A Forward Energy Market to Improve Resiliency: Frequently Asked Questions
Peter Cramton, University of Maryland

21 December 2023 [Click here for the latest version]

[Presentation] [Interactive Demo] [Sample Source Code] [Seminar 18 Dec 2023]

Abstract

Electricity markets use energy prices to balance supply and demand in real-time. Frequent shocks to supply and demand imply high volatility of spot prices. Prices that average about $50 per megawatt-hour may vary from -200 to 5,000 dollars, depending on market circumstances. Market participants trade forward energy to manage risk from the high spot-price volatility. Demanders typically buy ahead a quantity roughly equal to their real-time consumption, and suppliers sell forward an amount about equal to their real-time production. This note describes how the system operator can facilitate forward trading with a forward energy market. The products are financial derivatives of the day-ahead energy product. (The day-ahead market already manages risk between day-ahead and real-time products.) Monthly forward energy is traded up to 12 × 4 = 48 months ahead by day type (weekday, weekend) and hour. Hourly forward energy is traded up to 24 × 30 = 720 hours ahead. The products also differ by load zone to hedge congestion risk and include renewable energy certificates (RECs) to manage renewable standards. These monthly and hourly products enable both sides of the market to establish forward positions to manage risk better. Trade occurs without friction with hourly clearing using the Budish-Cramton-Lee-Kyle-Malec flow trading methodology. Flow trading allows participants to adjust portfolio positions efficiently as information changes. The approach identifies unique prices and quantities for the products that maximize as-bid social welfare. The system operator performs the settlement and optimizes collateral requirements. There is transparency about price, quantity, and forward positions. Load-serving entities have a mandatory schedule of purchase obligations, increasing linearly from 0 percent 48 months ahead to 100 percent one month ahead; dominant suppliers have a symmetric obligation to sell; in other respects, the market is voluntary. The forward energy market replaces any capacity market or capacity requirements. The forward energy market has several advantages: it gives participants flexibility in adjusting positions to manage risk better, ties collateral requirements to default exposure, mitigates market power, and provides robust price signals to guide behavior. A full-scale simulation of a US market provides proof-of-concept of the forward energy market.

1 In collaboration with Simon Brandkamp and Axel Ockenfels at the University of Cologne, Albert S. Kyle and David Malec at the University of Maryland, and Jason Dark, Darrell Hoy, Pcharasut Sujaritanonta, and Chris Wilkens at Cramton Associates. I thank researchers, regulators, system operators, and market participants for their questions and comments. The extension to intraday rolling settlement is in collaboration with Hung-Po Chao and Robert Wilson. Support by the German Science Foundation through Germany's Excellence Strategy (EXC 2126/1 390838866) is gratefully acknowledged.
Contents

Introduction.................................................................................................................................................3
What are the benefits of the forward energy market to regulators, system operators, and participants? ..........6
Who are the winners and losers of this proposal? .....................................................................................7
You emphasize resiliency rather than reliability. Can you clarify the difference? ..................................7
Is the market physical or financial? .............................................................................................................8
Is the market voluntary or mandatory? .......................................................................................................8
Can you provide sample market rules? .......................................................................................................10
Does the forward energy market limit the system operator's ability to improve the spot market? .............12
How can a market tied to the spot market work in a world with many zero-marginal-cost renewables? .........13
Is reliability reduced because the forward energy obligation does not include a reserve margin? ..............13
Can you provide an example of preference expression and market clearing? .............................................13
How should the system operator conduct the market? ..............................................................................25
Isn’t the system operator taking on more counterparty risk in conducting the forward energy market? .......27
Being in a balanced position eliminates price risk and market power, but what about physical risk? .........27
If the purchase of forward energy is mandatory, who establishes the requirement? .................................28
How do load-serving entities deal with customers switching with retail choice? .......................................28
By focusing solely on energy, are you missing other essential services? ....................................................29
What about oversupply? Are there times when renewable production exceeds load? .............................29
How does the forward energy market differ from a capacity market? ......................................................29
I see how energy price risk is addressed, but what about risk from reserves and unanticipated demand? ....33
If flow trading is so great, why has it not been adopted in financial markets? ...........................................34
How does your proposal motivate flexibility? ..............................................................................................34
How would participation in the market work for resources with variable renewables, such as wind? ........34
Is a forward energy market pro-competitive? ...........................................................................................36
Do all contracts have to be disclosed? Are price and quantity disclosed? Can the contracts be options? ....36
Can the market address the states’ preferences to set renewables portfolios and clean energy standards? ...36
Can this approach reduce the need for reliability must-run contracts? ....................................................36
How will the benefits of the forward energy market change through the energy transition? ......................36
How do you see the market evolving? .......................................................................................................36
Is four years in advance enough? ..............................................................................................................36
Are there complementary reforms to improve the spot market? ..............................................................38
How would the adoption of a forward energy market proceed in different markets? .................................38
Is the forward energy market desirable in places that do not have the US two-settlement system? .............39
What are the next steps in this research? ...................................................................................................41
Closing thoughts ..........................................................................................................................................43
References ....................................................................................................................................................44

