

[WORKING DRAFT]

On Market Designs for a Future with a High Penetration of Variable Renewable Generation

Submitted to the U.S. Department of Energy Future Markets Workshop, September 8, 2017

Eric Gimon, Ph.D.

Senior Fellow, Energy Innovation LLC

Abstract:

Restructured wholesale electricity markets serve stakeholders by providing three principle services: economically efficient real-time dispatch, incentives for resource adequacy and long-term cost recovery. Because at least two of these functions happened dynamically through market forces, future scenarios for clean, affordable and reliable power with a high fraction of variable resources challenge today's market design. This paper presents two possible paths for future designs to continue serving these three functions and principles for policy-makers working on future market design.

Introduction and Motivation

This paper asks the question of what kind of wholesale electricity market design can adequately support a clean, affordable and reliable power system with a major fraction of electricity (70%+) coming from variable renewable electricity (VRE) generation like wind, solar and run-of-river hydro. This is a big ask on our imagination and future vision, as today the most advanced systems top out around 33% VRE and are connected to other systems with much lower fractions that they can lean on. High penetration systems are unlikely to occur before 2040, so we still lack a lot of information about which technologies will be available and how much they will cost. More importantly, without any experience we cannot yet draw on lessons that could help on the way to higher penetrations.

Yet, it seems important to ask the question nonetheless. Market operators, regulators, participants, and other stakeholders are already having to work hard on existing markets to adapt to current and near-term changes in the grid in a constructive way. Given the likely role of low-cost variable renewables in an *affordable, reliable, and clean* future power system, it would be extremely helpful to have some vision of how they fit in such a system. Policy-makers could then focus on the most important policy changes for the long-term and avoid getting stuck on dead-end near-term fixes. The purpose of this paper is to sketch possible futures for market designs that can accommodate high fractions of variable renewable energy in service of the goals above. We do not seek to identify an ideal market design for high VRE penetration power systems, but rather look to point out some of key principles at play in redesigning and adapting today's wholesale electricity market designs (the restructured markets) that can best inform policymakers going forward.

The paper begins by looking at some key elements of power markets today, and describe a key tool for understanding these – the price-duration curve. Next, it describes how high penetration of renewable generation and its cousin the rise of distributed resources will create challenges for the current construct. The paper then identifies two different families for a viable market design, defined as a stable market providing incentives for resource adequacy and long-term cost recovery while driving economically efficient dispatch decisions. It then ties things off by discussing how these families of

solutions relate, touches briefly on how they might emerge out of today's markets, and then highlights key lessons for future market designs.

Today's Markets

Today's grid is a system of long-lived assets, power plants as well as poles and wires, for delivering electric power to end-consumers. Historically it was managed by set of vertical monopolies which mostly provided electricity by burning fuel. Their main concerns were providing a sufficient, affordable supply of electricity year-round by maintaining sufficient but not excessive resources to meet peak demand and reserves to manage predictable variations and occasional contingencies. In this traditional model, electricity rates are set to recover long-term investments and pass through fuel costs. The main advantages from this model are a stable investment environment for financing necessary capital intensive assets and accessing economies of scale.

In the US since the 1990's, however, a large swath of the country seeking to improve the economic efficiency of the grid and shift financial risk back onto generators has restructured its wholesale electricity markets. These markets have changed the meaning of resource adequacy to include a stronger focus on flexibility in ways that are very relevant for variable generation and their future is the subject of this paper.

The key features of a restructured wholesale electricity market are an independent system operator managing a non-discriminatory, open-access transmission network and a competitive energy-market (spot market) that together guide grid operators in dispatching connected assets¹ to maximum efficiency. Price signals from the market, e.g. generator offers for the power they can commit at a given price, allow grid operators to do least cost dispatch – also called security-constrained economic dispatch (SCED) – to minimize cost to consumers. This optimization in a competitive and transparent market does an excellent job of setting near-term prices so that all the dispatched assets recover their marginal costs and consumers pay as little as possible, performing one of its main functions. But the other two functions, providing incentives for resource adequacy and long-term cost recovery, require the market to provide a mechanism for generators and other grid assets (like storage or demand response programs) to recover their fixed operations and maintenance costs, including paying back their investors.

The ability for market forces to get prices “right” over the long term is an emergent property of market design, and cannot be taken for granted as the market shifts to a new, drastically different resource mix. For adequate resources to exist and long-term cost recovery to happen, an equilibrium must emerge which delivers the right mix of resources to match the whole gamut of supply and demand conditions at the lowest long-term cost. Simplistically, this equilibrium around the “right” long-term prices emerges because when prices are too high new resources are built, and when prices are too low inefficient resources retire.

