumption of the existence of an operating reserve demand curve simplifies the analysis. The demand curve gives rise to a reserve benefit function that can be included in the objective function for economic dispatch. The basic framework approximates the complex problem with a wide range of uncertainties and applies a pricing logic to match the actions of system operators. The main features include:

- **Single Period Model**. There is a static representation of the underlying dynamic problem. This static formulation is a conventional building block for a multi-period framework.
- **Deterministic Representation**. The single period dispatch formulation is based on bids, offers, and expected network conditions as in standard economic dispatch models. The operating reserve demand curve represents the value of uncertain uses of reserves without explicitly representing the uncertainty in the optimization model.
- **Security Constrained**. The economic dispatch model includes the usual formulation of N-1 contingency constraints to preclude cascading failures.
- **Ex Ante Dispatch**. The dispatch is determined before uncertainty about net load relative to forecast is revealed.
- **Expected Value for Reserves**. The reserve benefit function represents the expected value of avoiding involuntary load curtailments and similar emergency actions.
- **Multiple Reserve Types**. The model of the operating reserve demand allows for a typical cascade model of different reserve types. On line spinning reserves and fast start standby reserves interact to provide complementary reserve prices.
- **Administrative Balancing**. Subsequent uncertain events are treated according to administrative rules to utilize operating reserves to maintain system balance and minimize load curtailments.
- **Consistent Prices**. The model co-optimizes the dispatch of energy and reserves and produces a consistent set of prices for the period.

The framework allows for a variety of implementations with multiple zones, forward markets and other common aspects of electricity markets.

### *Multiple Zones*

The assumption that there is a single system wide use of operating reserve benefits may not apply in all regions. Although the steady-state constraints of transmission limits and loop flows apply to the base dispatch do not apply necessarily to the short-term use of operating reserves in stressed situation, it may not be that there are no location restrictions. The usual approach for operating reserves is to define a zonal requirement and interface constraints that limit the emergency movement of power.

The task is to define a locational operating reserve model that approximates and prices the dispatch decisions made by operators. To illustrate, consider the simplest case with one constrained zone and the rest of the system. The reserves are defined separately and there is a known transfer limit for the closed interface between the constrained zone and the rest of the system. The zonal requirements for operating reserves that interact with energy and economic dispatch, incorporate local interface constraints, and provide compatible short-term prices for operating reserves and interface capacity. This basic argument leads to a simple numerical model that can incorporate multiple embedded zones and interface constraints and be implemented with the co-optimization framework for energy and reserves. (Hogan, 2010)

### Forward Markets and Settlements

The ORDC framework outlined here illustrates the model that is most naturally thought of as the real-time dispatch formulation. The same idea would extend to forward markets such as a day-ahead market with the associated economic dispatch formulation. The principal modification would be to include virtual offers and bids for reserves as well as energy. The day-ahead market would incorporate an ORDC consistent with the real-time ORDC. The day-ahead market would settle with payments for both reserves and energy at the day-ahead prices.

The real-time model would be as outlined above. Settlements in real time would be based on real-time prices for energy and reserves. The payments would be made for deviations from the day-ahead schedules. The settlement rules including payment for reserves that did not provide energy would maintain the necessary indifference at the margin between providing energy and providing reserves.

### Back Cast ERCOT Case Study

An application of the model for the case of ERCOT illustrates the possible scale of the impacts. The purpose of the back cast was to suggest the scale of the scarcity prices that would have been relevant under the tight conditions that existed in 2011 and the greater abundance of capacity in 2012. The charge was not to simulate the full system to include changes in behavior and dispatch, which could be expected to occur. Rather the mandate was to assume the same offers, bids and dispatch that actually occurred, and then recalculate the energy and reserve prices. This provides a first order approximation of the effects of scarcity pricing.

“The back cast analysis of the price adder shows that the energy-weighted average energy price increases over a range of $7/MWh to $26.08/MWh in 2011 and $1.08/MWh to $4.5/MWh in 2012. This range results from different parameter settings that were used in the back cast. The back cast results for the average energy price increase with minimum contingency levels (X) of 1375 MW and 1750 MW are presented in Table 1. At the minimum contingency level, scarcity prices achieve the maximum allowed value.” (ERCOT Staff & Hogan, 2013)

**Table 1 : Energy-weighted average energy price adder (and Online reserve price) ($/MWh) for 2011 & 2012 for different VOLLs and minimum contingency levels (X)**

| VOLL | Energy-weighted average price increase with X at 1375 MW ($/MWh) | | | Energy-weighted average price increase with X at 1750 MW ($/MWh) | | |
|---|---|---|---|---|---|---|
| | 2011 | 2012 | 2011 & 2012 combined | 2011 | 2012 | 2011 & 2012 combined |
| **$5000/MWh** | 7.00 | 1.08 | 4.08 | 12.03 | 2.40 | 7.28 |
| **$7000/MWh** | 11.27 | 1.56 | 6.48 | 19.06 | 3.45 | 11.35 |
| **$9000/MWh** | 15.54 | 2.05 | 8.87 | 26.08 | 4.50 | 15.42 |

Source:(ERCOT Staff & Hogan, 2013)

By way of comparison, the ERCOT market monitor reports the "ERCOT-wide load-weighted average real-time energy price was $53.23 per MWh in 2011, a 35 percent increase from $39.40 per MWh in 2010." (Potomac Economics, 2012) Apparently co-optimization of energy and reserves would have produced material scarcity prices with a significant increment to the revenues in the energy market in Texas.

## Resource Adequacy and Scarcity Pricing

Better scarcity pricing would improve many aspects of market efficiency, and can be recommended independent of the problem of resource adequacy. In addition, better scarcity pricing would contribute towards making up the missing money and supporting resource adequacy.