Appendix ....................................................................................................................................................46

Can you provide sample source code that implements flow trading? ......................................................46
Consumers like flat rates. Why would they tolerate dynamic rates? .........................................................47
Don't dynamic rates expose consumers to unacceptable levels of risk? ....................................................48
How much price-responsive demand does one need to improve resiliency substantially? ....................48
How does your proposal differ from Frank Wolak's and Australia's mandatory contracting proposals? ......49
Are there ways to amplify the positive resiliency benefit of the forward energy market? .........................49
Do you need to worry about a short squeeze as in other forward commodity markets? ..........................50
Will clearing be network-constrained and require specification of sources and sinks? ...........................51
Is the forward energy market consistent with strategic reserves? ............................................................52
Are you asking the participants in the forward market to anticipate all risks? ........................................53
Introduction
As the world transitions to net zero, it is necessary to electrify nearly everything. The electricity sector will double in the next three decades and see rapid innovation in both supply and demand. As the share of intermittent renewable resources grows, balancing supply and demand every second becomes more challenging. A transparent and efficient forward energy market is needed. The market provides robust price signals to guide participants' investment and operating decisions. Flow trading (Budish et al. 2023) allows this to happen even though the forward market, because of network and storage constraints, requires trading thousands of interrelated products simultaneously every hour. The forward energy market is a promising solution to an essential problem every market faces as we decarbonize.

The forward energy market supports three goals: resiliency, innovation, and efficient investment. A reliable electricity market provides electricity during system peaks with independent outages. A resilient electricity market is reliable and avoids or limits shortages during systemic events, such as extreme weather. The costliest twenty-first century failures in North America and Europe have been resiliency events: California 2000-2001, the Northeast 2003, Texas 2021, and Europe 2022. A focus on resiliency is appropriate. Encouraging innovation is especially important in the years ahead. Thousands of supply and demand resource types will facilitate energy balance during the energy transition. Efficient energy investment requires innovation. Innovation also fosters competition. Together, resiliency, innovation, and efficient investment bring what consumers want: resilient electricity at the least cost.

Forward trade and the virtuous cycle that fosters innovation, resiliency, and efficient investment

How does the forward energy market foster innovation?
Transparent forward energy prices, updated hourly with ample liquidity and competition, provide the price information needed by market participants for efficient operation and investment. The prices enable demand and supply resources to create maximum value in their operation throughout the day, season, and year. This value motivates efficient investment, the main driver of competition. The prices also stimulate innovation in new resource types, especially resources that create value from flexibly responding
to prices. Batteries and other low-carbon technologies are good examples. Demand-side innovation benefits from transparent prices. Service providers can offer valuable services that optimize the use of low-carbon technologies to maximize consumer welfare. Energy efficiency programs also benefit from the multi-year forward price information.

**How does the forward energy market encourage resiliency?**

The forward energy prices provide regulators, system operators, and market participants with the essential information to respond to market issues. As forward prices rise, participants are encouraged to make additional resource investments. If high prices raise concerns about reliability, regulators and system operators can look for and address the source of the high prices. For example, interconnection entry barriers may limit competition.

