The Brattle Group

A Comparison of PJM's RPM with Alternative Energy and Capacity Market Designs

September 2009

Johannes Pfeifenberger
Kathleen Spees
Adam Schumacher

Prepared for

PJM Interconnection, L.L.C.

TABLE OF CONTENTS

Acknowledgements

The authors would like to acknowledge the contributions of Sam Newell, Jose Garcia, and Serena Hesmondhalgh. Opinions expressed in this article, as well as any errors or missions, are the authors' alone.

I. EXECUTIVE SUMMARY

The Brattle Group has been commissioned by PJM Interconnection L.L.C. ("PJM") to produce a whitepaper that compares and contrasts the PJM Reliability Pricing Model ("RPM") with alternative power market designs and evaluates each design's ability to maintain resource adequacy. We examine six alternative market designs, beginning with pure energy-only markets and progressing to more highly-structured designs that rely on administratively-determined capacity payments or explicit current or forward reserve requirements in possible combination with centralized capacity markets. For each of these market designs, we summarize relevant U.S. and international experience and discuss the advantages and disadvantages of the various approaches.

- *Energy-Only Markets.* In energy-only markets, resources obtain revenues solely through markets for energy and ancillary services without additional payments for capacity. Suppliers recover their investment costs through periodic severe price spikes. *Pure* energy-only markets do not rely on resource adequacy requirements to ensure a desired level of reliability; the level of reliability is determined solely through the marketplace. However, such a design has not been implemented, as *real-world* examples of energy-only markets rely on various out-of-market mechanisms to maintain reliability. Most energy-only market designs in the U.S. and internationally also cap market prices at levels well below the values required to let market forces choose the desired level of reliability. *Examples*: Alberta, Great Britain, Australia's National Energy Market ("NEM"), Electric Reliability Council of Texas ("ERCOT"), Ontario.

- *Energy Markets with Administratively-Determined Capacity Payments.* Adding an administratively-determined capacity payment to an energy market allows the market operator to maintain a desired level of capacity investment, while imposing price caps and other mitigation measures on energy markets to avoid severe price spikes. This market design does not directly impose reliability standards (*e.g.*, reserve requirements) on load serving entities ("LSEs"); the achieved reliability level is the result of suppliers' response to the administratively-set capacity payments. The absence of reserve requirements makes this the only market design that allows for differentiation of capacity payments to new and existing resources, though doing so introduces significant market distortions. By recovering the cost of capacity payments as an uplift charge on energy consumed rather than based on peak loads, capacity payment mechanisms also often distort price signals to consumers, thereby undermining demand response. *Examples*: Chile, Colombia, South Korea, Spain, Peru.

- *Energy Markets with Reserve Requirements (without Centralized Capacity Markets).* This market design imposes an administratively-determined level of capacity reserves—ideally on a locational basis in transmission constrained market areas—that LSEs must maintain through either resource ownership or bilateral contracting. The reserve requirement imposed on LSEs—which is monitored and enforced during or immediately prior to the delivery period—creates a bilateral capacity market in which both demand- and supply-side resources, as well as both existing and new capacity resources, are equally valuable. A version of this market design is predominant in non-restructured U.S. power markets where utilities' compliance with reserve requirements and reliability standards is monitored and enforced through the state regulatory process. *Examples*: Southwest Power Pool ("SPP"), former Eastern U.S. power pools (NYPP, PJM, NEPOOL), some Canadian markets.

- *Energy Markets with Reserve Requirements and Centralized Capacity Markets.* Adding a centralized capacity market to reserve requirements provides a transparent backstop procurement mechanism for the system operator and, in addition to asset ownership and bilateral contracting, offers LSEs a third option to satisfy or adjust their mandated reserve requirements. The capacity market also standardizes the capacity product, facilitates market monitoring, and provides greater liquidity and transparency than a bilateral market. While adding complexity in market design, these features are particularly beneficial in market environments with many small LSEs, retail competition, and migrating customer loads. *Examples*: New York Independent System Operator ("NYISO"); Midwest Independent Transmission System Operator ("MISO").

- *Energy Markets with Forward Reserve Requirements (without Centralized Capacity Markets).* In markets with a forward resource requirement (also often referred to as a forward resource adequacy standard), LSEs must demonstrate that they have acquired sufficient reserves one or several years in advance. The forward requirement—which should be imposed on a locational basis in transmission constrained market areas—creates a bilateral capacity market that allows sufficient time for additional capacity resources to come online. This increases competition, mitigates risk of market power abuses, reduces price volatility, and provides the system operator with sufficient time to contract bilaterally for backstop capacity on behalf of any deficient LSEs. *Example*: California ISO.

- *Energy Markets with Forward Reserve Requirements and Centralized Capacity Markets.* This "forward capacity market" design combines the advantages of forward reserve requirements (including increased competition and lower price volatility) with those of centralized capacity markets (transparent market prices and backstop procurement mechanism, reduced transactions costs, and improved market monitoring). The considerable complexity of the market design increases the risk of design flaws, but offers features that are particularly valuable in market environments with many small LSEs, retail competition, and migrating customer loads. *Examples*: PJM, ISO New England ("ISO-NE"), Brazil.

Each of these market designs has advantages in certain market and regulatory environments, which explain their implementation in various power markets around the world. In some cases, the market designs have existed long enough that we are able to evaluate some of their successes and failures. More often, the market designs are still relatively new and rapidly evolving, and have not yet been tested over the full investment cycle of capacity resources. Nevertheless, the experience to date provides some indication of how well these market designs will likely function over time.

The need for out-of-market mechanisms in many market designs is a step back in the effort to create competitive conditions in restructured power markets. These out-of-market mechanisms include payments under reliability must-run ("RMR") contracts, government investments in generating capacity, and backstop mechanisms based on regulated cost recovery. Such interventions are attractive because they allow regulators or system operators to maintain reliability even when the market design has otherwise failed to attract sufficient capacity. However, these out-of-market mechanisms suppress market prices and create market distortions, which perpetuate and accelerate the need to expand the scope of out-of-market solutions to maintain reliability.

A well-functioning market will attract enough capacity to provide an efficient level of reliability. In most power market designs, the determination of what constitutes an adequate level of capacity is based on traditional engineering and reliability standards. This is the case even in capacity market designs that rely on downward-sloping demand curves, which are developed around a reliability target. Pure energy-only markets are fundamentally different from many other market designs in that they do not have predetermined reliability standards. Instead, they rely on customers to choose their desired level of reliability through the market mechanisms of interruptible rates and demand response. However, because these market mechanisms have not yet developed sufficiently to bring supply and demand into equilibrium and differentiate reliability across customers during shortage periods, real-world energy-only markets tend to rely on regulatory solutions for ensuring reliability, including out-of-market incentives and administratively-determined scarcity pricing mechanisms. In fact, to achieve efficient price signals in any of the discussed market designs—including designs with capacity payments, reserve requirements, and centralized capacity markets—a regulatory solution generally is needed both to set proper prices during scarcity periods and to facilitate the development and integration of demand-response resources.

Power system operators and regulators have struggled to find mechanisms for setting scarcity prices at efficient levels, which is a particular challenge in the absence of significant demand response. Many energy-only markets have imposed price caps at levels set below the likely value of lost load ("VOLL") and therefore do not allow for prices that can sustain needed investments. None of the power markets utilize mechanisms that increase scarcity prices gradually to the full level of VOLL as operating reserves diminish and demand curtailment becomes necessary. This creates considerable uncertainties about long-term resource adequacy in markets without explicit reserve requirements.

The liquidity and transparency of capacity markets is an important design consideration. Sole reliance on bilateral contracts and resource ownership can impose significant transaction costs on participants—particularly in markets with small LSEs, retail competition, and load migration. Bilateral markets are also more difficult to monitor for the exercise of market power. Without transparent market prices, new and small participants will have insufficient information to inform their investment and contracting decisions. The introduction of a standardized capacity product and centralized capacity market provides transparency, facilitates market monitoring and mitigation, and allows participants greater flexibility in meeting their reserve requirements and evaluating their investment decisions. The combination of forward procurement and capacity markets with transparent price signals also reduces price volatility and revenue uncertainty which, at high levels, can reduce or delay needed investments.

Limiting capacity payments only to new resources and existing resources that otherwise would retire is sometimes offered as a solution to mitigate the cost of maintaining reliability. However, this is not feasible in market designs that impose resource adequacy requirements on LSEs. Even in the absence of a capacity market, whenever such resource requirements are imposed in restructured markets, all capacity resources that can be used to satisfy the requirement will have the same capacity value. Even though an LSE's embedded cost of existing resources or fixed-priced contracts may differ, all other resources would be able to obtain the full capacity value. Differentiating capacity payments for new and existing resources is possible only if LSE resource adequacy standards are abandoned and replaced with administratively-determined capacity payments. In these market designs, the level of reliability achieved is solely a function of investment response to the offered capacity payment. This approach, currently used in Spain, introduces significant distortions of market prices that perpetuate the need for and can quickly expand the scope of capacity payments as existing resources would retire. (The expanding need for such payments can be significant because the all-in costs of retaining existing plants are high; they often are surprisingly close to, and can even exceed, the cost of new plants.) The long-term cost of an arrangement that limits capacity payments to new and about-to-retire resources will likely be higher due to market distortions and associated short-term and long-term inefficiencies.

PJM's locational forward capacity market, RPM, has several advantages compared with the other market designs, considering that PJM operates in an environment that includes deregulated generation, transmission constraints, retail competition, and many (often small) LSEs with migrating customer loads. In this environment, centralized capacity markets, which also include the ISO-NE's forward capacity market, provide significant benefits by facilitating the wholesale market participation of demand response resources and allowing for greater transparency, liquidity, market monitoring, backstop capacity procurement, and flexibility to address load migration compared to bilateral markets. In addition, the forward reserve requirement and sloped demand curve of RPM facilitate competition, mitigate market power, and reduce investment and reliability risks.

The most significant drawback of the RPM and other forward capacity market designs is their complexity, which increases implementation costs and the risk of design flaws. However, in the case of RPM, most of the design and implementation costs have been incurred already, many concerns about initial design parameters have been addressed, and the market has been shown to attract and retain capacity resources, including substantial amounts of demand-side resources. RPM and other market designs with centralized capacity markets also have the disadvantage that the clearly visible capacity price draws attention to the high cost of maintaining existing reliability targets. However, replacing the forward capacity market with a purely bilateral resource requirement would not likely offer a lower-cost solution to maintaining reliability targets. Rather, reliance on self provision of LSEs' resource requirements through bilateral contracts and resource ownership would simply make these costs less visible. To the extent that resource ownership or long-term bilateral contracts offer a lower-cost solution to maintain LSEs' resource requirements and system-wide reliability standards, these self-provision options also exist under RPM and other centralized capacity market designs.

II. INTRODUCTION: RPM AND ALTERNATIVE MARKET DESIGNS

A. SUMMARY AND PURPOSE OF RPM

On June 1, 2007, PJM began its first delivery year under its new forward capacity market construct, the Reliability Pricing Model ("RPM"). The stated purpose of RPM is to enable PJM to obtain sufficient resources to reliably meet the needs of consumers within PJM while:[1]

- Supporting LSEs using self-supply to satisfy their capacity obligations for future years;
- Administering competitive auctions to secure additional capacity resources, demand response, and qualifying transmission upgrades to satisfy LSEs' unforced capacity obligations that are not satisfied through self-supply;
- Recognizing the locational value of capacity resources in the auctions; and
- Using a backstop mechanism to ensure that sufficient generation, transmission, and demand response solutions will be available to preserve system reliability.

The Brattle Group's June 30, 2008 report *Review of PJM's Reliability Pricing Model (RPM),* assessed the performance of RPM to date, evaluated how well RPM is addressing the infrastructure investment needs that it was intended to address, analyzed the key RPM design elements for their effectiveness in achieving RPM goals, and recommended modifications to numerous RPM design elements.[2] As we noted in that report, the key design elements of RPM are:

- A mandatory three-year forward resource adequacy requirement, based on target reserve margins for all LSEs;
- A downward sloping (rather than a vertical) demand curve, called the Variable Resource Requirement ("VRR") curve, that is anchored at an administratively-determined value for the Net Cost of New Entry (Net "CONE") and that determines required reserve margins as a function of capacity prices;
- Locational Deliverability Areas ("LDAs") and locational capacity prices that are able to reflect the greater need for capacity in import-constrained areas;
- Provisions that allow demand-side resources and new transmission projects to compete with generating capacity;
- Accommodation of self-supply and bilateral procurement of capacity;
- Annual, PJM-administered "base-residual" and "incremental" auctions that clear capacity supplies against the downward sloping demand curve and procure residual

[1] PJM (2008), Ch. 1.
[2] Pfeifenberger, *et al*. (2008).

capacity needs (*i.e.*, capacity not already self-supplied through resource ownership or bilateral contracts) on a forward basis;

- Explicit market power mitigation rules, including a must-offer requirement for existing generating resources, and Market Monitoring Unit ("MMU") review of new entrant offers;
- Performance metrics during the delivery year and peak periods; and
- An opt-out mechanism under the Fixed Resource Requirement ("FRR") alternative.

Requiring capacity obligations to be self-supplied or procured through the capacity auction on a three-year forward basis is meant to roughly match the minimum lead time needed to bring new capacity resources online. It also allows for sufficient time to delay or cancel projects before irreversible major financial commitments have been made. The forward resource adequacy requirement in concert with clearing supply against a sloping demand curve is meant to improve price stability and reliability by providing meaningfully predicable forward market signals that can help avoid periods of extreme scarcity or excess capacity. It also forces existing resources to compete with a potentially large supply of new resources that can be brought online within three years.

In our 2008 RPM Report we found that RPM has been successful in encouraging investment in new and existing capacity to maintain resource adequacy both on an RTO-wide as well as on a local basis.[3] We found that since RPM was implemented: (1) approximately 4,600 MW of capacity has been retained that otherwise would have retired; (2) almost 10,000 MW of incremental capacity has been committed; and (3) the volume of generation interconnection requests has grown to make an additional 33,000 MW of new generation projects eligible to participate in future RPM auctions. However, we also identified a number of concerns and recommended improvements to the RPM design, many of which have been addressed through recent modification of the RPM design.[4]

B. Overview of Alternative Market Designs

Table 1 provides a brief summary of six market designs we will review in detail in Sections IV through IX. We present these alternatives in order of increasingly structured market designs that are targeted to ensure long-term resource adequacy, beginning with a discussion of the energy-only market and continuing with market structures that progressively add design components.

---

3 This analysis included the first five base residual auctions conducted within the RPM, for delivery years 2007/08 through 2011/2012. Because of the transitional period after RPM was implemented, only the last of these auctions was held a full three years ahead of the delivery year. See Pfeifenberger, *et al*. (2008).

4 FERC (2009a).

Table 1
Overview of Alternative Market Designs

Market Design	*Description*	*Examples*
1. Energy-only markets	• The design is limited to spot markets for energy and, possibly, ancillary services (*e.g.*, operating reserves). • There are no (or very high) caps on market prices or supply bids, allowing for very high energy prices during periods of scarce supply. • No resource adequacy standards or planning reserve margin requirements are enforced, although regulators or system operators may identify target reserve margins. Resource adequacy is not guaranteed. • In <u>pure</u> energy-only markets, payments for supplied energy (and ancillary services) would be the only source of revenues from which suppliers can cover their fixed and variable costs. • In <u>actual</u> energy-only markets, additional revenue sources may be available to certain suppliers through out-of-market contracts or cost-of-service payments.	Alberta; Australia's NEM; Initial California market; ERCOT; Great Britain; Nordpool; Ontario.
2. Energy markets with administratively-determined capacity payments	• Explicit capacity payments based on an administratively-determined capacity price are added to the energy-only market design. • Administrative capacity payments supplement capped energy market revenues, with the intention of adding sufficient revenues to attract adequate generation investment.	Argentina; Chile; Colombia; Previous Great Britain Pool; Peru; Spain; South Korea.
3. Energy markets with reserve requirements (but without centralized capacity markets)	• Explicit reserve margin requirements for LSEs are added to energy-only market design. • Reserve requirements are satisfied by LSEs through bilateral contracts or resource ownership. • The system operator monitors and enforces reserve requirements during delivery period, but there is no centralized capacity market.	SPP; former power pools (NYPP, PJM, NEPOOL); other NERC reliability regions, including some Canadian markets.

Table 1 (Continued)
Overview of Alternative Market Designs

Market Design	*Description*	*Examples*
4. Energy markets with reserve requirements and centralized capacity markets	• RTO-administered centralized capacity market is added to "energy market with reserve requirement" (design No. 3). • Centralized capacity market facilitates capacity procurement by small LSEs and on behalf of deficient LSEs. • Various auction designs (*e.g.*, vertical vs. sloped demand curve) are used to clear resources and determine capacity prices.	Prior PJM; MISO; Current NYISO; Australia's SWIS.
5. Energy markets with <u>forward</u> reserve requirements (but without centralized capacity markets)	• Similar to design No. 3 above, but reserve requirement is specified and enforced on a forward basis. • LSEs must commit resources for delivery period ahead of time (*e.g.*, 1 to 4 years) through bilateral contracts or direct resource ownership. • No centralized capacity market, but reserve requirements monitored and enforced by system operator both on a forward basis and during the delivery period.	California (1 year).
6. Energy markets with <u>forward</u> reserve requirements and centralized capacity markets	• RTO-administered centralized capacity market is added to "energy market with forward reserve requirement" (design No. 5). • Centralized capacity market facilitates forward capacity procurement by small LSEs and on behalf of deficient LSEs. • Forward commitment periods of centralized capacity markets (*e.g.*, 3-4 years) are meant to roughly match the minimum lead time needed to bring new resources online (or the time needed to commit to, delay, or cancel planned projects before irreversible major financial commitments have been made).	ISO-NE FCM (3 years); PJM RPM (3 years); NYISO proposal (4 years); Brazil (up to 30 years).

As Table 1 shows, the most basic market design alternative to RPM is the "*energy only market*" (discussed in Section IV) in which resources recover their investment and operating costs solely through revenues from energy and ancillary services markets. Although regulators or market operators may set reserve margin targets, there is no enforcement of resource adequacy standards or planning reserve margin requirements. Investment costs are recovered and resource adequacy is achieved through very high prices during occasional periods of scarce supply. Though several examples of energy-only market exist (*e.g.*, ERCOT), none of these markets actually rely on a "pure" energy-only design. Rather, in most of these markets additional revenue sources are available to suppliers through mechanisms such as reliability-related backstop procurement contracts or cost-of-service-regulated resources. Because energy-only market designs tend to yield highly volatile market prices and occasional severe price spikes without guaranteeing resource adequacy, additional design components are often added in attempts to mitigate price volatility and ensure resource adequacy.

For example, there are several international examples of "*energy markets with capacity payments*" (*e.g.*, Spain, as discussed in Section V), in which administratively-determined capacity payments are added to energy-only market designs in an attempt to attract needed supplies and facilitate resource adequacy. The added payments are intended to stimulate investments needed to maintain resources adequacy without imposing any reserve requirements.

In most U.S. market areas, however, resource adequacy is ensured by imposing explicit reserve requirements on LSEs that are monitored and/or enforced during or immediately prior to each delivery period (*e.g.*, peak season). For example, in the Southwest Power Pool ("SPP") and other "*energy markets with reserve requirements*" which have no centralized capacity markets, resource adequacy is ensured via reserve requirements satisfied by LSEs through bilateral contracts or resource ownership (as discussed in Section VI). In the majority of restructured markets that impose reserve requirements, however, RTO-administered centralized capacity markets have been added to improve pricing transparency, reduce transactions costs, facilitate market monitoring and mitigation, and provide a backstop capacity procurement by the system operator on behalf of deficient load serving entities. Examples of such "*energy markets with reserve requirements and centralized capacity markets*" (Section VII) include NYISO and the Midwest ISO, which has recently adopted both a reserve requirement and a centralized capacity market.

In a number of power markets, the reserve requirement is monitored and enforced on a forward basis: one to several years ahead of the actual delivery period. In such "*energy markets with forward reserve requirements*" (Section VIII), load serving entities must show that resources have been procured one or several years prior to the actual delivery period. For example, CAISO currently imposes a one-year-ahead resource adequacy standard. Without a centralized capacity market, the resource adequacy standard creates a strictly bilateral market for reserve capacity in which LSEs own or contract for capacity resources.

Finally, in the most structured design we discuss, centralized capacity markets also have been added to market designs with forward resource requirements. Examples of such "*energy markets with forward reserve requirement and centralized capacity markets*" (Section IX) are PJM's RPM and ISO New England's Forward Capacity Market ("FCM"). Both of these designs enforce a 3-year forward resource adequacy standard that can be met through self-supply, bilateral contracts, or capacity market purchases. Similar forward resource adequacy and capacity market designs are currently being evaluated in both New York and California.

These six market designs, as summarized in Table 1, are examined in detail in Sections IV through IX. However, we first discuss in Section III several topics useful for understanding the electric reliability and market design issues analyzed for the alternative market designs.

III. ELECTRICITY MARKET FUNDAMENTALS AND DESIGN CONSIDERATIONS

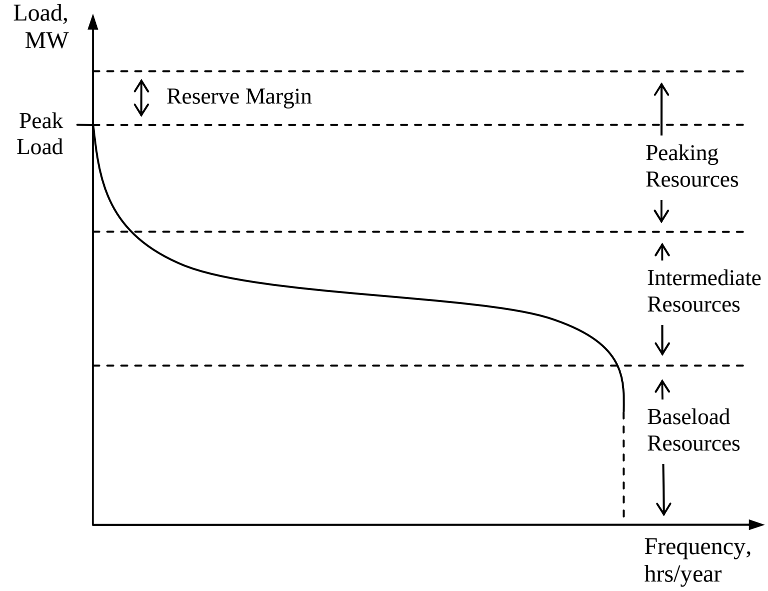
In this section we first review the cost structure of base-load and peaking resources. This is followed by a short discussion of reliability concepts and the need for reserve generating capacity. We then highlight differences in the cost recovery and timing of capacity prices in restructured and traditionally-regulated markets. Finally, we discuss implications of price volatility and revenue uncertainty, pay-as-bid and uniform price auctions as alternative mechanisms to determine market clearing prices, and the feasibility of differentiating capacity payments to new and existing resources.

A. LOAD PROFILES AND THE COST STRUCTURE OF BASE-LOAD AND PEAKING RESOURCES

Customers' demand for power is cyclical on a daily, weekly, and seasonal basis, affected by patterns of commerce and weather. In most systems, the most extreme demands for power occur over the hottest few days each summer when air conditioning loads are greatest. Most hours of the year have power requirements within a much more moderate range. Figure 1 shows a typical load duration curve, illustrating how many hours of the year various load levels are reached.

The total quantity of generating capacity that must be installed in the system is determined by the maximum load expected. The mix of generation assets that can meet peak demand at the lowest cost depends on how often each unit would be economic to run. Baseload assets such as nuclear and coal have high fixed costs and low operating costs, potentially making them the lowest cost sources only if they are expected to operate at full capacity most of the time. Peaking assets such as combustion turbines ("CTs") have low fixed costs and high operating costs, potentially making them the lowest cost sources only if they are expected to operate a small number of hours. Intermediate resources are those in between the two extremes. Figure 1 also shows what portions of total capacity requirements would be supplied by these types of resources.

Figure 1
Illustrative Load Duration Curve, Showing Type of Generating Resources that can Supply Demand at or Above a Certain Level at Lowest Cost



The market design discussions in this report focus on the means of attracting and retaining the resources that ensure system reliability, with particular attention to resource adequacy during peak hours.[5] Peaking resources are rarely called upon to run. Traditionally, peaking capacity is supplied by CTs, but increasingly peaking resources are provided through demand-side resources: instead of supplying more power from generating plants during times of scarcity, some customers curtail their consumption according to a contractual agreement or in response to time-varying prices. These demand-side resources can reduce system costs by reducing the need for peaking generators. Although some markets have been designed specifically to accommodate these demand-side resources because significant participation is expected and desired in future, many market designs have not yet incorporated these resources.