Long-term cost recovery implies a pattern of prices where during substantial intervals, revenues need to exceed short-term marginal costs for most generators. Generators barely being paid over their marginal costs in some intervals depend on more expensive units to set the higher prices in other intervals. At

¹ Ancillary markets are also an important feature of restructured markets but for simplicity's sake we will avoid them for now.

equilibrium, resources accumulate sufficient economic rents when prices exceed marginal costs to cover long-term costs.

Restructured markets in equilibrium deliver resource adequacy through a slightly different price signal. Constraints like peak demand or other shortages, ramp rates and startup times drive higher prices which then attract resources. The highest prices are often referred to as “scarcity prices².”

Restructured markets have changed the meaning of resource adequacy to include a stronger focus on flexibility through the emergence of resource adequacy from scarcity pricing combined with the drive for economic efficiency. Flexible resources are resources that can maintain a high availability and can quickly change their output in so that they can collect maximum revenues when prices spike. The more economically efficient the fleet becomes the less likely an unnecessary resource is going to be available during some time interval. It might be turned off, be on maintenance or be providing some other function. But as markets increasingly provide exactly the resources needed to cover typical conditions in each interval, more flexibility is required to cover variations from interval to interval as well as deviations from expectations. Flexibility becomes a core component of resource adequacy and reliability. The advent of variable resources like wind and solar only increases the need for flexibility as intervals with scarcity pricing become more disconnected and variable.

In summary, apart from providing the right price signals for the optimum dispatch of available resources, wholesale electricity market designs must provide prices that allow for long-term cost recovery for all resources and pay for system reliability – especially flexible resources.

The Price-Duration Curve

The fact that long-term cost recovery and resource adequacy are only emergent characteristics of restructured markets, not a direct result of market design, exposes one of their main weaknesses. Structural changes in the market, like new disruptive technologies, policy changes, or large swings in relative fuel costs, can disrupt the market equilibrium. For a future system with zero marginal cost variable generation driving flexibility needs and pushing down prices, the question becomes: will a new stable equilibrium manifest under current market structures?

A key tool for answering this question is the **price-duration curve**, which charts energy prices sorted from high to low during the year. In today’s markets this curve is shaped by the marginal cost of generation – prices in each period almost always set by a generator’s fuel costs. As examples, Figure 1 below shows price duration curves for the Houston zone in ERCOT for 2014, 2015, and 2016.

² The term scarcity pricing is sometimes used to describe administratively set prices (or price caps) during times of system stress. Here we use it in the broader sense of “high prices occurring when resources for matching supply and demand become scarce.”

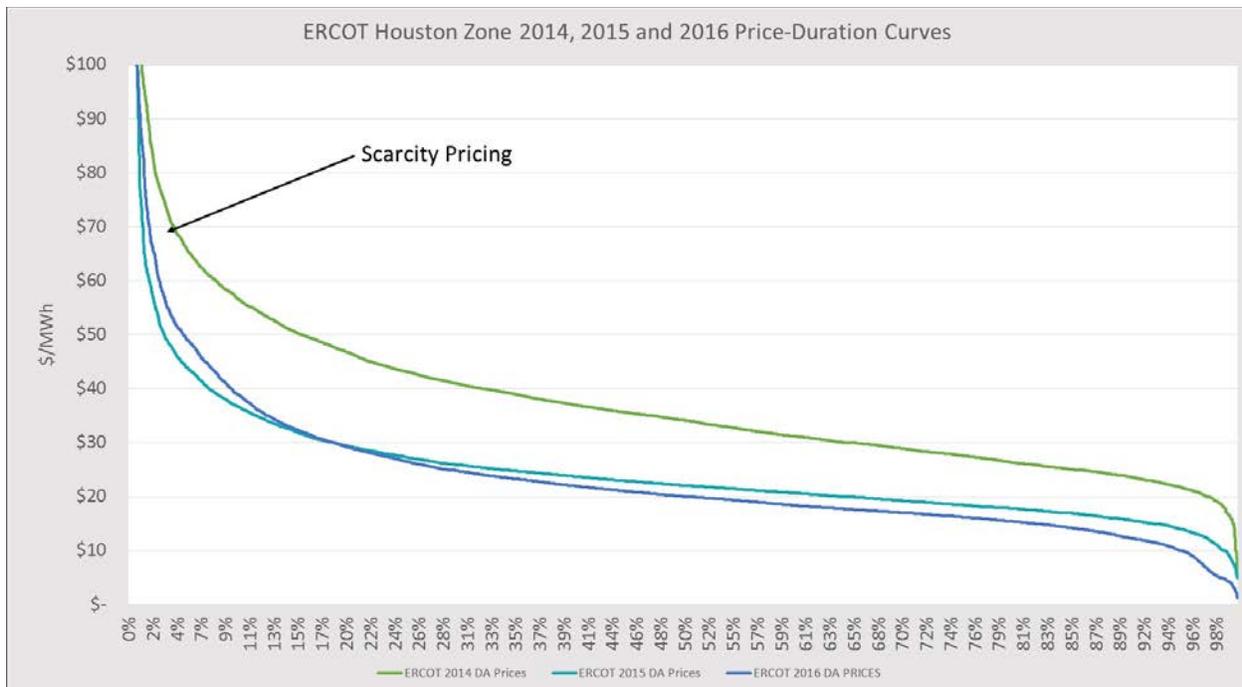


Figure 1 - ERCOT Houston Zone Price-Duration Curves (day-ahead prices)

This price duration curve has two key features typical of today’s markets: a sharp rise on the left during hours of scarcity pricing and a gradual but steady decline through the middle which then drops off on the right.