Rapid innovation and competition fuel resiliency. Adopting low-carbon technologies like electric vehicles promotes price-responsive demand (Bobbio et al. 2023). These technologies shrink potential gaps between supply and demand in stressful times, such as extreme weather. Even a minority of price-responsive customers can eliminate a shortage. Energy efficiency also improves resiliency.

Fostering demand-side innovation to improve resiliency may require multiple steps. For example, consider a market like Germany that has no smart meters. Demand-side innovation is impossible.

The first step is to install meters systemwide that measure electricity use at each time and location. Suppose Germany takes this first step but does not improve its spot market. There is still a single "German" price for electricity at every time despite persistent congestion. Utilities could introduce time-of-use rates that reflect systemwide scarcity but not local scarcity.

Suppose Germany introduces a nodal market that correctly prices wholesale energy at every time and location, but the existing monopoly utilities still deliver electricity to consumers. It is now possible for a utility to offer dynamic rates that reflect marginal social cost, but there is no incentive for a monopoly utility to do so. Only the consumer benefits from dynamic rates, and they do not know they need it. How many consumers thought they needed a mobile phone with a camera in the early days of mobile phones? Consumers dislike change. The utility might conduct consumer polls to confirm to regulators that German customers do not want dynamic rates. Innovation would not occur. This was the experience in California. The 2000-2001 California energy crisis led to systemwide installation of smart meters. It was many years before the meters were used to improve consumer behavior. Even today, some 23 years after the crisis, the use of the meters is limited to time-of-use rates.

But what if German regulators take the third step and allow retail choice? Innovative service providers enter the market and offer rate plans tailored to the needs of consumers. The innovative service provider offers the EV owner a dynamic rate so the owner can capture the value of her battery resource through optimized charging and discharging. The gains to the EV owner are dramatic. The net cost of EV ownership falls, and EV adoption accelerates. Improved price information stimulates this virtuous cycle.

The forward energy market encourages participants to be balanced as we get close to real-time. Balanced positions mean that little volume will trade at the volatile real-time prices, yet the high scarcity prices can be seen and felt on the margin, providing ideal performance incentives.

Two of this century's four costliest electricity crises in North America and Europe—California 2000-2001 and Europe 2022—would have been averted with the forward energy market. Both resulted from
participants' imbalanced positions during a long period of high prices. Thanks to robust price-responsive demand and energy efficiency, the Texas 2021 crisis would have been substantially mitigated or eliminated with the forward energy market. The forward energy market would not address the poor tree trimming and software bugs that caused the Northeast 2003 blackout. Still, the forward energy market's three-out-of-four score on these major resiliency events is far superior to the zero-out-of-four score for traditional resource adequacy—none of the costliest crises were caused by a shortfall of capacity.

The forward energy market embraces the notion that resilience is essential. With the system operator’s expert analysis, the regulator tunes the market parameters to enhance resilience to systemic events like extreme weather. The regulator adopts its preferred reliability metrics. The desired level of reliability at the least cost is achieved by 1) providing ideal performance incentives, 2) optimizing collateral through knowledge of participant positions, 3) improving interconnection challenges, 4) encouraging innovation, especially on the demand side, and 4) establishing an appropriate price cap or operating reserve demand curve to motivate investment.

Point 4 is worth emphasizing. The price cap and operating reserve demand curve are critical controls for the regulator to encourage investment if the resiliency assessments suggest inadequate resources. A higher price cap or operating reserve demand curve strengthens energy revenues and performance incentives. Both contribute to resiliency.

Adjusting the price cap is a more direct and effective dynamic control than adjusting the parameters of a capacity market. And the forward energy market's 48-month prices provide ample information and time to change.

*How does the forward energy market promote efficient investment?*

Electricity market design has two broad goals: spot pricing to induce efficient operation of existing resources and longer-term pricing to motivate efficient investment in durable assets to attain the reliability and resiliency goals. The forward market provides pricing and trading opportunities to support efficient investment. The time and location products complete the market; see Hart (1975), Radner (1972), and Stiglitz (1982) for an analysis of inefficiencies arising from incomplete markets.

