RAP
Energy solutions
for a changing world

What Lies “Beyond Capacity Markets”?

Delivering Least-Cost Reliability Under the New Resource Paradigm

A “straw man” proposal for discussion, issued August 14, 2012[1]

I. Introduction

Around the world, the ultimate aim of those involved in regulating a monopoly power sector or designing and overseeing competitive electricity markets is to find the set of rules and practices that efficiently and reliably delivers the right amount and the right mix of resources. Many different approaches have been taken and all have been subject to multiple revisions. The next challenge is to understand and address how the growing share of variable renewable production will require us to rethink our current practices.[2] While many of the discussion points of this paper apply equally to all industry structures, our primary focus is on adapting competitive wholesale power markets to deliver their intended economic efficiency and reliability outcomes under this new resource paradigm.

There has long been debate in competitive wholesale power markets over how to be confident of the resource investments needed to ensure reliability. While some regions remained committed to the energy-only model[3] other regions adopted mechanisms of various types to pay for capacity.[4] These mechanisms were intended to address the

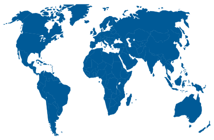
---

[1] Lead author: Mike Hogan (RAP Senior Advisor); Project coordinator and substantive contributor to this paper: Meg Gottstein (RAP Principal). Valuable editorial input and feedback on earlier drafts also provided by David Moskovitz (RAP Principal), Phil Baker (University of Exeter), Simon Skillings (Trilemma LTD), Kevin Porter (Exeter Associates) and Elizabeth Watson (RAP Energy and Environment Fellow).

[2] “Variable” as used in this paper refers to any source of electricity production where the availability to produce electricity is largely beyond the direct control of operators. It can be simply variable – changing production independently of changes in demand – or variable and uncertain – variable and, in relevant timeframes, unpredictable. Another term for this latter category of sources is “intermittent.”

[3] “Energy-only” markets do not explicitly value the capacity to produce energy; compensation for capacity is implicit in the price of energy. They do often include mechanisms or “services markets” to pay for specific non-energy system services, such as operating reserves and reactive power.

[4] Wherever this paper refers to “capacity markets” or “capacity mechanisms” it should be assumed, unless stated otherwise, that this encompasses any energy market interventions similarly intended to provide financial support for investment in a desired quantity of firm productive capacity.



Global US China EU
The Regulatory Assistance Project
Rue Stassart 48
Brussels 1050
Belgium
phone: +32 2-894-9300
fax: +32 2-894-9301
www.raponline.org

concern that markets without a capacity mechanism were not supporting enough investment in resources to meet peak demand. This is often referred to as "resource adequacy" and is discussed below.

The impressive rise in variable renewable production in many markets has once again focused attention on the investment question. Many are asking if capacity mechanisms are the answer. But the rise in variable renewables raises a new type of reliability question: Will there be investment in the right type of resources? This is what we describe as the "system quality" challenge.[5] Much of this paper is focused on system quality and how to address it.

At the outset, two points are clear:

- Addressing both aspects of investment (adequacy and quality) is essential to maintaining reliability at least-cost while the power sector shifts from being dominated by fossil fuels to relying more heavily on renewables.
- Investment incentives aimed at resource adequacy will no longer suffice to ensure system quality. The demand for specific resource capabilities must also be reflected properly in decisions at investment timescales.

Reliability has always involved these two dimensions, but they have traditionally functioned in different timescales. ***Resource adequacy*** **–** access to enough firm resources[6] to be able to meet the highest expected level of demand – has dominated planning at investment timescales. In contrast, ***system quality*** **–** the right mix of resource capabilities deployed to ensure that in every moment supply can be balanced with demand – has been the focus of services markets that have functioned primarily at operational timescales. It is the system quality dimension that is fundamentally transformed by rising shares of variable renewables, making resource flexibility increasingly an investment consideration as well as an operational one.

The operational needs are best revealed by forecasts of ***net demand*** (gross demand minus demand served by variable resources). Net demand forecasts are essential tools for quantifying gaps between the need for resource flexibility over investment timescales and the capabilities of the current and prospective resource portfolio to meet that demand cost-effectively. In other words, net demand forecasts are used to see if the system has, and will have resources with the mix of operating characteristics needed to deliver least-cost reliability.

Pure capacity markets tend to exacerbate the gap between the resource capabilities needed and the capabilities of existing and planned resources. Energy-only markets may do so as well. Regrettably, both ignore the system quality implications of changing net demand profiles. In so doing they are at risk of driving needlessly costly and unstable outcomes.

We propose a framework for addressing reliability investment concerns in ways that are mutually reinforcing for cost-effective reliability. In doing so, we build upon a companion paper[7] that suggests our collective thinking will need to evolve "beyond capacity markets" in order to address the reliability challenges associated with meeting power sector de-carbonization imperatives. We intend this "straw man" proposal to serve as a catalyst for further, productive discussion on this topic.

---

[5] There is no universally recognized term for this – in many markets the concept is denoted by the term "system security."

[6] "Firm" as used in this paper refers to that portion of the maximum capacity of a resource that can be relied upon with a high level of confidence to deliver whenever needed. It is a resource-specific determination based on statistical analysis.

[7] See *Beyond Capacity Markets: Delivering Capability Resources to Europe's Decarbonized Power Sector,* Regulatory Assistance Project, http://www.raponline.org/featured-work/beyond-capacity-markets-delivering-capability-resources-to-europes-decarbonised-power

II. Variable Renewables and Reliability

The advent in many regions of significant and increasing contributions to the resource mix by variable renewables has a number of important implications for reliability. Among these is how it affects traditional resource planning and, in competitive wholesale power markets, the ability of existing market mechanisms to deliver desired outcomes. The policy-driven addition of renewable supply resources to markets that were in most cases already fully served creates an excess of supply, depressing wholesale market prices and placing pressure on marginal producers. (In regulated monopoly areas this same phenomenon can occur creating surplus capacity, lower utilization rates, and possibly stranded assets.) The consequent threat of disinvestment is sometimes offered as evidence that the simple fact of a large share of variable renewables production distorts energy market price signals for all other generation. Yet it is only natural under conditions of excess supply, and indeed it is desirable, for economically obsolete production capacity to be retired from the market.

Nonetheless, increased reliance on variable resources does change the reliability calculus for resource planners and market operators. The key is to recognize that investment in resources to ensure reliability is no longer simply a question of quantity. To quote a leading reliability authority: "Traditionally (and primarily for simplicity), resource planning has been a capacity-focused process. However, with high penetrations of variable generation resources in the system, existing planning methods will have to adapt to ensure that adequate resources are available to maintain bulk power system reliability."[8] Where markets are meant to substitute for resource planning, it is only logical to expect a similar need to adapt.[9]

III. Resource Adequacy vs. System Quality

Reliability has traditionally rested on two distinct dimensions operating in different timescales. Resource adequacy has dominated investment timescales and looks at whether the system has access to enough firm resources to serve the highest expected level of demand. System quality has focused almost exclusively at operational timescales and asks whether the system has the right resources deployed at any given moment to ensure that supply can be balanced with demand. This simplification worked well enough most of the time because the need for resource flexibility was commonly bounded and predictable. It was generally assumed that sufficient capabilities could be sourced most cost-effectively from whatever supply portfolio emerged from the competitive energy markets or from an administered least-cost resource adequacy process.

Current discussions about reliability often slide into this historical rut. The variability and uncertainty associated with certain renewable resources is perceived as a challenge for investors in the balance of the supply portfolio. The discussion therefore turns to how to ensure a sufficient quantity of investment in firm resources. But optimizing decisions at investment timescales is no longer quite so simple. This new investment challenge is not principally to do with the total quantity of resources but rather with a marked shift in the demand for some operational capabilities relative to others. The key differentiator is resource flexibility. Flexible resources can respond to system needs by

---

[8] The North American Electric Reliability Corporation ("NERC"), "Accommodating High Levels of Variable Generation", April 2009, pg. 38. NERC is the organization federally certified to establish and enforce reliability standards for the bulk power system in the United States.

[9] We note that the technical and institutional aspects of integrating large shares of variable renewables – larger balancing areas, expanded transmission capacity, shorter scheduling intervals, better forecasting, more flexible resources, responsive demand, energy storage – are virtually identical in both competitive wholesale market and regulated monopoly environments. The relevant difference between these environments lies in the extent to which investment is driven by resource planning vs. market forces.

ramping up, ramping down, and turning on and off quickly and often. If resources cannot respond quickly to system needs, customers will pay the price in higher operating costs, unnecessary capital investment and less reliability.

The increased need for flexibility is illustrated in a particularly vivid manner in Figures 1 and 2 below. These graphs show total (gross) and residual (net) demand at the level of the Danish system in the first two months of 2007, at a time when Denmark was generating the equivalent of ~20% of its annual demand from variable renewable sources.