Would better scarcity pricing be enough to resolve the resource adequacy problem? This is a larger topic, but a few observations connect to the resource adequacy question. First, posing a choice between capacity markets and better scarcity pricing is a false dichotomy. Even if the scarcity pricing is not enough and a long-term capacity market is necessary, better scarcity pricing would make the capacity market less important and thereby mitigate some of the unintended consequences.

Second, since the work of Telson there has been a recognition that the reliability planning standards that drive resource adequacy policy do not derive from a conventional cost benefit analysis. (Telson, 1975) The development of planning reserve margins starts with criteria such as the 1-event-in-10-years standard that appears to be a rule of thumb rather than a result derived from first principles. (Carden & Wintermantel, 2013) Depending on the details of filling in missing pieces in the economic analysis, the VOLL implied by the reliability standard is at least an order of magnitude larger than the range that would be consistent with actual choices and technology opportunities. (Telson, 1975) (Wilson, 2010) There is general agreement that applying reasonable estimates of VOLL and the cost-benefit criterion of welfare maximization would not support the typical planning reliability standards. (Newell, 2012) (Carden & Wintermantel, 2013)

Furthermore, in the absence of adequate scarcity pricing, the gap between the implications of conservative reliability standards and conventional economic analysis is obscured. If suppressed market prices are unavoidable, them something else like a capacity market is needed. But with better scarcity pricing the gap between economic cost benefit analysis and reliability planning standards will be more transparent. If better scarcity pricing, using a realistic estimate of the average VOLL, still leaves us with missing money and a resource adequacy gap, then new questions will arise about how best to close the gap. One approach could be to question the premise of the gap. If the probability and consequences of inadequate capacity are accounted for in scarcity pricing, this at least raises the possibility that the problem is not with the markets but with too conservative a reliability planning standard.

If the conservative planning standard is to be maintained and justified, then the justification would depend on a more nuanced argument for market failure that goes well beyond suppressed scarcity prices. One proposal is to include transfer payments in the benefits calculation, such as those that follow from excess capacity, depressed market prices and socialization of the cost of capacity payments. (Carden, Pfeifenberger, & Wintermantel, 2011) (Carden & Wintermantel,

2013) This may appeal to the beneficiaries of such market manipulation, but it can hardly be the foundation for a regulatory mandate that espouses support for economic efficiency and non-discrimination.

A more complicated argument might address dynamic issues about the credibility of future market returns versus future regulatory mandates. The volatility and uncertainty of market forces might tip the argument one way or the other. Or a different engineering argument might call for efforts to compensate for the errors of approximation in the engineering models that underpin both the reliability planning studies and the cost-benefit analyses. These efforts might include a margin of safety beyond the already conservative assumptions of security constrained N-1 contingency analysis. The final analysis may flow from the other work in this volume. But simple appeal to a missing money argument without further addressing the basics of market design would no longer suffice. At a minimum, a more transparent cost-benefit analysis for the resource adequacy requirement would be indicated to inform the policy decisions.

## Conclusion

The fundamental problem of inadequate real-time scarcity pricing has a fundamental solution. The need for a well-designed operating reserve demand curve was overlooked in early wholesale electricity market designs. There is today a much better understanding of why better scarcity pricing is essential, why demand bidding has not solved the problem, and how a well-designed operating reserve demand curve could fill in this significant gap. An operating reserve demand curve derived from first principles would address, in part, the missing money problem. In addition, the added benefits for improved operating conditions would warrant a better design for the demand curve independent of the needs for resource adequacy.

## Appendix

An outline of the basic framework illustrates the representation of an operating reserve demand curve.

### *Modeling Economic Dispatch and Operating Reserves*

The model presented here is a one-period DC-load model with co-optimization of reserves and energy. The notion is that the dispatch set at the beginning of the period must include some operating reserves that could deal with uncertain events over the period. The emphasis is on the co-optimization of energy and reserves to illustrate the major interactions with energy prices. The canonical example assumes the existence of a separable non-locational benefit function for reserves.

Here the various variables and functions include:

$d$ : Vector of locational demands

$g_R$ : Vector of locational responsive generation

$r_R$ : Vector of locational responsive reserves

$r_{NS}$ : Vector of locational non-spin reserves

$r_R^0$ : Aggregate responsive reserves

$r_{NS}^0$ : Aggregate non-spin reserves

$g_{NR}$ : Vector of locational generation not providing reserves

$B(d)$: Benefit function for demand

$C_k(g_k)$: Cost function for generation offers

$K_k$ : Generation Capacity

$R_k(r_k)$: Reserve value function integrating demand curves

$r_k^{\max}$ : Maximum Ramp Rate

$H, b$ : Transmission Constraint Parameters

$i$ : Vector of ones.

Assuming that unit commitment is determined, the stylized economic dispatch model is:

$$
(1) \quad
\begin{array}{lcc}
\underset{d, g_R, g_{NR}, r_R^0, r_R, r_{NS}^0, r_{NS} \geq 0; y}{Max} & B(d) - C_R(g_R) - C_{NR}(g_{NR}) + R_I(r_R^0) + R_{II}(r_{NS}^0) & \\
d - g_R - g_{NR} = y & \text{Net Loads} & \rho \\
i^t y = 0 & \text{Load Balance} & \lambda \\
Hy \leq b & \text{Transmission Limits} & \mu \\
g_R + r_R \leq K_R & \text{Responsive Capacity} & \theta_R \\
g_{NR} \leq K_{NR} & \text{Generation Only Capacity} & \theta_{NR} \\
r_{NS} \leq K_{NS} & \text{Non-Spin Capacity} & \theta_{NS} \\
i^t r_R = r_R^0 & \text{Responsive Reserves} & \gamma_R \\
i^t r_R + i^t r_{NS} = r_{NS}^0 & \text{Non-Spin Reserves} & \gamma_{NS} \\
r_R \leq r_R^{\max} & \text{Responsive Ramp Limit} & \eta_R \\
r_{NS} \leq r_{NS}^{\max} & \text{Non-Spin Ramp Limit} & \eta_{NS}.
\end{array}
$$

This formulation assumes that the non-spin reserve generators are not spinning and, therefore, cannot provide energy for the dispatch. The Non-Spin Reserve equation implements a cascade model for reserves, where both responsive and non-spin reserves contribute to the aggregate non-spin supply.