The potential for incomplete markets is pronounced when significant technological and regulatory uncertainties exist. Investors must guess how climate policy and energy technologies will evolve over the life of their investments. Such forecasts are challenging at best. Still, forward trade can help complete the market and provide essential information to improve predictions and investment decisions. The forward prices are more meaningful if they reflect the competitive consensus. The forward energy market effectively addresses market power through gradual, coordinated trade.

The improvement in investment efficiency may be especially large on the demand side. The forward prices are the essential input for innovative service providers to unleash the full value of batteries and other flexible resources on behalf of their customers. Competition among service providers ensures that the customers capture this value. This reduces the consumer’s effective cost of an electric vehicle and other low-carbon technologies, accelerating the energy transition and adding flexibility in a virtuous cycle. Intermittent renewables and storage are complements.

The forward energy market is closely tied to the mission of system operators. For example, ERCOT’s mission is “We serve the public by ensuring a reliable grid, efficient electricity markets, open access, and retail
choice.” The mission can be restated in economic language as: 1) The system operator seeks to address potential market failures, including incomplete markets, incomplete information, market power, entry barriers, and systemic risk. 2) The system operator conducts transparent and efficient markets by pricing energy and ancillary services to maximize social welfare subject to network and resource constraints. Conducting the forward energy market is consistent with the system operator’s mission. Doing 2 on a forward basis mitigates potential market failures.

The incentives for forward trade are great in electricity given the large price and quantity uncertainties. About 90 percent of electricity is hedged before real-time in today’s markets. However, this does not mean the hedging problem has been solved. It just means that there is a huge demand for hedging.

This paper answers frequently asked questions about the forward energy market. A complete paper, including preliminary modeling, will be available in early 2024. I begin with a summary of the benefits.

What are the benefits of the forward energy market to regulators, system operators, and participants?

- It provides transparent and efficient forward prices to guide investment in and operation of supply and demand resources. The gains from efficient pricing are empirically significant; for example, see Gowrisankaran et al. (2023) for an empirical analysis of the distortions from mispricing in non-restructured markets.
- The trading method allows fine granularity of time and space. This fine granularity encourages responsive resources when and where they are needed most. The energy products include monthly forwards, 48 to 1 month ahead, by hour, load zone, and type of day (weekday and weekend), and hourly forwards, 30 to 1 day ahead by load zone. The complexity of the forward energy market is robust to finer granularity from a computational, liquidity, and behavioral perspective.
- Analogous renewable energy certificates (RECs) allow market participants to manage jurisdictional renewable requirements efficiently.
- The forward energy market replaces the contentious capacity auctions or capacity requirements with a more straightforward and effective instrument.
- The accreditation of capacity value is used for resource adequacy assessments rather than in market rules that directly tie a resource’s payments to the administrative certification. This limited role encourages resource innovation and avoids costly fights over accreditation rules. Estimating capacity values becomes a technical exercise like forecasting the weather.
- The forward energy market embraces rapid resource innovation through technology-neutral rules and payments. The system rewards each resource for the value it creates for the system. The playing field is level and transparent.
- The market relies on a few administrative parameters. The chief parameter is the price ceiling. Regulators set the price ceiling to achieve the reliability and resiliency standard. A higher price ceiling induces more investment and achieves a higher standard. With time, this parameter becomes less critical as improved flexibility reduces the frequency and duration of shortages.
- The forward prices provide regulators and system operators with detailed information to better understand and manage resource adequacy. Forward price information improves the analysis in resource adequacy assessments.
- With the complementary reform of an intraday rolling settlement, the system operator can extend forward prices to intraday operations, improving incentives and efficiency.
• The flow trading approach enables market participants to express preferences and trade in a way consistent with their interests, much as they do in the day-ahead market. The method lets market participants efficiently manage risk, create value, and avoid adverse price impact.

• The transparency of positions enables the independent market monitor, the system operator, and regulators to understand and manage market power concerns.

• Position transparency lets the system operator optimize collateral to reduce counterparty risk and reduce participants’ costs of satisfying collateral requirements.

• The forward energy market improves trading opportunities across markets and provides valuable information for inter-market transmission planning.