Figure 1[10]

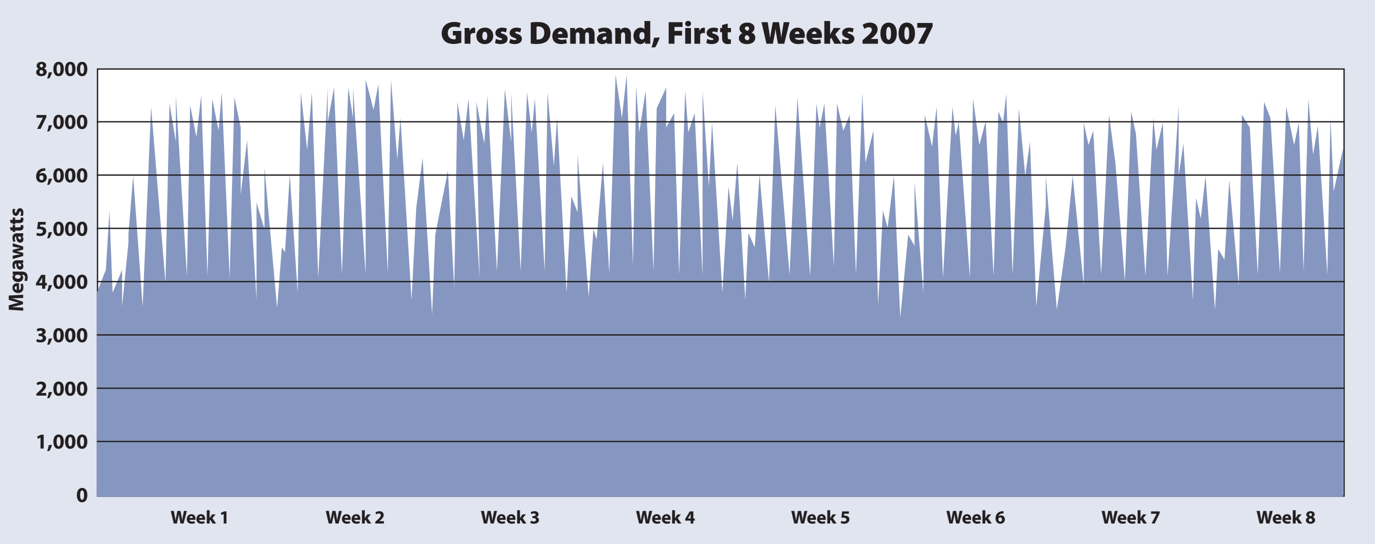
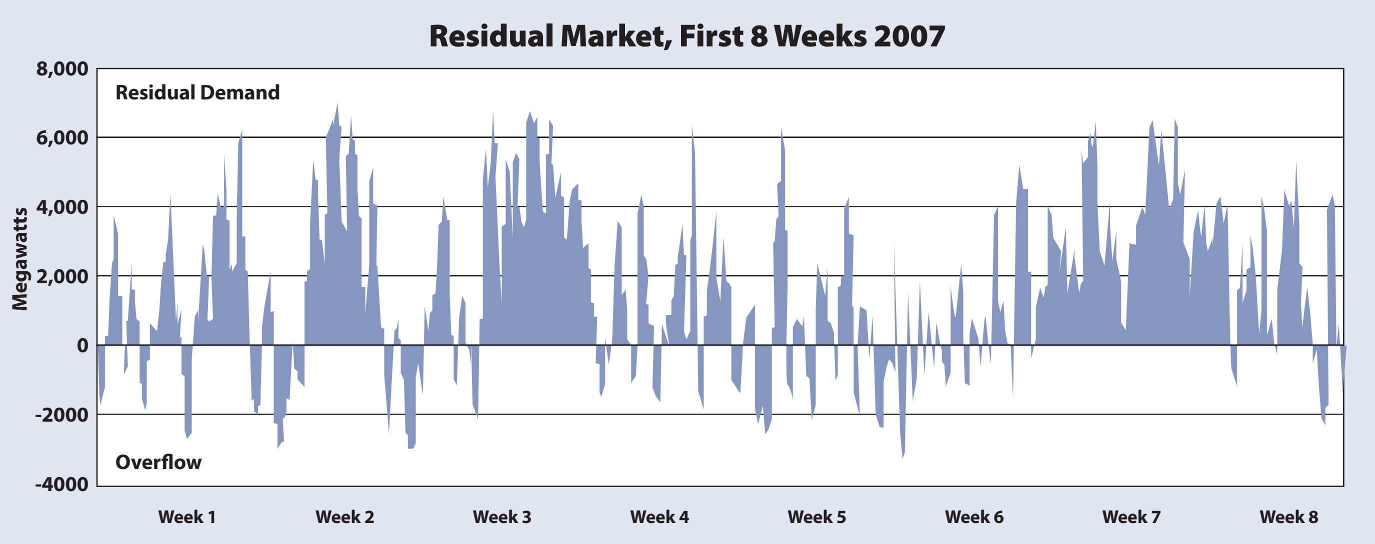


Figure 2

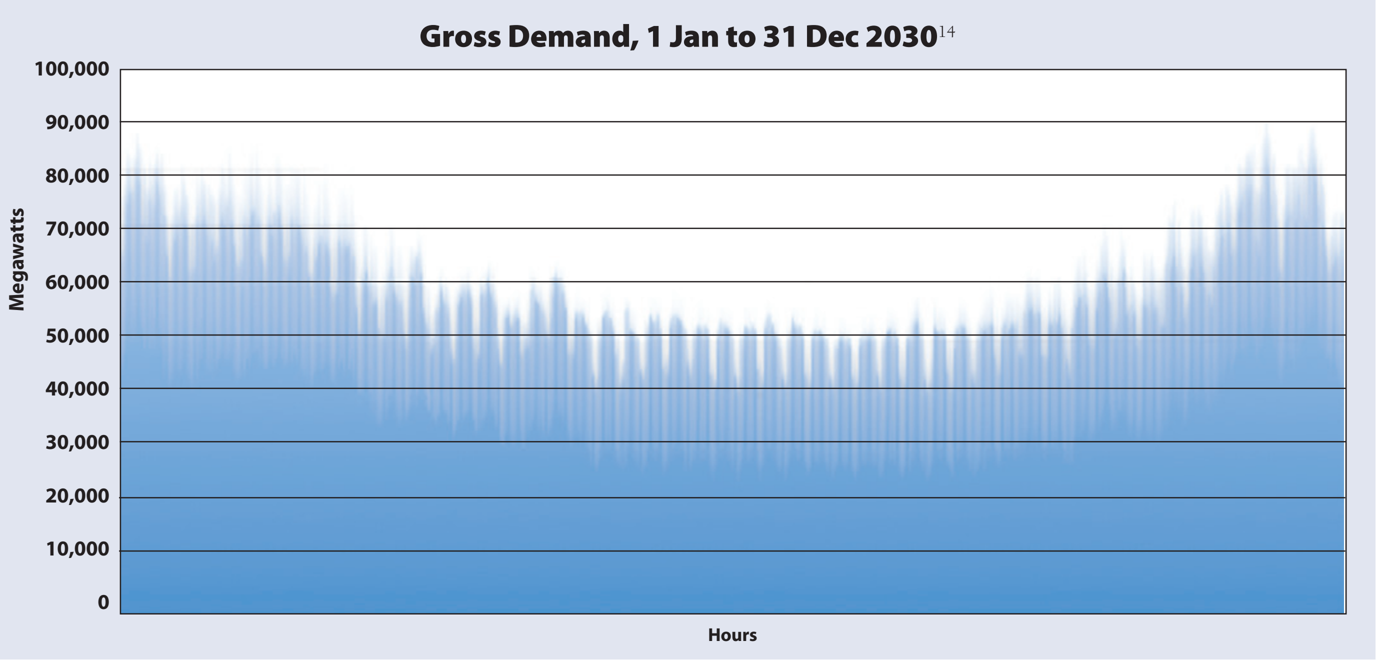


---

[10] Figures 1 and 2 are from a presentation by Kim Behnke, energinet.dk, at the Workshop on Sustainable Development, Technical University, Lyngby, Denmark 14-15 January 2009, available at http://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CEUQFjAA&url=http%3A%2F%2Fwww.risoe.dtu.dk%2FConferences%2FWorkshop_Sustainable_Energies%2F~%2Fmedia%2FRisoe_dk%2FConferences%2FEnergyconf%2FDocuments%2FStorage%2FKim_Behnke_tilladelse.ashx&ei=z7slULSKH4nl0gGNyYGoDg&usg=AFQjCNFFoP059ccB2Cl0XsfOs-4Ha7QczA

Because this analysis takes no account of measures available to policy makers to mitigate these effects[11] — and because Denmark relies to an unusual extent on uncontrollable electricity production from combined-heat-and-power facilities — Figure 2 portrays a net demand profile that is unnecessarily severe. Yet even with a more optimized institutional and physical infrastructure (such as the regional transmission organization model deployed in Scandinavia and parts of North America[12] or that envisioned by the Target Model proposed for Europe’s Integrated Electricity Market[13]) the effect can still be quite significant. Figures 3 and 4 depict gross demand and net demand for a specific market area in Europe in 2030 with approximately 27% of annual electricity from variable renewables. In contrast to the figures above, this analysis assumes substantially all cost-effective transmission expansion, a more diversified resource base, high-fidelity weather forecasting, and a unified balancing area that encompasses at least the proximate regional member states.