B. System Reliability Planning and the Need for Reserves

The reliability of the electrical system is measured by the frequency, duration, and magnitude of service interruptions or power quality problems. Overall reliability depends on two separate properties of the system: resource adequacy and system security.[6] Resource adequacy means there is sufficient generation and transmission capacity in the electric system to supply the aggregate electrical demand at all times, taking into account scheduled and reasonably expected unscheduled outages of system elements. If there is insufficient generating or transmission

---

5 A well-functioning electric market should attract the mix of resources with the lowest overall cost of supply, when accounting for both capital and operating costs. This will include a combination of baseload, intermediate, and peaking resources.

6 For an overview of how NERC defines these two aspects of reliability, see NERC (2008a) pp. 273-274.

capacity to meet demand, then the system operator will enact emergency measures to maintain a stable power system, such as voltage reductions or load shedding. These measures result in disruptions of some customers' service, but do not constitute a breakdown of system stability as long as the measures are implemented in a planned and controlled fashion.

A system can have adequate generation resources but still be unreliable due to insufficient system security. System security (or stability) is the ability of the electric system to withstand sudden disturbances, such as electric short circuits or unanticipated loss of system elements. For example, if the system has a single line outage that results in an uncontrollable cascade of additional outages, then a large number of customers will experience interruptions in service because the system was not sufficiently secure. Although security problems can have many possible causes not related to resource adequacy, shortage conditions exacerbate stability problems and make an electrical system more likely to experience cascading blackouts.

Forward system planning is an attempt to ensure resource adequacy. When determining how much capacity must be installed in order to reliably meet peak demands, system planners must forecast zonal loads and probable generation outage levels. These forecasts are made for future years, with high uncertainty around expected peak demands. System planners cannot be *certain* that the target capacity level will be sufficient to meet peak demand; their capacity target can only be anticipated to meet peak demand with a certain probability. Thus, with any particular target level of capacity, there will be a certain loss of load expectation ("LOLE"). All else being equal, a lower target level of capacity would correspond with a higher LOLE, and vice versa.

The North American Electric Reliability Council ("NERC") has developed guidelines for conducting resource adequacy assessment, and the NERC regions implement their assessments in different ways.[7] Most NERC regions have a LOLE reliability target of one day in ten years.[8] Once the relationship between the levels of capacity and LOLE is estimated, the targeted LOLE can be translated into a target level of necessary capacity. If high reliability is desired, then the target capacity will be greater than the estimated peak load by some margin. The planning reserve margin or capacity margin is the amount by which available capacity is expected to exceed normal future peak load as shown in Figure 1.[9]

As others have also pointed out, a consensus has not yet emerged on which market designs are best suited to ensure long-term resource adequacy in restructured power markets.[10] The six designs discussed in Sections IV through IX are a representative range of options currently employed in power markets internationally.

---

7 NERC (2008c-d).

8 The exact meaning of this varies. For example, it can mean one event in ten years or 0.1 days each year.

9 Reserve margin is expressed as a percent of peak load while capacity margin is expressed as a percent of available capacity.

10 For example, see Roques (2007).

C. COST RECOVERY IMPACTS OF RETAIL AND WHOLESALE RESTRUCTURING

Since the Public Utilities Holding Company Act ("PUHCA") of 1935, and until a wave of industry restructuring began with the Public Utilities Resources Purchasing Act ("PURPA") of 1978, the electric industry was regulated under a quite stable regulatory compact which granted electric utilities exclusive rights of supply in their service area, including all generation, transmission, distribution, and retail services. The provision of these services was based on rates that reflected the utilities' full cost of service, including all prudently-incurred costs related to generating facilities, reliability requirements, and reserve margins held by the utility.

This traditional regulatory framework still provides full cost recovery for utilities operating generating plants in most U.S. power markets without retail competition. In these cost-of-service regulated settings, the role of power markets is limited to wholesale purchases or sales of power, which utilities undertake to supplement their cost-based generation activities. These wholesale power transactions contribute to but are generally not needed to allow regulated utilities to recover capacity-related costs.

The role of power markets is quite different in regions with a restructured utility industry and retail competition.[11] In such restructured markets, most utilities are no longer vertically integrated and, to the extent they own generating facilities, the cost of power plants is no longer recovered through cost-based rates. Rather, generating facilities need to operate on a "merchant" basis and recover costs through market-based (short- or long-term) bilateral contracts or spot market sales. Spot market revenues can be obtained in energy markets, ancillary service markets, and (where available) capacity markets. The value of these revenue streams would also be reflected in bilateral contracts with load serving entities ("LSEs") who rely on purchases from the wholesale market.

In fully restructured markets, investments in new generating facilities and retention of aging existing facilities will only be achieved if total market-based revenues are expected to equal or exceed the facilities' total operating and (forward-looking) investment costs. This means that system-wide reliability and resource adequacy directly depend on the level of market-based revenues available to suppliers. If such revenues are insufficient to cover the total forward-looking costs, new capacity will not be built and existing capacity will not be retained.

This link between resource adequacy and the level of market-based revenues is much less pronounced in power markets without retail competition. Many of these "wholesale-only" power markets remain dominated by traditionally-regulated, vertically integrated utilities, which recover most of their capacity costs through regulated retail rates. In these regions, reliability and resource adequacy typically are not dependent on the level of expected market-based revenues. Rather, resource sufficiency is ensured through the utilities' long-term planning processes, which typically are subject to regulatory review processes and targeted to meet

[11] In Europe, restructured power markets with retail competition are generally referred to as "liberalized" power markets.

applicable reserve requirements and reliability standards (*e.g.*, as specified in market design No. 3, which is analyzed in Section VI below).

D. TIME PROFILE OF CAPACITY PRICES IN REGULATED AND RESTRUCTURED POWER MARKETS

An important but often ignored or misunderstood difference between regulated and fully-restructured power markets relates to the time profile of capacity prices. In a cost-of-service regulated environment, retail rates will reflect the cost of generating capacity only *after* new generating resources are placed in service and reflected in utilities' ratebases. This means there can be a lag of several years before regulated retail rates reflect the addition of expensive new capacity resources. This lag causes a significant misalignment of retail prices and investment signals. Because demand continues to grow due to low rates, more new resources may be added to the system than will ultimately be needed when retail prices increase to reflect the added costs. This can lead to excess capacity, high regulated rates, and the risk of stranded costs or regulatory disallowances.

The time profile of capacity prices is quite different in restructured power markets. As in all other competitive markets, in restructured markets the market price for capacity will increase *before* new generating capacity needs to be added. As market participants perceive an approaching scarcity of generating capacity, market prices for capacity will increase and, in response, market participants will identify the lowest-cost resources that can operate profitably at the anticipated market prices. If capacity prices are reflected in retail rates or are otherwise made available to demand-side resources, this market-determined portfolio of resources will also include demand-response resources. The fact that capacity prices increase *before* new resources are actually added to the system will dampen demand growth, which will reduce the resource need.

The fact that prices in restructured markets increase before any new capacity is added, can quickly raise public policy questions about the value and effectiveness of restructuring. Sharp anticipated increases in the market price for capacity in restructured markets can also lead to over-investment and the subsequent collapse of market prices (*e.g.,* the boom-bust cycle experienced 5-10 years ago). Market designs based on forward resource requirements (market designs Nos. 5 and 6) have been developed specifically to reduce that type of price volatility and mitigate the risk of such boom-bust cycles.

E. IMPLICATIONS OF PRICE VOLATILITY AND REVENUE UNCERTAINTY

The market designs we examine have varying degrees of price volatility and corresponding uncertainty of revenues and cost recovery. Uncertainty in suppliers' revenues increases the financial risks and thus the cost of financing capital projects.[12] In turn, investors require a higher rate of return on their investment, which will directly increase the price paid by customers. Caballero and Pindyck have found, for example, that "a doubling of industry-wide uncertainty

---

[12] For example, see Minton and Schrand (1999), pp. 423-26.

raises the required rate of return on new capital by about 20 percent."[13] Market designs that help reduce price volatility and revenue uncertainty consequently will tend to reduce the cost of capacity investments.

Uncertainty also decreases available discretionary capital[14] and creates a preference for more flexible, less capital-intensive generation assets. These assets with lower investment costs may have a higher total costs, but are still attractive to investors if faced with considerable uncertainty of recovering their fixed investment costs. Finally, since some of the uncertainty is anticipated to be resolved over time, investors will have an incentive to delay their investments. In other words, there is an "option value of waiting" because new information that becomes available over time increases the chances of making better investment decisions.[15] The higher the uncertainty and the more of this uncertainty is anticipated to be resolved over time, the higher will be the value of delaying investments. Market designs that reduce this uncertainty, will consequently also tend to facilitate investments on a more timely basis.

Even though uncertainty and price volatility come with these costs, this does not mean that policy-makers should attempt to completely eliminate the risk exposure of market participants. Rather, total system costs will be minimized when the risks are assigned to the party most able to control that risk.[16] Cost-effective public policy should not completely protect suppliers or customers from financial risks if they have some control over the causes of risk, because that would eliminate the incentives to reduce these risk-related costs. However, risks beyond the control of market participants, such as weather and economic risks, should be fully reflected in market prices.

Price volatility and financial uncertainty result from several underlying risk factors. One factor is regulatory risk. Regulatory risk can be created by changing market designs. Regulatory risk can also be associated and exacerbated by perception of excessive regulatory discretion to intervene in markets as well as arbitrary or insufficiently transparent market rules.

Other sources of uncertainty are the inherent price volatility in energy or capacity markets, the future amount of available supply, macroeconomic factors affecting demand, fuel and equipment costs, and weather. Generators are the parties most able to control overall generation investment levels and the availability of their resources during peak times. This suggests that it is cost-effective public policy to subject suppliers to at least some financial risk associated with these factors.

Finally, a large portion of the volatility in capacity prices and uncertainty in the amount of capacity needed is a function of customer demand uncertainty. Because customers are most able to control their demand levels, it would be sound policy to assign them associated risks. Traditionally, customers have had no exposure to the price volatility caused by the variance of

---

[13] The finding is not specific to the electric industry. See Caballero and Pindyck (1996).
[14] Minton and Schrand (1999), pp. 423-26.
[15] Bernanke (1983), pp. 85-87.
[16] Priest (2007).

their demand; instead, they paid a flat price for power no matter when they used it. This is now changing with the introduction of demand response and dynamic pricing.

F. MECHANISMS TO DETERMINE MARKET PRICES

Many discussions of energy and capacity market designs also explore the advantages and disadvantages of different "auction designs" used to determine market prices and supplier compensation. These auction design alternatives include pay-as-bid approaches (prevalent in bilateral markets and RFP-based procurement processes) and uniform clearing price approaches (prevalent in RTO-administered energy markets and centralized capacity markets).

High or low market prices for energy or capacity are sometimes attributed to the type of procurement mechanism that is used by the market administrator. For example, New York legislators are currently exploring whether a pay-as-bid approach to clearing the NYISO energy market would result in lower prices than the currently used uniform market clearing price approach.[17] The uniform pricing format has also been blamed for perceived high prices in centralized capacity markets. Similarly, a 30 to 50 percent retail rate increase associated with the end of a 10-year price cap period covering Illinois' transition to deregulation was attributed to the "declining clock" auction format used to procure standard offer service. And some even attributed the high prices during the California power crisis to the uniform pricing design. We believe these discussions over the advantages or harm done by different auction designs are either unfounded or greatly exaggerate their likely impacts and importance. It is the case that each auction design has offers practical or theoretical advantages over others under some circumstances, but the format of the auction design—including whether to rely on centralized capacity markets or rely solely on bilateral contracts to satisfy resource adequacy requirements—will not generally be a major factor influencing the observed market prices.

It may seem at first blush that a pay-as-bid auction should reduce prices below uniform clearing price auctions because each supplier receives only the price they offered to sell at, rather than receiving the highest price that cleared the market. This is not the case, however. Because the pay-as-bid market design does not force suppliers to bid their incremental costs, each supplier will submit an offer based on their best estimate of what the clearing price will be, and sales prices will converge to a range above or below the actual "market price." This range will be very narrow in competitive markets (*i.e.*, close to a uniform price), but wider and more uncertain in illiquid power markets.

The extent to which pay-as-bid approaches could lower prices compared to uniform pricing approaches was explored in great detail with respect to energy markets by a "blue ribbon panel" of experts with input from market participants and economists in the aftermath of the California power crisis. The panel concluded that switching from uniform to pay-as-bid pricing would not reduce prices and could even be harmful.[18] A more recent, similarly detailed analysis of this

---

17 For example, see Megawatt Daily (2009a).

18 Kahn *et al.* (2001), p. 2. ("The expectation behind the proposal to shift from uniform to as-bid pricing—that it would provide purchasers of electric power substantial relief from the soaring prices of the electric

question arrived at the same conclusion.[19] Some studies, however, have also suggested that uniform auctions in electricity markets *can* result in higher prices than those from pay-as-bid auctions.[20] Yet others find that, based on a literature review and independent analysis, a switch from uniform pricing to pay-as-bid pricing in electricity markets could either increase or decrease market prices.[21]

The extent to which market designs rely on pay-as-bid or uniform prices also has some implications for how suppliers bid and the extent to which market power can be monitored and mitigated. For example, several studies note that uniform pricing can reduce competition in un-mitigated power markets while it facilitates mitigation of market power.[22] Similarly, while pay-as-bid market designs can help increase competition between existing suppliers, the design makes it much more difficult to monitor and mitigate market power compared to uniform pricing market designs that allow suppliers to bid their marginal costs.

In our discussion of capacity market designs we contrast market designs that solely rely on bilateral markets with designs that also include RTO-administered centralized capacity markets. However, we do not focus specifically on different auction formats that are used (or could be used) in these markets. While some auction formats (such as multi-round auctions) will tend to be more efficient than others (*e.g.*, single-round auctions), these are design features that, along with many other design details, are beyond the scope of our analysis.[23] Overall, however, the available research suggests that it is unlikely that switching RTO-administered energy or capacity markets from the widely-used uniform pricing approach to a pay-as-bid approach would markedly reduce (or increase) market prices.

G. DIFFERENTIATING CAPACITY PAYMENTS TO NEW AND EXISTING RESOURCES

The design of capacity markets or capacity payment mechanisms raises the question of whether all resources should receive capacity payments, or whether such payments should be limited to new resources and resources which would retire otherwise. As discussed in Section V.B, the capacity payment design in Spain does just that: it provides a fixed, administratively-determined capacity payment to new resources for a 10-year period. These capacity payments are also available for significant investments in upgrading existing capacity that might otherwise be

---

power, such as they have recently experienced—is simply mistaken. In our view it would do consumers more harm than good.")

19 Baldick (2009), p.1. ("There is no empirical or experimental evidence that pay-as-bid or other alternatives would reduce prices significantly compared to a single market-clearing price design. In fact, some evidence suggests that pay-as-bid would increase prices compared to explicitly setting the single market-clearing price. Moreover, pay-as-bid has some significant drawbacks.")

20 For example, see Fabra *et al*. (2004).

21 Federico and Rahman (2003).

22 See Kahn, *et al*. (2001), Baldick (2009), Fabra *et al*. (2004), Federico and Rahman (2003).

23 For a detailed discussion of specific capacity market features, see Pfeifenberger *et al*. (2008), which also presents a number of recommended improvements to PJM's capacity market design.

retired. Limiting capacity payments to new resources or resources that would otherwise retire has the appeal it can reduce the total costs associated with such capacity payments.

The market design based on administratively-determined capacity payments (as discussed in Section V) is the only market design in which capacity payments can be limited to new and existing units that would otherwise retire. Energy only markets (discussed in Section IV) do not offer capacity payments and in market designs with reserve requirements (discussed in Sections VI through IX), the reserve requirement creates a capacity value that is available to both new and existing units.

Capacity revenues available to existing and new resources cannot be differentiated in markets that impose reserve requirements on LSEs because, as soon as reserve requirements are imposed, all capacity resources that can be used to satisfy the requirement will have a capacity market value. As a result, even if RTO-administered capacity markets were limited only to new resources, the full market value of capacity would still be captured by all existing resources through bilateral contracts, assuming that the resources are not cost-of-service regulated or under existing fixed-priced contract.

Note, however, that even in market designs with broad-based capacity markets, a large portion of the total resource requirement is generally self-supplied through owned or bilaterally-contracted resources.[24] The average cost of self-supplied capacity consequently may differ from the current market clearing capacity price. For example, the market price for capacity will only affect the cost of service of regulated, vertically-integrated utilities to the extent the utility is a net buyer or seller of capacity. In restructured power markets with retail competition, however, the market price of capacity will tend to be passed through in retail rates, regardless of whether the market design is based on a centralized capacity market or a bilaterally-satisfied reserve requirement.

Only in power markets that do not impose resource adequacy requirements on LSEs can capacity payments be targeted specifically to new resources or the retention of existing resources. Conceptually, targeting capacity payments to a selected group of resources differs little from the out-of-market reliability backstop mechanisms that have been used under many market designs in the absence of a localized resource adequacy requirement. However, as discussed further in Section IV.B, such payments can cause significant market distortions that, while potentially saving costs in the short-term, can result in substantial inefficiencies in the long term.

When limiting capacity payment to new resources or existing resources that would otherwise retire, it is also necessary to recognize that a sizeable portion of the existing pool of resources could be forced to retire in the absence of capacity revenues. For example, in the six years before RPM was introduced in PJM, between 500 MW and 3,500 MW of generating resources

[24] In PJM, for example, approximately 50 percent of the RTO's total forward resource requirement is self supplied. This includes 40 percent of approximately 130,000 MW of capacity cleared through the RPM framework and an additional 25,000 MW of PJM-internal capacity that is self supplied under the Fixed Resource Requirement ("FRR") option, which allows certain vertically-integrated LSEs to satisfy the 3-year forward resource adequacy requirement outside the RPM capacity clearing framework.

retired each year.[25] After RPM was introduced, annual retirement dropped to a range of zero to 500 MW. More importantly, however, an analysis of market monitoring data showed that at least 30,000 MW of PJM's capacity resources are at risk for retirement in the absence of capacity payments due to revenue deficiencies in PJM's energy and ancillary services markets.[26] This is not surprising considering that the going-forward costs of many existing resources can be high even in comparison to new resources. As a result, capacity auctions may select new capacity resources even when cost-based bids for many of the existing resources do not clear. For example, in PJM's auction for the 2011-12 planning year, a total of 2,337 MW of new capacity cleared in the auction, while 496 MW of new capacity did not clear.[27] In comparison, 4,600 MW of capacity from existing resources did not clear, even though the bid prices for the existing resources were mitigated to reflect their incremental costs. These data show that the all-in costs of retaining existing plants can even exceed the costs of new plants. This is the case because existing plants are often less efficient and keeping them operational may require significant ongoing costs (*e.g.*, high annual repair, refurbishment, and maintenance costs) as well as occasional substantial investments (*e.g.*, environmental retrofits or replacements of major plant components).

IV. ENERGY-ONLY MARKETS

The remainder of this report presents a detailed discussion of each market design summarized in Table 1 of Section II. This section lays out the theoretical foundation for energy-only markets and the real-world implementations of this market design. The remaining sections then build upon that foundation by discussing specific design components that are added to the energy-only design to make up each of the other alternative designs. This includes the justifications for the additional design elements, explaining how essential market functions are carried out in the new design, and examining real-world experience with these alternative market designs.[28]

We divide our discussion of energy-only market designs into separate discussions of "pure" energy-only markets and actually implemented "real-world" energy-only markets. The pure energy-only market is a theoretical model, set out by some in academia, industry, and government as the possible end-state toward which current markets could transition.[29] So far, pure energy-only markets have not been implemented because the design depends on market structures that do not yet exist, such as widespread demand response or the ability to deliver different levels of reliability to different groups of customers. Real-world energy-only markets

---

25 Pfeifenberger *et al*. (2008), p. 20.

26 Pfeifenberger *et al*. (2008), pp. 22-23.

27 Pfeifenberger *et al*. (2008), p. 36.

28 For a discussion and evaluation of various alternative market designs, see also Crampton and Stoft (2006) and Roques (2007).

29 For example, see Hogan (2005), p. 33. ("The sketch of an 'energy-only' market design sets a destination but does not define the path… With a common understanding of the objective, it would be easier to make choices along the way.")

have included various modifications to prevent price spikes or to attract and retain sufficient capacity resources.

A. Basic Market Design Elements and Considerations of "Pure" Energy-Only Markets

In a pure energy-only market, the capital costs of a generating unit must be recovered exclusively through energy market prices and associated ancillary services.[30] In this model, suppliers selling solely into the spot market would not receive payments for undispatched capacity. The energy-only model, however, does not limit suppliers from entering into bilateral contracts for capacity if load serving entities or individual customers are willing to enter into such contracts (*e.g.,* as a hedge against high price volatility in the energy market).

In most hours of the year and under most circumstances there will be more generating capacity than is needed to meet hourly demand. During these hours, assuming workably competitive market conditions, the market price for energy will tend to be approximately equal to the marginal operating cost of the most expensive unit dispatched. On some occasions, however, supply will be scarce. Under these occasional scarcity conditions energy prices will climb above marginal operating costs to include a "scarcity premium."

In other words, a pure energy-only market is characterized by moderate energy prices punctuated by occasional severe price spikes. During most hours, base-load and intermediate-load generators with low operating costs can recover both their variable operating costs and part of their fixed costs, while peaking generators are able to recover little more than their variable operating costs. During the rare occasions of capacity shortage, the system experiences price spikes during which both peaking and base-load plants can recover contributions toward their fixed costs.

Occasional capacity shortages and scarcity-related price spikes are an explicit and intentional design feature of energy-only markets. Scarcity prices must be sufficiently high and sufficiently frequent to attract and recover investment in capacity.[31] In the absence of those spikes, existing resources would exit the market without being replaced through new entry. This would reduce supply and increase the frequency of scarcity conditions and price spikes.

Occasional high scarcity prices will also motivate demand reductions in two ways:

- *Price-Based Demand Response* – If customers are exposed to spot energy prices, they will adjust electric usage patterns. This price response would be small during normal system conditions and moderate prices. Under a scarcity event, however, the hourly price would be allowed to spike to levels necessary to induce the required reduction

---

[30] For a full discussion of the theoretical basis for pure energy-only markets, see Hogan (2005) and Joskow and Tirole (2004).

[31] Alternately, some generation capacity might also be attracted by bilateral capacity contracts with customers that demand a higher level of reliability or who are willing to pay a premium for a financial hedge against price spikes.

in demand. If customers are not very responsive to price, then these price spikes would be very high. If customers are more responsive to price, then scarcity prices would be more moderate.[32]

- *Reliability-Based Retail Pricing (Interruptible Rates)* – LSE may be able to curtail interruptible service customers during scarcity conditions. Interruptible customers pay a lower average price for power in exchange for their willingness to endure interruptions in service. The more frequent the anticipated interruptions, the lower the price. However, the extent to which this option exists today is too limited to allow resolving reliability challenges in every circumstance. Broad-based differentiation of service quality that would allow some customers to pay less for lower reliability while allowing other customers to pay more for higher reliability is not yet possible.

Through this interaction between available supply and demand under normal and scarcity conditions, customers and suppliers mutually determine the economically optimal level of installed capacity via the prices established in the market. This means that there are no administratively-determined levels of required capacity or reliability. Rather, the level of reliability is determined by the marketplace.

The concept of energy-only markets becomes more complicated when applied to today's typical power markets that do not have significant demand response nor an ability to differentiate reliability across broad classes of customers. Without demand response, a shortage condition will not induce load reductions, no matter how high prices rise. Because market forces will not be able to bring supply and demand into equilibrium during scarcity events, a regulatory solution for "scarcity pricing" and involuntary load reductions must be implemented to avoid uncontrolled blackouts.

Without sufficient demand response, excess demand must be reduced system-wide through controlled temporary voltage reductions ("brownouts")[33] and rotating curtailments of service on different portions of the grid ("rolling blackouts").[34] Because such rationing events have no associated market-determined prices, prices must be determined through regulatory means. The appropriate regulated price to apply in such circumstances is the estimated value of lost load ("VOLL"). The VOLL is determined based on either an estimate of the cost that customers incur due to a service interruption or the payments that customers would be willing to make to avoid

---

32 Demand response does not yet exist at the scale necessary to support the "pure" energy-only market functions as described here, as discussed in the following sections.

33 Reducing system voltage reduces the power drawn by end-users because of the physical relationship between voltage V and power P; $P = V^2/R$, with resistance R being close to constant (for example, see Pai (2003), Ch. 2). Voltage is reduced only as a near-last resort before enacting rolling blackouts because this can cause equipment damage to electric customers.