Who are the winners and losers of this proposal?
Forward trade benefits the market through reduced risk and more efficient operation and investment. The gains to consumers are seen in lower prices and reduced risk. Suppliers and many load-serving entities benefit from reduced risk and improved opportunities from innovation. The one group harmed is legacy utilities—those ill-suited to innovate. Strengthened competition will push these laggards aside. Such is the nature of competition.

You emphasize resiliency rather than reliability. Can you clarify the difference?
Reliability is resource adequacy 1.0—the electricity system’s ability to satisfy 100 percent of demand. It measures the frequency, duration, and magnitude of shortage events and looks at the system’s average interruption duration and frequency. Outages are short and localized, caused by routine events that cause demand to spike and supply to drop. Events are triggered, for example, by the failure of large units on a windless hot summer day. A reliable system, according to the North American Electric Reliability Corporation (NERC), has two components: 1) operational reliability—the ability of the system to withstand disturbances, such as outage of the largest resource, and 2) adequacy—the ability of the system to always meet aggregate demand.

Resilience is resource adequacy 2.0. It examines the system’s robustness to a wide range of environments. Events are rare and involve systemic failure of many elements. An example is extreme cold. The same event triggers a drop in supply and a spike in demand. The events are system-wide, long, and have implications for other critical infrastructure. They are front-page news.

There are four stages to resilience: 1) preparing for events before they happen, 2) alleviating problems during the event, 3) recovering quickly after the event, and 4) learning from the experience to improve for next time. The forward energy market is primarily about 1 and 2 but also helps with 3 and 4.

While reliability will fast become a private good with improved demand engagement, discussed in the Appendix, resiliency will remain largely a public good. Regulators and system operators can improve resiliency through better markets and preparation for systemic events (Cramton 2022).

Resiliency is essential in avoiding system outages. Historically, system outages are resiliency failures:

• The 2000-2001 California energy crisis was caused by an arid year, resulting in low hydro production. The tight market, combined with a design flaw, caused market failure. The flaw was preventing the utilities from managing risk. The utilities entered the tight year without adequate hedging, purchasing their energy in the day-ahead and real-time markets. When the tight market
caused sustained high prices, the utilities neared bankruptcy, interrupting trade from excessive counterparty risk.

- The 2003 Northeast blackout's proximate cause was a software bug in FirstEnergy's alarm system that led to an inadequate response when foliage fell onto transmission lines. The cascading outage caused 55 million people to lose power, many for seven hours or more.

- The 2021 Texas crisis was caused by extreme cold. Electricity demand spiked, driven by electric heat, and supply tanked, primarily from a lack of gas, creating a gap of about 20 Gigawatts in a system with a 66-Gigawatt winter peak. Millions of households lost power for multiple days in extreme cold. At least 246 people died, and damages were about $130 billion.

These three failures include all the major North American failures since 2000. Including Europe, we would add the 2022 European electricity crisis caused by the Russian invasion of Ukraine, resulting in sustained high natural gas prices and the failure of many inadequately hedged service providers.

Systemic events caused all four failures. Resource adequacy and the traditional reliability measures had nothing to do with these costly events. To focus on reliability—non-systemic events—rather than resiliency is to ignore the essential lessons of history. Systemic events from extreme weather and high renewable variability are rising, and our dependence on electricity is increasing. A focus on resiliency is critical.

Rather than a resource adequacy assessment, we need a resiliency assessment. The assessment finds the most effective ways to improve resiliency, reducing the system's vulnerability to a broad set of systemic events: extreme weather, high variability in renewable production, cyber-attack, fuel supply disruption, and critical system failure. The forward energy market improves resiliency to each of these vulnerabilities.

Is the market physical or financial?
Like the day-ahead market, the forward energy products are financial. Deviations between forward and day-ahead positions settle financially. Efficient settlement rests on robust day-ahead pricing and the elimination of counterparty risk. The system operator manages counterparty risk with collateral obligations that depend on deviations between the participant's forward positions and its capabilities or needs.