Figure 3



---

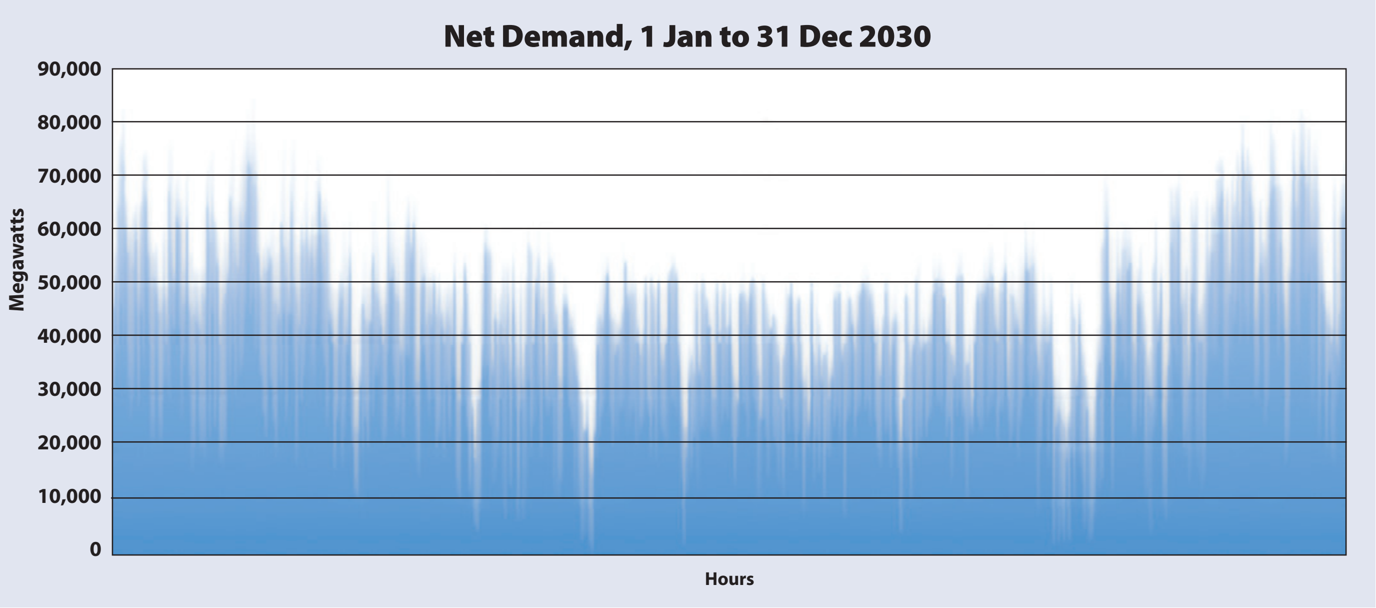
[11] The magnitude of difference between gross and net demand is highly dependent on answers to a number of choices facing regional market authorities, including inter alia: size of area over which the system is balanced in real time; inter- and intra-regional power transfer capacities; diversity of variable resources; length of scheduling intervals; quality of day-ahead and intra-day resource forecasts; and extent to which consumers are enlisted as resources in response to uncontrollable changes in supply. In reality the Nordic region already employs a number of these measures and is moving forward on others, so that this depiction, while technically accurate, imposes artificial constraints.

[12] For information about regional transmission organizations see http://www.ferc.gov/industries/electric/indus-act/rto.asp

[13] For a description of this “market coupling” Target Model and related implementation steps envisioned for Europe’s Integrated Electricity Market, see: *Advancing Both European Market Integration and Power Sector Decarbonisation: Key Issues to Consider* (RAP, 2011) at: http://www.raponline.org/document/download/id/879

[14] Figures 3 and 4 are derived from analysis by RAP and Imperial College London of data from Power Perspectives 2030, Nov. 2011. They depict the “UK South” market area in 2030.

Figure 4



As all of these graphs demonstrate, the demand for the kind of resource flexibility traditionally associated with peaking and cycling plants is no longer either bounded or predictable but rather extends erratically across most of the non-renewable resource portfolio. Exactly how this would affect underlying resource investment – and disinvestment – depends on a number of things including the feasibility, required investment, and operating/maintenance cost to operate existing resources more flexibly. But one thing is clear: Questions of system quality can no longer be confined to operational timescales. Flexibility will be in greater demand, it will acquire greater value, and that value needs to be reflected properly in decisions at investment timescales.

IV. Short-Term vs. Long-Term

We are proposing that there will be both short-term and long-term aspects to the system quality dimension as variable renewable production expands. In the short term the question is whether the system has effective access to all of the cost-effective flexibility available from the existing resource portfolio, including untapped existing demand-side potential. The longer-term aspect is whether the market supports investment in a portfolio of new and existing supply- and demand-side resources capable of efficiently and cost-effectively meeting the projected need for flexible resource capabilities over investment time horizons.

1) Short-Term Responses

All competitive wholesale power markets, even those that have held firmly to the "energy-only" model, have instituted various forms of separate short-term ancillary service[15] mechanisms with the intention to create an incentive for resource owners to offer whatever relevant capability options they have to system operators. This task is complicated in some regions by a preponderance of bilateral trading arrangements involving self-scheduling by

---

[15] "Ancillary services" is one way of describing the various balancing tools and reserves system operators rely upon to maintain operational stability on the power system.

supply resources.[16] As the need for system flexibility increases it will be prudent as a first step to explore options to reduce the extent to which self-scheduling of resources obscures or obstructs access to the range of flexibility available in the existing resource portfolio.[17]

In vertically integrated monopoly systems these services are bundled as part of the monopoly providers' overall supply-side operations. This affords system operators more direct access to supply-side resource capabilities, but it also gives cause to be concerned about whether the potential for cost-effective demand-side flexibility is being adequately exploited. In this environment the responsibility rests with regulators to ensure that the monopoly provider is using all cost-effective options, including demand-side alternatives, to minimize the cost of integrating larger shares of variable renewable supply. The success of regulators in this regard can benefit from a variety of measures, including improved integrated resource planning, competitive procurement of the identified resource needs, transparency in the planning and procurement process as well as effective stakeholder engagement.

In addition, some markets have begun to recognize that rising shares of variable resources are creating the need for a new class of flexibility services. These mechanisms are designed to ensure there is sufficient capability to stop-start and ramp supply resources up and down (or, in the case of demand resources, down and up) fast enough, far enough and frequently enough over multiple scheduling intervals to complement in a least-cost manner the variability and uncertainty characteristic of some renewable resources. In response some have begun to develop markets for various types of ramping services, e.g., the California and Midwest ISOs.[18] So far the available examples appear to be short-term market mechanisms operating alongside the market for wholesale energy.

2) Long-Term Responses

The long-term picture is different. As regards resource adequacy, capacity mechanisms are long-term responses adopted in some regions to operate alongside energy markets. They were triggered by concerns that investors, for various reasons, may be unable to foresee with sufficient confidence a combination of volume and price they can expect to realize in the energy market to support investment in long-lived assets. These concerns apply equally to new investment and to the economic viability of existing investments. All capacity mechanisms require a multi-year forecast of maximum gross demand. That forecast then forms the basis for periodic auctions for, or administratively determined payments to, that quantity of firm capacity resources needed to meet established resource adequacy standards.

To date there are no fully comparable examples addressing system quality, where requirements are determined for periods as short as within-day to, for example, the two-year-forward assessment for Great Britain's Short Term

---

[16] "Self-scheduling" refers to the fact that in a bilateral arrangement the decision as to when to operate and in what mode is driven by the provisions of the bilateral contract rather than by the matching of offers to sell with offers to purchase in an open market. System operators can be constrained in setting the next day's schedule by operational decisions taken under these bilateral arrangements.

[17] The Australian and Irish market operators have each imposed minimum ramping requirements on effectively all wholesale generators (see for Australia http://www.aemc.gov.au/Electricity/National-Electricity-Rules/Current-Rules.html, ch. 3, pgs. 116-127; and for Ireland see http://www.eirgrid.com/media/Grid%20Code%20Version%204.pdf, section CC.7). These requirements address only one aspect of resource flexibility in a rather blunt fashion but illustrate one way markets are dealing with this issue.

[18] The California ISO is developing a 5-minute upward or downward ramping product procured day-ahead and dispatched in real-time. See the most recent proposals and notices at: http://www.caiso.com/Documents/Flexible%20ramping%20product%20-%20relevant%20market%20notices. The Midwest ISO is also developing a product to secure additional up ramp and down ramp capabilities. The mechanism issues payments to resources cleared for ramp capability regardless of real-time dispatch instructions. See *Ramp Capability for Load Following in the MISO Markets* at https://www.midwestiso.org/Library/Repository/Communication%20Material/Key%20Presentations%20and%20Whitepapers/Ramp%20Capability%20for%20Load%20Following%20in%20MISO%20Markets%20White%20Paper.pdf.

Operating Reserve. Few if any examples exist of a standing long-term (multi-year) procurement mechanism for flexibility services comparable to the capacity mechanisms that have been implemented in some markets.[19] This may be because system operators have traditionally been able to extract sufficient flexibility from legacy resource portfolios to meet their requirements and expect to continue to be able to do so as new resources are added to their system.

A handful of markets have reached the point where this can no longer be taken for granted (markets confronting this issue include Denmark, the Pacific Northwest, Ireland, Germany and California), but many other markets are not far behind. Markets that reach this stage of evolution come face-to-face with the need to ensure not only that the quantity of firm resources meets resource adequacy requirements but also that the resource portfolio is capable of efficiently addressing emerging system quality needs. As with resource adequacy, questions about the services capabilities of system resources apply equally both to new investment and to existing resources. While the value of greater resource flexibility is expected to increase significantly over investment timescales as the share of supply from variable resources grows, there is no more reason to expect investors to rely on price signals from the short-term services markets than there is to expect they will rely on price signals from the short-term energy market. In other words, if some find it implausible that investors will decide *whether* to invest strictly on the basis of short-term energy markets, it is at least equally implausible that investors will choose *what* to invest in strictly on the basis of short-term energy and services markets.

To put this in concrete terms, consider that leading manufacturers of combined-cycle gas turbine (CCGT) power plants have developed highly flexible product offerings with precisely these future system needs in mind.[20] The new designs are capable of fast-start, short-cycling, and ramping down and stabilizing at a very low fraction of maximum output for hours at a time. These manufacturers also still offer the more traditional, far less flexible CCGT product offerings. Not surprisingly, the more flexible options will require higher initial investment. Yet there is now concern that current energy and capacity market designs will not provide the forward-looking information investors and suppliers will need to evaluate properly the trade-offs between higher initial cost and life cycle consumer benefits.[21]

V. Undervaluation of Flexibility in Current Market Designs

The absence of long-term market mechanisms for flexible resource capabilities can be problematic under competitive wholesale markets that adopt capacity mechanisms as well as those that do not, for the reasons discussed below.