34 Joskow and Tirole are careful to point out that rationing is an orderly interruption in service and that network collapses differ from other forms of energy shortages and rationing in a fundamental way. While scarcity makes available generation (extremely) valuable under orderly rationing, it makes it valueless when the network collapses. (See Joskow and Tirole (2004), p. 5).

curtailment. Various studies indicate that the average VOLL tends to be at least in the \$5,000/MWh to \$10,000/MWh range.[35]

For energy-only markets to be able to work as intended and attract sufficient capacity, market prices must consequently be allowed to reach the VOLL during scarcity events. As long as the VOLL is estimated accurately, an energy-only market would provide a price incentive for building capacity that reflects the true demand of customers for reliability and associated capacity additions.[36] If VOLL is higher, then more capacity will be attracted by higher scarcity prices or bilateral contracts intended to hedge against such scarcity prices. If VOLL is lower, then less capacity will be built and reliability will be lower. This model differs substantially from traditional U.S. reliability and resource adequacy requirements based on engineering standards. Rather than engineering standards, market forces would be allowed to determine the level of reliability.

B. EXPERIENCE WITH REAL-WORLD ENERGY-ONLY MARKETS

A pure energy-only market design is difficult to implement for both policy and technical reasons. First, policy makers and regulators generally are unwilling to accept the potentially severe price spikes and demand rationing (possibly including rolling blackouts) associated with pure energy-only markets. Second, without an ability to differentiate reliability across customers, system reliability, including reserve capacity, is a "common good" that will not be provided at sufficient levels through energy-only markets without regulatory intervention, including carefully-crafted scarcity pricing mechanisms.

Because of these realities, actual examples of energy-only markets have involved significant alterations of the conceptual design. We examine several such energy-only markets in varying degrees of detail—those of Great Britain, Nordpool, Australia's National Electricity Market ("NEM"),[37] ERCOT, the Alberta Electric System Operator ("AESO"), and the Ontario Independent Electricity System Operator ("IESO")—to address several policy considerations and challenges associated with energy-only market designs and to show how the conceptual energy-only market design has been modified in practice.

---

35 VOLL is difficult to estimate because it varies widely depending on customer class, business sector, duration of outage, and advanced warning of the outage. For example, for a 1-hour outage, MISO has estimated VOLL at \$730-\$2510/MWh for residential, \$15,000-\$50,000/MWh for small commercial and industrial ("C&I"), and \$16,000-\$78,000/MWh for large C&I customers. The range in estimates shows the range across industries, where, for example the mining sector has a much larger VOLL than the services sector. (See MISO (2006).)

36 See Joskow and Tirole (2004) and Hogan (2005).

37 The NEM covers the six southeastern jurisdictions of Australia with most of the country's electric energy demand: Queensland, New South Wales, the Australian Capital Territory, Victoria, South Australia, and Tasmania. Sometimes referred to simply as the "Australian market," the NEM is not administratively or physically interconnected with Australia's South West Interconnected System ("SWIS"), North West Interconnected System ("NWIS"), or the 29 other small, non-interconnected regions. None of these other, smaller Australian markets are "energy-only" markets. (See AER (2007), pp. 59, 204.)

1. High Price Caps and the Relationship to VOLL

Consistent with the pure energy-only design, some real-world energy-only markets have high or no price caps. However, most energy-only markets impose price caps below the VOLL. Only the power market in Great Britain is not capped.[38] While Nordpool (the Scandinavian power market) does not cap its intra-day and real-time balancing market, it imposes a price cap of $2,800/MWh (€2,000/MWh) in the day-ahead market.[39] Similarly, the energy market in Alberta and Ontario are capped at C$1,000/MWh and C$2,000/MWh, respectively.[40]

The only market that has implemented a price cap based on the estimated VOLL is the Australian NEM. There the price cap is $7,850/MWh ($10,000 AUD),[41] in line with other estimates of VOLL.[42] This estimate of VOLL is reviewed and updated periodically and will be increased to $9,820/MWh ($12,500 AUD) in 2010.[43] In comparison, the price cap in ERCOT, the only U.S. energy-only market, is much lower. ERCOT's offer cap was increased to $2,250 in 2008, and is planned to increase further to $3,000/MWh after the transition to nodal pricing.[44] This offer cap is not specifically related to VOLL and is set at a level lower than common estimates of VOLL.[45]

Both NEM and ERCOT also have additional price mitigation measures that limit the duration of elevated scarcity prices, although both the size and purpose of these measures are different. The NEM has a seven-day cumulative price threshold ("CPT") of $118,000 ($150,000 AUD), increasing to $147,000 ($187,000 AUD) in 2010. If the sum of all prices over a seven-day period exceeds the CPT, amounting to an average half-hour price of $350/MWh ($446 AUD), then a price cap of $236/MWh ($300 AUD) is enacted until the sustained high prices have ceased. The CPT is intended to limit risk of prolonged high price levels to buyers. Original proposals had set the CPT at $236,000 ($300,000 AUD), high enough such that a peaking plant could earn approximately three years of fixed costs during one week of scarcity.[46] Although the current CPT is lower than that, the construct does not have an annual limit, leaving open the possibility that generators could earn multiples of their annualized fixed costs in one year—which is necessary as an offset to years during which generators recover less than their average annual fixed costs.

---

38 The Great Britain market does not have a price cap; the market is a bilateral forward market up to one day in advance of delivery. Only the balancing market is within a centralized market. This market is a discriminatory or pay-as-bid market. See IEA (2005), pp. 175-179.

39 Nordpool (2008). Currency conversion of 0.714 €/USD from FRB (2009).

40 Adib, *et al*. (2008) p. 332.

41 All dollar values have been converted to USD based on an exchange rate of 1.273 AUD/USD. Currency conversion from FRB (2009).

42 AER (2007), p. 43.

43 AEMC (2009).

44 See ERCOT (2008) and Potomac (2008), p. vi.

45 This offer cap is conceptually different from a price cap because it is a cap on suppliers' bids into the market rather than a cap on market prices themselves.

46 AEMC (2008) and AEMC (2009).

ERCOT uses a Peaker Net Margin ("PNM") which estimates the money a hypothetical peaking plant could apply toward fixed costs. ERCOT imposes an annual PNM cap of $175,000/MW. If the PNM is reached within any calendar year, then the ERCOT offer cap is reduced to the higher of $500/MWh or 50 times a daily gas price index.[47] The PNM is high enough such that during a year of significant scarcity, a peaking plant could recover in that year approximately two or three times its annualized capital cost.

2. Revenue Deficiencies and Reliability Concerns in Energy-Only Markets

Scarcity prices in an energy-only market must be high enough and frequent enough to attract sufficient investment in capacity when needed. If prices are capped at too low of a level and no other revenue sources exist, revenues from energy-only markets will be insufficient to retain generation or attract entry of new generating facilities. This revenue deficiency is what is referred to as the "missing money" problem.[48] When revenue deficiencies become a structural part of the market, the result will either be a drop in installed capacity and system reliability or more likely, regulatory interventions that result in the implementation of various out-of-market mechanisms to prevent such declines in reliability. While such out-of-market mechanisms often appear to be a low-cost solution to ensuring reliability compared with market-based options, the volume and cost of these mechanisms can increase quickly while simultaneously distorting market prices and reducing market efficiency and competition.

It appears that the missing money problem is present in all existing energy-only markets. The primary reasons for this revenue deficiency are: (1) price caps set at levels below the VOLL (which appears to be the case in all markets but NEM);[49] (2) out-of-market mechanisms preventing scarcity events and associated price signals; and (3) the "common good" aspect of reliability that results in under-pricing and under-investment.

The Australian NEM is often held up as a good example of an energy-only market in which sufficient capacity has been and will be added, possibly because the cap is explicitly based on the VOLL and scarcity prices are allowed to rise to this level, which is much higher than in other markets.[50] Since its inception in 1999, NEM has seen net increases of 3,353 MW in base-load capacity and 1,198 MW in peaking capacity, for a total increase in capacity equal to about

---

47 Potomac (2008), p. xxxvii.

48 For example, see Cramton and Stoft (2006).

49 Although there is no price cap in Great Britain's market, the comparison with other energy-only markets is quite difficult given the discriminatory pay-as-bid system. Nordpool is a mixed case theoretically; the day-ahead price cap is lower that VOLL which should result in under-investment in capacity. With no price cap in the smaller intra-day and balancing markets, the prices could be efficiently set if there is sufficient demand response in the system. In any case, the initial Nordpool condition of excess capacity and ongoing subsidies for renewable capacity investments, have clouded examination of the outcome. (See Botterud 2008, pp. 13-14).

50 See, for example, Edwell (2008) and Moran (2000).

14 percent of peak load.[51] Almost all of these increases in capacity occurred within the first 2-3 years of the market, after which installed capacity leveled off.[52]

These capacity investments, however, are not solely the result of a well-designed energy-only market. Government-owned and government-controlled assets represent 63 percent of the total installed capacity in the NEM and government ownership and public-private partnerships appear to have accounted for much of the capacity additions during that period.[53] We believe that government ownership meant that sub-commercial returns on investments have been accepted on both existing and new capacity. If government investments are made despite low rates of return from energy-only markets, then the ability of energy-only markets to attract private investment will be undermined and private investors will fare poorly. This appears to have been the case in Australia. For example, Moran (2000) discusses two generation developments accounting for 1,770 MW of capacity built by a consortium of investors led by Shell, one of which was a partnership with the government of Queensland.[54] Soon after this capacity came on line, "Shell clearly felt its investment had turned sour and steadily sold down its interest … [selling] investments that had cost $2.2 billion at only $1.2 billion."[55]

Even the British Electricity Trading and Transmission Agreements ("BETTA"), implemented in 2005 as a geographically expanded version of the New Electricity Trading Arrangements ("NETA"), may be experiencing revenue deficiencies despite the absence of a price cap.[56] Although NETA started in 2001 with significant excess reserve margins of around 25 percent, many existing capacity resources were mothballed as energy market prices dropped dramatically.[57] In fact, during its first year of operation, wholesale energy market prices dropped by 20 percent for on-peak hours and 27 percent for off-peak hours.[58] Reductions in wholesale electricity prices were attributed in part to decreased ability to manipulate market prices under the pay-as-bid structure of the market and the elimination of the capacity payment. Under the current British market design, generators are able to engage in long-term over-the-counter ("OTC") bilateral contracts with electric distributors, but these agreements are voluntary, with retailers being under no obligation to obtain contracts for sufficient capacity to meet their peak demand. However, capacity retirements have now become a significant concern, as not enough generation has been attracted into the system to keep pace with demand. The CEO of National Grid, the system operator, stated that "Britain would fail to attract enough investment in new plants and would lack sufficient generating capacity to meet peak demand around 2015."[59] Similarly, *The Economist* (2009) recently reported that a shortage of power-generation capacity due to the likely retirements of 20,000 MW of aging generating plants could lead to blackouts

---

51 Based on a peak load of 32,000 MW in 2006 (AER (2007), p. 85).
52 AER (2007), p. 71-77.
53 AER (2007), p. 67.
54 Note that these particular investments amount to 40 percent of the total net increase in capacity observed since NEM began.
55 Moran (2000), p. 410.
56 For BETTA start date, see NGET (2009) p. 5.
57 Ofgem (2002), p. 3.
58 Hesmondhalgh (2003), p. 552.
59 Pagnamenta (2008).

across Britain between 2013 and 2016. In part to address these anticipated shortages, the market regulator, Ofgem, has announced that it is considering adding investment incentives for renewable generation projects. [60]

Similar concerns have been raised in ERCOT.[61] Despite reliability challenges during the summer of 2009 and alarming forecasts of inadequate reserve margins only two years ago, ERCOT currently projects that the combination of the economic downturn and an apparent power plant boom will keep reserve margins above resource adequacy targets. However, this optimistic forecast of resource adequacy levels has been questioned by industry participants noting that current projections consider neither that many of the older coal-fired power plants will likely retire nor that many of the currently planned generation additions may not be realized due to financing problems. The fact that retirement plans need to be announced only three months before the plants are taken off line aggravates challenges because there may be insufficient time to replace retired capacity. In addition to current economic conditions, financing new generating capacity has become more difficult due to a significant reduction in current and projected energy prices caused in part by the massive construction of wind power plants. To address these resource adequacy questions and uncertainties, including the need for additional regulation capacity to balance wind generation, ERCOT has reactivated its generation adequacy task force.

The two Canadian markets we reviewed (Alberta and Ontario) have demonstrated opposite results with respect to revenue adequacy for new generation capacity. In Alberta, market prices appear to be high enough to support new entry for a variety of generation technologies. Annual net revenues as a percentage of total capital costs are reported to range from 11 to 20 percent for various baseload generation resources, leading the Alberta Market Surveillance Administrator to conclude that the investment climate was attractive. While a price cap of only $890/MWh (C$1,000/MWh)[62] is in place, average monthly market prices in the AESO have reached levels as high as $160/MWh (C$180/MWh).[63]

In contrast to Alberta, market prices in Ontario do not appear to support sufficient investment. Using an estimate of revenue requirements of approximately $76/kW-year (C$85/kW-year) for a combined cycle plant and $58/kW-year (C$65/kW-year) for a combustion turbine peaking plant, the Ontario Energy Board found that average net revenues derived from energy market sales by generating plants over the past five years have fallen well below levels that would support investment.[64] As discussed in more detail below, the Ontario Power Authority was granted the authority in 2005 to enter into long-term contracts at the government's directive in order to maintain adequate system resources.

The attractiveness of the energy-only market design is in part that market forces, rather than regulatory mandates, would determine the appropriate level of reliability. This reliance on

---

60 Pagnamenta (2008).
61 Megawatt Daily (2009b), including "ERCOT wrestles with how to spur generation."
62 Converted to USD based on an exchange rate of 1.1234 CD/USD. Currency conversion from FRB (2009).
63 Alberta Market Surveillance Administrator (2009), pp. 6-12.
64 Ontario Energy Board (2009), pp. 64-65.

market forces, however, also creates considerable risk and public policy concerns related to the question whether the market design will indeed be able to provide for adequate (*i.e.*, publicly and politically acceptable) levels of reliability. As the California power crisis has shown, resource adequacy is far from assured in restructured energy-only markets without regulatory backstop mechanisms or specific resource adequacy requirements. As discussed further below, long-term resource adequacy may not be achievable in fully restructured energy-only markets due to factors such as mitigation of necessary price spikes, the "common good" aspect of reliability, and associated concerns related to "free ridership." In fact, as the above discussions show, considerable concerns exist even in successful energy-only markets over whether generating capacity will continue to be adequate to ensure sufficient reliability and whether investment signals are sufficient to retain existing capacity and ensure timely investment in new capacity.

3. The "Common Good" Aspect of Reliability

Another reason that energy-only markets can result in insufficient retention and entry of capacity is that reliability cannot generally be differentiated between those customers who are willing to pay for higher reliability and those who are not willing to pay for such reliability. This creates a "common good" problem that invites "free ridership" by customers (or LSEs) who hope to benefit from high reliability paid for by the common pool of market participants without incurring such payments themselves. In other words, in today's power markets, reliability, like all common goods, is subject to the classic "tragedy of the commons" problem, where customers and load serving entities have individual incentives to over-use the resource or free ride on the system.[65]

Joskow and Tirole identify this under-investment in peaking capacity for reliability as a systematic part of the energy-only market.[66] To address this shortcoming, they propose a regulatory solution under which scarcity pricing is improved during scarcity events, even in cases when customers are not curtailed.

4. Out-of-Market Reliability Backstops

The inability to differentiate reliability among customers means that the system operator must act on behalf of the entire system in its role of ensuring reliability. This means operating reserves need to be maintained during normal conditions and shortage events. If a capacity shortage event is imminent, the system operator will need to identify the situation and engage in a series of out-of-market measures in an attempt to prevent brownouts, rolling blackouts, or cascading

---

65 For a classic treatment of the tragedy of the commons, see Hardin (1968). A common good, or common pool resource, is one that is rivalrous and non-excludable. For a discussion of rivalry and excludability, see Nicholson (2005) pp. 595-597. With respect to reliability, it is rivalrous because the more peaking capacity one customer uses (i.e. the higher their peak load), the less peaking capacity will be available to other customers in the system. It is non-excludable because individual customer cannot generally be prevented from using more than their paid-for share of reliability. With the exception of some load control programs, the technical capability to differentiated reliability (e.g., by targeting rolling blackouts to customers who have not paid for higher levels of reliability) does not yet exist.

66 Joskow and Tirole (2004), pp. 44-48.

outages. These include both measures such as long-term out-of-market capacity contracts and out-of-market actions taken on an operational time scale.

For example, system operators might purchase emergency energy from neighboring regions on behalf of the market, request curtailment from generators on all non-critical plant loads, call all emergency curtailment available from LSEs, dispatch power from generators that are normally shut down for environmental reasons, or dispatch generators at their maximum emergency generation levels[67] before enacting voltage reductions (brownouts) and load curtailments (rolling blackouts). If these procedures are not reflected in the hourly market prices, then customers are not exposed to true costs of reliability during the scarcity event, nor do most of them face actual curtailment of service.

If high VOLL-based prices occur only during actual service curtailments but the system operator engages in out-of-market measures to prevent curtailments, this will eliminate the scarcity prices that are necessary to attract sufficient capacity investments. The result is a vicious cycle where more and more resources must be maintained by the system operator through out-of-market solutions that keep prices artificially low.

Almost all power markets, including but not limited to energy-only markets, have developed out-of-market backstop mechanisms for ensuring reliability and sufficient capacity. In most cases, market operators simply procure reserve capacity outside the energy-only market framework if they expect peak capacity to be short of their targeted reliability standard. For the most part, these capacity procurement mechanisms are intended to be temporary or transitional until the energy-only market is developed more fully. However, a pure energy-only market that does not employ these backstops has not yet materialized.

In Australia's NEM, for example, reserve capacity can be procured under the Reliability and Emergency Reserve Trader ("RERT") mechanism, which recently replaced the Reliability Safety Net.[68] Although this out-of-market capacity procurement mechanism was set to expire several times, it was extended each time.[69] Under this mechanism, when capacity is forecasted to be too low to meet an expectation of 0.002 percent unserved demand, the system operator will procure reserve capacity outside the market.[70] The NEM has generally had sufficient capacity, however, and this mechanism was only used *once*—in 2006 when it procured 375 MW, or about 1 percent of system peak demand,[71] for a period of 2 months.

---

67 Maximum emergency generation is higher than normal maximum generation and can be provided only for short periods of time before causing equipment damage due to the operating stress.

68 AEMC (2008).

69 AEMC (2005) and AEMC (2008).

70 AEMC (2007).

71 Based on a system peak demand of 32,000 MW in 2006 (see AER (2007), p. 85).

In Nordpool, reliability and capacity adequacy are ensured by the member Transmission System Operators ("TSOs")[72] based on a 0.1 percent loss-of-load probability ("LOLP").[73] When capacity is forecast to be insufficient on a 3-year forward basis to meet the reliability target, the TSO is authorized to procure peaking resources under long-term contracts with the costs of the procurement paid by the state. The potential for conflict between this peaking procurement backstop and attracting capacity through the market has been acknowledged explicitly. For this reason, Nordel, the TSOs' common organization, states that the length of the peak capacity procurement "should be based on an evaluation of when commercially-driven investments will be available to enter the market."[74] It is unclear, however, how it would be determined when such commercially-driven investments will become available in the presence of backstop procurement.

In the old British pool of the 1990s, capacity payments were an explicit part of the market price but were eliminated with the advent of the NETA in 2001.[75] Although capacity payments were eliminated, Great Britain does have a provision for the procurement of Short Term Operating Reserves ("STOR"). Although used only for short-term reserves, the capacity providing the reserves is procured by the system operator under a long-term contract of up to five years because such longer-term contracts are seen to be required to "provide sufficient investment signals to providers … and allow sufficient time for payback of a provider's investment" and can be tendered even prior to the installation of an asset.[76] As is typical in other markets for peak power and reserve arrangements, STOR is paid for with an availability charge and a usage charge, with the availability charge being quite similar to a capacity payment. Over 2007-08, National Grid Electricity Transmission[77] ("NGET") procured, on average, 1,926 MW of STOR from committed service,[78] or more than 3 percent of peak demand.[79] Average availability payments for flexible and committed STOR were $76/kW-year during the 2007-2008 operating

---

72 Each country has its own TSO, which controls the grid and is responsible for maintaining reliability in the short-run and long-run (see Nordpool (2009)). These TSOs each maintain their own emergency reserve capacity (see Botterud (2008)). See also Roques (2007)

73 Nordel (2007), p. 5.

74 Nordel (2007).

75 IEA (2005), pp. 176-177. NETA was replaced with its successor BETTA on April 1, 2005, see NGET (2009), p. 5.

76 NGET (2008a), p. 2.

77 NGET is the transmission owner and grid operator responsible for ensuring reliability; it is also the whole owner, but not controller, of ELEXON, the market operator (see ELEXON (2007)).

78 Along with an additional 436 MW from flexible service. The difference is that flexible service can opt out of providing reserves and therefore the contract is not as similar to a capacity contract. For the same reason the availability payment paid for flexible service is lower than that for committed service.

79 Based on a peak load of 61 GW (see NGET (2008b) p.1 and NGET (2008c), Ch. 2.

year.[80] These STOR costs are recovered from retailers along with other ancillary services costs as uplift charges calculated separately for each half hour.[81]

In ERCOT, out-of-market capacity purchases have been made in the form of reliability-must-run ("RMR") contracts. These contracts are signed to retain capacity resources that might otherwise be retired or mothballed. In the 2003-2004 operating year, the total capacity procured through this mechanism was 1,850 MW (about 3 percent of peak load)[82] at a cost of $66/kW-year.[83] ERCOT has reduced payments under RMR contracts from $122 million in 2004 to $33 million in 2007. This decline is a result of ERCOT's effort to reduce reliance on out-of-market contracts.[84] These efforts are identified in ERCOT's planning processes as any alternatives to RMR contracts that could lower system cost, and consist primarily of transmission system upgrades.[85]

RMR contracts were also relied upon heavily by the CAISO in its initial energy-only market design. In particular after the 2000-01 power crisis, RMR contracts were used to maintain reliability by retaining old, inefficient generating units in zones with local transmission constraints where significant revenue deficiencies were caused in part through stringent mitigation of market prices. For example, in 2004, a year when reliability-related expenditures peaked, the CAISO spent about $649 million on RMR contracts, including fixed contract payments, pre-dispatch costs, and real-time dispatch costs.[86] Retaining old, inefficient generating facilities through out-of-market procurement mechanisms, however, further increased revenue deficiency in the California energy-only market. The need for these specific out-of-market measures has declined substantially since 2006, however, after the California Public Utilities Commission ("CPUC") implemented a 1-year forward resource adequacy standard for all load serving entities under its jurisdiction, as discussed in Section VIII.[87] By 2008, total RMR expenditures had declined to $71 million.[88]

---

80 Based on an average availability payment of 5.89£/MWh and an exchange rate of 1.589 USD/£ (see NGET (2008b), p. 1; currency conversion based on FRB (2009)). Both committed and flexible availability payments are included in the number because the payments are not reported separately; the availability payment number would be higher if only committed resources were considered.

81 Transmission charges are assessed in a similar manner, but are based on the retailer's annual peak load. See NGET (2009) for full details.

82 Based on a system peak load of 58.5 GW in 2004 (see Potomac (2008), p. 71).

83 Calculated from the RMR net cost number of $122M and the total of all MW procured from 13 RMR contracts (see Potomac (2008), p. 111; ERCOT (2002), p.4; and ERCOT (2004), p. 45). Total RMR contract payments differ between two of these sources ($138M from ERCOT 2004 and $122M from Potomac 2008), possibly related to the "balancing energy neutrality account" credit subtracted from the total contract payment.

84 Potomac (2008), pp. 109-113.

85 ERCOT (2002).

86 For details behind CAISO 2004 reliability expenditures, see CAISO (2005), Ch. 6.

87 CAISO (2008), p. 5.

88 For details behind CAISO 2008 reliability expenditures, see CAISO (2009a), Ch. 6.

In Ontario, the lack of sufficient investment incentives is addressed through centralized, long-term purchases by the Ontario Power Authority at the request of the government.[89] This government agency currently manages contracts of about 10,500 MW, a third of which involved resources already in commercial operation. By 2013, the contracts entered into by the Ontario Power Authority will represent more than $14 billion of infrastructure investments made since 2005.[90] While the government directs these purchases, the Ontario Power Authority generally uses competitive solicitation for selecting individual resources. Recently, a large RFP for nuclear resources was suspended due to concerns about the pricing of the lead bidder.[91]

Creation of such targeted backstop measures to assure reliability and prevent the retirement of power plants to maintain reliability are often attractive initially. The measures can avoid the severe price spikes that would otherwise need to occur in energy-only markets to attract investments. By limiting additional capacity-like payments to only a few power plants, the measures also appear to be a less expensive solution than other options, such as system-wide capacity payments. The disadvantage of such backstop measures, however, is that they suppress market prices, which increases retirements of other existing plants, reduces the entry of new plants, and undermines or delays development of demand-response measures. These distortions grow over time and the need for backstop payments increases quickly as more existing resources retire and the development of new supply and demand side resources is delayed. Left in place over a number of years, these distortions decrease system efficiency, making it more difficult and expensive to eliminate the market distortions and transition to market-based solutions. Furthermore, such out-of-market solutions can be more costly even in the short-term; for example, system operators may opt to rely on RMR contracts with old generating units even in cases when a move toward demand response might have been a more cost effective alternative.