For the present discussion, the prices relationships follow from the usual interpretation of this economic dispatch model. This could be expanded to include unit commitment and extended LMP formulations (ELMP), but the basic insights would be similar. (Gribik, Hogan, & Pope, 2007)

An interpretation of the prices follows from analysis of the dual variables and the complementarity conditions. For an interior solution, the locational prices ($\rho$) are equal to the demand prices for load.

(2) $$\rho = \nabla B(d).$$

The same locational prices connect to the system lambda and the cost of congestion for the binding transmission constraints in the usual way.

(3) $$\rho = \lambda i + \mu^t H.$$

In addition, the locational prices equate with the marginal cost of generation plus the cost of scarcity.

(4) $$\rho = \nabla C_R(g_R) + \theta_R.$$

A similar relation applies for the value of non-reserve related generation.

(5) $$\rho = \nabla C_{NR}(g_{NR}) + \theta_{NR}.$$

The marginal value of responsive reserves connects to the scarcity costs of capacity and ramping limits.

(6) $$\theta_R + \eta_R = \gamma_R i + \gamma_{NS} i = \frac{dR_I(r_R^0)}{dr} i + \frac{dR_{II}(r_{NS}^0)}{dr} i.$$

The corresponding marginal value of non-spin reserves reflects the scarcity value for capacity and ramping limits.

(7) $$\theta_{NS} + \eta_{NS} = \gamma_{NS} i = \frac{dR_{II}\left(r_{NS}^{0}\right)}{dr} i.$$

If there are no binding ramp limits for responsive reserves, then $\eta_R = 0$ and from (6) we have $\theta_R$ as a vector where every element is the price of responsive reserves. Similarly, for the ramping limits on non-spin reserves, if these are not binding, then $\eta_{NS} = 0$ and from (7) we have $\theta_{NS}$ as a vector where every element is the price of non-spinning reserves.

### *An Approximate Operating Reserve Demand Curve*

This co-optimization model captures the principal interaction between energy offers and scarcity value. The assumption of a benefit function for reserves simplifies the analysis. Here, a derivation of a possible reserve benefit function provides a background for describing the form of an ORDC. To simplify the presentation, focus on the role of responsive reserves only. And consider only an aggregate requirement for reserves with no locational constraints.

To the various variables and functions add:

$f(x)$: Probability for net load change equal to $x$

Again, assume that unit commitment is determined. The stylized economic dispatch model includes an explicit description of the expected value of the use of reserves. For reserves here, only aggregate load matters. This reserve description allows for a one dimensional change in aggregate net load, $x$, and an asymmetric response where positive net load changes are costly and met with reserves and negative changes in net load are ignored. This model is too difficult to implement but it provides an interpretation of a set of assumptions that leads to an approximate ORDC. Here we ignore minimum reserve requirements to focus on the expected cost of the reserve dispatch.

The central formulation treats net load change $x$ and use of reserve, $\delta_x$ , to avoid involuntary curtailment. This produces a responsive benefit minus cost of $VOLL \cdot \left(i^t \delta_x\right) - \left(C_R\left(g_R + \delta_x\right) - C_R\left(g_R\right)\right)$ and this is weighted by the probability $f(x)$. This term enters the objective function summed for all non-negative values of $x$. The basic formulation includes:

$$\max_{d, g_R, g_{NR}, r_R, \delta_x \geq 0; y} \quad B(d) - C_R(g_R) - C_{NR}(g_{NR}) + \sum_{x \geq 0} \left( VOLLi^t \delta_x - \left( C_R(g_R + \delta_x) - C_R(g_R) \right) \right) f(x)$$

(8)

$$\begin{array}{lll}
d - g_R - g_{NR} = y & \text{Net Loads} & \rho \\
i^t y = 0 & \text{Load Balance} & \lambda \\
Hy \leq b & \text{Transmission Limits} & \mu \\
g_R + r_R \leq K_R & \text{Responsive Capacity} & \theta_R \\
i^t \delta_x \leq x, \forall x & \text{Responsive Utilization} & \gamma_x \\
\delta_x \leq r_R, \forall x & \text{Responsive Limit} & \varphi_x \\
g_{NR} \leq K_{NR} & \text{Generation Only Capacity} & \theta_{NR}.
\end{array}$$

This model accounts for all the uncertain net load changes weighted by the probability of outcome, and allows for the optimal utilization of reserve dispatch in each instance. This problem could produce scarcity prices that could differ across locations.

To approach the assessment of how to approximate reserves with a common scarcity price across the system, further simplify this basic problem.

1. Treat the utilization of reserves as a one-dimensional aggregate variable.
2. Replace the responsive reserve limit vector with a corresponding aggregate constraint on total reserves.
3. Utilize an approximation of the cost function, $\hat{C}$, for the aggregate utilization of reserves, and further approximate the change in costs with the derivative of cost times the utilization of reserves.