For the purposes of long-term cost recovery and grid management the gradually declining middle feature (e.g. 30-70% of hours on the x-axis) is very important. In any given interval addressed by the energy market, there is a mix of resources available for the system operator to dispatch to meet load. Today, almost all these resources are on the supply side, and they can be arranged in order of cost along a supply curve showing marginal cost as a function of demand. It is mainly because this supply curve rises continuously and evenly as a function of typical demand, reflecting a variety of fuel costs and generator efficiencies, that we see a continuous and steadily declining slope in the middle part of the price duration curve.

This prosaic feature of the price-duration curve is the key to long-term cost recovery. As discussed above, generators depend on marginal prices being higher than their offer bids to generate economic rents. The size of these rents depends on how steep and broad the price-duration curve is. For example, imagine a natural gas fired generator whose offer price in the Houston Zone is accepted (clears the market) 60% of the time in 2014. Its marginal costs in the \$30/MWh range but 30% of the time prices are over \$40/MWh, 15% over \$50, 5% over \$66/MWh and so on. Looking at the area under the curve (minus short-term marginal costs) we can

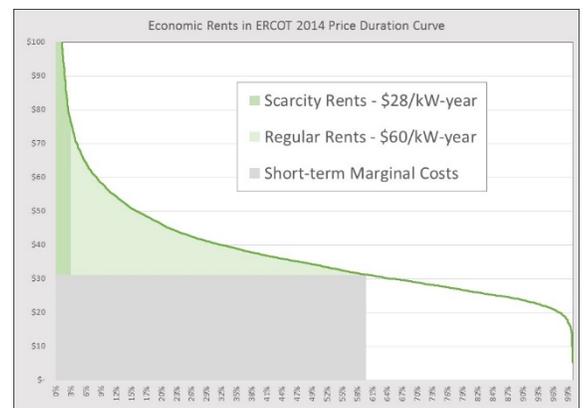


Figure 2 Calculating Annual Net Revenue (Rents) for a Sample kW of Generation.

calculate that it collects about \$28/kW-year in scarcity rents from the first 250 hours and \$60/kW-year in regular rents from the middle part of the curve.

To better understand the importance of the shape of the price-duration curve, consider the net revenues for the same gas plant in 2015. Because its natural gas fuel is significantly cheaper, its marginal costs have gone down (part of the cost drop seen in the chart). But prices 30% of the time are now only \$5 higher (instead of \$10), \$11 higher (rather than \$20) 15% of the time, \$24 higher (rather than \$36) 5% of the time. Now the scarcity rents are still quite similar at \$26/kW-year, but the regular rents have dropped to \$36/kW-year – the milder decline in the central slope can really be felt.

The continuity of the price-duration curve reflects another important feature of the underlying system. Its smoothness implies that for a given normal demand pattern a least-cost dispatch exists with plenty of nearby solutions at a slightly higher marginal cost. The algorithm which runs the SCED can pick out a unique set of dispatch orders that minimizes system cost; in engineering parlance, the total dispatch cost is a good “objective function” for selecting an optimal dispatch solution. The solution can easily be varied to accommodate slight changes in load and generator availability, and this variation will not change prices much.

Turning now to scarcity pricing, the other prominent feature of the price-duration curve, we see that roughly 3% of the time prices are more than double the average. During these hours there are fewer solutions for the SCED algorithm to find and thus small variations in supply and demand can make prices shoot up or down. For some suppliers, net revenue (revenue minus variable costs) from this segment of the price-duration curve is significant, e.g. covering 30% of long-term cost recovery needs for a combined cycle plant running at a 60% capacity factor modeled above, but can vary significantly from year to year based on system conditions. For peaker plants like gas turbines scarcity pricing is the main source of revenue, while plants that run more of the time tend to derive a smaller fraction of net revenue from scarcity pricing, but are still happy to collect revenue during these hours.

The link between scarcity pricing and resource adequacy is very strong. By supporting peaker plants, for example, high prices help maintain enough capacity to meet peak system demand. However, system stress and its related shortage pricing don't just happen at predictable times like sweltering summer days. Unusual weather conditions and other surprises while some units are down for maintenance can lead to tight supply conditions and shortage pricing as well.³ Flexible units are in a better position to collect rents from soaring or even just elevated prices. Hence, scarcity pricing is an important way to reward flexibility, one of the principle constraints we will be solving for in the future.