The temptation to rely on out-of-market contracts is not limited to energy-only markets. It is a particularly difficult problem in energy-only markets because of the lack of a market-based backstop procurement option. We discuss some of these same issues in relation to other market designs, and also address examples in which RMR contracts were replaced with market-based backstop mechanisms, such as centralized capacity markets.

5. Scarcity Pricing Mechanisms

Scarcity pricing mechanisms are regulatory tools to increase market prices to sufficiently high levels when system reliability is threatened, before customers are curtailed. One approach to scarcity pricing is to define a scarcity event as any period during which the system operator is unable to maintain its target level of operating reserves (*e.g.*, 6 percent depending on local reliability standards).[92] Hogan (2005), for example, recommends that prices are gradually increased above marginal-cost-based prices to the full value of VOLL as reserves reach the point where brownouts or rolling blackouts must be implemented (*e.g.*, at an operating reserves level

---

89 Harvey and Travers (2008), p. 28.
90 Ontario Power Authority (2009).
91 Barber (2009).
92 Hogan (2005), pp. 9-18 and Joskow and Tirole (2004), pp. 43-46.

of 2 percent, depending on local reliability standards). As Cramton and Stoft note, incorporation of such a "demand" function for reserves restores revenue sufficiency (*i.e.*, fixes the missing money problem). They note that this is in contrast to other energy-only market design options would not provide sufficient revenues to retain and attract the capacity necessary to maintain reliability.[93]

Setting scarcity prices at correct levels is vital for energy-only markets. If scarcity prices are too low, insufficient capacity will be retained or built. If they are too high, excess capacity will result. If the design of scarcity pricing is uncertain, meaningful investment signals will not be sent to market participants. In fact, several energy-only markets have experienced difficulties and have recognized problems with their scarcity pricing mechanisms. Each of the energy-only markets has developed its own method for scarcity pricing, although none has implemented the demand function approach recommended by Hogan (2005) at this point.

Nordpool's scarcity pricing mechanism is quite simple: if the level of available capacity is so low that TSOs must provide additional supply out of their capacity reserves, then the day-ahead price is increased to the price cap, and prices in the intra-day and balancing markets must be as high or higher.[94] Joskow (2006a) recommends this method of setting prices equal to the price cap as a "rough and ready" mechanism for scarcity pricing.[95] Nordpool sees its reliance on TSOs' non-market-based reserves as a transitional market failure, which justifies setting market prices equal to the price cap in an attempt to attract market-based investment.

In both ERCOT and NEM, scarcity prices are set largely as they are set during normal conditions: the highest accepted supply bid determines the market price. However, because large generators' bids are mitigated, these prices are set primarily by bids from small, unmitigated suppliers. Potomac Economics, the independent market monitor for ERCOT, noted that this scarcity pricing method of "relying upon the offers of small participants to ensure scarcity prices during legitimate shortage conditions produced unreliable results."[96] The market monitor notes that the level of small suppliers' bids does not accurately reflect system conditions, resulting in widely varied price outcomes during identical system conditions when balancing reserves are exhausted.[97] In this pricing system, the unmitigated small suppliers would need to be able to accurately predict the existence and severity of a system shortage, so that prices would be set correctly according to shortage conditions. The problem is that during a legitimate shortage, even assuming the absence of out-of-market interventions by the system operator, the marginal bid will be dispatched and determine the market price even if that offer does not accurately reflect severity of the scarcity event.

ERCOT's scarcity prices may also have been too low, resulting in average annual prices that, according to the market monitor's estimate, have "reached the level sufficient for new market

---

93 Cramton and Stoft (2006) pp. 32-34.
94 Nordel (2007).
95 Joskow (2006a), p. 45.
96 Potomac (2008), pp. xxxv and 46-52.
97 *Id.*

entry in only one of the last five years."[98] The market monitor has recommended moving toward a scarcity pricing model similar to the Hogan (2005) proposal after the market has implemented nodal pricing.

Energy-only markets are not the only designs under which scarcity pricing adds valuable incentives to the market. For example, NYISO, ISO-NE, and PJM have or are currently developing scarcity pricing mechanisms in addition to preexisting capacity markets, as discussed in Section IX.B.3.

6. Difficulties Distinguishing Market Power from Scarcity

Many energy-only markets, including NEM, BETTA, and ERCOT, rely on generators to increase their bid prices above marginal costs in order to set scarcity prices. This means that the markets implicitly rely on suppliers to accurately predict system shortages and make scarcity-priced offers if there is an anticipated, legitimate system shortage. If suppliers are inaccurate in predicting system shortages when they submit their bids (often hours to a day before actual operations) they may cause price spikes even when there are no shortages, or they may bid too low and fail to reflect scarcity when there are shortages. In addition, because some suppliers may also be able to exercise market power and manipulate prices through their bidding, they may inflate the market price to scarcity-type levels even in the absence of a physical shortage.

Because suppliers obtain market power only during shortage events (*e.g.*, by becoming a "pivotal" supplier), it can be almost impossible to distinguish between high prices that are consistent with efficient scarcity prices and high prices that reflect an abuse of market power. This issue was one of the more intractable issues during the 2000-01 California power crisis, when high loads and low hydro conditions combined to create shortages. When prices skyrocketed, some market participants argued that the prices were consistent with scarcity conditions, while others (including witnesses of *The Brattle Group*) concluded that the shortages created conditions that allowed for the abuse of market power, including the infamous Enron gaming strategies, which increased prices beyond those consistent with market conditions.[99] Similar examples of the difficulty of distinguishing between legitimate scarcity and manipulative market behavior exist elsewhere, such as an event currently under investigation in Australia.[100]

---

[98] The market monitor used estimated annualized capacity costs and simulated energy market revenues to make the determination of whether overall price levels were sufficient for new entry. See Potomac (2008), p. 46-47.

[99] Fox-Penner (2003).

[100] For example, one Australian supplier is currently under investigation. "[P]rices in South Australia averaged [$173/MWh ($243 AUD)] in the March quarter 2008 … Against a backdrop of high demand and tight supply AGL Energy, which owns 39 percent of South Australia's generation capacity, bid a significant portion of its capacity at close to the price cap. In combination these factors led to sustained high prices in South Australia over 15 consecutive days in March 2008 … The [Australian Energy Regulator] is investigating whether the generator bidding behaviour breached the National Electricity Rules during these high priced events." See Edwell (2008).

The challenge of distinguishing between supplier bidding behavior that is consistent with scarcity pricing and behavior that constitutes an abuse of market power also makes it difficult to mitigate energy-only markets without creating revenue deficiency for existing and new generating assets. ERCOT's method for mitigating possible market manipulation during scarcity events has been to allow only small suppliers that are not perceived to hold market power to submit high offer prices. The problem with this method is that, although it may prevent exercise of market power by larger suppliers with market power, it also places small suppliers in danger of inefficiently pricing themselves out of the market or incorrectly creating scarcity-level price spikes if they mistakenly predict a shortage event.

7. Price Volatility

Energy-only market designs rely on occasional price spikes to retain existing generation and attract new investment. Significant price volatility and uncertainty is thus an accepted and expected aspect of well-functioning energy-only markets. This volatility, however, comes at a cost—both political and financial. Politically, the price spikes of energy-only market are difficult to accept and explain to the public, and even in the absence of apparent market power abuses this can lead policy-markers to impose out-of-market solutions. Financially, the increased uncertainty in market revenues increases the return that investors require before they are willing to enter a market. Due to higher risks, a higher risk premium must be paid to attract investments compared to market designs that impose less uncertainty. As discussed in Section III.D, the high level of uncertainty will also tend to delay investments and create a bias toward resources with lower fixed but higher operating costs.

The higher uncertainty means that a higher risk premium must be paid to attract investments. As Professor Joskow notes, there is "no particular theoretical reason that price volatility or price uncertainty per se should make it impossible to finance new generating facilities if the 'missing money' problem is solved … [although] price uncertainty will affect the cost of capital used by investors to evaluate projects."[101] In addition, although higher volatility increases financing costs, this does not necessarily mean it is good public policy to attempt to shelter customers or suppliers from price uncertainty as they may be able to reduce system costs if they can respond to price signals (*e.g.*, through demand response programs) or otherwise manage that risk (*e.g.*, contractually with retail suppliers).

8. Energy Call Options and Forward Contracting Obligations

As discussed previously, bilateral contracts are not discouraged by or excluded from the energy-only market design. However, a number of energy market design proposals also *impose* obligations on LSEs to sign multi-year forward contracts for energy or energy options. These forward contracting obligations are an attempt to mitigate price volatility to consumers, reduce market power concerns, provide generators with additional revenues, or provide revenue

[101] Joskow (2006a), p. 40.

certainty to investors.[102] In our opinion, these forward contracting obligations, which are meant to address the shortcomings of pure energy-only markets, modify the energy market design in a way that essentially creates the type of market designs discussed in Sections VI through IX below.

Exactly what such mandatory forward contract obligations mean depends in part on whether they rely on physical or financial contracts. Physical forward contracts include arrangements in which customers pay the supplier a fixed annual or monthly payment in order to secure the option (whether or not that option is ever exercised) of having power supplied physically from a specific generating unit.[103] These contracts have additional provisions for determining the energy price (or "strike" price) for any power actually dispatched, and the size of the penalty for failure to generate. The supplier must have a physical asset backing the contract, which means it amounts to a physical "call option" for energy in which the supplier must sell to the buyer at a fixed price at all times even in times of shortage when market prices have spiked.

A financial forward contract can have the same payment structure, but does not need to be backed with a physical resource. In the case of a financial call option contract, the customer makes a fixed payment for the option of buying power at a given strike price, which the seller will guarantee. If the market price is below the strike price, the buyer will buy from the market; if the market price rises above the strike price, the buyer will exercise the option and the seller will have to pay the difference between the market price and the strike price.

Traditional capacity contracts are very similar to energy call option contracts. This is why some observers refer to physical or financial capacity contracts as "energy call options." In fact, many capacity contracts, such as tolling arrangements, are call options in which the strike price may be based on current plant operating costs. The contract payments for the physical call option contracts discussed earlier essentially are capacity payments.

Both physical and financial forward contracts can shelter risk-averse buyers from the volatility and price spikes inherent in energy-only markets. For example, Australia's NEM experiences great activity in financial forward contracts to hedge against price spikes.[104] However, because system operators generally are not yet able to provide differentiated levels of reliability for different customers, neither a physical nor financial forward contract can provide buyers with greater reliability than is available in the system as a whole. For example, suppose an LSE desiring high reliability contracts enough physical capacity to meet their expected peak load even under extreme peak load conditions. If other LSEs have *not* contracted for similar levels of capacity, then the system may still be short overall during extreme peak load conditions, forcing the system operator to enact rolling blackouts. Because the system operator is not able to

---

[102] See discussion in Crampton and Stoft (2006). As they note (p. 34), Wolak (2004) focuses on the importance of long-term contracting; Chao and Wilson (2004) focus on options and risk management; and Bidwell (2005), Bidwell and Henney (2004), and Oren (2005) focus on energy options.

[103] Note that in other commodities markets, a "physical" contract is a promise to deliver a physical good like corn when called; within the context of power markets, a "physical" contract refers not only to the good (the energy produced), but also usually to the capacity asset that creates the good.

[104] AER (2007), Ch. 3.

distinguish among LSEs for reliability purposes, the high-reliability LSE will experience interruptions in service just like all other LSEs. The overall result is that the rest of the system "free rides" on the reliability value provided by the LSE with sufficient capacity contracts.

Regulatory mandates forcing all LSEs to own physical energy options contracts are intended to remedy this incentive to free ride on the greater reliability provided by others' bilateral contracts. Oren (2005), for example, proposed that *all* LSEs be required to obtain *physical* call-option contracts to cover their location-specific peak demand, including an administratively-determined reserve margin, on a multi-year forward basis. Adib *et al*. (2008) argue that such mandatory locational call-option contracts are superior to capacity markets because the pricing of the energy call option is directly tied to location-specific energy prices. Crampton and Stoft (2006), however, point out that imposing such forward call option contracts covering LSEs' peak load and an administratively-determined reserve margin no longer constitutes an energy-only market design. Rather, they note, such obligations create a "convergence" between energy-only and capacity market designs. We agree. Physical call option contracting requirements create a bilateral market for physical capacity, very similar to the capacity markets created by the market designs discussed in Sections VI and VIII.

Others propose that imposing only *financial* forward contracting requirements would be sufficient to reduce price risk, mitigate market power, address revenue deficiency and, thereby, create investment incentives that ensure resource adequacy. Wolak (2004), for example, argues that mandatory financial forward contracts would provide resource adequacy, because if locational hourly market prices are sufficiently high (indicating a need for new capacity at that location), then the supplier will have the incentive to build physical assets to back the financial agreement. We agree that mandatory financial forward contracts (including call option contracts) would create price stability for customers, but we do not agree that this approach would achieve resource adequacy unless the energy-only market already had sufficient investment incentives for resource adequacy. If the missing money problem persists in the energy-only market for any of the previously-discussed reasons, then the seller of a financial contract will prefer to pay the spot price during times of scarcity rather than building a physical asset.

C. ADVANTAGES AND DISADVANTAGES OF ENERGY-ONLY MARKETS

Pure energy-only markets are the simplest market design we examine. They are appealing because, at least theoretically, they require the fewest regulatory interventions. In pure energy-only markets, customers can choose the level of reliability they would like to purchase. The model does not need to rely on additional mechanisms to pay for capacity.[105]

There have been a number of challenges in translating the theoretical concept of energy-only markets into practice. Both technical and political realities have delayed significant penetration of demand response and do not allow the differentiation of reliability across different classes of customers. This prevents the market from achieving efficient outcomes during scarcity events.

[105] As Wolak (2002) put it, customers should not have to buy the entire bakery, only the loaves of bread.

This also means that the scarcity prices vital for ensuring proper investment incentives are undefined during shortage events and regulatory interventions are required both for setting prices and for maintaining sufficient reliability.

Advantages and disadvantages of energy-only markets are summarized in Table 2. The theoretical advantages of simplicity and minimal need for regulatory determination of reliability levels are accompanied by a number of distinct disadvantages. They include significant uncertainty about whether generating capacity will remain adequate, high price volatility that is politically unpopular and increases investment costs, challenges in distinguishing between true scarcity and exercise of market power, a need for regulatory mechanisms to define efficient prices during scarcity events in the absence of wide-spread demand response, and challenges in differentiating reliability for different customer classes. Without scarcity prices that are allowed to rise sufficiently to attract new investment, real-world energy markets routinely rely on out-of-market mechanisms for purchasing capacity and ensuring reliability, thereby undermining competitive pricing signals and perpetuating and increasing the need for out-of-market solutions. Not surprisingly, revenue deficiency (*i.e.*, the "missing money" problem) has become a barrier to generation investments in most real-world examples of energy-only markets. Various approaches to addressing the "missing money" problem are discussed in the remainder of this report, beginning in Section V with the addition of *administratively-determined capacity payments*.

Table 2
Advantages and Disadvantages of the Energy-Only Market

Advantages	***Disadvantages***
• Simple and transparent market design. • Reliability and investment levels are determined by market participants based on market conditions, not regulatory mandates. • "Solves" missing money problem if scarcity pricing is done carefully, prices are not capped at levels below VOLL, and reliability is allowed to match market demands (even if that means allowing reliability to drop). • No prescribed amount of capacity necessary. Proper incentives are put in place and customers choose the amount of capacity (*i.e.*, reliability) they are willing to buy in the aggregate. • Does not require but also does not preclude long-term bilateral contracting for capacity as hedge against scarcity prices.	• Demand response and reliability-based pricing do not yet exist on the scale necessary for efficient pricing. Demand curtailment (brownouts or rolling blackouts) are politically unpopular. • Price volatility causes a risk premium for investments, possibly inefficiently increasing the overall cost to customers. • Severe price spikes during scarcity events are politically unpopular. Difficult to distinguish between high prices due to legitimate scarcity conditions and high prices due to market manipulation. • Substantial uncertainties and public policy risks remain about the market design's ability to achieve resource adequacy in the long-term. • Common good character of reliability and associated incentives to "free ride" mean that scarcity pricing mechanisms are vital to proper market function, but markets have struggled with getting it right. • In the absence of providing different levels of reliability to different customers, VOLL and appropriate scarcity pricing mechanisms still need to be determined administratively. VOLL is uncertain and its calculation is difficult. • Most energy-only markets utilize out-of-market mechanisms for purchasing capacity and ensuring reliability, thereby undermining competitive pricing signals and perpetuating the need for out-of-market mechanisms. • Revenue deficiency ("missing money") is a problem in most energy-only markets.

V. ENERGY MARKETS WITH ADMINISTRATIVELY-DETERMINED CAPACITY PAYMENTS

A. ADDITIONAL MARKET DESIGN ELEMENTS AND CONSIDERATIONS

1. Design of Administrative Capacity Payments

Energy markets with capacity payments include most of the elements of energy-only markets, but add an explicit, administratively-determined capacity payment on top of market-based energy revenues. These capacity payments provide additional revenues to suppliers to allow the full recovery of their fixed costs. Unlike the other capacity market designs discussed in this report, which rely at least partly on market forces to determine capacity prices, the price and allocation of capacity payments under this market design is purely a matter of administrative judgment.

The addition of capacity payments allows for caps on energy prices that are much lower than the VOLL-based cap in the energy-only design discussed in Section IV, although the cap still needs to be higher than the marginal operating costs of peak generators at peak times. The capacity payment and lower energy price cap has the advantage that it reduces investment costs and protects buyers from the much higher volatility and uncertainty associated with energy-only markets. Capping prices is also a tool for market monitors to prevent price spikes that could have been caused by the exercise of market power. In the presence of the lower price caps, the addition of administratively-determined capacity payments is needed to make up for revenues otherwise missing in recovering fixed resource costs.

There are many variations of the design, but all include similar components. Regulators or market administrators determine the approximate level of capacity desired in the market, including the reserve margin desired for resource adequacy. A capacity payment structure is then developed for generating units serving peak loads and providing reserves that would otherwise be unable to recover their fixed costs from the capped or otherwise mitigated energy market.

There are also mechanisms for either making some generators ineligible for capacity payments or reducing the size of the capacity payments in the case of over-capacity.[106] Because the primary purpose of these capacity payments is to supply missing revenues in order to facilitate the construction of new capacity resources, the actual size of the payments is often determined based on the portion of the annualized fixed costs of a peaking unit that is not recoverable through energy and ancillary service revenues.[107] This approach is used because peak load must be met by generating capacity that will operate only infrequently. Some of these peaking units would run only a few hours per year or would not run at all in some years. The least expensive

[106] Oren (2000), p. 5.

[107] One notable exception was the design for capacity payments based on the VOLL used in the first British power pool design discussed in V.B.4, which has since been replaced by NETA and BETTA as discussed in Section IV.

way to meet these peaking requirements is generally a combustion turbine ("CT"), which is thus used as the cost proxy to determine capacity payments.

An equally important but often overlooked design feature of capacity payment mechanisms is how the costs associated with these capacity payments are recovered from customers. In terms of sending proper price signals, these capacity payments should be assigned to customers as a demand charge based on their load during peak periods. In practice, however, this theoretically efficient cost recovery is not generally used in these market designs. Rather, capacity payment costs tend to be recovered as an uplift charge spread out over all energy (MWh) supplied in the market.[108] This essentially eliminates demand response as one of the main factor that can reduce system-wide capacity costs.

This market design with administratively-determined capacity payments differs significantly from designs that rely on market-based mechanisms, including bids from demand and supply resources, to determine market prices for capacity. Rather than relying on market forces to determine the value of capacity and the level of capacity payments to suppliers, this design simply relies on administratively-determined payments that may or may not be consistent with actual market conditions and capacity prices consistent with needed investment signals. While capacity payments can be adjusted administratively over time in response to observed market conditions, these adjustments can be too infrequent or too *ad hoc* to be consistent with efficient market signals. This approach may distort markets and ultimately differ little from those achieved with other out-of-market mechanisms.

2. Payments to New Versus Existing Resources

In market designs without resource adequacy requirements for LSEs, administratively-determined capacity payments can be targeted to new resources or the retention of existing resources.[109] Administratively limiting payments to new plants or plants that would otherwise retire constitutes an out-of-market mechanism, similar to RMR and other reliability backstop mechanisms used under other market designs. While attractive because such limitations reduce the total amount of capacity payments, the selective payments will cause significant market distortions.

The additional capacity maintained through such payments will create or prolong the "missing money" problem for all resources that are not receiving special payments. No new capacity will be attracted (not even bilaterally) unless it receives a capacity payment. Over time, an ever-larger proportion of installed capacity will tend to receive such payments. In addition, if capacity payments are limited to a period less than the economically useful life of a resource (*e.g.*, to the 10 years used in Spain as discussed below), the payments will need to *exceed* the annualized capacity costs of the resource before entry occurs. New plants simply will not be built unless total investment costs can be recovered through the combination of (1) margins

---

[108] Adib, *et al*. (2008) pp. 336-37.

[109] As discussed in Section III.G, imposing resource adequacy requirements on LSEs creates equal market value for both existing and new resources.

earned in energy markets over the economic life of the facility; and (2) capacity payments earned over the facility's economic life. In addition to payments for new capacity, the need to extend such payments to existing generation will quickly grow as the suppression of energy market prices forces more existing resources into retirement unless they too receive the additional payment. Market distortions increase further as the administrator, rather than market participants, chooses which resources should be retained and which resources should be allowed to retire. This can result in the inefficient retention of high-cost resources that would retire in a market environment.

Costs likely are higher in the long-term due to the difficulties and market distortions associated with such an approach. The resulting market distortions may also make it increasingly difficult to transition from the administrative determination of capacity payments to market-based solutions.

B. Experience With Administratively-Determined Capacity Payments

There are several international power markets where administratively-determined capacity payments have been used to address the missing money problem and restore revenue sufficiency in an effort to attract and retain generating resources. For example, administratively-determined capacity payments have been used in Chile, Colombia, Italy, Peru, Spain, and South Korea.[110] These markets have employed various methods to determine the exact size and allocation of these payments, giving regulators or system administrators varying levels of discretion and judgment in influencing the overall investment outcomes.

1. Fixed, Availability-Based Capacity Payments

Availability-based capacity payments were first introduced in Chile in 1982, followed by Colombia with essentially the same design.[111] The payment is an availability-based system, where generators offering capacity into the energy market would receive the administratively-determined payment as compensation for having capacity available, regardless of whether it was dispatched to run. Generators receive a daily capacity payment based on the quantity of capacity available during months of peak demand or capacity shortage.[112] The Chilean system is based on a two-year planning horizon and includes a penalty for failure to deliver based on the VOLL.[113] Over the year, the capacity payments are intended to cover the fixed costs of a peaking unit that has demonstrated sufficient availability.

South Korea's system has a similar availability-based capacity payment for all capacity offered into the market, but it has a two-tiered structure that applies separately to base-load and peaking

---

[110] Adib, *et al*. (2008), pp. 336-37.
[111] Batlle (2007), p. 4547; Larson (2004); Rudnick (2002).
[112] In Chile, the peak demand months are May-September; in Colombia, the payments are made during the dry season of December-April when hydro capacity is limited (Rudnick (2002), p. 161).
[113] Huber (2006), pp. 4-5.

generators.[114] Both the hourly energy price and the capacity adder are determined separately for baseload and peaking plants, and the classification of each unit as "baseload" or "peaking" is administratively determined ahead of time. The capacity payment for baseload generators is derived from the gross capital and O&M costs of a coal unit while the capacity payment for peaking plants is based on the costs of a combustion turbine.[115]

The exact method for determining these capacity payments is a matter of administrative judgment and is generally adjusted over time based on investment outcomes and other market factors. In Chile, for example, the Argentinean financial crisis resulted in interruptions in natural gas imports, meaning that generators that had received availability payments were no longer able to produce reliably. In the aftermath of this gas shortage, Chile's assessment method now "derates combined cycle plants that do not have alternative fuel arrangements and, therefore, reduces their capacity payments."[116]

In both South Korea and Colombia, the capacity payment mechanisms have been criticized for lack of transparency and predictability. According to Park *et al*. (2007), the Korean "fixed payment is controversial… as it is generally set in an *ad hoc* way for a period of time, not precisely and transparently linked to actual market conditions, and therefore presents an additional element of regulatory risk to market participants."[117] Similar initial concerns in Colombia led to revisions in 1999 that were meant to stabilize the mechanism and provide investors more information about how these payments would be determined.[118]

All three of these markets, South Korea, Colombia, and Chile, are characterized by rapid demand growth, where significant investments in new generation capacity are vital for resource adequacy. If the capacity payment mechanisms were not attracting sufficient generation capacity, these markets would quickly face rationing. It appears, however, that these markets *have* generally attracted sufficient generation capacity and rationing has been avoided outside periods of severe drought in the hydro-dominated markets.[119]

Early results from the Korean market have shown significant investments from independent power producers ("IPPs"), although Park *et al*. (2007) report that, as in most other markets, these investments have slowed following the initial investment rush after deregulation. More recent plans for natural gas combined cycle plants have been delayed or cancelled, which is attributed to regulatory uncertainty and the government's stalled plans to divest itself completely from the six main generation companies.[120] Any lack of investment from IPPs will likely be compensated by investments from the government-affiliated generating companies, which still follow the official prescriptions for investments by the Ministry of Commerce, Industry, and Energy.