This set of assumptions produces a representation for the use of a single aggregate level of reserves for the system:

$$\max_{d, g_R, g_{NR}, r_R, \delta_x \geq 0; y} \quad B(d) - C_R(g_R) - C_{NR}(g_{NR}) + \sum_{x \geq 0} \left( VOLL\delta_x - \partial \hat{C}_R\left(i^t g_R\right) \delta_x \right) f(x)$$

(9)

$$\begin{array}{lll}
d - g_R - g_{NR} = y & \text{Net Loads} & \rho \\
i^t y = 0 & \text{Load Balance} & \lambda \\
Hy \leq b & \text{Transmission Limits} & \mu \\
g_R + r_R \leq K_R & \text{Responsive Capacity} & \theta_R \\
\delta_x \leq x, \forall x & \text{Responsive Utilization} & \gamma_x \\
\delta_x \leq i^t r_R, \forall x & \text{Responsive Limit} & \varphi_x \\
0 \leq r_R, & \text{Explicit Sign Constraint} & \omega_R \\
g_{NR} \leq K_{NR} & \text{Generation Only Capacity} & \theta_{NR}.
\end{array}$$

This formulation provides a reasonably transparent interpretation of the implied prices. Focusing on an interior solution for all the variables except $r_R$, we would have locational prices related to the marginal benefits of load:

(10) $\rho = \nabla B(d).$

The same locational prices connect to the system lambda and the cost of congestion for the binding transmission constraints.

(11) $\rho = \lambda i + H^t \mu.$

The locational prices equate with the marginal cost of generation-only plus the cost of scarcity when this generation is at capacity, which appears in the usual form.

(12) $\rho = \nabla C_{NR}\left(g_{NR}\right) + \theta_{NR}.$

The locational prices equate with the marginal cost of responsive generation and display the impact of reserve scarcity. First, the impact of changing the base dispatch of responsive generation implies:

$$\rho = \nabla C_R\left(g_R\right) + \sum_{x \geq 0}\left(\partial^2 \hat{C}_R\left(i^t g_R\right)\delta_x i\right) f\left(x\right) + \theta_R.$$

The second order term captures the effect of the base dispatch of responsive dispatch on the expected cost of meeting the reserve utilization. This term is likely to be small. For example, if we assume that the derivative $\partial \hat{C}_R$ is constant, then the second order term is zero.

When we account for the base dispatch of reserves, we have:

$$\theta_R = \sum_{x \geq 0} \varphi_x i + \omega_R .$$

When accounting for utilization of the reserves, we have:

$$\gamma_x + \varphi_x = \left(VOLL - \partial \hat{C}_R\left(i^t g_R\right)\right) f\left(x\right).$$

Let $r = i^t r_R$. Then for $x \leq r, \quad \varphi_x = 0; \quad x \geq r, \quad \gamma_x = 0$. Hence,

$$\theta_R = \sum_{x \geq r} \varphi_x i + \omega_R = \left(VOLL - \partial \hat{C}_R\left(i^t g_R\right)\right)\left(1 - F\left(r\right)\right) i + \omega_R .$$

Combining these, we can rewrite the locational price as:

(13) $\rho = \nabla C_R\left(g_R\right) + \sum_{x \geq 0}\left(\partial^2 \hat{C}_R\left(i^t g_R\right) i \delta_x\right) f\left(x\right) + \left(VOLL - \partial \hat{C}_R\left(i^t g_R\right)\right)\left(1 - F\left(r\right)\right) i + \omega_R.$

Equations (2) thru (13) capture our approximating model for aggregate responsive reserves. Here $1 - F\left(r\right) = Lolp\left(r\right)$. The term $\left(VOLL - \partial \hat{C}_R\left(i^t g_R\right)\right)\left(1 - F\left(r\right)\right)$ in (13) is the scarcity price of the ORDC. If the second order terms in (13) are dropped, then the scarcity price is the only change from the conventional generation only model. In practice, we would have to update this model to account for minimum reserve levels, non-spin, and so on, to include an estimate of $\overline{c} \approx \partial \hat{C}_R$ in defining the net value of operating reserves $v \approx VOLL - \overline{c}$ .

Note that under these assumptions the scarcity price is set according to the opportunity cost using $\hat{C}$ for the marginal responsive generator in the base dispatch. Depending on the accuracy of the estimate in $\hat{C}$, this seeks to maintain that the energy price plus scarcity price never exceeds the value of lost load.

Providing a reasonable estimate for $\hat{C}$ could be done either as an (i) exogenous constant, (ii) through a two pass procedure, or (iii) approximately inthe dispatch. For example, a possible procedure would define the approximating cost function as the least unconstrained cost,

$$\hat{C}\left(\hat{g}_R\right) = Min\left\{C\left(g_R\right) \middle| \hat{g}_R = i^t g_R\right\}.$$

This information would be easy to evaluate before the dispatch.

Construct the ORDC for responsive reserves that modifies (13) to incorporate the minimum or last resort reserves $X$ priced at $v$. Here $Lolp\left(r\right) = \text{Probability}\left(\text{Net Load Change} \geq r\right)$. For a candidate value of the aggregate responsive reserves define the corresponding value on the operating reserve demand curve:

$$\pi_R\left(r_R\right) = \begin{Bmatrix} Lolp\left(i^t r_R - X\right), & i^t r_R - X \geq 0 \\ 1, & i^t r_R - X < 0 \end{Bmatrix}$$

$$P_R\left(r_R\right) = v\pi_R\left(r_R\right).$$

This defines the ORDC for responsive reserves.

### *Multiple Reserve Types*

The organized market practice distinguishes several types of reserves. Setting aside regulation, the principal distinction is between "responsive" reserves (R) and "non-spin" reserves (NS). The ORDC framework can be adapted to include multiple reserves. This section summarizes one such modeling approach and relates it to the co-optimization examples above. The main distinction is that "responsive" reserves are spinning and have a quick reaction time. These reserves would be available almost immediately and could provide energy to meet increases in net load over the whole of the operating reserve period. By comparison, non-spin reserves are slower to respond and would not be available for the entire period.