In summary, the “scarcity prices” in the high end of the price-duration curve help wholesale markets maintain resource adequacy. Because these rents depend on extreme conditions which can be very different from one year to the next, this is a volatile source of annual revenue. Resources that run most of the time collect these revenues but in general depend on the economic rents reflected in the central gradient for most of their long-term cost recovery. This is a more stable revenue stream, except, of course, when structural changes in the market disrupt the market equilibrium associated with a given supply curve.

Challenges for Future Markets

³ See e.g., E. Gimon. *How a Cold Day in Texas Exposed the Value of Grid Flexibility*. Greentech Media. April 2017.

One advantage of future fuel-free or fuel-light grids is that markets and consumers will no longer have to adjust to big swings in fuel commodity prices. But lower gas prices not the only causal factors behind the changes price-duration curves and generators' net revenue woes. For example, Texas has seen remarkable growth in wind generation, with wind providing respectively 11%, 12%, and 15% of total ERCOT generation in 2014, 2015, and 2016. Solar is also starting to come on strong. These variable resources are only predictable in the near to medium term, leading to less certainty about when system stress will happen. Even if these resources were completely predictable, their inherent variability – as when the sun goes down – also leads to bigger swings in net demand (actual demand minus variable generation) leading to less clustering of scarcity hours. In markets like ERCOT, the size and distribution of scarcity hours is changing to reflect the need for more flexible resources to accommodate cheap new variable generation.

New wind generation in Texas also affects long-term cost recovery by depressing prices when the wind blows (wind generation shifts the supply curve). By 2016 this last effect accounted for an average drop of around \$5-6/MWh in average wholesale electricity prices compared to a wind-less fleet.

The effects of wind and solar on the Texas fleet beget interesting challenges in the near term as markets adjust to these new entrants and struggle to maintain their emergent ability to provide resource adequacy and long-term cost recovery through market forces during a major energy transition. While these challenges are not the subject of this paper, they give us clues as to what challenges future markets will have to meet for a transition to a clean future grid to function properly.

A clean future grid with a substantial fraction of generation from variable renewable energy sources like wind and solar must confront the fact the fuel costs will no longer be the main driver of wholesale market pricing and dispatch. As a result, the wholesale market will need to find a way to finance assets with a cost structure dominated by up-front capital but the main means of doing this in today's markets depends mostly on a central gradient feature in the price-duration curve which seems predicated on having many market participants with varying fuel costs and fuel efficiency. Put simply, the first challenge for future markets is: if the marginal cost of electricity floats most of the time around zero, how will the market pay for the long-term provision of electricity?

A second challenge has to do with the use of SCED when most resources have no marginal cost and produce more as a function of the weather and natural cycles than an operator's signal. How does the grid operator know which resources to dispatch or curtail when there are more than enough resources providing an equivalent service at the same cost?

Finally, a clean, affordable and reliable future grid will likely feature not just variable renewables at the bulk transmission or "utility" scale but will involve a tremendous number of distributed energy resources (DERs). What will be the roles of these resources, especially the controllable ones? What price signals will they follow and how will they be dispatched?

Nobody has the answer to all these questions, and there is likely no unique answer in any case. Short of re-regulating wholesale markets, if visionaries for a clean energy future want wholesale markets to continue as a principle tool for economic dispatch and asset compensation they will need to develop some resolution for the issues above. From conversations with experts, regulators and practitioners in wholesale markets two broad paths forward emerged. Loosely speaking they fall under the rubrics of

“continue building on spot markets” and “bifurcate the long-term commodity electricity market from real-time balancing of the system.”