---

114 Adib, *et al*. (2008), pp. 336-37.
115 Park (2007), pp. 5821-22.
116 Rudnick (2005), p. 58.
117 Park (2007), pp. 5821-22.
118 Larsen (2004), p. 1772.
119 Pollitt (2004), pp. 10-12.
120 Park (2007), pp. 5820-22.

Chile's capacity payment system has been in operation in various forms for almost 30 years, which makes it the most time-tested market design we examine in this report. Pollitt (2004) describes Chile's generation investments as being robust, allowing Chile to meet peak demand through private investments in generation capacity, and allowing generators to earn sufficient rates of return. Even so, it appears that Chile is now moving away from the administratively-determined capacity payment model in favor of a centralized capacity market following the Brazilian forward capacity market design, as summarized in Section IX below.[121]

2. Capacity Payments that Decline with Increasing Reserve Margins

Spain's capacity payment mechanism is similar to the designs in South Korea, Colombia, and Chile in that generators receive payments for available generation capacity regardless of whether it was dispatched. But the Spanish market differs in several important respects. Most fundamentally, as of 2007, the Spanish market design relies on a downward-sloping capacity payment curve, shown in Figure 2 below, which is referred to as the "investment incentive" curve.[122] This additional mechanism provides capacity payments to new capacity or significant upgrades to existing capacity. The investment incentive applies in addition to availability payments of approximately $64/kW-year for a combined-cycle plant.[123] The investment incentive curve is a publicly-posted description of how the size of the capacity payment will be set as a function of the system-wide peak reserve margin.[124] This means that the greater the need for capacity (*i.e.*, the lower the reserve margin), the higher the capacity payment.

Spain's system operator can enter into agreements with generators making both (1) investments in *new capacity* or (2) significant investments in *upgrading existing capacity* that might otherwise be retired. The determination of the size of the capacity payment is made at the time of the investment, and the system operator pays the determined annual "investment incentive" to the generator for 10 years. In contrast to the capacity payments in the previously discussed markets, this means that existing generators in Spain *are not* eligible to receive these investment incentives unless they have recently entered under this agreement.[125] As with most other capacity payment mechanisms, the costs of these payments are recovered through an uplift charge for all energy purchased by load. Again, as noted earlier, such cost recovery based on uplift charges does not provide meaningful price signals that would reduce the need for generating capacity by stimulating demand response.

---

121 Rudnick (2005), pp. 58-59.

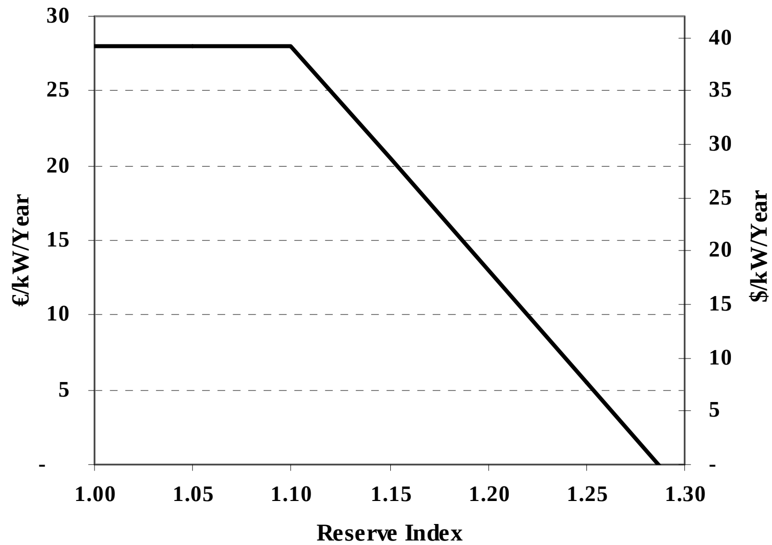
122 For a background discussion related to these recent revisions of the Spanish market design, see Battle *et al*. (2008). The current capacity investment incentive mechanism was added to the previous design that already rewarded generators for availability.

123 In March 2008, the system operator had proposed availability payments of $3.43/MWh (2.45 €) for hydro, $7.67/MWh (5.48 €) for NGCC, and $1.13/MWh 0.81 €) for coal. At an availability of 95 percent, this would amount to annual payments of $29/kW-year, $64/kW-year, and $9/kW-year respectively. See Federico and Vives (2008), p. 59. Currency conversion of €1 = $1.4 from FRB (2009).

124 Federico and Vives (2008), pp. 59-60, 124-129.

125 Federico and Vives (2008), pp. 59-60. For a brief discussion of market distortions and poor incentives related to the Spanish capacity payment design, see Roques (2007).

Figure 2
Spanish Capacity Payment Curve



Sources and Notes:
From Federico and Vives (2008), pp. 59-60.
Based on exchange rate of 1.400 $/€ from FRB (2009).

While this curve looks similar to the downward sloping demand curves used in some U.S. capacity markets, such as PJM and NYISO, capacity payments are determined solely based on current reserve margins, without considering actual costs or bids from new capacity.

3. Dispatch-Based Capacity Payments

Another version of the capacity-payment method was used in Argentina and Peru. There generators were paid for capacity only if they were actually dispatched. The capacity payment in Argentina was a fixed adder, $10/MWh in 1994, which was paid in addition to the energy market price during the day-time hours from 6:00 a.m. to 11:00 p.m.[126] This design created incentives for generators to underbid their energy costs so that they could be dispatched and receive the capacity payment.[127] This underbidding behavior was observed in Argentina, but was prevented in Peru by the enforcement of strictly cost-based energy bids.

Despite a structure that did not compensate generators for availability unless they were dispatched, the Argentinean electric market experienced significant private investments in capacity under this design. These investments amounted to a 4.9 percent annual increase in peak generating capacity between 1992, when deregulation began, and 2002, when the country was

---

[126] Pollitt (2004), p. 7.
[127] Adib (2008), p. 337.

rocked by a devastating financial crisis.[128] In the aftermath of the financial crisis, Argentina has experienced severe underinvestment and capacity shortages, prompting a move toward a government-controlled financial pool for making capacity investments.[129]

4. Capacity Payments Based on VOLL

The design or the original British pool—which was subsequently replaced by the NETA and BETTA energy-only market designs—also had a capacity payment component that was paid to all generators available for dispatch. The payment was conceptually different from nearly all other administratively-determined capacity payment designs in that the size of the payment for each half-hour dispatch period was calculated based on three factors: (1) the estimated VOLL, (2) the dynamically-calculated loss of LOLP, and (3) the system marginal price (SMP).[130] The design was developed so that capacity payments were low when available capacity was high compared to load and payments increased as the reserve margins dwindled.[131]

While many proposals for properly compensating available capacity follow this design conceptually, the British model was prone to manipulation that inflated the capacity payments made to generators. For example, Chuang (2000) states:

> Generators found early on that capacity payments were particularly sensitive to the amount of spare capacity declared in the pool. The method used to compute LOLP exaggerated the probability that plants would not be available, and led to magnified capacity payments. Before the problematic LOLP computation scheme was revised, generators could mis-report unavailability and collect capacity payments based on an invalid predication of scarce capacity. Even after revision, rules for computing LOLP led to other perverse affects, such as encouraging generators under certain circumstances to delay redeclaring availability after experiencing default.

These problems with manipulation and over-payments for capacity were some of the reasons that the system was discontinued and replaced by NETA, which is based on an energy-only market design as discussed in Section IV.

C. Advantages and Disadvantages of Adding Administratively-Determined Capacity Payments to Energy Markets

Table 3 summarizes the advantages and disadvantages of adding administratively-determined capacity payments to energy markets. Capacity payments are flexible tools that can solve the missing money problem, stimulate investments needed to achieve resource adequacy and, as a result, reduce price volatility and associated investment risk premiums. Given the absence of

---

128 Pollitt (2004), pp. 11-13 and Nagayama and Kashiwagi (2005), pp. 115, 123.
129 Nagayama and Kashiwagi (2005), pp. 124-25.
130 Chuang (2000), p. 5. Capacity payments were calculated as LOLP x (VOLL-SMP).
131 IEA (2005).

reserve requirements for LSEs under this market design, capacity payments can also be differentiated for new and existing resources.

The disadvantages of administratively-determined capacity payments include possible substantial deviations from efficient price signals and resources adequacy targets because of the lack of market feedback. Without market feedback, it is more likely to achieve outcomes under which capacity payments may be insufficient to attract needed entry to maintain resource adequacy or may over-compensate suppliers. Thus, resource adequacy levels remain uncertain. In addition, costs associated with capacity payments are also often assigned to energy rather than peak load, which blunts price signals and undermines demand response. Finally, frequent regulatory adjustments and revisions associated with administratively-determined capacity payments increase regulatory risk and investment costs.

Imposing *reserve requirements* instead of utilizing administratively determined capacity payments is an alternative market design that addresses some but not all of these problems, as discussed in Section VI.

Table 3
Advantages and Disadvantages of Adding
Administratively-Determined Capacity Payments to Energy Markets

Advantages	***Disadvantages***
• Flexible tool for administrators to retain and attract needed capacity. • Can differentiate incentives for new investment and retention of existing resources. • If capacity payments are broadly available to all market participants, energy market price caps can be lower, allowing market operators to mitigate energy market price shocks and possible market power. Payments can restore revenue sufficiency (solve the missing money problem) in mitigated energy markets with low price caps. • Stable capacity payments and improved resource adequacy reduce price risk and the risk premium required for new investments.	• Price and quantity are administratively determined, meaning that substantial deviations from efficient price signals are possible. • Generally no market feedback: capacity payments that are too low will fail to attract investments while payments that are too high will inefficiently burden customers; resource adequacy levels remain uncertain. • Capacity payments are often recovered from customers are through uplift charges based on their energy consumption rather than based on peak loads, meaning that economically-efficient market signals for reducing peak load through demand response are lost. • Capacity payments tend to be changed in an *ad hoc* or non-transparent manner by regulators, adding regulatory risk to potential suppliers and increasing the risk premium for investments. • Risk of significant market distortions if payments are differentiated for new and existing capacity.

VI. ENERGY MARKETS WITH RESERVE REQUIREMENTS BUT WITHOUT CENTRALIZED CAPACITY MARKETS

A. ADDITIONAL MARKET DESIGN ELEMENTS AND CONSIDERATIONS

The prior market design offered administratively-determined capacity payments to attract and retain capacity resources. An alternative approach is to ensure resource adequacy by imposing a reserve margin requirement on all LSEs operating in the market.

Under these "resource adequacy standards," each LSE is responsible for acquiring capacity rights that exceed its predicted peak load by the required reserve margin.[132] The required resources must be obtained by LSEs either through self-supply or bilateral contracts for capacity. This creates a demand for capacity, yielding a market price for capacity and associated capacity payments that restore revenue sufficiency (*i.e.*, solves the missing money problem) even in mitigated energy markets.

This market design avoids the "common good" problem discussed in the context of energy-only markets by imposing reliability standards directly on each load serving entity and preventing LSEs or their customers from "free riding" on system-wide reliability paid for by others. By imposing reserve requirements on each LSE, the design also provides a direct incentive to reduce peak demand and reduce required capacity reserves through implementation of efficiency and demand response programs.

Rather than relying on an administratively-determined capacity payment to facilitate the investments needed to maintain reliability, this market design relies on an administratively-determined reliability level and a corresponding reserve margin requirement that is imposed on LSEs based on their peak load. As is the case with other market designs that rely on administratively-determined reserve requirements, this design generally does not explicitly consider the cost of achieving that level of reliability and whether these costs are justified by the value of that reliability. Customer preferences and their use of electricity have changed and the cost of adding capacity has increased significantly over the last decade. It is consequently unclear that current reliability standards, which have remained unchanged for a decade or two, still appropriately balance the tradeoff between the cost of achieving mandated levels of reliability and the value provided by that reliability. Today, the "optimal" reliability standard may be higher than a decade ago, if customers are more sensitive to service interruptions. It may also be lower than a decade ago, considering the significant increases in the cost of new generating facilities. This issue is present in all markets with reserve requirements, including those discussed in Sections VII through IX below.

Adding an enforceable reserve requirement to energy markets creates a bilateral capacity market. This means that even in the absence of a centralized capacity market, the market design is no

---

[132] The imposed resource adequacy standard is also sometimes referred to as mandatory load hedging ("MLH") if the required contracts are financial rather than physical. (See Hogan (2005), pp. 27-33.).

longer energy-only. The capacity-related costs that LSEs and their customers incur in such bilateral markets, however, are not immediately visible. This lack of cost visibility can make the imposition of reserve requirements more acceptable politically than market designs based on administratively-determined capacity payments or centralized capacity markets in which the potentially high cost of satisfying the imposed reliability standards is much more obvious. Of course, the fact that these costs are less visible if they are incurred solely in bilateral capacity markets does not mean that the costs are lower. In fact, if bilateral capacity markets are not sufficiently liquid or competitive, which is the case in most of today's power markets, the costs of satisfying reliability requirements may actually be higher.

This market design neither requires nor explicitly provides for backstop capacity procurement by the system operator to prevent deterioration of system reliability caused by capacity deficiencies of non-compliant LSEs. This means the design does not provide a clear role for system operators to enforce procurement of capacity in case of system-wide deficiency. While the system operator is usually able to impose financial penalties on deficient LSEs, there may not be sufficient time to procure additional capacity once capacity deficiencies are detected. Such capacity deficiencies can be difficult to anticipate, particularly in restructured markets, given (1) the uncertainty in the development and timely ultimate construction of planned new resources and, perhaps even more importantly, (2) the fact that generation retirement decisions often become known only months before a plant shuts down.

The market design can also suffer from a lack of procurement transparency in cases when the system operator is forced to procure capacity in order to maintain reliability. The desire for such backstop procurement transparency often leads to the addition of explicit capacity procurement mechanisms, generally in the form of a centralized capacity auction that is administered by the system operator, as discussed in the next section of this report.

This design, which relies solely on self provision and bilateral capacity transactions, poses additional challenges in restructured markets. Bilateral capacity markets also tend to be fairly illiquid and, as a result, offer only limited pricing transparency. As a result, bilateral markets also tend to impose higher transactions costs and risks on small load serving entities, which need to transact in small quantities and who generally do not have the resources and risk management expertise of larger market participants. These transactions costs can be particularly challenging because of load migration in restructured power markets with retail competition. With load migration, frequent adjustments are needed to individual LSEs' capacity obligations, requiring frequent transfers of existing capacity resources.

Sole reliance on self-supply and bilateral contracts also creates challenges with respect to the enforcement of reserve requirements and monitoring and mitigation of market power. Large numbers of small LSEs with variety of contractual terms and bilateral arrangements can make it difficult and time consuming for system operators to determine whether an individual LSE actually satisfies its capacity obligation. The potentially wide variety of price and non-price terms will also make market monitoring and mitigation more difficult, if such monitoring and mitigation activities are even attempted in a purely bilateral setting. These challenges also exist for the bilateral market design with forward reserve requirements discussed in Section VIII.

B. Experience with Reserve Requirements without Centralized Capacity Markets

This market design has been the standard design of most *traditionally-regulated* U.S. power markets. In these traditionally regulated markets, vertically integrated utilities "voluntarily" comply with regional reliability standards or state-specific planning standards by adding sufficient resources, the costs of which can be recovered through regulated retail rates that reflect the utilities' cost of service. Voluntary compliance is monitored by the regional reliability organization and state regulators, but (until recently) generally without explicit verification or penalties for non-compliance. This approach has performed well in a traditional industry structure with relatively few vertically-integrated utilities and exclusive service areas.

This market design also existed in the traditional power pools and reserve sharing pools of the Northeastern U.S., which have since evolved into the current system operators ISO-NE, NYISO, and PJM. It is also still used in SPP. Under this pool-based design, several neighboring utilities "pool" generating resources, including those devoted primarily to reserves. By pooling, these utilities are able to reduce both total system dispatch costs and the total cost of shared reserves. Participating utilities are under a contractual obligation to provide sufficient capacity to the pool either through self-supply or bilateral contracts. For example, before it became an RTO, PJM relied upon the PJM Installed Capacity Accounting ("PICA") process, in which installed capacity obligations were established for annual planning periods starting June 1 each year. The installed capacity margin was apportioned to each of the participating utilities based on their forecast and historic loads. An after-the-fact review of actual load and unit performance assessed penalties to deficient members.[133]

The current SPP reserve requirement is similar to the traditional pool design. Member utilities of the RTO are required to fulfill the mandated 12 percent capacity margin through bilateral contracts or self-supply. Bilateral contracts must be physical contracts with firm transmission reservations.[134] Although the SPP governing documents give the pool broad authority to verify that each member utility has fulfilled its capacity obligation, SPP has no backstop procurement mechanism and no penalties for deficiency.[135] The lack of specific enforcement provisions is not detrimental, however, because SPP members are traditionally regulated, vertically-integrated utilities with resource planning requirements overseen or enforced by state regulatory commissions. SPP has been able to attract significant capacity over time, including merchant generators, increasing installed capacity by 33 percent between 2000 and 2007.[136] In 2006 and 2007, SPP had reserve margins of 34 percent and 33 percent respectively, although this large number overstates resource adequacy since the region is internally transmission constrained.[137]

---

[133] Hinkel (2001), p. 1.
[134] NERC (2008a), p. 222.
[135] SPP (2009), pp. 2.2-2.4.
[136] NERC (2008b). The SPP membership requirements, including those for capacity requirements, predate its RTO status and initiation of market operations (SPP (2009), p. i). SPP was granted RTO status by the FERC in 2004 and launched its energy balancing market in February 2007.
[137] Roach (2008), p. 19.

In today's <u>*restructured power markets*</u>, the system operator generally has the central role of verifying both the size of LSEs' reserve requirements (a function of their peak load) and the sufficiency of the LSEs' available capacity resources. The primary means of ensuring that these entities have procured sufficient capacity is through financial penalties for non-compliance. Penalties for non-compliance are generally assessed at the beginning of the peak season or during the delivery period. This function is particularly important in restructured markets with active retail competition that involve a larger number of small LSEs that are not rate regulated.

C. ADVANTAGES AND DISADVANTAGES OF ADDING RESERVE REQUIREMENTS TO ENERGY MARKETS WITHOUT A CENTRALIZED CAPACITY MARKET

Table 4 summarizes the advantages and disadvantages of the added market design components in energy markets with reserve requirements and sole reliance on bilateral capacity transactions. The added reserve requirement (*e.g.,* in the form of an installed capacity obligation on all load serving entities) ensures reliability and creates a bilateral market for capacity that solves the missing money problem experienced in mitigated energy-only markets. The focus on reserve margins—*i.e.*, the difference between available resources and peak load—also allows for integration of demand response resources.

Several challenges are associated with this approach. As is the case with any other market design that sets reserve requirements, the approach does not generally balance the value of the mandated reliability level against the cost of providing the required reserves. Setting any kind of reserve requirements requires monitoring and enforcement of the requirement, which can be difficult due to the great variety of terms and conditions in bilateral arrangements. In addition, if these reserve requirements are enforced only shortly prior to the delivery period, there may be little that can be done to resolve a deficiency by the time it is detected. Solely bilateral models often leave it unclear how to provide backstop procurement options to system operators in cases of such unexpected deficiencies.

The sole reliance on often illiquid bilateral capacity markets also imposes additional transactions costs and pricing uncertainty on market participants, in particular small LSEs and LSEs with migrating customer demand. Finally, sole reliance on bilateral capacity markets can lead to a lack of pricing transparency, difficulty in market monitoring and mitigation, and high transactions costs associated with potentially illiquid bilateral capacity markets. The bilateral market model also increases the difficulty of monitoring and mitigating capacity-related market power.

While these disadvantages may not be insignificant in traditionally-regulated markets where resource adequacy is maintained through the regulation of integrated utilities, including long-term planning processes, they can be substantial in restructured markets with retail competition and small LSEs. The addition of *centralized capacity markets* to energy markets with reserve requirements addresses these concerns by adding a backstop procurement mechanism, reducing transactions costs, providing greater liquidity and pricing transparency, and improving the effectiveness of market monitoring as discussed in Section VII.

Table 4
Advantages and Disadvantages of Adding Reserve Requirements to Energy Markets without Centralized Capacity Markets

Advantages	***Disadvantages***
• Clearly defines and enforces existing reliability standards that LSEs can satisfy through ownership or bilateral contracts. • Directly assigns the costs of capacity to LSEs whose customers are causing reserve requirements based on peak load. • Reserve requirement imposed on LSEs creates bilateral capacity market that restores revenue sufficiency in mitigated energy markets with low price caps. • Capacity obligations based on peak load facilitates participation by demand response resources. • Long track record in traditionally-regulated U.S. power markets, with energy markets for purchases and sales among regulated utilities and reserve requirements that are monitored and enforced through state regulators and integrated resource planning processes.	• Required reserve margins may not appropriately balance the value of increased reliability against the cost of providing the reserves. • "Voluntary" compliance not reliable in restructured markets with many (and often small) load serving entities; enforcement of and penalties for non-compliance can be difficult due to range of bilateral contract terms and potentially large number of LSEs. • Lack of a forward requirement means there may be too little time to make alternative arrangements once a system-wide reserve deficiency is discovered. • Potential lack of transparent mechanisms for backstop capacity procurement by system operator in case of deficiencies. • Lack of liquid, transparent capacity markets imposes additional costs and uncertainties, particularly on LSEs with migrating customer demand. • Small LSEs often face higher transaction costs and may find it more difficult to meet the requirements bilaterally or through self-supply. • Bilateral market structure makes it more difficult to monitor and mitigate market power, which can be costly in restructured markets with mostly unregulated generation and retail competition.

VII. ENERGY MARKETS WITH RESERVE REQUIREMENTS AND CENTRALIZED CAPACITY MARKETS

A. ADDITIONAL MARKET DESIGN ELEMENTS AND CONSIDERATIONS

This market design imposes a reserve requirement and adds a centralized capacity market administered by the system operator for the purposes of purchasing sufficient capacity, procuring backstop capacity, and facilitating capacity transactions among market participants. Market participants are still able to satisfy all of their capacity obligations through self supply or bilateral transactions, including through long-term capacity contracts. This means the centralized capacity market is only a *residual* capacity market for settling uncommitted resources and load obligations not already met through self-supply or bilateral contracts. For this reason, the addition of a centralized capacity market increases the options available to market participants, but does not preclude options open to them in a solely bilateral model.

The market operator determines the resources necessary to satisfy reliability requirements, verifies the residual capacity requirement that must be met by each LSE, and verifies capacity available from each supplier. The capacity market, which can be cleared through various auction formats, balances market-wide reserve requirements against bids for incremental supply, resulting in a transparent market price for capacity. The transparency of the capacity price, in combination with the additional liquidity provided by the centralized market, also facilitates bilateral contracting. In many market designs with centralized capacity markets, resources that are self-provided or that clear through the capacity market are obligated to bid into the energy market as well, although some exemptions exist, particularly for demand-response resources.

Committed resources that are no longer available to the market need to be replaced or they face penalties imposed by the system operator. The process of specifying the terms and conditions of how resources can participate in the capacity market, along with the specified terms and conditions of how the system operator will verify that capacity requirements have been met by LSEs, also can be used to create a standardized capacity product that can be traded more readily through bilateral transactions. The centralized capacity market format similarly facilitates the monitoring and mitigation of market power concerns, both on an *ex ante* or *ex post* basis.