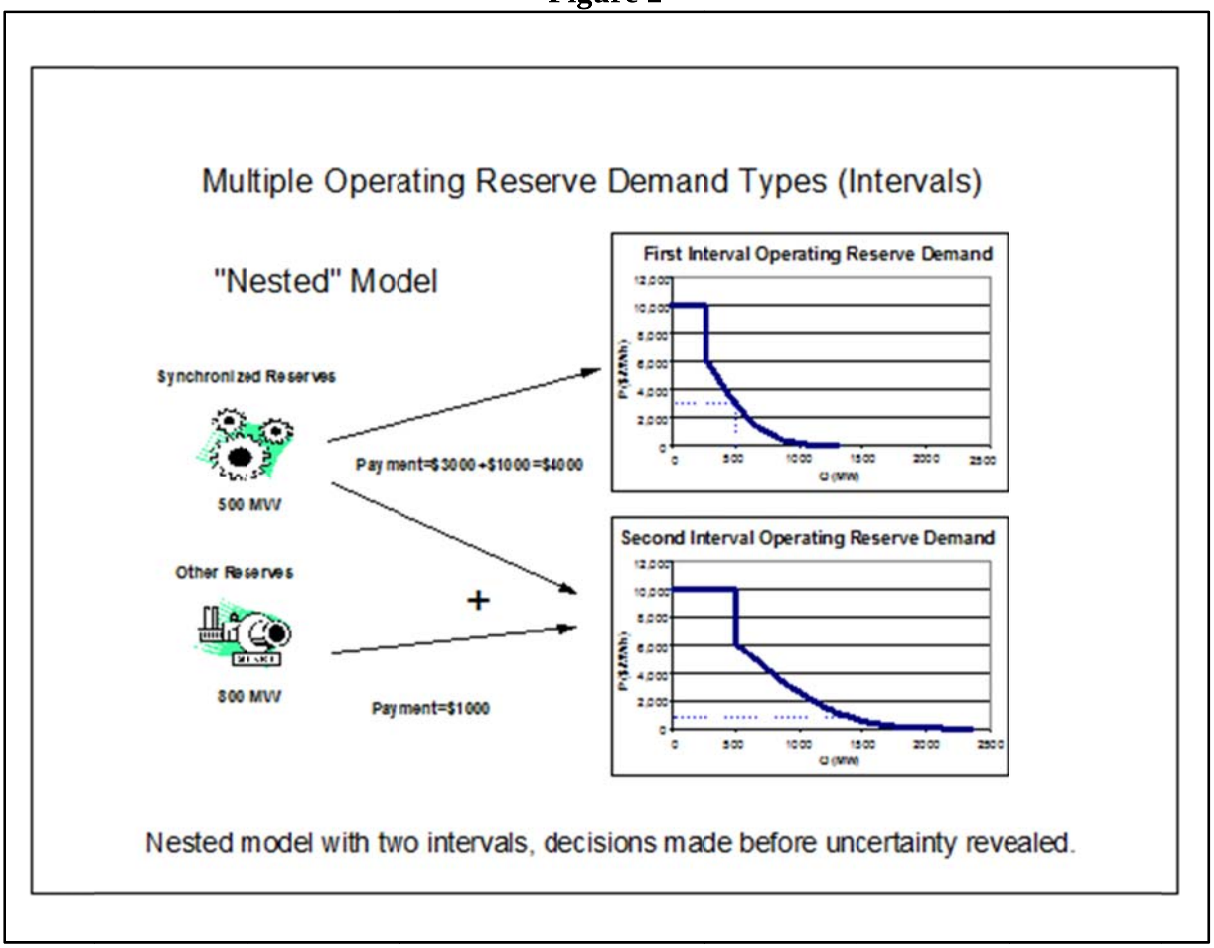
The proposed model of operating reserves approximates the complex dynamics by assuming that the uncertainty about the unpredicted change in net load is revealed after the basic dispatch is determined. The probability distribution of change in net load is interpreted as applying the change over the uncertain reserve period, say the next hour, divided into two intervals. Over the first interval, of duration ($\delta$), only the responsive reserves can avoid curtailments. Over the second interval of duration (1-$\delta$), both the responsive and non-spin reserves can avoid involuntary load shedding.

This formulation produces different values for the responsive and non-spin reserves. Let $v$ be the net value of load curtailment, defined as the value of lost load less the avoided cost of energy dispatch offer for the marginal reserve. The interpretation of the prices of reserves, $P_R$ and $P_{NS}$, is the marginal impact on the load curtailment times $Lolp$, the probability of the net change in load being greater that the level of reserves, $r_R$ and $r_{NS}$. This marginal value differs for the two intervals, as shown in the following table:

| Marginal Reserve Values | | |
|---|---|---|
| | Interval I | Interval II |
| Duration | $\delta$ | 1-$\delta$ |
| $P_R$ | $vLolp_I\left(r_R\right)$ | $vLolp_{I+II}\left(r_R + r_{NS}\right)$ |
| $P_{NS}$ | 0 | $vLolp_{I+II}\left(r_R + r_{NS}\right)$ |

This formulation lends itself to the interpretation of Figure 2 where there are two periods with different demand curves and the models are nested. In other words, responsive reserves $r_R$ can meet the needs in both intervals and the non-spin reserves $r_{NS}$ can only meet the needs for the second interval.