1) Markets with Capacity Mechanisms

The limitations of short-term capabilities markets in the face of rising shares of variable renewables are likely to become most apparent in those markets that adopt capacity mechanisms. This is due to three factors:

---

[19] Australia operates an energy-only market but has imposed minimum ramp requirements on all registered generators. Ireland imposes a less stringent minimum ramp requirement on all generators and separately administers a capacity payment mechanism. Neither Ireland nor Australia use market mechanisms to determine the need for and value of flexible resource capabilities. PJM's original (August 2005) filing for its current forward capacity market proposed to apportion the market for capacity into three parts based on specified resource capabilities; that feature was dropped in the final market design as a result of the stakeholder settlement process.

[20] See "Fast starts and flexibility: Let the gas turbine battle commence" (T. Probert, *Power Engineering International,* June 2011); also http://www.ge-flexibility.com/solutions/flexefficiency-50-combined-cycle-power-plant/index.html); see also "Flexible Future for Combined Cycle" (L. Balling, Siemens, December 2010).

[21] See e.g. "Contributions of Flexible Energy Resources for Renewable Energy Scenarios" (G. Hinkle et al., GE Energy, 8 March 2011)

1) Capacity mechanisms provide long-term visibility uniquely to the value of firm capacity, indirectly devaluing other resource attributes.
2) Firm capacity is a more visible and more tangible commodity to value.
3) Competitive markets for firm capacity favor low-cost (and thus, most likely, less flexible) firm capacity.

As a result it is highly likely that price signals from shorter-term capabilities markets will be overwhelmed in any new investment decisions as well as in the valuation of existing resources. In that context it becomes even more important that some means be found to compensate investors in resource flexibility in a manner at least as compelling as that by which investors are compensated for firm capacity. Failure to do so may compromise the very reliability benefits that motivated the adoption of a capacity mechanism in the first place.[22]

2) Energy-Only Markets

There may well be reason to be concerned about the valuation of flexibility even in markets that do not adopt capacity mechanisms. It is not uncommon even in traditional power systems for less flexible resources occasionally to operate at below marginal cost through periods of low demand rather than incur the cost of shutting down and restarting hours later or, alternatively, forego the opportunity to earn higher prices during subsequent scheduling periods. When this occurs the result is that lower marginal cost (and more flexible) resources are pushed out of the market uneconomically, while those higher-marginal-cost resources that have chosen to operate at a loss during the low-demand hours must earn prices above their marginal cost during tight supply periods in order to remain in business. As swings between tight supply and ample supply become more frequent and unpredictable, this raises the possibility that more flexible resources in energy-only markets will see their operations uneconomically constrained for a significant number of hours without offsetting compensation. In effect, energy-only markets may systemically undervalue flexibility. Legacy resource portfolios in most cases already lack the physical or financial capability to follow net demand efficiently at expected levels of variable renewables penetration, resulting in a tug-of-war between renewable curtailment and curtailment of more flexible non-renewable supply. Under this view of energy-only markets, energy market incentives would disproportionately disadvantage more flexible non-renewable supply. Without market mechanisms to value and pay specifically for investments in long-term flexibility there would be a systemic investment bias against more flexible resources.

Many proponents of energy-only markets will take a different view, maintaining that energy markets functioning properly will compensate more flexible resources commensurate with the value their added flexibility provides to the market. That may well be the case, just as it may be the case that energy-only markets functioning properly will adequately compensate investments in firm capacity. More work is required to favor strongly one view over the other. In the absence of compelling evidence, however, the case for an energy-only approach to investment in resource flexibility would appear to be more difficult to sustain.

3) Need for Long-Term Visibility to the Value of Flexibility

In any market or regulatory environment, rising shares of variable renewables give rise to the need to factor the value implications of the net demand forecast into the revenue models available to investors in system resources. In the first instance it may be sufficient to strengthen and expand the suite of short-term ancillary services markets that compensate all resources, including demand-side resources, for making their flexibility options available to system operators as needed. Such short-term markets may no longer suffice, however, when integrating significantly higher

---

[22] For a potentially relevant discussion of the situation faced by one market operator with an existing capacity market, see PJM's "Whitepaper on Operational Reliability Metrics for Generation Capacity in the Reliability Pricing Model" (Jan. 2005), pgs. 1-7.

shares of variable production. Investors' ability to foresee the future value proposition for investing in, or sustaining investment in, resources with enhanced operational capabilities is at least as limited as their ability to foresee the future value proposition for investments in firm productive capacity.

The following sections present concrete steps for creating long-term visibility to the value of flexibility in power systems expecting a high share of production to come from variable renewables.

VI. Setting Key Reliability Metrics Under the New Resource Paradigm

Essential to an assessment of reliability is the development of a multi-year forecast of maximum gross demand. This forecast provides the basis for setting resource adequacy targets (in MWs of firm capacity). In vertically integrated monopoly systems this forecast is part of a long-term resource planning process, led by the regulator or other responsible agency, used to inform decisions relating to new investment. In regions with competitive wholesale power markets such forecasts are used in monitoring the effectiveness of existing markets and to help determine the need for and design of new mechanisms.

As an example of the latter, assessments of resource investment patterns relative to forecasts of gross demand led to decisions by some regions in the US to supplement energy markets with forward (long-term) capacity markets that set specific quantity targets for firm resources. That decision carries with it the need to carry out an on-going process of long-term gross demand forecasting. This embrace of long-term forecasting and target setting in competitive wholesale power markets has typically stopped short of a return to any form of long-term resource planning. Where the resource mix is expected to include only an insignificant quantity of variable resources there may be little reason, at least on the basis of reliability concerns, to extend the role of or modify the approach to forecasting any further.

On the other hand, if there is reason to expect the resource mix will include significant shares of variable resources, two additional steps become important: development of a multi-year forecast of net demand, and the development of analytical tools appropriate to the particular market circumstances for revealing the need for and value of certain relevant categories of resource capabilities. These steps form the basis for setting system quality targets. The metrics for such targets are perhaps less generic than the simple metric for resource adequacy, but they are likely to include metrics like MW of various classes of operating reserves, MWh of up- and down-ramping capability, and numbers and frequency of stop-start cycles.

1) Forecasting Net Demand

Net demand is derived from two directly measurable quantities – gross demand and variable renewable production.[23] Focusing on net demand can give a fuller picture of resource needs in markets with significant shares of variable renewables. For such markets it is thus useful to create forecasts of not only gross demand but also expected energy production from variable resources. While this does not constitute a return to long-term resource planning, it does require market operators to take a view on at least one segment of the future resource mix.

Given the periods typically involved (three to ten years forward) it should be possible to develop a view of the quantity of variable renewable plant with a confidence level comparable to the forecast of gross demand. On the other hand one would expect more uncertainty around forecasts of quantity and shape of production. The level of uncertainty in the resulting net demand forecast will reflect the combined uncertainties in the gross demand forecast and the variable resource production forecast.

---

[23] In calculating net demand it is also necessary to consider other forms of 'must run' generation such as combined-heat-and-power plants whose production is tied to the provision of heat.

2) Valuing Flexibility Options

In the same way that gross demand forecasts provide a benchmark for the quantity of firm capacity required to meet resource adequacy standards, the purpose of net demand forecasts is to provide a basis for estimating the mix of resource capabilities that can most efficiently deliver the desired level of system quality. The net demand forecast on its own will only indicate the desired quantities of the relevant resource capabilities. Additional analysis is required to determine the value to the system of an incremental unit of those capabilities. In other words, additional flexibility is not desirable at any cost, it is only desirable if the cost to obtain it is less than the cost of the alternative, e.g., the full lifecycle cost of curtailing renewable production or the investment and operational costs of procuring or committing additional back-up generating capacity. The goal of a capabilities market would be to seek the optimal balance between greater resource flexibility and least cost.[24] An ideal market design would be capable of discovering that optimum balance through a highly dynamic, competitive process—for example, the approach described in Section VII below. In reality the degree of complexity and precision that is useful, at least as an entry point, in setting target quantities and values will vary for different markets depending on the particular circumstances and policy objectives.

This will depend in part on the pace of transformation in the resource mix, particularly in the share of variable resources. In markets where the share of variable renewables remains relatively modest or is expected to grow slowly over the planning period, a simple deterministic net demand forecast based on a statistical analysis of recent experience (including experience with forecasting error) may suffice. A "demand curve" to bound the price to be paid for flexibility services can likewise be extrapolated from historical data on the cost of inflexibility.[25] The Midwest Independent System Operator's recent proposal for a ramping service presents a good example of this approach.[26]

In markets where variable renewable production is expected to increase significantly relative to the overall resource mix it is likely that past experience is a poor guide to future requirements. In this case a more complex process may be required. A sufficiently robust picture of net demand will require a probabilistic approach based on Monte Carlo simulations of variable production.[27] Given the significant increase in demand for system flexibility over historical norms and a reasonable confidence level around the net demand forecast, multiple iterations of a detailed system production model may be required to derive a reliable value for a unit of additional resource flexibility. At what point the "high" market penetration threshold is crossed will vary across markets based on factors such as geographic reach, internal grid congestion, interconnectedness with neighboring markets, and diversity of resources, but empirical evidence suggests an indicative threshold for production from variable renewables is in the range of 10-15% of total annual production.[28]

---

[24] As the impact of variable resources grows, a key benefit of long-term capability mechanisms is that they will be more likely than short-term mechanisms to reflect the full costs of inflexibility, e.g., by recognizing the that inflexibility imposes higher lifecycle costs for variable renewables due to higher levels of curtailment, or increases the investment and operating costs due to a greater need for, and more frequent deployment of back-up peaking resources.