B. EXPERIENCE WITH RESERVE REQUIREMENTS AND CENTRALIZED CAPACITY MARKETS

1. Capacity Markets with Fixed Resource Requirements

The simplest way to administer a capacity requirement is to determine a set level of resources required and ensure that exactly that requirement is met. The MISO Resource Adequacy Requirement (“RAR”) works essentially this way. MISO develops a planning reserve margin based on a once in ten years LOLE, and that reserve margin is then imposed as a requirement on all LSEs unless their state PUC has already stipulated a different reserve margin requirement.[138] All LSEs are obligated to self supply that quantity of reserves (through ownership or bilateral contracts) or obtain the reserves through the capacity market. An LSE wishing to rely on the MISO capacity market can fulfill its capacity obligation through purchases in the monthly auctions that occur a few days before each deadline for demonstrating resource adequacy.[139]

The MISO capacity market is entirely voluntary, meaning that there is no obligation for either LSEs or suppliers to participate in the capacity market.[140] Any LSEs deficient in meeting their requirements during a particular month will be assessed monthly penalties for that deficiency based on a percentage of MISO's estimated CONE, initially set at $80/kW-year.[141] These penalties are modest if deficiencies occur only occasionally but increase significantly if deficiencies are detected consistently.[142]

The voluntary nature of the MISO capacity market design has two important possible limitations. First, since generators are not obligated to offer their residual capacity into the market, the market is vulnerable to the exercise of strategic withholding. Suppliers with sufficient market power may be able to ensure that the capacity market price is always close to its high limit. Second, the voluntary nature of MISO capacity markets may invite free ridership if paying the financial penalty is lower than the cost of procuring physical capacity, although a penalty that is higher than the cost of procuring in the market should discourage this behavior. While some market participants suggested that these concerns justify a mandatory capacity market, the FERC disagreed.[143] The initial experience with MISO's capacity market shows little trading activity and substantial price uncertainty. For example, for the August 2009 planning month, only 110 MW of capacity (out of an installed capacity of almost 130,000 MW) cleared at a price of $1/MW-month. In comparison, 864 MW of capacity cleared for the June 2009 planning month at a price of $50/MW-month, while the July auction cleared 364 MW at a price of $10,015/MW-month.[144]

The market design in Australia's smaller South-West Interconnected System (SWIS) has a similar residual capacity market with largely the same features, but enforces and procures capacity on an annual basis. Each peak-load resource wishing to supply capacity can apply to be

---

[138] MISO (2009), pp. 3.6-3.9. The state PUC's requirement would apply for those LSEs regardless of whether it was more lenient or more strict than the MISO requirement.

[139] MISO (2009), p. 6.86.

[140] MISO (2009), pp. 6.87-6.88.

[141] MISO (2009), pp. 6.91-6.93.

[142] The penalty is 100 percent of the CONE for the each MW of the LSE's capacity holding deficiency for the month of greatest deficiency in a single planning year. If the LSE is deficient in more than one month in the planning year, then an additional monthly penalty of 25 percent of the CONE will apply on the MW deficiency for each summer or winter month, or an additional 8.3 percent of the CONE on the MW deficiency for each spring or fall month.

[143] "We reject arguments that a mandatory auction or a mandatory centralized capacity market is necessary to ensure resource adequacy. Well-structured financial settlement provisions can cerate appropriate incentives for LSEs to invest in and contract for sufficient capacity to meet their resource adequacy needs." (See FERC (2008a), p. 12.)

[144] "MISO capacity auction on swinging pendulum," *Megawatt Daily*, June 25, 2009.

assigned capacity credits, which are assigned by the SWIS system operator based on analyses of how many credits the resource can provide. The SWIS design also assigns capacity obligations to each retailer that they must purchase either bilaterally from suppliers or through the centralized market. Any LSE capacity obligations unfulfilled bilaterally will then be obtained through centralized residual capacity auctions. If retailers have all acquired sufficient capacity credits bilaterally, no auctions are conducted. Capacity credits for the 2007/08 delivery year cleared the auction at $100/kW-year ($127.5 AUD).[145]

2. Capacity Markets with Downward-Sloping Demand Curves

A disadvantage of imposing fixed resource requirements that are enforced only days or months before the applicable delivery period is that this can easily lead to substantial price volatility and "bipolar" capacity prices that are either at the cap (when there is insufficient capacity) or close to zero (when there is excess capacity). The fixed resource requirement is also more prone to market power abuses when the system is close to the target resource requirement, because suppliers may be able to move prices from close to zero to the cap even by withholding relatively minor amounts of capacity. Buyers may similarly be able to move capacity prices levels close to zero by adding only modest amounts of capacity through self-supply. The fixed resource adequacy requirement also has the disadvantage that it does not reflect the incremental value of increased reliability that is associated with adding resources beyond the required target level.

To address these concerns, several capacity markets are based on a "downward sloping demand curve" that varies resource adequacy requirements as a function of capacity prices. If capacity prices are low, more than the target reserve capacity is procured. If capacity prices are high, less than the target is procured. Such a downward sloping demand curve is included in the NYISO unforced capacity ("UCAP") market as well as PJM's forward capacity market.

A downward-sloping demand curve stabilizes capacity market prices over time by avoiding the "bipolar" nature of capacity prices under a fixed capacity requirement and inelastic supply.[146] Figure 3 illustrates how a sloping demand curve accomplishes this. The chart on the left shows capacity market results when a fixed resource is required from the market. The supply curve offered into the market can be very steep, because existing capacity may bid in at close to zero while it may be impossible, at almost any price, to supply new capacity only days or months before delivery is required.[147] As shown in the chart, a small movement of the vertical required capacity line to the right, or a shift of the offer curve to the left, would quickly increase the

---

145 IMO (2006). Conversion from FRB (2009).

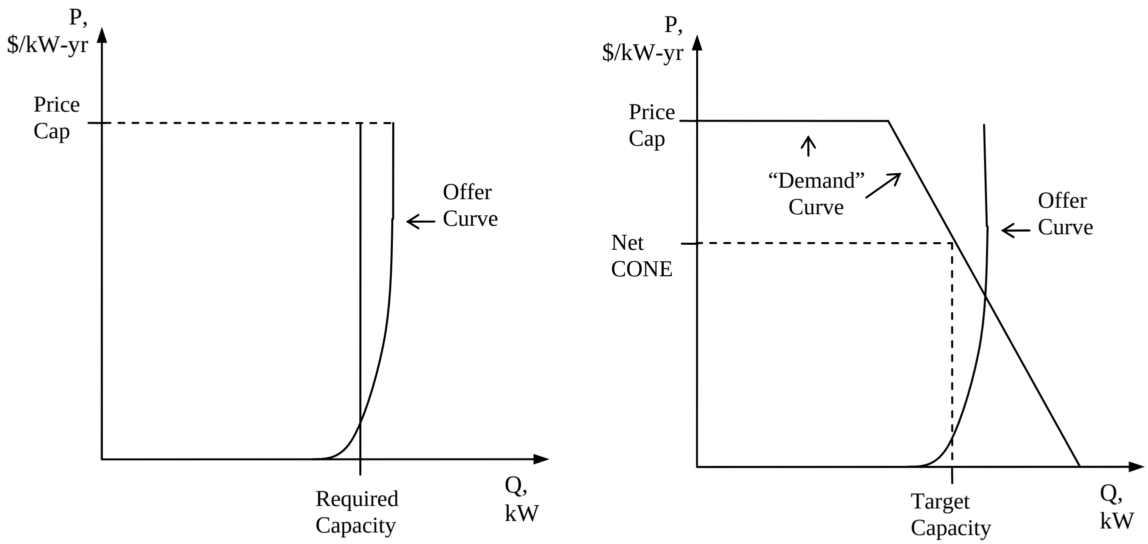
146 Oren (2005), p. 30.

147 As we discuss in the context of the following market designs, this problem is somewhat ameliorated in forward capacity markets, where more new capacity resources *would* be able to enter the market in time for delivery. However, probabilistic simulations have also shown that in market designs that impose and enforce a 3-year forward commitment period, a sloping demand curve still reduces price volatility and risk premiums for potential investments. (Hobbs (2005), pp. 61-64.)

capacity market price to the cap. This can result in substantial uncertainty in capacity prices that imposes a risk premium on market participants.[148]

The right chart in Figure 3 shows how the introduction of a downward-sloping "demand" curve for capacity serves to reduce price volatility. In this case, moving the sloped demand curve or the offer curve a similar amount results in a much smaller increase or decrease in market prices. Rather than falling close to zero with only modest levels of excess capacity, the market price decreases more gradually. Similarly, if supplies fall short of target capacity levels, market prices also increase gradually, recognizing that modest shortfalls may not warrant market prices equal to the price cap. This also reduces the incentive for suppliers with market power to withhold capacity, as withholding produces much smaller movements in market prices. It similarly reduces the ability of buyers to collapse capacity market prices by adding only small amounts of capacity to the system.

Figure 3
Capacity Market with a Fixed Capacity Requirement (left) or a Downward-Sloping Demand Curve (right)



(Adapted from Bidwell (2005), pp. 13-14.)

Note, however, that the shape of the downward sloping curve is not actually a representation of buyers' bids or actual customer demands. One point on the curve is defined simply by the target level of capacity (*e.g.*, the reserve margin needed to meet the regional reliability requirement) and the estimated net costs of new entry ("Net CONE") for peaking resources. The slope of the curve is typically based on multi-party settlements or administrative judgment. It is not generally related to the incremental or decremental value of reliability. Although this may lead

[148] For example, see Bidwell (2005), pp. 11-14.

to a somewhat arbitrary shape of the demand curve, the slope will nevertheless mean that shifts in either the demand curve or the offer curve will result in more gradual price movements.[149]

3. Locational Requirements for Capacity

Many regional power markets are not only capacity constrained, but also contain transmission import constrained locations. This means that simply making sure there is enough generation capacity in the entire footprint does not ensure resource adequacy because the power may not be deliverable to these locations. In order to make sure that capacity is deliverable to specific load centers, several U.S. power markets have introduced capacity requirements that are imposed on a zonal or “locational” basis. In any markets with import-constrained subregions, these locational requirements—referred to as locational installed capacity markets or locational resource adequacy standards—are a critical market design feature without which capacity payments or reserve requirements will not be able to attract or retain resources in locations where they are needed the most. This lack of proper locational investment signals will require and perpetuate out-of-market solutions to maintain system reliability, which undermines the very purpose of the market design.

The NYISO was the first market in the U.S. to introduce a locational capacity requirement.[150] It has three different capacity zones: New York City (“NYC”), Long Island (“LI”), and Rest of State. Because the two zones in southeastern New York have been perennially import-constrained zones with insufficient internal supply, creating these separate regions has the advantage of allowing capacity prices to climb in NYC and LI in response to localized shortages, while capacity prices in the Rest-of-State zone remain at lower levels commensurate with the greater available resources.[151] Note, however, that the NYC capacity zone has experienced market power concerns related to high bids, leading to unsold capacity.[152] In response to these concerns, in March 2008, the FERC ordered the NYISO to implement market mitigation measures that appear to have reduced or eliminated these market power concerns.[153]

Locational capacity requirements have the advantage of providing clear incentives to increase capacity in locations where it is needed most. The disadvantage is that workably competitive conditions are less likely to exist and are more difficult to establish in these small geographic zones, which leads to market power concerns and the need to monitor and mitigate such market power. The presence of a mandatory RTO-administered capacity market, however, greatly facilitates such monitoring and mitigation of market power. Even if most of the capacity obligations are self supplied through asset ownership or bilateral contracts, market monitors can

---

[149] NYISO (2008), pp. 5.4-5.9; PJM (2008), pp. 23-27.

[150] Patton (2007), pp. i-ii.

[151] Patton (2007), pp. 107-114.

[152] Patton (2007), pp. 109-110, stated that high prices “persisted even after a surplus emerged in New York City when approximately 1,000 MW were added in 2006. Prices remained near the revenue caps because a significant amount of existing capacity was not sold in the UCAP market due to the suppliers’ offer prices.”

[153] Patton (2007), p. 110.

observe all capacity resources in the region as well as the bid prices of any residual resources that are not already self supplied by market participants.[154]

PJM's RPM and California's resource adequacy standard (though not combined with a centralized capacity market) are similarly based on locational capacity requirements. In PJM's market, some transmission upgrades into constrained locations are also eligible to receive payments through the locational capacity market.[155]

C. Advantages and Disadvantages of Adding Centralized Capacity Markets to Energy Markets with Reserve Requirements

The introduction of a capacity market for LSEs to acquire their resource requirements and for suppliers to offer their capacity gives market participants more options and more flexibility. This increases market liquidity by creating a standardized product that can be bought and sold in almost any quantity, rather than relying purely on bilateral contracts that may have a variety of different terms and often do not accommodate the small quantities needed by LSEs in retail access markets. The transaction costs associated with such small, residual transactions are reduced as well.

The standardized capacity product and liquid market similarly make it easier for LSEs to adjust capacity resources in response to load migration, which is generally a challenge in markets with retail competition. An LSE that is just entering the market may find that their customer base is too uncertain to justify a long-term bilateral contract for capacity, but would still be able to acquire sufficient capacity resources via the centralized market. The challenge of customer migration is addressed by periodically adjusting each LSE's resource obligation in the capacity market to its current customer base.

A centralized capacity market further provides a transparent price signal for all market participants. In solely bilateral markets, only the largest market participants may have a clear understanding of the prices at which potential counterparties would be willing to transact. In contrast, the public posting of capacity market clearing prices allows all market participants, including potential new market entrants, to observe prices and market results over time and make more informed decisions about when to invest and what to contract bilaterally.

Centralized market administration allows market monitors to observe and mitigate potential market power abuses. Any participant that has the ability to exercise market power in the centralized capacity market would also have that ability in the bilateral market, but such abuses would be difficult to discover.

Many restructured power markets rely on out-of-market stop-gap measures to acquire or retain sufficient capacity when the system is short on reserves. This has been used as an option of last

---

[154] Note that locational capacity needs in import constrained zones exist irrespective of market design. Sole reliance on bilateral markets does not eliminate the fact that reserve requirements may need to be imposed on individual import-constrained zone within the regional power market.

[155] PJM (2008), p. 41.

resort to ensure reliability. These out-of-market contracts, however, are problematic because they do not provide, and often interfere with providing, useful market signals for attracting new capacity. A centralized capacity market, particularly one with effective locational requirements and prices, provides an in-market physical backstop for the system operator to enforce LSEs' resource requirements. Because the resource responsibility of each LSE is known, any residual amount not purchased by the LSE bilaterally can be procured via mandatory participation in the residual capacity market. Further, instead of the costs of these out-of-market contracts being paid by all LSEs, they would be paid only by the deficient LSE. Thus, while localized capacity markets may not entirely eliminate out-of-market solutions to ensuring reliability, they will generally significantly reduce the need for such solutions and avoid the distortions and inefficiencies they create in power markets.

A clear disadvantage of centralized capacity markets is the complexity they add to the overall power market design. This complexity creates a greater risk of design deficiencies that can impose regulatory risks on market participants. This added complexity may not be justified in markets with mostly traditionally-regulated utilities in which resource adequacy is already ensured and enforced through the regulatory process, such as traditional resource planning processes. In power markets with retail competition, small LSEs, and unregulated generation companies, however, the disadvantage of added complexity likely will be more than offset by the discussed advantages. In addition, the overall complexity associated with capacity markets may not be greater than the complexity of out-of-market reliability backstop mechanisms required in the absence of capacity markets.

Another disadvantage of capacity markets with reserve requirements that are verified and enforced only days or months before the delivery period is that the short lead time can make it impossible to address capacity deficiencies at almost any price. As discussed above, this leads to a steep offer curve, which leads to higher price volatility and higher risk of market power abuses. This inability to do much about identified capacity deficiencies on short notice has led to the introduction of *forward* reserve requirement, which is discussed in Section VIII (without a centralized capacity market) and Section IX (with a centralized capacity market). Table 5 summarizes the advantages and disadvantages of adding centralized capacity markets to energy markets with reserve requirements.

Table 5
Advantages and Disadvantages of Adding Centralized Capacity Markets to Energy Markets with Reserve Requirements

Advantages	***Disadvantages***
• Centralized capacity markets provide a transparent price and standardized capacity product for both buyers and sellers and help facilitate efficient bilateral transactions, including long-term contracts. • Small LSEs are able to fulfill their resource requirements with better information and at lower transaction costs. • Supports retail competition by facilitating capacity transactions of small LSEs and allowing adjustments to reflect load migration across LSEs. • Capacity market prices are determined through market forces, rather than solely through administrative judgment. • Provides market monitors with the information necessary to monitor and mitigate market power. • Creates an in-market mechanism for the market operator to acquire necessary resources on behalf of any deficient LSEs, thereby reducing the need for out-of-market contracts. • Allows including a downward-sloping demand curve to help stabilize capacity prices, reflect the value of incremental resources, and reduce incentives to exercise market power. • Capacity markets with locational capacity requirements improve the pricing and deliverability of capacity in transmission-constrained systems.	• Small changes in fixed resource requirements can result in large changes in capacity market prices. • Lack of forward resource requirement leaves little time (only days or months) to respond to identified capacity deficiencies and can lead to price volatility and market power concerns because it may be impossible, at almost any price, to supply new capacity on short notice. • Increased complexity of market design increases risk of initial design flaws. Changing capacity market rules imposes regulatory risk. • Can create political backlash because clearly visible capacity prices draw attention to the high cost of ensuring reliability at current target reserve margins (a cost that may be even higher but less visible in market designs that rely solely on bilateral arrangements to satisfy reserve requirements).

VIII. ENERGY MARKETS WITH FORWARD RESERVE REQUIREMENTS BUT WITHOUT CENTRALIZED CAPACITY MARKETS

A. ADDITIONAL MARKET DESIGN ELEMENTS AND CONSIDERATIONS

To address concerns associated with short lead times between when capacity deficiencies are detected and when additional resources need to be available to assure reliability, forward reserve requirements have been introduced in several power markets. Forward reserve requirements require that LSEs procure sufficient resources one or several years ahead of the delivery year. Both forward commitments of resources and the actual availability of the committed resources during the delivery period are verified by the system operator.

By requiring resource commitments sufficiently prior to delivery, enough time is left for either market participants or the system operator to procure additional resources if deficiencies are detected. With a lead-time of one or several years, this also gives suppliers enough time to modify their resource development plans, for example, by bringing back online mothballed plants, by making the capital investments necessary to defer retirement of other plants, by speeding up the development of a new power plant, or by developing additional demand response resources. Importantly, this gives the market operator enough time to purchase physical resources as a backstop mechanism on behalf of deficient LSEs, rather than risking that—other than imposing financial penalties on deficient LSEs—nothing can be done to rectify a detected reliability problem.

Suppliers' ability to bring additional resources on line for a multi-year forward delivery period also "flattens" the supply curve discussed in the context of Figure 3 of Section VII. This increases competition, reduces the risk of market power abuses, and reduces price volatility, including the risk of bipolar capacity prices. This reduction of price volatility is, however, at least partially offset by the risks associated with forward commitments. Unexpected changes in permitting, equipment and construction markets, and transmission interconnection processes can delay project development schedules or increase project development costs between the time when the forward commitment needs to be made and when the capacity needs to be operational. As a result, the "optimal" forward commitment period will need to consider not only the time to develop capacity resources but also the uncertainty that longer forward commitment periods impose on suppliers.

It is not necessary for a forward commitment period to cover the entire power plant development time frame. It is important, however, that a sufficient number of resources are able to respond and modify their development timelines within the given timeframe. This may mean, for example, that the forward commitment period should be long enough to cover, for a sufficiently large proportion of potentially available resources, the period between when major irreversible financial commitments are made during development and when the plant becomes operational. Shorter forward commitment periods may be warranted in markets with a substantial amount of resources with short lead times, such as demand-side resource, gas turbine plants, or existing resources that are able to modify pending retirement plans. The differences in deployment time frames across various types of resources may also warrant staged forward commitment schedules under which varying amounts of resources are committed incrementally over a multi-year period.

As is the case with reserve requirements that are imposed only going into the delivery period, imposing reserve requirement on a forward basis does not require implementation of a centralized RTO-operated capacity market. Rather, it is possible to allow market participants to rely solely on self supply and bilateral transactions to satisfy the forward reserve requirement. Imposing such a forward reserve requirement, however, will create a bilateral forward capacity market, impose on load serving entities the full cost of the associated capacity resources, and provide suppliers with capacity revenues in addition to energy market and ancillary service revenues.

B. Experience with Forward Reserve Requirements without Centralized Capacity markets

California's forward resource adequacy requirement ("RAR") is the primary example of an explicit forward reserve requirement that, at this time, has not been combined with a centralized forward capacity market. This resource adequacy standard imposes a 1-year forward requirement. The system was implemented in 2006 by the California Public Utility Commission ("CPUC") and was expanded in 2007 to include a locational resource adequacy standard that applies to several import-constrained load zones within the CAISO.[156]

The Local Resource Adequacy Requirement ("LRAR") helps ensure the availability and deliverability of contracted capacity within the constrained areas, called Local Capacity Area ("LCAs"). The implementation of these localized resource requirements greatly reduced the out-of-market RMR contracts needed to meet local reliability requirements; with RMR contracts dropping from covering 9,300 MW in 2006 to 3,300 MW in 2007.[157]

Enforcement of the forward locational resource adequacy standard includes both a financial penalty for deficiency and a physical backstop. If an LSE is found to be deficient in acquiring resources and has not been granted a waiver, then it will be fined $40/kW-year and will be further charged with the cost of procuring replacement capacity on its behalf.[158] The combination of the financial penalty and the physical backstop not only provide a financial incentive for the LSEs to acquire the necessary capacity, but also protect the system against system-wide deficiencies. Backstop procurement is undertaken through the CAISO's Interim Capacity Procurement Mechanism ("ICPM"),[159] which is used both when LSEs fail to meet their forward resource adequacy requirements and in response to a "significant event" causing a material change in system resource adequacy requirements.

In order to make this bilateral forward capacity market design more efficient, the CAISO is currently in the process of revising its tariff to create a standard capacity product ("SCP") that will allow the system operator to assign clearly-defined equivalent capacity values to various

---

[156] CAISO (2008), p. 5.

[157] CAISO (2008), pp. 2, 4.

[158] The size of the penalty was based on 100 percent of the CAISO's standard price estimate for capacity procurement at the time (see CPUC (2006), p. 4, 65-69).

[159] For the FERC order describing and approving ICPM, see FERC (2008b).

resource types.[160] The proposal also assigns penalties for resources with less availability and performance incentives for resources that have greater availability.[161] Retailers that must purchase capacity resources bilaterally to meet the requirement have supported the effort, stating that the SCP will allow them to reduce the transaction costs associated with bilateral contracting and speed up their contracting processes. The stakeholder process for developing the SCP has, however, encountered significant debate over exemptions to the availability standards for demand response, renewables, existing contracts, and small resources below 1 MW.

The addition of a standardized capacity product to the CAISO resource adequacy requirement was also supported in an opinion by the CAISO's independent Market Surveillance Committee, which evaluated the need and desirability of additional modifications under discussion, including the implementation of a centralized forward capacity market by the California Public Utilities Commission (CPUC).[162] In this report, the surveillance committee recommended against implementation of a centralized capacity market at this time given that (1) there are already significant changes in California market design currently underway; (2) there is little need for additional non-renewable generation through 2020; and (3) retail choice is not currently available to most electricity consumers in California.

The first two points mean that it is possible to delay a decision on significant capacity market redesign until the currently high degree of regulatory uncertainty is reduced, while allowing California to learn from other markets and preserving the option for a redesign at a later date. The third point, the absence of retail competition, means that the vast majority of load continues to be served by CPUC-jurisdictional utilities which are able to recover in regulated retail rates the costs associated with their regulatory obligation to procure the generating capacity needed to meet the State's resource adequacy requirements.[163] The committee, however, also recommended that California's current resource adequacy requirement can be refined, supporting the addition of a standard capacity product and suggesting that a 3 to 4 year forward requirement would be more beneficial than the current one-year requirement.[164]

Other examples of market designs with forward resource requirements combine the forward requirements with RTO-administered capacity markets as discussed in Section IX.

C. ADVANTAGES AND DISADVANTAGES OF ADDING FORWARD RESERVE REQUIREMENTS WITHOUT CENTRALIZED CAPACITY MARKETS

Table 6 summarizes the advantages and disadvantages of specifying and enforcing resource requirements on a forward basis but without implementing a centralized capacity market. A primary advantage of specifying reserve requirements on a forward basis is that inadequate reserves can be discovered with enough time to allow for procurement of needed capacity

---

160 See the draft tariff language for the SCP from CAISO (2009b).
161 *Restructuring Today* (2009b).
162 Wolak *et al*. (2007).
163 Wolak *et al*. (2007), pp. 1-3.
164 Wolak *et al*. (2007), pp. 6-9.

resources. The additional time facilitates entry of and competition from resources that could not be deployed on short notice.

In the absence of a centralized capacity market, exclusive reliance on self provision or bilateral arrangements will tend to reduce liquidity and transparency of the capacity market, which will increase risks and transaction costs, particularly for small LSEs and in a retail access environment with migrating customer loads. As already discussed in Section VI (short-term resource requirements without centralized capacity markets), exclusive reliance on self supply and bilateral transactions also increases the challenges related to the monitoring and mitigation of market power. However, forward procurement requirements will mitigate market power concerns by allowing market participation of more resource options and suppliers. The lack of a centralized capacity market will nevertheless reduce transparency of backstop procurements of capacity by system operators in case of resource deficiencies. These disadvantages can be substantial in markets with retail competition and small LSEs but which may be insignificant in traditionally-regulated markets where resource adequacy is maintained through integrated utilities. They are addressed by market designs in which centralized capacity markets are added to forward reserve requirements as discussed in Section IX.