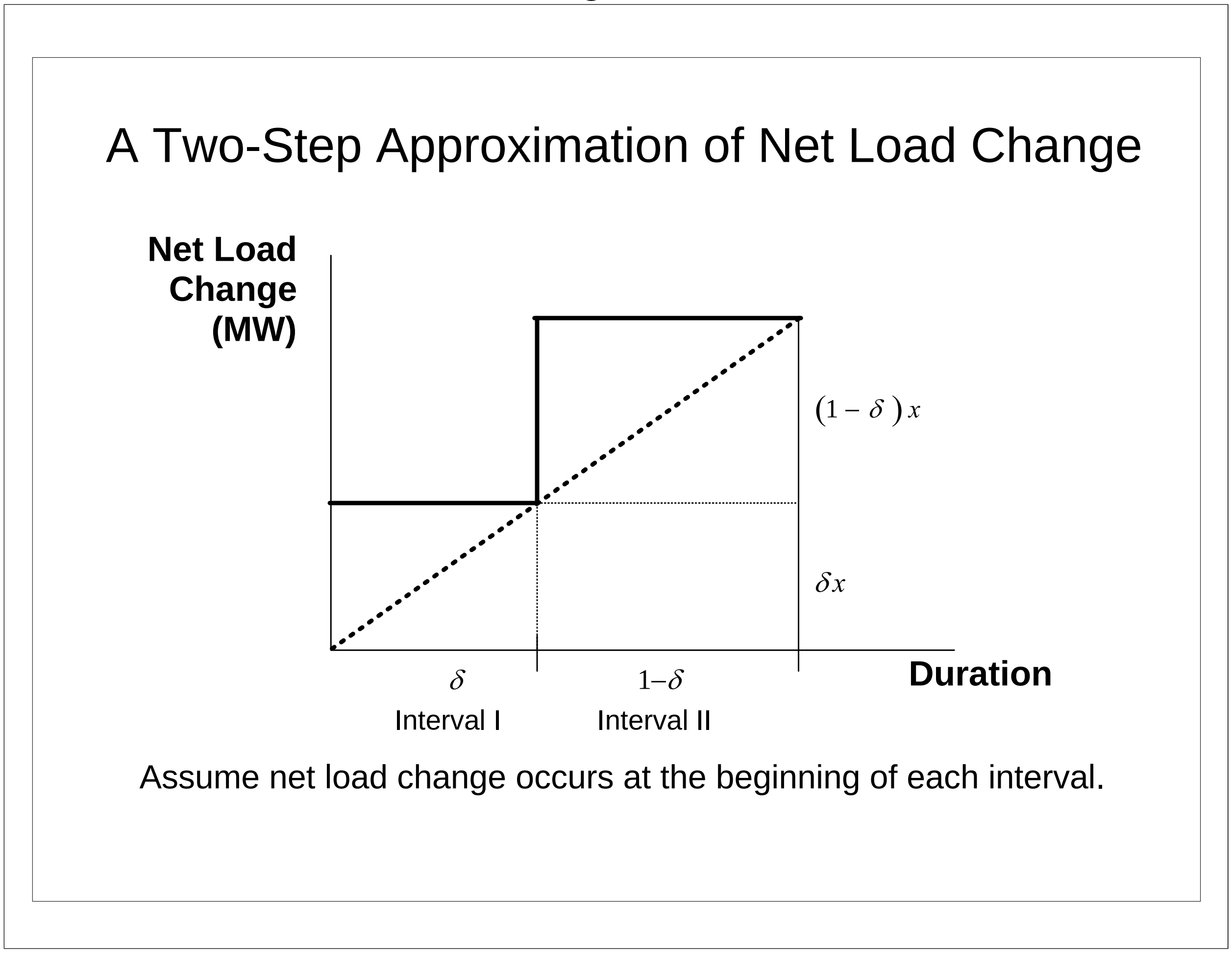
**Figure 2**




In order to keep the analysis of the marginal benefits of more reserves simple, there is an advantage of utilizing a step function approximation for the net load change. (This keeps the marginal value in an interval constant, and we don't have to compute expectations over the varying net load change. We only need the total LOLP over that interval.)

The standard deviation of the change in net load is for the total over the period. If the change is spread out over the period, then on average it would be more like the diagonal dashed line in Figure 3. An alternative two-step approximation in Figure 3 assumes that the net load change in the first interval, when only responsive reserves can respond, is proportional to the total load change, and the second step captures the total change at the beginning of the second interval.

**Figure 3**




During the first interval, only the responsive reserves apply. In the second interval, both responsive and non-spin reserves have been made available to help meet the net change in load. Suppose that there are two variables $y_I, y_{II}$ representing the incremental net load change in the two intervals. Further assume that the two variables have a common underlying distribution for a variable $z$ but are proportional to the size of the interval. Then, assuming independence and with $x$ the net load change over the full two intervals, we have:

$$
\begin{aligned}
&E\left(y_I\right)=E\left(\delta z\right)=\delta E\left(z\right),\\
&E\left(y_{II}\right)=E\left(\left(1-\delta\right)z\right)=\left(1-\delta\right)E\left(z\right).\\
&Var\left(y_I\right)=Var\left(\delta z\right)=\delta^2 Var\left(z\right),\\
&Var\left(y_{II}\right)=Var\left(\left(1-\delta\right)z\right)=\left(1-\delta\right)^2 Var\left(z\right).\\
&E\left(z\right)=E\left(y_I+y_{II}\right)=E\left(x\right)=\mu.\\
&Var\left(x\right)=Var\left(y_I+y_{II}\right)=Var\left(y_I\right)+Var\left(y_{II}\right)=\left(\delta^2+\left(1-\delta\right)^2\right)Var\left(z\right).\\
&Var\left(z\right)=\frac{Var\left(x\right)}{\delta^2+\left(1-\delta\right)^2}=\frac{\sigma^2}{\delta^2+\left(1-\delta\right)^2}.
\end{aligned}
$$

The distinction here is that the implied variance of the individual intervals is greater compared to the one-draw assumption, even though the total variance of the sum over the two intervals is the same. This is simply an impact square root law for the standard deviation of the sums of independent random variables.

Hence, for the first interval, the standard deviation is $\frac{\delta\sigma}{\sqrt{\delta^2+\left(1-\delta\right)^2}}$, where $\sigma$ is the standard deviation of the net change in load over both intervals.

Here the different distributions refer to the net change in load over the first interval, and over the sum of the two intervals. The distribution over the sum is just the same distribution for the whole period that was used above. Then $y_I \sim Lolp_I, y_I+y_{II} \sim Lolp_{I+II}$. A workable approximation would be to utilize the normal distribution for the net load change.