[25] A "demand curve" charts the value one would be prepared to pay for an incremental unit of the relevant capability, up to the target quantity (beyond which point the value would go to zero). Supply offers are then graphed from lowest to highest, and the point at which this "supply curve" crosses the demand curve establishes the market-clearing quantity and price.

[26] See https://www.midwestiso.org/_layouts/miso/ecm/redirect.aspx?id=112806 for a summary description the rationale for and design of MISO's proposed ramping service.

[27] Monte Carlo simulation is a tool used to characterize the behavior of random events in an analytically useful manner. It has been used widely in reliability planning for many years. See e.g. "Reliability of Electric Power Systems Using Monte Carlo Methods" (Billinton & Li, 1994).

[28] The ERCOT market experience is particularly instructive on this point; see Sioshansi, R. and Hurlbut, D., "Market Protocols in ERCOT and their Effects on Wind Generation" (2009), particularly the discussion at pages 5-7.

Implementing some combination of the foregoing processes should produce a more valid picture of the future need for and value of various resource capabilities. This brings us to the question of how the market can be designed to act on those findings.

VII. Capability Market Design Options

The improved methods described above can be used to amend existing resource planning and procurement practices in regions that employ them. In some situations this will include competitive acquisition of supply and demand-side resources that meet the identified need in a least-cost fashion.

In regions with competitive wholesale power markets new mechanisms may be needed. The level of sophistication in a market designed to deliver capability targets will depend in part on whether the market faces an urgent need to invest in new resources; additional considerations will be the maturity and capacity of existing market institutions and processes. The reality is that few markets have the time or the institutional capacity to set as their primary objective a market that will deliver exactly the amount of flexibility needed, for not a penny more than it's worth. Short of such an ideal market there are two potential avenues of simplification that we put forth in this "straw man" proposal: (1) enhanced services market mechanisms and (2) apportioned forward capacity mechanisms.

1) Enhanced Services Market Mechanisms

This approach utilizes a long-term services market (essentially an investment timescale adaptation of existing ancillary services mechanisms, with new services added as necessary) to procure the target mix of resource capabilities derived from the net demand forecast. Capabilities of interest would most likely include traditional system operator functions such as ten-minute spinning and non-spinning reserves and perhaps a thirty-minute operating reserve. Obligations to secure such services would likely remain with the system operator. At least as important but more difficult to specify are less traditional balancing functions. These may include short-cycle stop-start and aggressive dispatch or ramping options, parameters meant to reflect how fast and how frequently, across multiple scheduling intervals, a resource can be turned off and on, as well as the up-ramp and down-ramp rates and ranges.

For both traditional ancillary services as well as these less traditional balancing services, their value could be set by periodic "forward" auctions and paid to all new and existing resources capable of providing them. As with existing capacity mechanisms, the "forward period" (the lead time before the service must commence) as well as the "commitment period" (the period of time over which winning bidders receive the cleared market payment for the service) would need to be designed with nominal investment lead times and investment horizons in mind. An enhanced services market mechanism, and in particular establishing separate market mechanisms for non-traditional balancing services, is conceived primarily to operate in the absence of a capacity mechanism. However, it is possible (if a bit cumbersome) to envision this solution operating alongside a capacity mechanism, if desired. This may be the case if market stakeholders would prefer to keep the capacity mechanism strictly focused on resource adequacy concerns.[29]

In either case, with or without a capacity mechanism, this approach would seek to realign the mix of system resources by providing those resources with the desired capabilities access to a stable, long-term revenue stream that is unavailable to less flexible resources. This would afford more flexible resources a competitive advantage in the

---

[29] For a relevant case study, see http://pjm.com/markets-and-operations/~/media/markets-ops/rpm/20060929-er05-1410-el05-148-part1.ashx, pages 18-19 for a summary description of the ultimate resolution of PJM's original proposal to apportion its forward capacity market based on certain resource capabilities; the parties ultimately agreed to proceed with a single-clearing-price capacity auction and address any requirements for "operational reliability" capabilities through separate mechanisms.

energy and (if applicable) capacity markets. Where a capacity market exists, the effect would be to reduce the market value of undifferentiated firm capacity (See Appendix A).

Examples of how such services might be defined and procured can be found in the California ISO's proposed Flexible Ramping Product and the Midwest ISO's proposed Ramp Capability Product (see footnote 18 for links to more information). In both of these cases the mechanisms have, for now, been conceived as short-term market mechanisms. That said, extending the terms over which they operate is entirely feasible should these markets opt for an enhanced services model for ensuring long-term access to increased resource flexibility. More traditional ancillary services markets are likewise typically short-term in nature with ISO New England's one-year-forward operating reserves market a relative rarity. Yet an example of where one system operator has extended the term of a more traditional ancillary services market can be found in Great Britain, where the system operator (National Grid) has until recently procured short-term operating reserves ("STOR") via auction for commitment periods of up to 15 years.[30]

This approach has the benefit of decoupling the long-term procurement of system services from processes designed around firm production capacity, allowing greater flexibility in targeting specific services (e.g., energy storage). For this same reason it may take longer to see the desired impact on the pattern of supply-side investments: Until these services markets establish a track record investors may be slow to incorporate the relevant capabilities into long-term resource investment plans unless they have a more immediate motive to do so (such as the apportioned capacity mechanism described below). Therefore pursuing an enhanced services market may be more appropriate for markets where there is no perceived urgency to invest in a significant amount of new firm supply resources. Nonetheless, this approach represents a viable option for regions experiencing a growing share of variable renewables where creating a separate forward capacity payment mechanism may not be desirable.

2) Apportioned Forward Capacity Mechanisms

An alternative approach, in markets where capacity mechanisms have been deployed or are under active consideration, involves simply apportioning the capacity mechanism into tranches based on the target mix of resource capabilities derived from the net demand forecast. This option leverages whatever resource adequacy mechanism is in place by breaking the total quantity of firm resources required into successive tranches based on specified resource attributes. All firm resources, including qualifying demand-response and end-use energy efficiency resources, would bid into the highest-value tranche for which they could qualify. The most flexible tranche of firm resources is cleared first, followed by the next most flexible tranche, and so on. The least flexible firm resource tranche would be cleared last at whatever residual quantity of resource requirement remains unfilled. The demand curves for each tranche would reflect the relative values of the resources specified, with the clearing price for each successive tranche also expected to be lower than the last, until the final tranche which would be expected to clear at a very low price in both relative and absolute terms. The desired realignment among resources would be driven by the size of each tranche, with value set by the relationship between the size of the tranche and the supply and costs of appropriate resources. Appendix A illustrates with figures and numerical examples how this type of "capabilities" mechanism would work in comparison to a "prototypical" capacity mechanism.

A good example of how an apportioned forward capacity mechanism might be deployed in practice can be found in the proposal by PJM to include "operational reliability metrics" in the original design they filed in August 2005 for their

---

[30] The original STOR contract terms included a long-term product of up to 10 years. This was extended in 2009 to 15 years following requests by some providers considering new plant construction. More recently National Grid withdrew the long-term option because of uncertainty over the future of electricity market reform in Great Britain. Two-year STOR contracts are still procured via auction three times each year. For more information on the STOR service see www.nationalgrid.com/uk/Electricity/Balancing/services/balanceserv/reserve_serv/stor/

current forward capacity market.[31] (The proposal was dropped in the final 2006 market design apparently in response to stakeholder concerns about complexity and market liquidity.) Interestingly, PJM considered such an approach beneficial despite the fact that variable renewable production in PJM was relatively modest at the time. Because of this the specifications put forward by PJM for the capacity tranches may be less stringent than would be the case today. Nonetheless PJM specified four categories of resources – dispatchable (i.e., rampable), flexible cycling (i.e., fast and frequent stop-start), supplemental reserves, and everything else – and proposed to clear the capacity market in stages based on the desired quantities of each type of resource. PJM has subsequently instigated short-term markets targeted at these capabilities but has yet to revisit formally the possibility of an investment timescale market mechanism.

It is not known whether PJM will revisit this concept, though it is perhaps instructive that PJM has recently adopted a three-tranche structure in place of what was previously a single-clearing-price auction for the demand response portion of its forward capacity market, with encouraging early results.[32] ISO New England has recently published a conceptual proposal[33] for addressing reliability challenges they are facing despite nearly a decade of experience operating their forward capacity market.[34] While ISO New England today has a more than sufficient quantity of firm capacity, they have pointed to a lack of needed flexibility in their current resource portfolio as a looming reliability issue for the region. Their forward capacity market does not currently provide for differentiated valuation of firm resources, and while ISO New England is currently the only North American ISO that conducts a forward operating reserves market, it involves a look-forward of only one year (far shorter than the capacity market) and covers only one of the flexibility services likely to be of concern. The response to this challenge outlined in their recent proposal is to apportion their forward capacity auction into several tranches based on specified resource capabilities. The timing of this deliberation has been precipitated in part by the expected retirement of a number of older supply resources, raising the question of what criteria ISO New England will be able to deploy in attracting the new resources eventually needed to replace them.

It is important to keep in mind that capacity mechanisms are not intended to provide additional revenues to system resources over and above what they would expect to earn in a properly functioning energy-only market. Rather they are designed to substitute a more stable, predictable stream of payments for capacity in place of a portion of the more variable, less predictable revenues that would otherwise have been earned through the sale of energy. With that in mind, the apportioned approach to capacity mechanisms described here allows market operators to differentiate the value of capacity payment streams available to system resources based on a set of critical operational capabilities. As a result more flexible resources can realize a higher proportion of their earnings from stable, long-term, predictable capacity (or "capability") revenues, which should afford them an overall competitive advantage over less flexible resources in the energy, capacity and ancillary services markets.