Path 1: Continue Building on Spot Markets

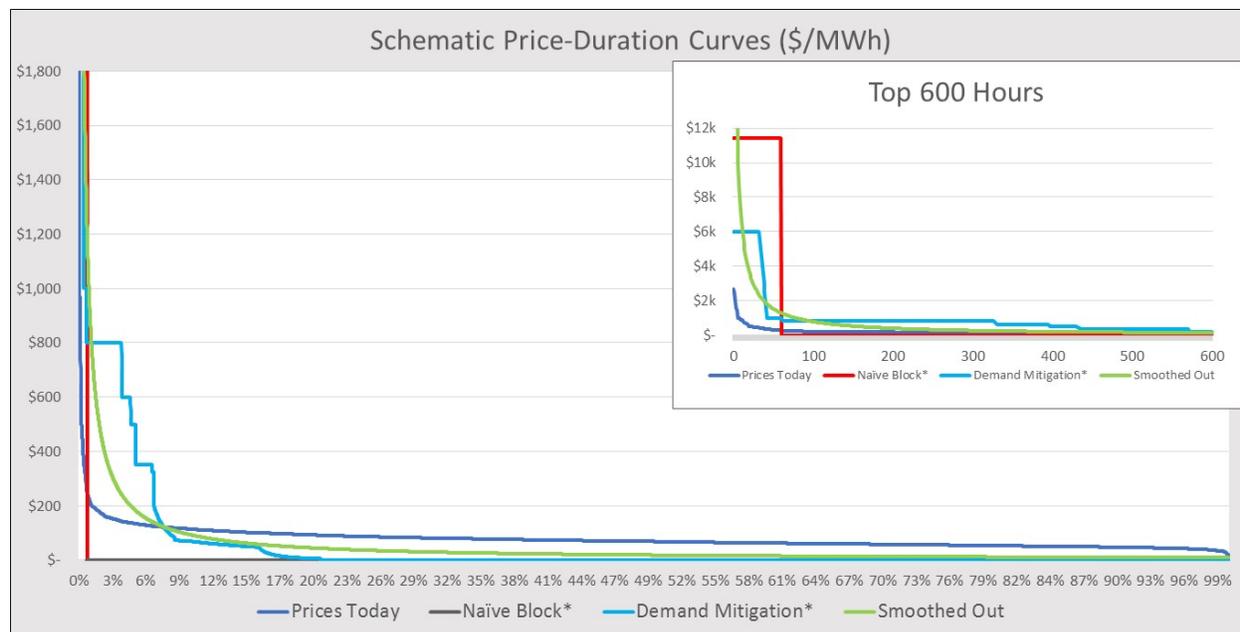


Figure 3 Examples of Potential Price-Duration Curves in Path 1 (* for curves from Jesse Jenkins et al.)

This approach advocates the continuance of a wholesale marketplace built primarily upon spot markets. Philosophically, the core task of the grid manager is to balance supply and demand for electricity in real-time; everything else follows. The master market is still a spot market regulated by derivative products⁴. However, this foundational spot market has price-duration curves that look very different from today (see Fig 3).

In the spot-market driven vision of the future, price-duration curves start by looking like a sharp step-shape (red line in Fig 3): scarcity pricing up to \$12,000/MWh for a few percent of hours and then a steep drop off to zero prices (sometimes even negative) most of the time⁵. There is no longer a steady slope in the middle and all long-term revenue recovery happens solely through scarcity pricing. Enough slight differences in supply offers presumably remain to resolve dispatch decisions.

There are some obvious issues with this vision:

- **Volatility:** Under today’s conditions scarcity rents already can vary a lot from year to year. In the scenario above, we can only expect more volatility as the system reckons with more variable resources. Using hedging tools like long-term bilateral contracts,

⁴ And with ancillary markets.

⁵ See for example, the June 2016 presentation by Jenkins, Sepulveda and de Sisternes: https://www.dropbox.com/s/o5moay9f2mask67/Zero Variable Cost Power Systems - IAEI Bergen - 06_20_2016.pdf?dl=0

revenue uncertainty can be managed, but this hedging will add to costs for buyers of electricity.

- Market manipulation: For scarcity prices to reach such highs, it is unlikely that they will reflect strictly marginal costs (unless carbon prices reach extreme highs), so how will offer prices be set? During times of scarcity the grid is likely to be under some type of stress and so large price swings can happen due to small swings in supply or demand. This creates an environment ripe for market manipulation, and the need for long-term cost recovery means you cannot use simple “fixes” like capping prices.
- What about optimal dispatch? During many hours of the year, the grid is likely to have large surpluses. In the simplest versions of this vision, prices get pegged to zero during any time of surplus. Since many different grid configurations can lead to surplus, how does least cost dispatch optimize fairly under such a situation?
- Cost Recovery for Variable Generation: Clearly dispatchable resources, or “flexible baseload⁶,” can arrange to generate during the few hours of scarcity pricing. However, the variable generation that makes up so much of the grid mix isn’t guaranteed to hit enough of those hours, will they be able to recover long-term costs enough to stimulate investment?
- What role for DERs? Today’s SCED is based on programs which must take into account the discrete operational constraints of generators. These are computationally intensive algorithms that do not scale well with an increased number of participants. It may be difficult to dispatch millions of distributed devices rather than just the hundreds of generators of today under a standard market design. Even if aggregators simplify the process, the increased number and variety of participants might still gum up the works for the SCED.

For certain, the spot-market based vision will have to expose load-serving entities or their customers to the possibility of sky-high prices. But the vision is not as dire as it seems at first.

Today in Texas, contrasting 2014 with 2015 and 2016 in Fig 1., we can see the big effects of natural gas fuel cost swings, and of new gas-fired generating capacity coming online, on the price-duration curve. However, generators can manage revenue collection risks with hedging contracts like long-term bilateral contracts to sell electricity at a fixed price. Derivative products like these, and even shorter-term ones like day-ahead obligations, play an important and often under-appreciated role in regulating today’s wholesale electricity market.