Table 6
Advantages and Disadvantages of Adding Forward Reserve Requirements without Centralized Capacity Markets

Advantages	*Disadvantages*
• Forward requirement increases ability of suppliers to bring new units online and adjust construction plans to meet the forward needs of their contract partners. • Inadequate reserves are discovered with sufficient time for the system operator to contract for needed resources, increasing the operator's ability to physically and economically remedy the deficiency. • Forward procurement increases competition by facilitating entry from new resources, thereby also mitigating market power. • Capacity value of bilateral contracts restores revenue sufficiency in mitigated energy markets • Forward procurement generally reduces capacity price uncertainty and risk premium on generation investment.	• Limited liquidity and transparency of bilateral capacity market imposes additional costs and uncertainties, particularly in retail access markets with many LSEs and migrating customer demand. • Backstop procurement on behalf of deficient LSEs, may not be transparent. • Long forward commitment periods can increase risks for some resources, in particular for short-lead-time resources such as demand response. • Other disadvantages of bilateral market structure, including less effective market monitoring and higher transaction costs for small LSEs, similar to those already summarized in Table 4.

IX. ENERGY MARKETS WITH FORWARD RESERVE REQUIREMENTS AND CENTRALIZED CAPACITY MARKETS

A. ADDITIONAL MARKET DESIGN ELEMENTS AND CONSIDERATIONS

Forward capacity markets combine the forward reserve requirement discussed in Section VIII with a centralized capacity market administered by the system operator. Much of the discussion related to the addition of centralized capacity market to a short-term reserve requirement as discussed in Section VII also applies to adding centralized capacity markets to forward reserve requirements. In other words, the forward capacity market design adds a forward reserve obligation to the elements of centralized capacity market discussed in Section VII.

In this market design, the resource requirements of LSEs must be met via self-supply, bilateral contracts, or through purchases from the RTO-administered capacity auctions on a one or multi-year forward basis. Intermediate auctions in which market participants can purchase or sell capacity are generally provided to adjust reserve obligations prior to the delivery period. These intermediate capacity auctions are referred to as incremental, balancing, or reconfiguration auctions. The intermediate auctions allow supplier and LSE to adjust their obligations in response to any updated load forecasts and changes in expected resource availability. The forward capacity market model also allows market participants to satisfy their forward capacity obligations through self-supply from physically-owned or bilaterally-contracted resources.

The main design elements of forward capacity auctions differ with respect to how the quantity of forward reserve requirements is determined (*e.g.*, vertical versus sloped demand curve), the specific auction format used to clear the market (*e.g.*, clearing bid-based supply curves against demand or use of descending clock auctions), and the length of the forward commitment period (*e.g.*, 3 to 5 years).

The use of centralized capacity markets provides transparent market information regarding quantities and market prices, which improves both buyers' and suppliers' ability to refine and adjust their investment decisions. As discussed in Section VIII, a multi-year forward commitment period has the advantage that more new capacity resources can be brought online in time for delivery. Generators in various stages of the development process, those formerly mothballed and those that might otherwise retire, all have more time to respond to a commitment to be available during the delivery period. This "flattens" the supply curve, increases competition, mitigates market power, and enhances price stability and predictability relative to short-term capacity markets.

Further reductions in price uncertainty can be achieved by combining the forward supply curve with a sloped demand curve rather than a fixed reserve requirement. This additional reduction in price volatility further reduces investment risks and the risk premium required by investors. A sloped downward curve also recognizes the additional reliability value that procurement of reserves in excess of the target amount will provide. In addition, the sloped downward curve further mitigates both buyer and supplier market power, although that will be less important due to the fact that a flattened-out forward supply curve already helps mitigate market power concerns.

Like in all other market designs with administrative capacity payments or reserve requirements, the capacity revenues received during the delivery period (assuming the resource is actually available) restore revenue sufficiency (*i.e.*, solves the missing money problem) in mitigated energy markets with low price caps.

B. Experience with Centralized Forward Capacity Markets

In addition to PJM's Reliability Pricing Model, which we summarized at the outset of this report, other examples of forward capacity market designs include the ISO-NE's Forward Capacity Market ("FCM") and the forward capacity market in Brazil. The NYISO is currently in the process of evaluating whether to replace its short-term locational capacity market design with a three to four year forward capacity market.[165]

1. Basic Forward Capacity Market Design Features

Both PJM and ISO-NE have implemented a three-year forward commitment period. LSEs can choose to self-supply capacity through physical ownership or bilateral contracts. The residual supply needed to satisfy system-wide procurement targets is procured by the RTOs in an initial auction three years in advance of the delivery year. While PJM clears a bid-based supply curve against a sloped demand curve (the variable resource requirement or VRR curve), ISO-NE uses a descending clock auction in which market prices are reduced until the quantity of available supply bids matches its fixed forward reserve requirement. This ISO-NE auction process is equivalent to clearing a bid-based supply curve against a vertical demand curve, although the multi-round nature of the auction process provides additional efficiencies and pricing transparency, including during the auction process itself. Both PJM and ISO-NE forward capacity market designs provide additional forward price stability by allowing certain suppliers of new capacity to "lock in" capacity prices for three to five years.[166] In addition, both PJM and ISO-NE conduct incremental auctions in which participants or the ISO can adjust capacity commitments in response to new circumstances, including expected increases in load or changes in the expected availability of previously-committed resources.

The forward capacity market design in Brazil differs markedly from the PJM and ISO-NE designs. Brazil engaged in a series of market restructuring activities during the 1990s, which ultimately failed to attract the necessary generation and transmission investments. After facing shortages and rationing during 2001 and 2002, the market was redesigned into a forward capacity market model, which held its first auctions in 2004.[167] The stability and predictability provided by forward capacity markets is eminently attractive in Brazil, where the hydro-dominated system can yield energy market prices that average only \$4/MWh several years in a row, while in a year of water scarcity average annual prices can rise to \$60/MWh.[168] In the Brazilian forward capacity market design, the distribution companies each declare their own energy demand and conduct a joint auction for capacity procurement.[169] The distribution companies then individually engage in bilateral contracts with each of the winning suppliers in proportion to their own declared demand. Each year there are several auctions with various forward durations, ranging from durations of 15 to 30 years beginning 5 years prior to delivery

---

165 For example, see Newell *et al*. (2009).

166 In ISO-NE, new capacity has the option of locking in a commitment period of one to five years (with the supplier deciding the number of years desired), see FERC (2006a) p. 7; FERC (2006b), A1 p. 6. In PJM the "New Entry Pricing" provisions, certain new suppliers in transmission constrained zones can lock in market clearing prices for up to three years. See PJM (2008), pp. 61-62.

167 de Oliviera and de Moraes Marreco (2005), p. 2361 and Barroso, *et al*. (2006), p.1.

168 Rudnick, *et al*. (2005), p. 52.

169 Brazil's distribution companies are regulated monopolies that do not face retail competition, although these companies have been privatized via public auctions, see Tankha (2009), pp. 75-76.

for new capacity, to contracts with durations as short as 5 years and beginning only 1 year prior to delivery for existing capacity.[170]

2. Incorporating Demand Response into Capacity Markets

Finding a way to include demand-side options in the provision of capacity resources involves a set of issues that are not unique to the forward capacity market design. While integration of demand response resources into wholesale market design has been dealt with most directly in the context of PJM's and ISO-NE's forward capacity markets, many of the same issues presented in this section also apply to all other market designs.[171]

Both the PJM RPM and the ISO-NE FCM allow demand-side resources to offer "supply" into the capacity market.[172] These services can take the form of contracts with customers on interruptible rates or direct load control. Participation of demand–side resources has grown dramatically in both forward capacity markets. In PJM, for example, participation of demand-side resources increased from less than 2,000 MW and 4,000 MW for the 2007/08 and 2008/09 planning years, to 9,848 MW offered (7,047 MW cleared) for the 2012/2013 planning year, much of which was focused on import-constrained regions.[173] Similarly, in ISO-NE demand-side resources of approximately 2,500 MW cleared it its first full FCM auction in February 2008 and approximately 2,900 MW in its second auction in December 20008.[174]

Integration of DR resources into capacity market designs faces a number of challenges, including how to monitor and verify availability of the committed resources and how to adjust load forecasts for any such commitments without understating or double-counting their impacts.[175] Where such integration is accomplished successfully, however, capacity markets can significantly reduce market barriers, improve price transparency, and reduce investment risk for demand-side resources. This facilitation of market participation by demand-side resource can also greatly improve pricing and market efficiency in the underlying wholesale energy markets, including during scarcity periods.

Technical requirements on deliverability and reliability of demand-side capacity resources are intended to be equivalent to the requirements placed on generating units that supply capacity resources. Some participants are nevertheless concerned that demand-side resources are compensated at the same level as firm generation, pointing out, for example, that a DR resource committed to a single year could quickly disappear if customers became dissatisfied with their interruptible rates. In contrast, a new generating unit exists for many years. The counterargument offered in this discussion is that supply-side options also could disappear through retirements or export commitments. Discussions addressing these concerns have been

170 Barroso, *et al*. (2006), pp. 5-6.
171 For example, see Earle *et al*. (2008), a Brattle Group report analyzing options to improve integration of demand response resources into the Midwest ISO's wholesale energy market.
172 PJM (2008), pp. 33-40; ISO-NE (2008a), p. 12.
173 PJM (2009b), pp. 5-10.
174 These numbers include both existing and new demand-side resources. ISO-NE (2008b).
175 Pfeifenberger, *et al*. (2008), pp. 115-121.

particularly active in ISO-NE, because demand response accounted for roughly two thirds of all incremental resources offered in the February 2008 auction.[176] This makes the reliability of the ISO-NE system critically dependent on the actual and continued availability of demand-side resources. If customers withdraw their demand side commitments after interruptions and curtailment are experienced, the system could end up short on reserves. However, if these resources *do* deliver as promised, then the entire framework of ensuring reliability and supplying peak load will have been changed, with potentially significant reductions in system-wide capacity costs.

3. Incorporating Scarcity Pricing in Capacity Market Designs

The interaction between scarcity pricing and capacity payments is another important design consideration. We discussed the need for scarcity pricing in Section IV within the context of an energy-only market. However, the efficiency benefits of proper scarcity pricing also apply to all other market designs, including those with forward capacity markets. Scarcity prices are high energy prices that should occur when the system is deficient in operating reserves. While scarcity conditions will necessarily occur less frequently in markets with capacity payments or reserve requirements than in energy-only markets, they cannot be avoided altogether. Appropriately high prices during such scarcity conditions create proper incentives for customers to reduce consumption, for local suppliers to increase their availability and output, and for additional imports from external market areas. Proper scarcity prices consequently provide valuable signals to market participants that improve overall market performance. For example, IPPs and other stakeholders in the Midwest ISO (which has only a voluntary short-term capacity market), recommended improving scarcity pricing signals, noting that out-of-market administrative actions during scarcity events have blunted efficient price signals and undermined demand response.[177]

The challenge of adding scarcity pricing to capacity markets is to avoid the possibility that generators get compensated twice for their capacity. In PJM, for example, a scarcity pricing mechanism has been proposed in addition to the existing RPM capacity market to improve real-time energy and ancillary service prices. This would improve the locational incentives for peak load reductions and attract additional supply (including imports) not already committed through the RPM. To avoid double-payment, it is proposed that resources committed that receive capacity market revenue under RPM would not receive scarcity revenues. Implementation of this scarcity pricing and reconciliation mechanism is a matter of current discussion within PJM's Capacity Market Evolution Committee (CMEC).[178]

ISO-NE also has a scarcity pricing mechanism, which is based on the reserve constraint penalty factor (RCPF). The RCPF is the maximum cost that can be imposed in order to meet reserve requirements; if the cost of meeting the reserve requirement exceeds the RCFP, reserve

---

[176] New resources (as opposed to existing resources) that cleared the market were made up of 1,188 MW of demand response resources and 626 MW of other resources. ISO-NE (2008a), p. 12.
[177] *Restructuring Today* (2009a).
[178] PJM (2009c). Monitoring Analytics (2009), pp. 2.116-2.118.

requirements will not be met.[179] ISO-NE reconciles the scarcity pricing mechanism with the forward capacity market by subtracting Peak Energy Rents ("PER") from the monthly capacity payments that could otherwise be received by suppliers through the real-time market.[180] The PER deduction is not explicitly tied to scarcity pricing revenues, but rather applies to all energy market revenues received above the strike price.

NYISO, which has a centralized capacity market without forward reserve requirements, enacts scarcity pricing based on the high bid prices of emergency and special case resources when they must be called.[181] The CAISO, with a forward resource adequacy standard but without a centralized capacity market, has just proposed a revision to their scarcity pricing mechanism which would include a demand curve for spinning and non-spinning reserves, which would allow prices for reserves and energy to increase automatically, even above the price cap, when there is not enough supply in the market to meet the target level of regulation and operating reserves.[182]

C. Advantages and Disadvantages of Adding Centralized Capacity Markets to Forward Reserve Requirements

Forward capacity markets allow suppliers to make investment decisions under conditions of greater certainty and transparency. For example, a previous *Brattle Group* study found that approximately 14,500 MW of PJM capacity commitments were added or retained through 2011-12 since RPM was implemented in early 2007.[183]

Although a forward capacity market does not provide complete certainty about the price that a capacity resource will fetch over its lifespan, it reduces that uncertainty. It provides certainty on the revenues that will be received in the first year(s) of operation and makes capacity revenues in subsequent years more predictable, especially if load growth and retirements are anticipated to continue to create a need for capacity additions over time. The annual pricing format of forward capacity market also provides incentives for LSEs to make longer-term forward bilateral commitments as a mean to hedge capacity costs relative to auction clearing prices.

Table 7 summarizes the advantages and disadvantages of forward capacity markets as discussed above. The combination of forward reserve requirements and centralized capacity markets provides transparent market price signals, addresses load migration challenges, facilitates participation by small LSEs, allows additional resources to compete on a multi-year forward bases, reduces price volatility and the associated investment risk premium, mitigate market power and facilitates market monitoring. Forward capacity markets can readily integrate

---

179 The RCFP is set at different administratively-determined levels for each type of reserve: $50/MWh for both ten-minute spinning reserves and local 30-minute reserves, $100/MWh for system-wide 30-minute reserves, and $850/MWh for system-wide 10-minute reserves. See p. 35, ISO-NE (2009b).

180 The PER deducted is based on the profits that a "proxy unit" with a heat rate of 22,000 BTU/kWh using the marginal fuel type would have made in the real-time energy market. FERC (2006b), A1 pp. 36-37.

181 NYISO (2005). pp. 331.01.02-07

182 *Megawatt Daily* (2009c). The current CAISO mechanism sets the real-time price at the $500/MWh bid cap if real-time offers are insufficient.

183 Pfeifenberger, *et al.* (2008). pp. 33-34.

demand response resources and can impose locational forward reserve requirements to address resource adequacy in transmission-constrained areas. These advantages are particularly pronounced in restructured markets with retail competition.

Forward commitment periods that are too long, however, will impose added risks on suppliers, could create market barriers for shorter lead-time resources, and increase the risk that suppliers default on their forward obligations. As is the case with all market designs that are based on pre-determined reserve requirements, the targeted reserve margins may not appropriately balance the value of increased reliability against the cost of providing the additional reserves. Centralized capacity markets can also create political backlash because clearly visible capacity prices draw attention to the high cost of ensuring reliability at current target reserve margins—although this cost may be even higher but less visible in market designs that rely solely on bilateral arrangements to satisfy reserve requirements. Locked-in forward commitment could appear unnecessarily high cost after changes in market conditions, such as economy-related declines in load, leads to excess capacity and reduced resource needs. Finally, the still limited experience and added complexity of forward capacity markets imposes high implementation costs for both RTOs and market participants with a considerable risk of costly design flaws and market inefficiencies.

Table 7
Advantages and Disadvantages of Adding Centralized Capacity Markets To Energy Markets with Forward Reserve Requirements

Advantages	*Disadvantages*
• Provides transparent, in-market mechanism for the system operator to acquire necessary resources on behalf of deficient LSEs. • Multi-year forward commitment period allows additional supply to compete in market, thereby reducing price volatility. • Supports retail competition by facilitating capacity transactions to address load migration and assist small LSEs. • Allows incorporation of demand-response in the forward capacity market design, which increases competition and helps reduce system-wide cost of ensuring reliability. • Allows for locational forward capacity requirements, which further improve pricing and deliverability of capacity in transmission-constrained areas. • Facilitates monitoring and mitigation of market power.	• Added complexity of market design imposes high implementation costs for RTO and market participants. • Complex market design also carries risks of initial design flaws and inefficiencies. • Lengthy forward commitment periods can increase supplier risks. Also increases risk that suppliers default on their forward obligations. • Can create political backlash because clearly visible capacity prices draw attention to the high cost of ensuring reliability at current target reserve margins; locked-in forward commitment could appear unnecessarily high cost after change in market conditions reduces resource needs.

X. CONCLUSIONS

This review of alternative energy and capacity market designs from a resource adequacy perspective shows that each design has advantages in certain market and regulatory environments. For example, simple energy markets with reserve requirements but without centralized capacity markets have excellent track records in power markets with mostly traditionally regulated, vertically-integrated utilities, where reliability is ensured through forward-looking regulated planning processes. However, maintaining adequate reserve margins under this design would be significantly more difficult in restructured power markets, such as PJM, with deregulated generation, retail competition, many (often small) LSEs, and migrating customer loads.

In such a restructured market environment, implementing reserve requirements on a multi-year forward basis offers improved reliability because inadequate reserves are discovered in time for the system operator to contract for capacity on behalf of the deficient LSEs. This also increases competition from suppliers who can bring new resources on line and adjust construction plans to meet the forward resource needs. The resulting "flattened" supply curve additionally reduces price volatility. Imposing reserve requirements creates a bilateral capacity market in which both demand- and supply-side resources, as well as both existing and new capacity resources are equally valuable. This eliminates the "missing money" problem of the energy-only market design and substantially reduces or eliminates the need to retain or attract resources through out-of-market mechanisms, such as RMR contracts.

Compared to relying solely on a bilateral capacity market structure to meet forward reserve requirements, adding a centralized capacity market offers increased price transparency, continued bilateral contracting options, a backstop procurement mechanism for the system operator, reduced transactions costs, and improved market monitoring. These benefits are particularly valuable in restructured markets with retail competition, many (often small) LSEs, and migrating customer loads.

A concern associated with implementing forward capacity markets is their complexity, which imposes considerable implementation costs and risk of unintended design flaws on both system operators and market participants. Transparent forward capacity markets also carry political risk because clearly visible capacity prices can draw attention to the high cost of the imposed reserve requirements. Eliminating the centralized capacity market would not reduce (and could even increase) the costs of meeting reliability requirements, although these costs would be less visible under a purely bilateral design. To the extent that resource ownership or long-term bilateral contracts offer a lower-cost solution to maintain LSEs' resource requirements and system-wide reliability standards, these self-provision options already exist under RPM and other centralized capacity market designs.

Both PJM and ISO-NE have implemented forward capacity markets to meet pre-determined reserve requirements in a market environment where forward capacity markets are most valuable: restructured power markets with deregulated generation, transmission constraints, retail competition, and many (often small) LSEs with migrating customer loads. The complexity of the market designs did impose considerable implementation costs and risk of design flaws. However, most of the implementation costs have now been incurred and many concerns about initial design parameters have already been addressed. To date, the experience in both markets shows that the forward capacity market design has been able to meet resource adequacy requirements by attracting and retaining capacity resources, including a substantial amount of demand-side resources.

Imposing reserve requirements is the most direct but not the only option to assure resource adequacy and pre-determined reliability standards. We also discussed two market designs that do not have pre-determined reliability standards: pure energy-only markets and energy-only markets with administratively-determined capacity payments. Pure energy-only markets rely on customers to choose their desired level of reliability through the market mechanisms of interruptible rates and demand response. Capacity investments and resource adequacy remain uncertain because they are solely a function of energy market prices, including occasional price spikes that appropriately reflect customers' value of lost load during scarcity conditions.

However, because demand response and mechanisms to differentiate reliability among customers have not yet developed sufficiently to bring supply and demand into equilibrium during shortage periods, energy-only markets need to rely on out-of-market mechanisms and administratively-determined scarcity pricing to provide proper investment signals. Further, price spikes must be sufficiently frequent and sufficiently high to attract investments. However, because price spikes often are not acceptable to policy makers, prices in many energy markets are mitigated or artificially depressed through out-of-market mechanism to ensure reliability. The resulting revenue deficiency leads to an inability to attract and retain sufficient resources to maintain the reliability standards desired by policy makers in the absence of out-of-market payments, including administratively-determined capacity payments. While market designs that rely on such payment options are often attractive initially because payments can be limited only to a subset of resources needed to maintain reliability in the short-term, this approach introduces significant distortions of market prices that perpetuate the need for such payments. The long-term cost of relying on out-of-market mechanisms or capacity payments that are limited to new and about-to-retire resources would likely be higher due to market distortions and associated short-term and long-term inefficiencies.

BIBLIOGRAPHY

Adib, P., Schubert, E., and Oren, S. (2008). "Chapter 9: Resource Adequacy: Alternate Perspectives and Divergent Paths." In F. P. Sioshansi (Ed.), *Competitive Electricity Markets: Design, Implementation, Performance*. Oxford: Elsevier.

Alberta Market Surveillance Administrator (2009.) *2008 Year in Review*. Retrieved from: http://www.albertamsa.ca/files/2008_Year_in_Review_amended_140509(1).pdf.

Australian Energy Market Commission (2005). *Draft Rule Determination: Reliability Safety Net Extension.*

Australian Energy Market Commission Reliability Panel (2008a). *Package Two. NEM Reliability Settings: VoLL, CPT and Future Reliability Review. Rule Change Proposal.* Retrieved from http://www.aemc.gov.au/Media/docs/Rule%20Change%20Proposal%20To%20The%20AEMC-df27e4cc-10d6-46e7-b9a4-36f559e8914c-0.pdf.

Australian Energy Market Commission Reliability Panel (2008b). *Review of VoLL 2008: Final Determination*. Retrieved from http://www.aemc.gov.au/Media/docs/Reliability%20Panel%20Review%20of%20VoLL%202008%20Final%20Determination-6103ad90-eea6-418c-af02-1d87055d912a-0.pdf.

Australian Energy Market Commission Reliability Panel (2009). *Reliability Settings: VoLL, CPT, and Future Reliability Review.* Retrieved from http://www.aemc.gov.au/electricity.php?r=20090122.095406.

Australian Energy Market Regulator, Reliability Panel (2007). *NEM Reliability Standard – Generation and Bulk Supply*. Retrieved from http://www.aemc.gov.au/pdfs/Reliability%20Panel/NEM%20Reliability%20Standard%20-%20Generation%20and%20Bulk%20Supply%20-%20December%202007.PDF.

Australian Energy Regulator (2007). *State of the Energy Market, 2007.* Retrieved from http://www.aer.gov.au/content/index.phtml/itemId/713232.

Baldick, R. (2009). *Single Clearing Price in Electricity Markets*. Retrieved from http://www.competecoalition.com/files/Baldick%20study.pdf

Barber, J. (2009). "Ontario suspends RFP for two nuclear plants," *Megawatt Daily*, June 30, 2009.

Barroso, L.A., Rosenblatt, J., Guimarães A., Bezerra, B., and Pereira, M.V. (2006). "Auctions of Contracts and Energy Call Options to Ensure Supply Adequacy in the Second Stage of the Brazilian Power Sector Reform." Power Engineering Society General Meeting, 2006, IEEE.

Batlle, C., Vázquez, C., Rivier, M., Pérez-Arriaga, I.J. (2007). "Enhancing Power Supply Adequacy in Spain: Migrating from Capacity Payments to Reliability Options." *Energy Policy*, Vol. 35, pp. 4545-4554.

Batlle, C., Solé, C., Rivier, M. (2008). "A New Security of Supply Mechanism for the Iberian Market." *The Electricity Journal*, Vol. 21, Issue 2.

Bernanke, B. S. (1983). "Irreversibility, Uncertainty, and Cyclical Investment," *The Quarterly Journal of Economics*, *98* (1).

Bidwell, M. (2005). "Reliability Options: A Market-Oriented Approach to Long-Term Adequacy," *The Electricity Journal*, *18* (5).

Bowring, J.E. (2008). "The Evolution of PJM's Capacity Market." In F.P. Sioshansi (Ed.), *Competitive Electricity Markets: Design, Implementation, Performance* (pp. 363-386). Oxford: Elsevier.