As before, there would be an adjustment to deal with the minimum reserve to meet the max contingency. The revised formulation would include:

$$
\begin{aligned}
&\pi_R\left(r_R\right)=\begin{Bmatrix} Lolp_I\left(i^t r_R - X\right), & i^t r_R - X \geq 0 \\ 1, & i^t r_R - X < 0 \end{Bmatrix}\\
&\pi_{NS}\left(r_R, r_{NS}\right)=\begin{Bmatrix} Lolp_{I+II}\left(i^t r_R + i^t r_{NS} - X\right), & i^t r_R + i^t r_{NS} - X \geq 0 \\ 1, & i^t r_R + i^t r_{NS} - X < 0 \end{Bmatrix}\\
&P_R\left(r_R, r_{NS}\right)=v*\left(\delta*\pi_R\left(r_R\right)+\left(1-\delta\right)*\pi_{NS}\left(r_R, r_{NS}\right)\right),\\
&P_{NS}\left(r_R, r_{NS}\right)=v*\left(1-\delta\right)*\pi_{NS}\left(r_R, r_{NS}\right).
\end{aligned}
$$

This formulation lends itself to implementation in the co-optimization model. For example, given benchmark estimates for each type of reserves, $\left(\hat{r}_R, \hat{r}_{NS}\right)$, the problem becomes separable in responsive and non-spin reserves. A numerical integration of $P_R\left(r_R, \hat{r}_{NS}\right)$ and $P_{NS}\left(\hat{r}_R, r_{NS}\right)$ would produce the counterpart benefit functions, $R_I\left(r_R^0\right), R_{II}\left(r_{NS}^0\right)$. With weak interactions between the types of reserves, the experience with this type of decomposition method suggest that updating the benchmark estimates in an iterative model could produce rapid convergence to the simultaneous solution. (Ahn & Hogan, 1982)

## References

Ahn, B., & Hogan, W. (1982). On convergence of the PIES algorithm for computing equilibria. *Operations Research*, *30*(2), 281–300. Retrieved from http://or.journal.informs.org/content/30/2/281.short

Carden, K., Pfeifenberger, J., & Wintermantel, N. (2011). The Value of Resource Adequacy. *Publci Utilities Fortnightly*, *149*(3), 34–40. Retrieved from http://proquest.umi.com.ezp-prod1.hul.harvard.edu/pqdlink?index=19&did=2297968081&SrchMode=3&sid=2&Fmt=6&VInst=PROD&VType=PQD&RQT=309&VName=PQD&TS=1307842201&clientId=18857&aid=2

Carden, K., & Wintermantel, N. (2013). *The Economic Ramifications of Resource Adequacy White Paper*. Astrape Consulting. Retrieved from http://www.naruc.org/RFP/NARUC2012-RFP009-DE0316_EISPC_RFP.pdf

ERCOT Staff, & Hogan, W. W. (2013). *Back Cast of Interim Solution B + to Improve Real-Time Scarcity Pricing White Paper*. Retrieved from http://www.ercot.com/content/news/presentations/2013/White Paper - Back Cast of Interim Solution B+ to Improve Re.pdf

Faruqui, A., Hledik, R., & Sergici, S. (2010). Rethinking Prices. *Public Utilities Fortnightly*, (January), 31–39. Retrieved from www.fortnightly.com

FERC. (2012a). PJM Compliance Order. Federal Energy Regulatory Commission, Docket No. ER09-1063-004. Retrieved from http://www.ferc.gov/whats-new/comm-meet/2012/041912/E-2.pdf

FERC. (2012b). SPP Order Conditionally Accepting Tariff Revisions to Establish Energy Markets. Federal Energy Regulatory Commission, Docket Nos. ER12-1179-000 ER12-1179-001. Retrieved from http://www.google.com/url?sa=t&rct=j&q=&esrc=s&frm=1&source=web&cd=1&cad=rja&ved=0CCEQFjAA&url=http%3A%2F%2Fwww.ferc.gov%2Fwhats-new%2Fcomm-meet%2F2012%2F101812%2FE-4.pdf&ei=BBWQUMmrLc3h0wH2soCIDw&usg=AFQjCNFBAeptP8181tVyyxOrEHAGmrPHuA&sig2=bvjKLAHXxAVALEvAbyZHRg

Gribik, P. R., Hogan, W. W., & Pope, S. L. (2007). Market-Clearing Electricity Prices and Energy Uplift. Retrieved from http://www.hks.harvard.edu/fs/whogan/Gribik_Hogan_Pope_Price_Uplift_123107.pdf

Harvey, S. M., Hogan, W. W., & Pope, S. L. (2013). *Evaluation of the New York Capacity Market*. Retrieved from http://www.nyiso.com/public/webdocs/markets_operations/documents/Studies_and_Reports/Studies/Market_Studies/Final_New_York_Capacity_Report_3-13-2013.pdf

Hogan, W. W. (2002). Electricity market restructuring: reforms of reforms. *Journal of Regulatory Economics*, *21*(1), 103–132. Retrieved from http://www.springerlink.com/index/M488771227KN3R02.pdf

Hogan, W. W. (2005). On an “Energy Only” Electricity Market Design for Resource Adequacy. Retrieved from http://www.hks.harvard.edu/fs/whogan/Hogan_Energy_Only_092305.pdf

Hogan, W. W. (2006). Resource Adequacy Mandates and Scarcity Pricing. Retrieved from http://www.hks.harvard.edu/fs/whogan/Hogan_PJM_Energy_Market_022306.pdf

Hogan, W. W. (2010). Scarcity Pricing and Locational Operating Reserve Demand Curves. Retrieved from http://www.hks.harvard.edu/fs/whogan/Hogan_FERC_060210.pdf

ISONE. (2012). Overview of New England’s Wholesale Electricity Markets and Market Oversight. Independent System Operator New England. Retrieved from http://www.iso-ne.com/pubs/spcl_rpts/2012/markets_overview_final_051512.pdf

Joskow, P. (2008). Capacity payments in imperfect electricity markets: Need and design. *Utilities Policy*, *16*(3), 159–170. doi:10.1016/j.jup.2007.10.003