The possibility that exposure to another dozen hours of scarcity might double the price of annual purchases will provide plenty of incentive for load-serving entities and others directly exposed to the wholesale market to enter into bilateral contracts, providing a stable investment environment for suppliers. In an enhanced version of their role today, retail providers could essentially act as insurers for customers, insulating them from the wild swings of the wholesale markets but further structuring their rates so that customers and DERs are incentivized to mitigate the retailers’ risks in the wholesale market through their demand management choices at the distribution level.

⁶ de Sisternes, Fernando J., Jesse D. Jenkins, and Audun Botterud. "The value of energy storage in decarbonizing the electricity sector." *Applied Energy* 175 (2016): 368-379.

In general, more elastic demand can play a key role in mitigating some of the worst drawbacks of a spot market-based vision of the future market. If elastic demand can shift its consumption away from scarcity hours, the aggregate effect will be to restore a broader gradient to the price-duration curve, tempering the effects of variable generation on prices (see light blue line). If demand is more elastic, then the system is less vulnerable in times of system stress because of this extra flexibility. It is even less prone to market manipulation: Holding consumers hostage by withholding resources just results in less consumption and fewer rents.

Energy storage can also help on both the supply and demand side. On the supply side, bulk storage (or aggregated resources acting similarly) motivated to arbitrage high prices will shift load from hours of scarcity to hours of plenty. Or on the demand side, storage can act as a risk mitigation tool for consumers, improving demand elasticity. In either case the price-duration curve gradient broadens, system and price stability improve, and perceived investment risk diminishes. If enough liquidity in moving energy around between hours develops, we can perhaps hope to see broader and smoother price-duration curve (like the green curve) that allow for better cost recovery and more certainty in dispatch.

So, it seems that a spot-market driven wholesale market design could still serve society well, even under very different conditions from today. This requires loads to be properly exposed to scarcity pricing: i.e. no price caps or free-ridership when power gets scarce. Consumers and investors mitigate their exposure to risk through hedging tools like bilateral contracts or by finding sources of flexibility to fit their needs. The structural rules to prevent manipulation and other market failures are not completely clear to us, but it seems likely that policy-makers will have to place a lot of trust in the market to achieve a clean, affordable and reliable grid through this spot-market driven vision.

Path 2: The Bifurcated Market

The bifurcated market involves splitting the wholesale market into two parts: the firm market and the real-time spot market.

The traditional spot-market driven wholesale electricity market design is a good example of *marginal-cost based valuation*. The strength of this design comes from the fact that it closely links an *economic construct*, the spot market, to an underlying *engineering constraint* to match supply and demand closely in each real-time interval for a stable grid. Yet realistically, matching supply and demand is a really a differential concept, meaning that in most markets today available supply and demand in any interval is usually known well in advance, and it is small changes in supply and demand that most affect marginal prices. So, we could retain much of the value of spot markets by trading in only the last few adjustments up and down around a firm amount of generation and demand traded in a firm market.

The Firm Market: How would this firm market price electricity? In some sense, a firm market already exists today. Generators and loads negotiate long-term contracts in separate bilateral markets. In this proposal, however, long-term procurement is no longer strictly bilateral but happens in a more centralized long-term market for the supply and purchase of electricity (not capacity). This opens the door for the *holistic valuation* of electricity. Much as the SCED optimizes load bids and supply offers in real-time to find a least cost but reliable dispatch, here an independent algorithm acts as the market-

maker matching resources and needs in the long-term purchase of electricity. Current programs like Vibrant Energy's WIS:DOM model⁷ lead us to believe this is feasible.

The independent firm market would periodically accept offers of various length via contracts from existing or new potential generators as well as offers from new and existing storage and transmission projects and demand-side resources. It would take in granular bids from load-serving entities for when and how much electricity they need in in each period with potential modifiers for factors like weather. It would integrate all this information with a variety of weather and outage scenarios and deliver an optimal long-term solution-set and electricity price in each period. This is *not* an integrated resource plan for a return to regulated cost-of-service long-term purchasing, but rather a structured way to incrementally pool and divvy up resources to match with long-term bids – it is a long-term *market*.

This firm market could evolve in many ways, but here are some features we imagine it would have:

- **Sequential:** In each subsequent period, the firm market would re-optimize while considering existing commitments from and to both supply and demand, as well as performance against contracts (for performance, construction schedules etc.). After integrating all additional information, it provides a new long-term optimal build and dispatch solution and new long-term prices.
- **Risk Management:** The firm market allows various resources to pool risk and it optimizes for the least overall risk of not satisfying its obligations. It can effectively price some risk by accepting, for example, different length contracts from renewable energy developers at different price points. Risks that the market is not well setup to mitigate, like fuel delivery costs, might have to be hedged separately by suppliers. Because of the probabilistic nature of any long-term forecast, risk management and evaluation will likely be at the core of the firm market design.
- **Adaptive:** One way for the firm market to remain adaptive to changing conditions is for it to only sell long-term contracts in tranches, what we have called in the past a *staircase market*⁸. It would only sell some fraction of anticipated demand at time, and loads would buy long-term power of various vintages. This adjusts for the fact that the market-making algorithm can only minimize some prediction risk and integrates new technologies and changing costs only as they offer into the firm-market.
- **Dispatch:** The objective function for the firm market is least cost delivery of electricity according to a pre-arranged schedule. It will have under its control transmission assets, storage, demand response, as well as some dispatchable generation. We imagine that the optimization goal of the market algorithm, the objective function, will be mostly dominated by minimizing electricity losses over time (e.g. round-trip efficiency for storage) and space (transmission losses) with reliability and environmental constraints.

There are plenty of questions to be resolved for such a long-term firm market to make sense, but many experts clearly see the common core of an idea (Climate Policy Initiative, Bloomberg New Energy Finance) and proto-typical elements already exist in the Brazilian market, Colorado's market-based IRP, and the socialization of transmission costs in RTOs like MISO's Multi-Value Projects.

⁷ http://www.vibrantcleanenergy.com/wp-content/uploads/2017/06/WISdom_July2017_A.pdf

⁸ See Gimon, Aggarwal and Harvey: <https://energy.gov/sites/prod/files/2015/03/f20/StaircaseMarkets.pdf>

The Residual Spot-Market: In the bifurcated market, even though most electricity would be purchased through the long-term firm market there is a key role for a residual spot market. The firm market is structured to deliver a fixed amount of electricity at prescribed times and places anticipated by load, but not to exactly match supply and demand. So, there will likely be a positive or negative gap between supply and demand in most time intervals.

Resources and loads will have surpluses or deficits relative to their firm commitments that they can manage through the residual spot market (and derivatives thereof). Loads, generation and storage can dynamically offer decrements or increments in supply or consumption at whatever price they want and the security constrained economic dispatch can select what clears and gets dispatched much as in today's market. This is a lot like today's spot market, but with much more negative pricing and, with its available flexible resources, a firm market that can act as a backstop when resources get too scarce.

Interestingly, if the residual spot-market ends up being mostly energy neutral (with as many decrements as increments over time) it could also organize itself around bids and offers for shifting electricity in time and place (load-shifting) as opposed to through up/down bids at each local node. In this future vision, many operational constraints that complicate unit commitment, like startup times or minimum power requirements, will either disappear with fuel-based generation or be taken care of by the firm market dispatch. This means the security constrained economic dispatch algorithm are much less complex and can accommodate many, many more dynamic market participants (see "Power Systems without Fuel"⁹) than today's SCED.

Putting Them Together: An important question for a bifurcated market is how the two halves of the market would share responsibility for reliability and work together to match supply and demand. On the first issue, we don't have much yet to offer, except that both markets would have some responsibilities. The firm market algorithm would incorporate secure dispatch in its calculations and dispatch decisions, but leave some of the final touch-up decisions to the spot-market.

On matching supply and demand, some of the supply resources committed in the firm market may over or under perform relative to expectations. These expectations could be a firm output or demand-response, or they could be contingent on external factors like the weather in ways that the market has already factored in, but in any case resources will not always deliver or might provide a surplus. For example, a wind farm might be contracted to supply power according to a power curve and prevailing wind conditions but fail to do so because of broken equipment. Some deviations within a dead-band would be ignored, others settled along a pre-agreed rate, and further deviations might incur penalties or rack up bonuses. Similar considerations apply for loads and other firm market participants such as utility-scale storage developers and operators. The aggregate of these deviations would be bid or offered in the residual spot market (deficits made up through purchases, surpluses sold) just like any other entity. The firm market can average deviations and react to weather via its available reserves, transmission rights, demand-side services and storage to deliver on its objectives, but it will also need to lean on the residual spot market.

The residual spot market doesn't just fill in for the firm market like regulation services do for today's energy markets, it can also act as a price referent: since the residual spot-market still defines a marginal-

⁹ Taylor, Josh A., Sairaj V. Dhople, and Duncan S. Callaway. "Power systems without fuel." *Renewable and Sustainable Energy Reviews* 57 (2016): 1322-1336.

cost based value for electricity, when the firm market needs to settle deviations from its committed resources it can use the price (positive or negative) in the spot market as an index.

On the flip side, it seems likely that for a firm market to come close to achieving its delivery objectives in a world dominated by variable supply resources it will need to hold some serious energy reserves – clean dispatchable resources or long-term storage – to cover seasonal differences or week-long lulls in the weather. These could act as a kind of supplier/buyer of last resort for the residual spot market, stabilizing prices and further ensuring reliability.