This approach avoids the trap of segregating capacity resources based on criteria that have no tangible reliability rationale (e.g., new vs. existing resources, or "strategic reserves" vs. all other firm capacity). Such measures inevitably

---

[31] See "Whitepaper on Operational Reliability Metrics for Generating Capacity in the Reliability Pricing Model" (January 2005), accessible at http://www.pjm.com/~/media/committees-groups/working-groups/pjmramwg/postings/whitepaper-rpm-reliability-metrics.ashx.

[32] See report at http://www.pjm.com/~/media/committees-groups/committees/mrc/20110818/20110826-brattle-report-second-performance-assessment-of-pjm-reliability-pricing-model.ashx.

[33] "Using the Forward Capacity Market to Meet Strategic Challenges" (11 May 2012) at http://www.iso-ne.com/committees/comm_wkgrps/strategic_planning_discussion/materials/index.html

[34] ISO New England currently has a comfortable market-wide reserve margin; what appeared in the early part of the last decade to be a looming resource adequacy crisis, particularly in its southwestern zones, has so far been effectively managed.

distort wholesale energy markets and, perhaps more importantly, most often fail to track the resource attributes that will increasingly be needed to deliver least-cost reliability as shares of variable production rise. Instead, a capacity mechanism that is apportioned in a manner similar to that described in the attachment, or as proposed by PJM in 2005 and by ISO New England in May 2012, properly rewards all firm capacity for its contribution to meeting resource adequacy requirements, but *only for the undifferentiated value of firm capacity.* Firm resources that contribute additional operational reliability benefits to the system – whether they be new or existing, supply- or demand-side resources – have the opportunity to clear the market first and earn higher capacity payments. In so doing, they drive less flexible resources to the margin and reduce the revenues such resources are likely to receive via the capacity mechanism. In addition, by allowing all resources to participate this approach not only avoids distortion of the wholesale energy market, but it also encourages access to lower cost opportunities that may exist to invest to gain additional flexibility from existing supply resources and in demand-side flexibility options.

This approach has the benefit of relative simplicity in those cases where a capacity mechanism is in place or under development. It also offers a more natural vehicle for valuing and deploying those non-traditional capabilities described under the enhanced services option by creating an auction process in which "whole resources" may compete. It may therefore be more likely to be adopted in markets facing an imminent need to add a significant quantity of new firm resources. At the same time, however, an apportioned forward capacity auction offers less flexibility in tailoring specific services and provides less opportunity for mid-course corrections. It also may not obviate the need for the system operator to adopt separate long-term mechanisms for certain more traditional ancillary services. But in many markets this approach will represent a more straightforward vehicle for increasing the flexibility of the non-renewable resource portfolio.

3) Common Considerations

In both of these approaches there is obviously a question of how many different capability products should be specified and procured. Accommodating the full range of ancillary services and a representative range of ramping, cycling and storage products might be the ideal, but the resulting market fragmentation could significantly compromise market liquidity and be prohibitively complicated to administer.

Fortunately, experience and a number of recent analyses are converging on a consensus view that demand for the shortest-duration ancillary services is not significantly impacted by high shares of variable renewables.[35] The categories of resource capability likely to be in short supply can be described broadly as:

- a "flex" option – the ability to shut down and re-start, or cycle, a resource multiple times within a reasonably short window of time and up to hundreds of times over the course of the year;
- a "dispatch" option – the ability to reduce a resource to a low level of stable operation and ramp it back up at a specified rate, not in a traditional operating reserve role but as a normal-course ramping capability; and,
- secondary reserves for regulation and load-following to address issues arising in the tens of minutes (e.g., forecasting error).

As mentioned earlier, the essential parameters at issue are how fast, how far and how frequently a resource can be started and stopped or ramped up and down (or, in the case of a demand-side resource, down and up) both within scheduling intervals and across multiple scheduling intervals.

---

[35] For a survey of recent analyses see "Meeting Renewable Energy Targets in the West at Least Coast: The Integration Challenge" (Various authors, 10 June 2012), prepared for the Western Governors' Association by the Regulatory Assistance Project. See also "Accommodating High Levels of Variable Generation" (NERC, April 2009); "How Do High Levels of Wind and Solar Impact the Grid? The Western Wind and Solar Integration Study" (Lew of NREL, Piwko et al. of GE Energy, December 2010); and "Cost-Causation and Integration Cost Analysis for Variable Generation" (Milligan et al. of NREL, Clark et al. of DoE, June 2011).

Consider two realistic examples. Envision a system with gross demand on a winter day ramping from 15,000 MW to 30,000 MW between 05:00 and 08:00. Add a large share of wind generation on the system, with wind production on the day ramping down to near zero in the morning. The ramp in *net* demand could be 2,000 MW to 30,000 MW during that same time period. This calls for much more and much steeper ramping capability than was previously needed. Now imagine the same system with gross demand on a summer day ramping from 10,000 MW at 05:00 to 40,000 MW at 18:00 and then ramping down over the rest of the day. Add a large share of PV generation, with PV ramping up from 07:00 but at a slower pace than the ramp in gross demand, peaking at 14:00 and then ramping down to zero by 19:00. Net demand would therefore peak once in mid-morning, subside dramatically during mid-day, then peak again in late afternoon/early evening. This creates a need for a significant share of the non-renewable resource portfolio being capable of starting up quickly, shutting down and starting up again within the space of only about nine hours. Operating a power system reliably and cost-effectively under these conditions is entirely feasible with commercially available supply- and demand-side resources, but it requires a different mix of such resources than in the past. Wholesale power markets will need to project and value these emerging investment needs more clearly than they do today.

While a wide range of such services could be specified in any market, the reality is that there are only a limited number of things even the most flexible supply resources can do well day in and day out. It would also be counterproductive to confront manufacturers with a profusion of inconsistent design requirements. Demand-side and storage resources may offer a wider range of flexibility options and under the enhanced services approach to long-term capabilities markets it may be practical to tailor a wide range of services markets to exploit any that prove to be cost-effective.

Under the apportioned capacity mechanism approach it is unlikely that the incremental value of specifying a wide range of resource tranches, or alternatively creating overly prescriptive specifications for a more limited number of tranches, would justify the costs in complexity and reduced liquidity. Furthermore, for traditional ancillary service capabilities that are already the target of short-term market mechanisms it may be preferable simply to adapt those mechanisms to long-term services auctions rather than try to address the issue via the capacity mechanism.[36] In the case of the apportioned capacity mechanism, therefore, the imperatives of preserving an acceptable level of market liquidity and minimizing administrative complexity dictate that market operators and stakeholders settle on a limited number of carefully specified resource tranches. For an instructive example we can look back to the categories addressed in PJM's 2005 proposal: one corresponding to the flex option described above (which could include energy storage options); one corresponding to the dispatch option; and one corresponding to secondary reserves. The remaining tranche of capacity would be open to all firm resources. Firm demand-side resources, including demand response and end-use energy efficiency, will still offer an attractive range of capabilities and must be permitted to participate equally in any tranche for which they qualify.

Under both the enhanced services and forward apportioned capacity market approaches, there will be important design parameters to consider for the forward period and commitment period. By "forward period", we refer to period of time between the auction and when the cleared resources (winning bidders) must deliver the services or capacity they committed to make available in the auction process. By "commitment period" we refer to the length of time (or term) over which they will receive the auction-clearing price as a fixed payment.

---

[36] For example, see the above-referenced ISO New England proposal "Using the Forward Capacity Market to Meet Strategic Challenges" at page 7: "In the case of 10- and 30-minute products, the ISO is evaluating whether it is preferable to reflect those needs in the FCM or continue to reflect them in the Forward Reserve Market. If the current mix of installed capacity is significantly deficient in these capabilities, then it may be desirable to reflect the needs in the capacity market to better time the price signal with the build decision, as well as providing a five-year capacity price guarantee. However, if there is sufficient capability but there are insufficient incentives for providing and maintaining that capability, then it may be preferable to maintain and modify (as needed) the current Forward Reserve Market construct."

While many permutations are possible, it is important to keep in mind that (i) auctions will be held on a regular basis providing current and prospective investors a rolling forward view of market supply and demand; (ii) these proposed mechanisms are designed to incentivize investors to provide what the market needs rather than remove development and investment risk altogether; and (iii) the objective is to give market participants a reasonable opportunity to respond to demand for critical resource capabilities. Experience suggests that annual or semi-annual mechanisms with forward periods of three to four years out from the auction and a range of available commitment periods from one to eight years (at bidders' option) would provide adequate liquidity and afford transparent and sufficient investment signals for both supply- and demand-side resource alternatives.