Joskow, P., & Tirole, J. (2007). Reliability and competitive electricity markets. *The RAND Journal of Economics*, *38*(1), 60–84. doi:10.1111/j.1756-2171.2007.tb00044.x

Midwest ISO. (2009). FERC Electric Tariff. Retrieved from http://www.midwestiso.org/publish/Document/1d44c3_11e1d03fcc5_-7cf70a48324a/Schedules.pdf?action=download&_property=Attachment

Monitoring Analytics. (2012). State of the Market Report for PJM, 2011. Monitoring Analytics, March 15. Retrieved from http://www.pjm.com/documents/reports/~/media/documents/reports/state-of-market/2011/2011-som-pjm-volume2.ashx

Newbery, D. (1995). Power markets and market power. *The Energy Journal*. Retrieved from http://www.jstor.org/stable/10.2307/41322605

Newell, S. (2012). *Resource Adequacy in ERCOT: “Composite” Policy Options*. Retrieved from http://www.ercot.com/content/news/presentations/2012/Newell Oct25 PUCT Workshop_Composite Options_10-19-2012.pdf

Newell, S., Spees, K., Pfeifenberger, J., Mudge, R., DeLucia, M., & Carlton, R. (2012). *ERCOT Investment Incentives and Resource Adequacy*. Retrieved from http://www.ercot.com/content/news/presentations/2012/Brattle ERCOT Resource Adequacy Review - 2012-06-01.pdf

NYISO. (2012). NYISO Ancillary Services Manual. New York Independent System Operator. Retrieved from http://www.nyiso.com/public/webdocs/documents/manuals/operations/ancserv.pdf

PJM. (2013). *PJM Manual 11 : Energy & Ancillary Services Market Operations Revision : 59 Effective Date : April 1 ,*. Retrieved from http://www.pjm.com/~/media/documents/manuals/m11.ashx

Potomac Economics. (2012). *2011 State of the Market Report for the ERCOT Wholesale Electricity Markets*. Retrieved from http://www.puc.texas.gov/industry/electric/reports/ERCOT_annual_reports/2011annualreport.pdf

Shanker, R. (2003). Comments on Standard Market Design: Resource Adequacy Requirement. *Federal Energy Regulatory Commission, Docket RM01-12-000*. Retrieved from http://elibrary.ferc.gov/idmws/common/opennat.asp?fileID=9619272

Telson, M. L. (1975). The economics of alternative levels of reliability for electric power generation systems. *The Bell Journal of Economics*, *6*(2), 679–694. Retrieved from http://www.jstor.org/stable/3003250

Wilson, J. (2010). Reconsidering Resource Adequacy: Part 1. *Public Utilities Fortnightly*, (April), 33–39. Retrieved from www.fortnightly.com

---

[i] William W. Hogan is the Raymond Plank Professor of Global Energy Policy, John F. Kennedy School of Government, Harvard University. This paper was prepared for a forthcoming issue of Economics of Energy and Environmental Policy. The work is based on research for the Harvard Electricity Policy Group and for the Harvard-Japan Project on Energy and the Environment. This paper draws joint work with John Dumas, David Maggio, Sai Moorty, Resmi Surendran of ERCOT. Support was provided by GDF SUEZ Energy Resources NA. The author is or has been a consultant on electric market reform and transmission issues for Allegheny Electric Global Market, American Electric Power, American National Power, Aquila, Atlantic Wind Connection, Australian Gas Light Company, Avista Energy, Barclays Bank PLC, Brazil Power Exchange Administrator (ASMAE), British National Grid Company, California Independent Energy Producers Association, California Independent System Operator, California Suppliers Group, Calpine Corporation, Canadian Imperial Bank of Commerce, Centerpoint Energy, Central Maine Power Company, Chubu Electric Power Company, Citigroup, Comision Reguladora De Energia (CRE, Mexico), Commonwealth Edison Company, COMPETE Coalition, Conectiv, Constellation Energy, Constellation Energy Commodities Group, Constellation Power Source, Coral Power, Credit First Suisse Boston, DC Energy, Detroit Edison Company, Deutsche Bank, Deutsche Bank Energy Trading LLC, Duquesne Light Company, Dynegy, Edison Electric Institute, Edison Mission Energy, Electricity Corporation of New Zealand, Electric Power Supply Association, El Paso Electric, Exelon, Financial Marketers Coalition, FTI Consulting, GenOn Energy, GPU Inc. (and the Supporting Companies of PJM), GPU PowerNet Pty Ltd., GDF SUEZ Energy Resources NA, GWF Energy, Independent Energy Producers Assn, ISO New England, LECG LLC, Luz del Sur, Maine Public Advocate, Maine Public Utilities Commission, Merrill Lynch, Midwest ISO, Mirant Corporation, MIT Grid Study, JP Morgan, Morgan Stanley Capital Group, National Independent Energy Producers, New England Power Company, New York Independent System Operator, New York Power Pool, New York Utilities Collaborative, Niagara Mohawk Corporation, NRG Energy, Inc., Ontario Attorney General, Ontario IMO, Pepco, Pinpoint Power, PJM Office of Interconnection, PJM Power Provider (P3) Group, PPL Corporation, Public Service Electric & Gas Company, Public Service New Mexico, PSEG Companies, Reliant Energy, Rhode Island Public Utilities Commission, San Diego Gas & Electric Company, Sempra Energy, SPP, Texas

Genco, Texas Utilities Co, Tokyo Electric Power Company, Toronto Dominion Bank, Transalta, Transcanada, TransÉnergie, Transpower of New Zealand, Tucson Electric Power, Westbrook Power, Western Power Trading Forum, Williams Energy Group, and Wisconsin Electric Power Company. The views presented here are not necessarily attributable to any of those mentioned, and any remaining errors are solely the responsibility of the author. (Related papers can be found on the web at www.whogan.com).