Finally, if the residual spot market ends up not being energy neutral, i.e. if you can end up buying power over the long term through the spot market, it is interesting to speculate what the average price difference might be between the firm market and the spot market. On the one hand buyers might expect to pay a premium for firm power in return for reduced exposure to volatility. On the other hand suppliers can achieve cheaper financing with long-term contracts and the firm market can extract efficiencies from its holistic design so the firm market should be able to offer more attractive pricing. It's unclear where or if the firm spread will settle.

Broader Themes for Future Market Design

In settling on and describing the two general paths above for wholesale market design to evolve for a future with a high penetration of variable renewable energy and distributed resources we see several important themes come up:

- (1) **Alignment:** In both visions, aligning the markets with physical and financial realities of the underlying assets is an important criterion for identifying a viable market design. Both paths connect the spot market with the physical need to match supply and demand in real time. The first path tries to create more predictable long-term cost-recovery through demand-side resources and storage spreading out the price-duration curve and reducing the volatility of scarcity rents, and addresses the remaining risk associated with financing capital-intensive asset synthetically through long-term bilateral contracts. The second path builds a firm market specifically to reflect the capital-heavy, fuel-light nature of a future clean system, dealing with the nature of financing and risk more explicitly than the first.
- (2) **Optimization:** Part of the value-add of restructured wholesale markets is the ability to optimize in a relatively transparent and information-rich fashion for a clean, affordable and reliable grid. Today's markets achieve this by optimizing around short-term marginal costs (mostly through minimizing fuel use) and integrating environmental and reliability criteria as cost adders or as constraints. It is unclear to us how well spot-market only market designs will be able to optimize without fuel costs driving price-formation, but it is quite possible that this can be achieved once elastic demand and storage factor significantly. In the bifurcated market, the firm market is predicated on optimizing around long-term costs, but just how price-based optimization will work in the residual market is less clear. The long-term price signal from the firm market might help by providing a measure of opportunity cost for this residual market, though.
- (3) **Risk Management:** The big value-add from restructured markets is the ability to manage risk: both by shifting risk from one set of parties to another (customers to generators), and by reducing risk through pooling (lowering costs as well). An early justification for restructured wholesale markets was that they would shift risk from consumers to suppliers. Suppliers were

seen a better able to manage and price risk... In the spot-market only path, risk is minimized through long-term contracts and through demand side management and storage by broadening the price-duration curve and giving consumers choice as to when to consume. Risk is still pooled by participating in the market, but this feature could be eroded through the more central role of bilateral long-term contracts in an opaque and nebulous market. For the bifurcated market, a lot of risk is managed and pooled through the firm market, but much of the actual risk exposure for market participants will be heavily contingent on just how the firm market is structured and obligations are enforced.

The paths we describe address the challenges above in diverse ways but they also overlap. They both lean on long-term contracting, a natural way to align with the investment needs of capital-heavy fuel-light assets. They also both avoid capacity remuneration mechanisms commonly seen today, aka capacity markets. This begs the question of how these paths described above might evolve from today's markets.

An answer to the evolution question is beyond the scope of this paper, but at least a few possibilities are worth highlighting here.

Today's "energy-only" designs, those with no capacity markets, could possibly evolve into a solely spot-market driven market design. They would gradually lean more and more on bilateral markets and would need to find ways to improve participation from storage and demand-side resources. If the bilateral markets became more formalized they could become firm markets and energy markets would devolve to the status of residual spot markets. Alternatively, energy markets could also evolve to become the firm markets and leave short-term balancing more and more to ancillary markets. Ancillary markets would could then become more market-based and birth residual spot-markets.

Markets with capacity markets might evolve differently. From our point of view capacity markets are really a substitute for scarcity pricing which allow the consumer to avoid exposure to high prices and give resources that mitigate against scarcity a steadier investment signal. In the spot-market driven design (Path 1) resources depend almost exclusively on scarcity pricing to finance themselves, so capacity markets either atrophy or eventually boil down to either re-regulation or evolve towards a firm market and a bi-furcated design. In the bifurcated design, the residual spot-market represents a much smaller fraction of electricity costs, so large swings don't affect consumers much, diminishing the rationale for a capacity market. The large reserves from the firm market acting as a backstop further erode the rationale. If capacity markets don't evolve into firm markets (a very different function) they could end up looking much like the regulation products purchased in ancillary markets today but on a longer time scale.

Concluding Thoughts

Interpolating current trends into the future is full of pitfalls: we have no idea exactly what technologies will be available and markets are dynamic things. Ramifications of rules and policies are hard to understand until seen in action and we have much yet to learn on how to operate markets with high penetrations of variable renewable energy. But in the words of Dwight Eisenhower: "In preparing for battle I have always found that plans are useless, but planning is indispensable." This exercise has pointed us at some common threads to focus on in our planning: alignment, optimization and risk management. Some no-regrets areas to look at are hedging markets and long-term contracting, as well

as the faster inclusion of sources of flexibility like demand-side resources and storage which will help us solve not only operational problems but also market design issues.