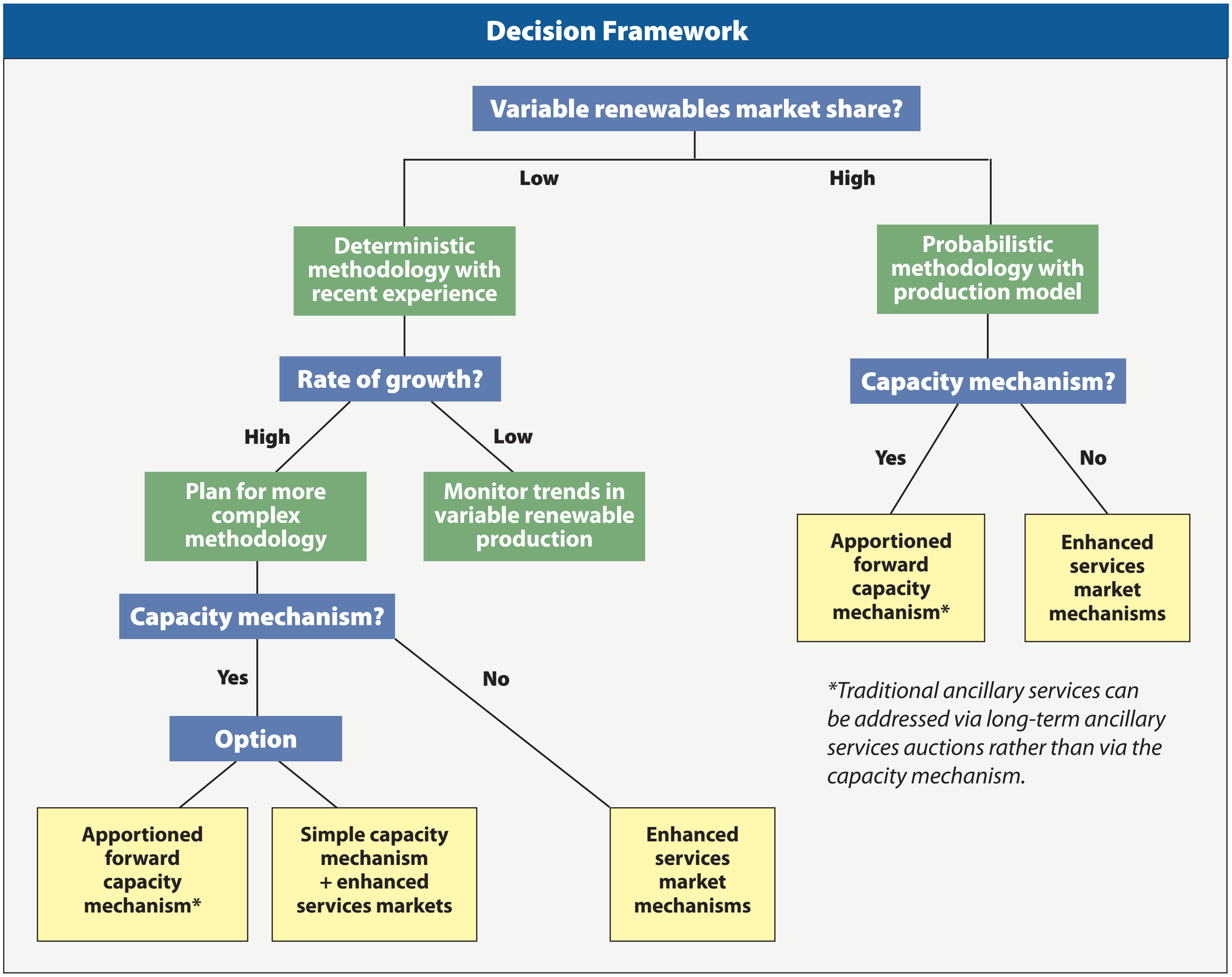
4) Progressive Improvement

Over time we may aspire to a more ideal reliability market design better able to incorporate cost/benefit trade-offs dynamically. One conception of such an ideal design (though surely not the only one) would start with a highly sophisticated Monte Carlo-based forecasting model for gross demand and net demand. On top of that forecast would be added a layer of analysis of the costs associated with issues like suboptimal dispatch and renewables curtailment and estimates of the costs to obtain the desired flexibility services. The model would be run to establish an initial estimate of the quantity of resources and the mix of resource attributes that would deliver a cost-optimized reliability solution. A resource auction would then be conducted and the market re-modeled with the resources that clear the auction. If the auction results in costs in excess of benefits the auction is re-run with the inputs reset to reflect the results from the first round. This process is iterated as many times as necessary to find equilibrium between the cost of additional flexibility and the consequent benefits. This is obviously complex, and it may prove so burdensome to whole classes of market stakeholders (e.g., demand response providers) that it would not be feasible in practice. But it provides one possible formulation of a theoretical ideal against which the implementation of more immediately practical market designs could be evaluated.

VIII. Wrap-Up

To reduce these recommendations and key observations to their simplest form, we present below a decision framework for creating long-term visibility to the value of flexibility discussed in Sections VI and VII together with key summary points:

Figure 5



**Traditional ancillary services can be addressed via long-term ancillary services auctions rather than via the capacity mechanism.*

1) **Get the broader context right:** While there will inevitably be a need for a greater degree of flexibility in the resource portfolio, the operational challenges posed by growing shares of variable production can be substantially mitigated through a number of valuable and relatively low-cost measures. These include:

- shorter scheduling intervals;
- aggregation of larger balancing areas;
- transmission investment to mitigate internal congestion;
- increased interconnectivity with neighboring markets;
- more accurate weather forecasting; and
- demand enabled to be more responsive to uncontrollable changes in supply.

2) Ensure that existing markets are designed and operated to **extract all cost-effective flexibility services available from all existing resources.**

3) **Ensure that all qualifying demand-side options are fully able to participate** in these markets, both directly and through aggregators.

4) Reliability operates on two dimensions – resource adequacy and system quality. Historically it has been possible to address these two faces of reliability at two separate timescales – resource adequacy in investment timescales and system quality in operational timescales. **Going forward, large shares of variable resources will drive system quality factors into investment timescales** as well.

5) Along with a forecast of gross demand, establish a procedure for forecasting variable resource production and combine the two to **derive a net demand forecast.**

6) Use net demand forecasts to **assess on a periodic basis the demand for critical flexibility services** relative to the range of capabilities available from all system resources to provide such services, just as gross demand forecasts are used periodically to measure the demand for firm capacity relative to the total quantity of firm resources available.

7) **Establish a methodology for setting the maximum value** to the system of each additional increment of capability up to the target quantity.

8) Depending on current and expected market conditions, **adopt either a simple deterministic approach based on recent experience, or a more complex probabilistic approach** using production modelling.

9) The **desired resource capabilities can be procured through either enhanced services markets or apportioned forward capacity mechanisms, depending on the individual market circumstances.** In either case, where there is a need to expand the capability to provide traditional ancillary services it may be most practical to do so via a long-term adaptation of existing ancillary services markets.

10) As shares of variable resources grow, experience is gained and institutional capacity is built, **markets may want to evolve over time toward a reliability market structure better able to incorporate cost/benefit trade-offs dynamically,** but this level of complexity is unlikely to be necessary or feasible in most markets in the immediate future.

As indicated in the Introduction, this paper represents a "straw man" proposal intended to build upon the body of work RAP has been developing along the "beyond capacity markets" theme. Power point presentations and related papers are available on our website (www.raponline.org) including the companion paper referenced in footnote 7. We encourage dialogue and feedback on this paper as it is intended to stimulate productive discussion on this topic. Please feel free to contact Mike Hogan (mhogan@raponline.org) or Meg Gottstein (mgottstein@raponline.org) with your comments or questions.

Appendix A

Comparison of "Prototypical" Capacity Mechanism vs. Capabilities Mechanism

The following graphs depict a hypothetical power system with a resource adequacy target of 100 GW of firm resources. Based on net demand forecasts it is also determined that the optimal apportionment of resources is at least 20 GW of flexible cycling resources (which could include energy storage services); at least 60 GW of highly dispatchable resources; with the balance (up to 20 GW) coming from firm resources of any type. Under this hypothetical, Figure 6 below depicts a prototypical capacity mechanism that addresses only the gross quantity of firm resources.

Figure 6

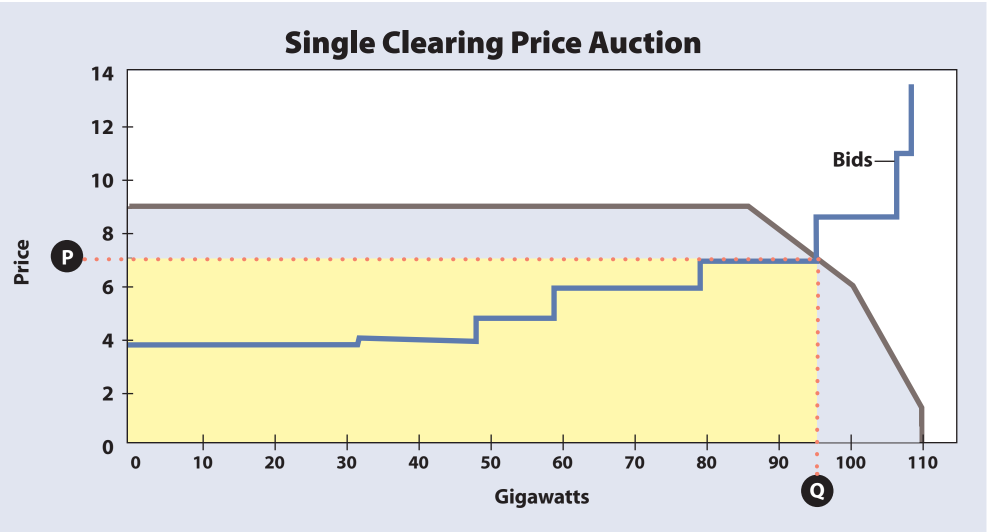
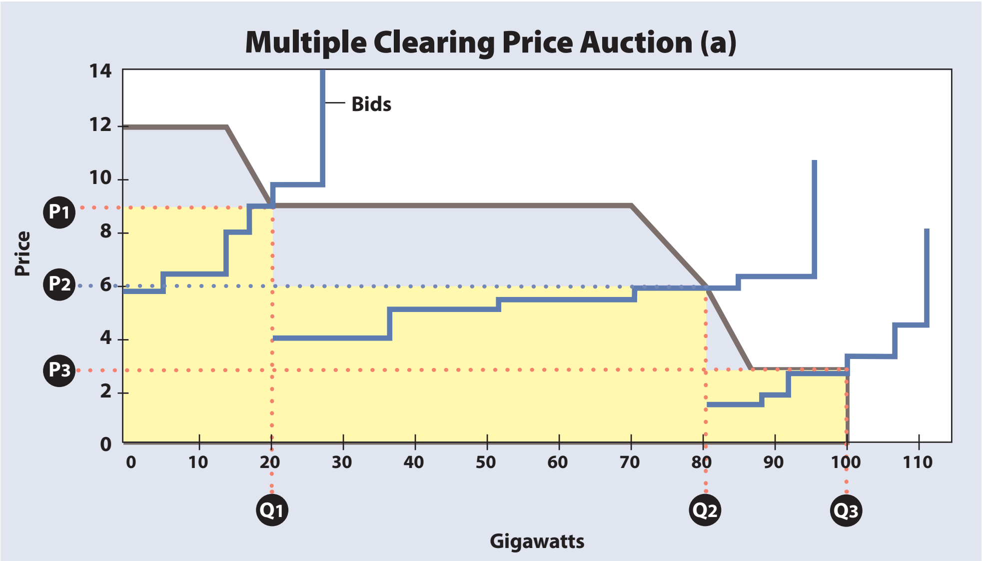