Bushnell, J. (2005). "Electricity Resource Adequacy: Matching Policies and Goals." Center for the Study of Energy Markets.

Botterud, A., and Doorman, G. (2008). "Generation Investment and Capacity Adequacy in Electricity Markets," International Association for Energy Economics. Second Quarter 2008. Retrieved from http://www.iaee.org/documents/newsletterarticles/208Botterud.pdf.

Caballero, R. J., and Pindyck, R.S. (1996). "Uncertainty, Investment, and Industry Evolution." *International Economic Review*, *37* (3).

California Independent System Operator (2009a). *Market Issues and Performance: 2008 Annual Report.* Retrieved from http://www.caiso.com/2390/2390818c3bc40.html.

California Independent System Operator (2009b). *Standard Capacity Product Amendment Filing: 4th Replacement CAISO Tariff.* Retrieved from http://www.caiso.com/238b/238baf4f3aa0.pdf

California Independent System Operator (2008). *Market Issues & Performance: 2007 Annual Report.* Retrieved from http://www.caiso.com/1f9c/1f9c8b49e9f0.pdf.

California Independent System Operator (2005). *2004 Annual Report on Market Issues and Performance.* Retrieved from http://www.caiso.com/docs/2005/04/28/2005042814343415812.html.

California Public Utility Commission (2006). *Opinion on Local Resource Adequacy Requirements* (Decision 06-06-064). Sacramento, CA: Retrieved from http://docs.cpuc.ca.gov/word_pdf/FINAL_DECISION/57644.pdf.

Chao, H., and Wilson. R. (2004). "Resource Adequacy and Market Power Mitigation via Option Contracts." Electric Power Research Institute and Stanford University.

Chuang, A. S., and Wu, F. (2000). "Capacity Payments and the Pricing of Reliability in Competitive Generation Markets," *Proceedings of the 33rd Hawaii International Conference on System Science.*

Cramton, P., and Stoft, S. (2006). "The Convergence of Market Designs for Adequate Generating Capacity with Special Attention to the CAISO's Resource Adequacy Problem." Retrieved from http://web.mit.edu/ceepr/www/publications/workingpapers/2006-007.pdf.

Earle, R.L., Newell, S.A., Faruqui, A., Hajos, A., and Hledik, R. (2008). "Fostering Economic Demand Response in the Midwest ISO," *The Brattle Group*, December 30, 2008.

The Economist (2009). "Dark Days Ahead: A shortage of power-generation capacity could lead to blackouts across Britain—and a dangerous reliance on foreign gas," August 6, 2009.

Edwell, S. (2008). "AEMC Review of the Effectiveness of South Australian Retail Markets," Retrieved from http://www.aemc.gov.au/pdfs/reviews/Review%20of%20the%20Effectiveness%20of%20Competition%20in%20the%20Electricity%20and%20Gas%20Retail%20Markets%20-%20South%20Australia/first%20draft%20submissions/000AER.PDF.

Electric Reliability Council of Texas (2008). "ERCOT News – February 21, 2008." Retrieved from http://www.ercot.com/news/press_releases/2008/nr-02-21-08.

Electric Reliability Council of Texas (2004) *Report on Existing and Potential Electric System Constraints and Needs within the ERCOT Region.* Retrieved from http://www.ercot.com/content/news/presentations/2004/tr-reportconstraints2004.doc.

Electric Reliability Council of Texas (2002). *ERCOT Reliability Must Run.* Retrieved from http://www.ercot.com/content/meetings/wms/keydocs/2002/1028/WMS10282002-7.ppt.

ELEXON (2007). *ELEXON Limited Report and Financial Statements.* Retrieved from http://www.elexon.co.uk/documents/Publications/Corporate_and_Financial_-_ELEXON_Limited_Annual_Report_and_Financial_Statements/Annual_Report_Financial_Statements.pdf.

Energy Information Administration (2000). *The Changing Structure of the Electric Power Industry 2000: An Update.* Retrieved from http://www.eia.doe.gov/cneaf/electricity/chg_stru_update/toc.html.

Energy Information Administration (2009). "Status of Electricity Restructuring by State." Retrieved from http://www.eia.doe.gov/cneaf/electricity/page/restructuring/restructure_elect.html.

Fabra, N. von der Fehr, N., and Harbord D. (2004) Designing Electricity Auctions. Retrieved from http://www.ucei.berkeley.edu/PDF/csemwp122.pdf

Federal Energy Regulatory Commission (2009a). *Order Accepting Tariff Provisions in Part, Rejecting Tariff Provisions in Part, Accepting Report, and Requiring Compliance Filings*, (Docket Nos. ER05-1410 *et al*.), March 26, 2009.

Federal Energy Regulatory Commission (2009b). "Electric Power Markets: National Overview," Retrieved from http://www.ferc.gov/market-oversight/mkt-electric/overview.asp.

Federal Energy Regulatory Commission (2008a). *Order Accepting Tariff Filing Subject to Modification* (Docket Nos. ER08-556-000; ER06-615-020). Retrieved from http://www.caiso.com/2066/2066e4883c7b0.pdf).

Federal Energy Regulatory Commission (2008b). *Order Conditionally Accepting Financial Settlement Provisions* (Docket No. ER08-394-003.). Retrieved from http://www.ferc.gov/whats-new/comm-meet/2008/101608/E-8.pdf.

Federal Energy Regulatory Commission (2006a). *Order Accepting Proposed Settlement Agreement.* (Docket Nos. ER03-563-030; ER03-563-055.) Retrieved from http://www.iso-ne.com/regulatory/ferc/orders/2006/jun/er03-563-030_er03-563-055_6-16-06.pdf

Federal Energy Regulatory Commission (2006b). *Explanatory Statement of the Settling Parties in Support of Settlement Agreement and Request for Expedited Consideration.* (Docket Nos. ER03-563-000; ER03-563-030; ER03-563-055).

Federal Reserve Board (2009). *Federal Reserve Statistical Release: H10 Foreign Exchange Rates (Weekly).* Retrieved from http://www.federalreserve.gov/releases/h10/20090526.

Federico, G. and Rahman, D. (2001) "Bidding in an Electricity Pay-as-Bid Auction." Retrieved from http://www.nuff.ox.ac.uk/Economics/papers/2001/w5/Federico-Rahman.pdf

Federico, G., and Vives, X. (2008). *Competition and Regulation in the Spanish Gas and Electricity Markets.* Retrieved from http://www.orkestra.deusto.es/images/pdf/competition%20and%20regulation%20in%20the%20spanish%20gas%20and%20electricity%20markets.pdf.

Fox-Penner, P. (2003). *Prepared Rebuttal Testimony of Dr. Peter Fox-Penner on Behalf of the California Parties* (Docket Nos. EL00-95 and EL00-98 [including sub dockets]). Retrieved from http://ag.ca.gov/antitrust/pdf/ferc/Fox-Penner.pdf.

Gallagher, S. (2005). *California Public Utilities Commission Capacity Markets White Paper*. Retrieved from http://stoft.com/metaPage/lib/CA-PUC-2005-08-ICAP-white-paper.pdf.

Hardin, G. (1968). "The Tragedy of the Commons," *Science, 162*. Retrieved from http://www.sciencemag.org/cgi/reprint/162/3859/1243.pdf.

Harvey, S. and Travers, S (2008). *Market Incentives for Generation Investment*. Retrieved from http://www.ieso.ca/imoweb/pubs/marketreports/omo/2009/generation_investment.pdf

Heffner, G., Goldman,C., Kintner-Meyer, M., and Kirby, B. (2007). *Loads Providing Ancillary Services: Review of International Experience.* Retrieved from http://certs.lbl.gov/pdf/62701-app.pdf.

Hesmondhalgh, S. (2003). "NETA the Blueprint for Wholesale Electricity Trading Arrangements of the Future?" *IEEE Transactions on Power Systems, 18* (2).

Hinkel, R. O. (2001). "Market Based generation Adequacy: How Do We Achieve It? An Independent System Operator's Perspective." IEEE.

Hirsch, R. F. (2002). *Power Loss: The Origins of Deregulation and Restructuring in the American Electric Utility System.* Cambridge: The MIT Press.

Hirst, E., and Hadley, S. (1999). "Generation Adequacy: Who Decides?" *The Electricity Journal.*

Hobbs, B. F. (2005). *Affidavit of Benjamin F. Hobbs on Behalf of PJM Interconnection, LLC* (FERC Docket ER05-___-000). Retrieved from http://stoft.com/metaPage/lib/Hobbs-2005-08-PJM-ICAP-affidavit.pdf.

Hogan, W. W. (2005). "On an 'Energy Only Electricity Market Design for Resource Adequacy.'" Retrieved from http://www.hks.harvard.edu/hepg/Papers/Hogan_Energy_Only_092305.pdf.

Huber, J., Ernesto, R., Rodrigo, M., Espinoza, V., and Palma-Behnke, R. (2006). "Hydrothermal Coordination and Capacity Payment Schemes in Chile: Current Discussion and Future Challenges." Retrieved from http://www.business.auckland.ac.nz/Portals/4/Research/ResearchCentresGroups/Energy%20centre/Espinoza-Huber-Hydrothermal%20Coordination%20and%20Capacity%20Payment%20Schemes%20in%20Chile_final.pdf

Independent Market Operator of Western Australia (2006). "Overview - An Introduction to the Reserve Capacity Mechanism." Retrieved from http://www.imowa.com.au/reserve_capacity_overview.htm.

International Energy Agency (2005). *Lessons from Liberalised Electricity Markets.* Retrieved from http://www.iea.org/textbase/nppdf/free/2005/lessons2005.pdf.

ISO-New England (2008a). "2007 Annual Markets Report," Retrieved from http://iso-ne.com/markets/mkt_anlys_rpts/annl_mkt_rpts/index.html.

ISO-New England (2008b). "New England's Second Power Resource Auction Produces Positive Outcomes for the Region." Retrieved from http://www.iso-ne.com/nwsiss/pr/2008/fca-2_updated_results_122308.pdf.

ISO-New England (2009a). *Internal Market Monitoring Unit Review of the Forward Capacity Market Auction Results and Design Elements*. Prepared by ISO-NE and *The Brattle Group*, June 5, 2009. Retrieved from: http://www.iso-ne.com/markets/mktmonmit/rpts/fcm_report_final.pdf

ISO-New England (2009b). *2008 Annual Markets Report.* Markets Monitoring and Mitigation. Retrieved from: http://www.iso-ne.com/markets/mkt_anlys_rpts/annl_mkt_rpts/index.html.

Joskow, P. L. (2006a). "Competitive Electricity Markets and Investment in New Generating Capacity." Retrieved from http://econ-www.mit.edu/files/1190.

Joskow, P. L. (2006b). "Markets for Power in the United States: An Interim Assessment." *The Energy Journal, 27* (1).

Joskow, P. L. (2005). "Why Capacity Obligations and Capacity Markets?" Presented at "The Economics of Electricity Markets," Toulouse, France.

Joskow, P. L., and Tirole, J. (2004). "Reliability and Competitive Electricity Markets." Retrieved from http://www.ucei.berkeley.edu/PDF/csemwp129.pdf.

Kahn, A.E., Cramton, P.C., Porter, R.H., Tabors, R.D. (2001). "Pricing the California Power Exchange Electricity Market: Should California Switch from Uniform Pricing to Pay-as-Bid Pricing?" Retrieved from http://www.cramton.umd.edu/papers2000-2004/kahn-cramton-porter-tabors-blue-ribbon-panel-report-to-calpx.pdf

Larses, E. R., Dyner, I., Bedoya V, L., and Franco, C.J. (2004). "Lessons from Deregulation in Colombia: Successes, Failures and the Way Ahead." *Energy Policy, 32*.

Midwest ISO (2009). *DRAFT Resource Adequacy Business Practices Manual*.

Midwest ISO (2006). "Value of Lost Load Final Report."

Minton, B. A., and Schrand, C. (1998) "The Impact of Cash Flow Volatility on Discretionary Investment and the Costs of Debt and Equity Financing," *Journal of Financial Economics.*

Monitoring Analytics (2009). *2008 State of the Market Report for PJM.* Retrieved from http://www.monitoringanalytics.com/reports/PJM_State_of_the_Market/2008.shtml.

Moran, A., and Skinner, B. (2008). "Chapter 11: Resource Adequacy and Efficient Infrastructure Investment." In F. P. Sioshansi (Ed.), *Competitive Electricity Markets: Design, Implementation, Performance*. Oxford: Elsevier.

Megawatt Daily (2009a). "N.Y. single market clearing price challenged" (by Eric Wieser) and "Single clearing price defended in new report" (by Craig Cano), March 5, 2009.

Megawatt Daily (2009b). "ERCOT wrestles with how to spur generation" (by H. Carr) August 31, 2009; "ERCOT Reserve Margin Promises Worry Some" (by H. Carr), August 13, 2009; "ERCOT awaits 15 plants under construction" (by J. Ryser), August 10, 2009; "Wind makes financing gas in ERCOT tougher" (by M. Watson), May 26, 2009; and "ERCOT enjoys power plant construction boom" (by H. Carr), May 13, 2009.

Megawatt Daily (2009c). "ISO Seeks Comments on Scarcity Pricing Plan," August 26, 2009.

National Economics Research Associates (2003). "Central Resource Adequacy Markets for PJM, NY-ISO, and NE-ISO." Final Report.

National Grid Electricity Transmission (2008a). *Short Term Operating Reserve Development Initial Proposals.* Retrieved from http://www.nationalgrid.com/NR/rdonlyres/9050F49A-DA88-432A-AF40-1F9C9EF6AD2B/28436/STOR_Development_InitialProposal_Final_v1.pdf.

National Grid Electricity Transmission (2008b). *Short Term Operating Reserve Review 2007/8.* Retrieved from http://www.nationalgrid.com/NR/rdonlyres/209E0BFA-17EB-4140-9CCF-3C92BE803191/27564/STOREndofYearReport0708_Final.pdf.

National Grid Electricity Transmission (2008c). *Seven Year Statement 2008.* Retrieved from http://www.nationalgrid.com/uk/Electricity/SYS/current.

National Grid Electricity Transmission (2009). *The Statement of the Use of System Charging Methodology*. Retrieved from http://www.nationalgrid.com/NR/rdonlyres/54443629-D9DC-4F98-AE10-E2D91E4A1BB8/33163/UoSCMI5R0Final.pdf.

Nagayama, H., and Kashiwagi, T. (2007). "Evaluating Electricity Sector Reforms in Argentina: Lessons for Developing Countries?" *Journal of Cleaner Production, 15.*

Newell, S., Bhattacharyya, A. and Madjarov, K. (2009). *Cost-Benefit Analysis of Replacing the NYISO's Existing ICAP Market with a Forward Capacity Market. The Brattle Group,* June 15, 2009. Retrieved from: http://www.nyiso.com/public/webdocs/committees/bic_icapwg/meeting_materials/2009-06-05/Brattle_FCM_Cost-Benefit_Analysis_for_NYISO_06-15-09.pdf

New York ISO (2008). *Manual 4: Installed Capacity Manual.* Version 6.8. Retrieved from http://www.nyiso.com/public/webdocs/products/icap/icap_manual/icap_mnl.pdf.

New York ISO (2005). *FERC Electric Tariff.* Original Volume No. 2, Attachment B. Effective July 18, 2005. Retrieved from http://www.nyiso.com/public/documents/tariffs/market_services.jsp.

Nicholson, W. (2005). *Microeconomic Theory: Basic Principles and Extensions* (9^{th} ed.) Mason, OH: Thomson.

Nordel (2007). *Guidelines for Implementation of Transitional Peak Load Arrangements: Proposal of Nordel.* Retrieved from http://www.svk.se/Global/01_Om_oss/Pdf/Elmarknadsradet/071115NordelGuidelines.pdf.

Nordpool Spot AS (2008). "No. 98/2008 Adjustment of Elspot Technical Price Ceiling in NOK and SEK." Retrieved from http://www.nordpoolspot.com/en/Market_Information/Exchange-information/No9808-Adjustment-of-Elspot-technical-price-ceiling-in-NOK-and-SEK.

Nordpool Spot AS (2009). "The System Operators," Retrieved from http://www.nordpoolspot.com/PowerMaket/The-Nordic-model-for-a-liberalised-power-market/The-system-operators.

North American Electric Reliability Council (2008a). *2008 Long-Term Reliability Assessment, 2008-2017.* Retrieved from http://www.nerc.com/page.php?cid=4|61.

North American Electric Reliability Council (2008b). *Historic Capacity and Demand.* Retrieved from http://www.nerc.com/page.php?cid=4|38|41.

North American Electric Reliability Council (2008c). *2008-2017 Regional & National Peak Demand and Energy Forecasts Bandwidths.* Retrieved from http://www.nerc.com/docs/pc/lfwg/NERC_2008-2017_Regional_Bandwidths.pdf.

North American Electric Reliability Council (2008d). "Inventory of NERC Regional & Sub-regional Resource Adequacy Assessment Criteria." Retrieved from http://www.nerc.com/docs/pc/ris/NERC_Res_Adeq_Criteria_Inventory_080130.xls.

Ofgem (2002). "New Electricity Trading Arrangements (NETA) - One Year Review." Retrieved from http://www.ofgem.gov.uk/MEDIA/FACTSHEETS/Documents1/1109-factsheet1102_24july.pdf.

de Oliveira, R.G., and de Moraes Marreco, J. (2005). "Natural Gas Power Generation in Brazil: New Window of Opportunity?" *Energy Policy, 34*.

Ontario Power Authority (2009). *A Progress Report On Electricity Supply*, First Quarter 2009. Retrieved from http://www.powerauthority.on.ca/Storage/100/9571_2009_Q1_A_Progress_Report_On_Electricity_Supply_(2).pdf.

Ontario Energy Board (2009). *Monitoring Report on the IESO-Administered Electricity Markets.* Retrieved from http://www.oeb.gov.on.ca/OEB/_Documents/MSP/msp_report_200901.pdf.

Oren, S. S. (2005). "Generation Adequacy via Call Options Obligations; Safe Passage to the Promised Land." *Electricity Journal.*

Oren, S. S. (2000). "Capacity Payments and Supply Adequacy in Competitive Electricity Markets." Retrieved from http://www.pserc.wisc.edu/ecow/get/publicatio/2000public/oren_capacity_payment.pdf.

Pagnamenta, R. (2008). "National Grid Chief Steve Holliday: Blackouts will be Common in 7 Years." TimesOnline.

Pai, M.A. (2003). *Power Circuits and Electromechanics.* Champaign, IL: Stipes Publishing LLC.

Park, J., Ahn, N. Yoon, Y., Koh, K., Bunn, D.W. (2007). "Investment Incentives in the Korean Electricity Market." *Energy Policy, 35*.

Patton, D. B., and LeeVanSchaick, P. (2007). *2007 State of the Market Report New York ISO*. Retrieved from http://www.nyiso.com/public/documents/studies_reports/market_advisor_reports.jsp?sort=date&order=ascending&=undefined.

Pfeifenberger, J., Newell, S., Earle, R., Hajos, A., and Geronimo, M. (2008). *Review of PJM's Reliability Pricing Model (RPM). The Brattle Group*, June 30, 2008. Retrieved from http://www.brattle.com/_documents/UploadLibrary/Upload696.pdf.

PJM (2008). *PJM Manual 18: PJM Capacity Market.* Retrieved from http://www.pjm.com/documents/~/media/documents/manuals/m18.ashx.

PJM (2009a). *PJM Manual 13: Emergency Operations, Revision 36.* Retrieved from http://www.pjm.com/documents/~/media/documents/manuals/m13.ashx.

PJM (2009b). *2012/2013 RPM Base Residual Auction Results.* Retrieved from http://www.pjm.com/markets-and-operations/rpm/~/media/markets-ops/rpm/rpm-auction-info/2012-13-base-residual-auction-report-document-pdf.ashx

PJM (2009c). *PJM Straw Proposal for Scarcity Pricing Offset.* Capacity Market Evolution Committee. Retrieved from http://www.pjm.com/~/media/committees-groups/committees/cmec/20090824/20090824-item-04e-pjm-proposal-for-scarcity-pricing-and-rpm.ashx.

Pollitt, M. (2004). "Electricity Reform in Chile: Lessons for Developing Countries." Retrieved from http://tisiphone.mit.edu/RePEc/mee/wpaper/2004-016.pdf.

Potomac Economics (2008). (Independent Market Monitor for the ERCOT Wholesale Market.) *2007 State of the Market Report for the ERCOT Wholesale Electricity Markets.* Retrieved from http://www.puc.state.tx.us/wmo/documents/index.cfm.

Priest, G. L. (1990). "The New Legal Structure of Risk Control," *Daedalus, 119,* (4).

Restructuring Today (2009a). "IPPs Urge MISO to Go Further on Scarcity Pricing Changes," May 27, 2009.

Restructuring Today (2009b). "AREM Supports Tradable Cal-ISO Capacity Product," May 20, 2009; "WPTF, NRG See Room for Improvement in Cal-ISO Proposal," May 21, 2009; "California ISO Defends Standard Capacity Product Proposal," June 5, 2009.

Roach, C. R., Rein, S., and Gottshall, K. (2008). *State of the Market Report 2007: Southwest Power Pool, Inc.* Prepared by Boston Pacific Company, Inc, External Market Advisor. Retrieved from http://www.spp.org/publications/2007_State_of_Market_Report.pdf.

Roques, F.A. (2007). "Market Design for Generation Adequacy: Healing Causes Rather Than Symptoms," University of Cambridge Working Paper in Economics 0821, December 2007. Retrieved from http://www.electricitypolicy.org.uk/pubs/wp/eprg0810.pdf

Roques, F. A., Newbury, D. M. and Nuttall, W. J. (2005). "Investment Incentives and Electricity Market Design: the British Experience." *Review of Network Economics*, *4* (2).

Rudnick, H., Barroso, L.A., Skerk, C., and Blanco, A. (2005). "South American Reform Lessons: Twenty Years of Restructuring and Reform in Argentina, Brazil, and Chile." *IEEE Power & Energy Magazine.*

Rudnick, H., and Montero, J. (2002). "Second Generation Electricity Reforms in Latin America and the California Paradigm," *Journal of Industry, Competition and Trade, 2* (1/2).

Southwest Power Pool (2009). *Criteria.* Retrieved from http://www.spp.org/publications/CurrentCriteria04282009-with%20AppendicesCurrent.pdf.

Tabors, R. D., and Cardell, J.B. (2002). "*Ex Ante* and *Ex Post* Designs for Electric Market Mitigation: Past and Present Experience and Lessons from California." *Proceedings of the 36th Hawaii International Conference on System Sciences*.

Tankha, S. (2009). "Lost in Translation: Interpreting the Failure of Privatization in the Brazilian Electric Power Industry," *Journal of Latin American Studies.* Vol. 41, pp. 59-90.

Tomain, J. P. and Cudahy, R.D. (2004). *Energy Law in a Nutshell.* Mason, OH: Thomson West.

Wolak, F. A. (2004). "What's Wrong with Capacity Markets?" Retrieved from http://stoft.com/metaPage/lib/WolaK-2004-06-contract-adequacy.pdf.

Wolak, F.A., Bushnell, J., Hobbs, B.F. (2007). "Final Opinion on 'Long-Term Resource Adequacy under MRTU'." Market Surveillance Committee of the California ISO, November 5, 2007.

LIST OF ACRONYMS

BETTA	British Electricity Trading and Transmission Arrangements
C&I	Commercial and Industrial
CONE	Costs of New Entry
CPT	Cumulative Price Threshold
CPUC	California Public Utility Commission
CCGT	Combined Cycle Gas Turbine
CT	Combustion Turbine
DLC	Direct Load Control
FCM	Forward Capacity Market
FRR	Fixed Resource Requirement
GENCO	Generation Company
ILR	Interruptible Load for Reliability
IPP	Independent Power Producer
LDA	Locational Deliverability Area
LOLE	Loss of Load Expectation
LOLP	Loss-of-Load Probability
LRAR	Local Resource Adequacy Requirement
LSE	Load-Serving Entity
NEM	Australia's National Electricity Market
NETA	Britain's New Electricity Trading Arrangements
NGCC	Natural Gas Combined Cycle
NGET	National Grid Electricity Transmission
PICA	PJM Installed Capacity Accounting
PJM	PJM Interconnection, L.L.C.

PUC	Public Utility Commission
RA	Resource Adequacy
RAR	Resource Adequacy Requirement
RAS	Resource Adequacy Standard
RFP	Request for Proposal
RMR	Reliability-Must-Run
RPM	Reliability Pricing Model
SCP	Standard Capacity Product
SMP	System Marginal Price
SPP	Southwest Power Pool
STOR	Short Term Operating Reserves
TSO	Transmission System Operators
UCAP	Unforced Capacity
VOLL	Value of Lost Load
VRR	Variable Resource Requirement