Figure 7



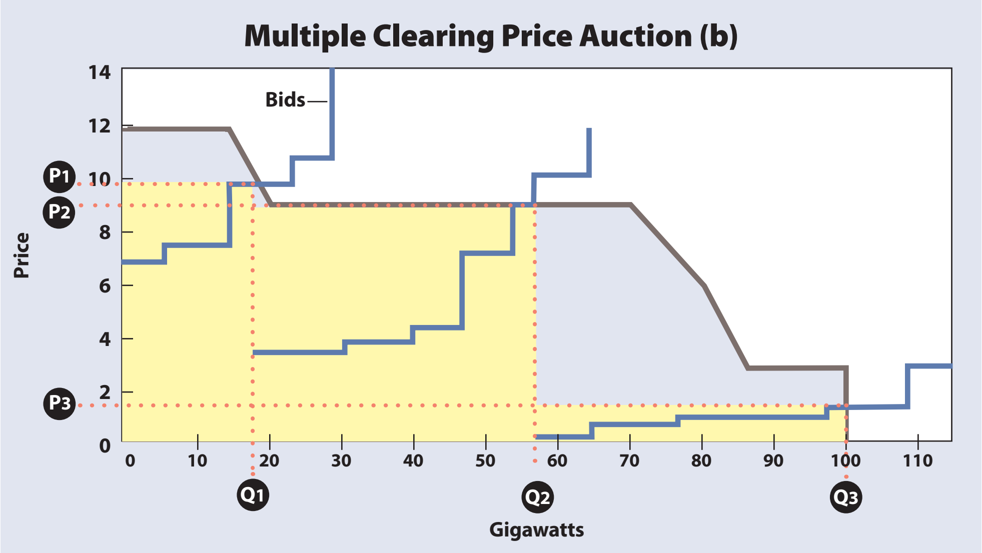
The value of capacity (the "demand curve" depicted by the dark grey line) is set based on the net cost of a new resource which is the all-in cost of a hypothetical new entrant (for example, a conventional gas-fired CCGT) minus the value such a resource would expect to realize from sales of energy and ancillary services. In this example, at the target quantity (100 GW) the maximum price set by the demand curve is 6, which is 100% of the net cost of a new resource. To the extent that less than 100 GW of firm resources are offered at or below the target price of 6 the clearing price (i.e., the price P at which the supply curve ("bids") crosses the demand curve) would gradually rise along the demand curve, up to but not beyond a predetermined level at which it is expected there would be adequate financial incentive for new investment in firm resources. (In the example depicted above the market clears at quantity Q=95 GW at a clearing price P=7, giving a total cost of 665.) Because the clearing price is higher than the net cost of a new resource there is a financial incentive to add new resources. As new resources are added to the system the market should clear closer to or at the target quantity and price in subsequent periods. (The mechanism also recognizes the value of buying additional resources if they are offered at sufficiently low price, with a price cap that declines steadily to zero.)

Figure 7 depicts the same system with an apportioned, three-tranche capacity mechanism. Rather than a uniform clearing price for all capacity resources, three different values for resources are determined based on the capabilities of the various resources offered. Under this variation, the mechanism operates on each tranche individually in the same way the prototypical capacity mechanism depicted in Figure 6 operated on the entire resource portfolio, but with some important differences. While this variation also recognizes the value of additional resources offered at sufficiently low prices, it does so only with respect to the first two (flexible) tranches. The first tranche (the most flexible firm resources) will clear when the bids cross the demand curve, wherever that occurs. The second tranche (the next most flexible firm resources) will then clear in a similar fashion. As long as the aggregate quantity of first and second-tranche resources totals at least 80 GW the mechanism will clear only enough third tranche resources (the least flexible firm resources) to fill the gap to 100 GW, regardless of the price offered.

Should the aggregate quantity of first and second tranche resources cleared be less than 80 GW the third tranche would be expanded as necessary to clear up to, but no more than, 100 GW. Should the aggregate quantity of first and second tranche resources cleared be more than 80 GW the quantity of third tranche resources procured will be reduced as necessary to meet the 100 GW target. In the event that less than 80 GW of first and second tranche resources clear the market, the availability of higher price caps for resources that can meet the specifications for the first or second tranches should provide sufficient incentives for investment either in existing resources to meet the required specifications, or in new flexible resources.

In the example shown in Figure 7 the mechanism procures exactly the target quantities in each tranche. In the example in Figure 6, were the market to have cleared at exactly the target quantity the total cost would have been 600 (P=6 times Q=100). In Figure 7, with all tranches clearing at the target quantities, the total cost is also 600 ((Q1 x P1) + ((Q2 - Q1) x P2) + ((Q3- Q2) x P3)). While the total costs of the two "on target" results are the same, in the latter case the payments are heavily skewed toward resources with the desired capabilities. The clearing prices in the first two tranches (P1 and P2) are correspondingly higher for the resources successfully bidding in those tranches and very low (P3) for the "all other resources" third tranche. This is a key feature of a sequenced auctioning process based on resource capabilities, and it is consistent with the desired outcome: To reflect more accurately the lower market value for inflexible resources currently on the system or planned for the future, and at the same time to reflect more accurately the higher market value of those resources with the flexible capabilities required to effectively integrate a growing share of renewables.

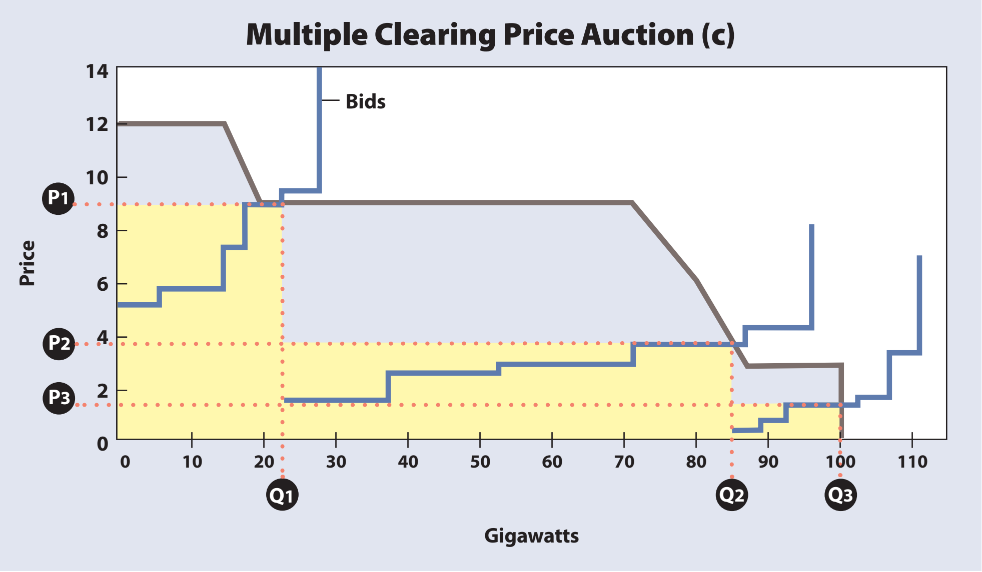
Figure 8



Figures 8 and 9 illustrate how the apportioned capacity mechanism works when the auctions for the first and second tranches do not clear at their target quantities. Here too, one can see how the mechanism prioritizes those resources with more valuable flexibility capabilities relative to those possessing the least.

In Figure 8, the first tranche clears 18 GW at a price of

Figure 9



10, the second tranche clears 40 GW at a price of 9 and the third tranche fills out the balance (in this case 42 GW) at a price of 1.5. In this period the total paid is 603 (only slightly less than the "on target" example), but there is significant incentive to add more second tranche resources. A uniform-price mechanism would have resulted in a range up to 810, depending on the aggregate quantity of resources offered and the price ranges, and while it may have offered an incentive to add new firm capacity it would have provided no financial incentive to re-align the resource portfolio.

In Figure 9, the first tranche clears 22 GW at a price of 9, the second tranche clears 63 GW at a price of 4 and 15 GW of resources are purchased in the third tranche at a price of 1.5 to fill out the 100 GW target. In this case the total paid is 472.5, with a surplus of attractively priced flexible resources being available over and above the targets.