Is the market voluntary or mandatory?
The main mandatory element is a modest purchase requirement on load-serving entities (LSEs) to help coordinate trade. The obligation starts at 0 percent 48 months ahead and increases linearly to 100 percent of expected demand one month ahead. This slow schedule gives the LSEs enormous flexibility in purchasing energy, allowing them to manage risk and any adverse price impact. Buying more than four years ahead is voluntary. Limited mandatory elements reap the benefits of consumer-centric resource adequacy discussed in Gramlich and Hogan (2019) and Hogan and Litell (2020).

The requirement is based on the LSE's reported expected load rather than a system operator estimate. The LSE knows its load much better than the system operator. It is straightforward for the LSE to motivate the LSE to report truthfully. Realized load is observed. The system operator can then discipline the LSE for any consistent or opportunistic bias in the reported load. The discipline can take the form of higher collateral and purchase obligations. Service provider reporting avoids inherent bias and complexity of system operator estimation (Hogan 2023 and Kavulla 2023).
A linear obligation, starting 48 months ahead, is reasonable to coordinate trade. However, a much longer horizon is possible, such as ten or fifteen years, and the obligation need not grow linearly. For example, the forward market may include annual or monthly products four to fifteen years ahead for market completeness. Whether the mandatory purchase begins at year fifteen, ten, or four years ahead is a subject for detailed modeling.

An advantage of four years is that it is modestly longer than the three-year term of existing capacity markets. Adoption may be easier with a four-year schedule of mandatory purchase. I support the inclusion of voluntary forwards for four to fifteen years ahead. No one can object to additional voluntary forward products. If price and volume are transparent, further forward trading may do some good and certainly no harm.

The service provider, not the system operator, is responsible for energy adequacy. If the LSE does not provide adequate energy to meet demand, it faces the economic consequences of purchasing at real-time prices. If the LSE consistently or opportunistically underperforms, leaning on the system, it is disciplined with higher collateral or purchase requirements.

A second mandatory component can be physical accreditation of forward energy using state-of-the-art methods of measuring forward-looking capacity value. Then, the market is much more like a capacity market. I do not recommend this additional obligation. It needlessly constrains trade.

There is little benefit from such physical assurance. Texas, for example, had plenty of "resource adequacy" in February 2021, but that did not mean there was enough energy. What good is it to have a physical guarantee that is not a guarantee? The best we can do is have a physical assessment, as all markets do, including Texas, using state-of-the-art estimation of capacity values. Yes, modeling physical capability is essential for the resource adequacy assessment.²

If the prospect of physical shortage appears 48 months ahead, the prices give investors and regulators ample opportunity to resolve any need. The best any market can do is provide the information to identify problems and the incentives to fix them. The forward energy market does this.

A third mandatory component is symmetric sale obligations for major incumbents with market shares exceeding a threshold, such as 15 percent. These generators would be obligated to sell an increasing amount of their accredited capacity as energy from 0 to 100 percent from 48 to 1 month ahead. This symmetric obligation addresses supplier market power by dominant incumbents. The obligation needs to be stated in accredited capacity since a company's selling strategy depends on the capability to generate energy, not the nameplate capacity.

The reason to limit the symmetric obligation to incumbents exceeding the 15 percent threshold is that it narrowly tailors the market power instrument to the few companies that exceed the threshold, thereby limiting the use of accreditation. I recommend this market power instrument. A company that wants to avoid the mitigation can divest resources to fall below the threshold.

² Resource adequacy assessments need to improve, as emphasized by Carvallo et al. (2023), "We find no systematic treatment of the costs of extreme weather and other hazards, the benefits of resilience, and resilience metrics in planning analyses..." Dynamic optimization of resiliency is essential (Rao et al. 2019).
Can you provide sample market rules?

Objective
The objective of the Forward Energy Market is to manage risk arising from high spot-price volatility and to facilitate energy trading to promote innovation, efficient investment, and resiliency.

Products
Each forward energy product is a derivative of the hourly day-ahead energy product and thus inherits the definition of delivered hourly day-ahead energy. The day-ahead energy product is a derivative of real-time energy and thus inherits the definition of real-time energy.

Day-ahead energy—the foundational product upon which other derivatives are based. The system operator publicly posts the results of the day-ahead optimization at 16.00 on the day before real-time. The optimization includes hourly prices at the nodal level. The nodal prices are aggregated to the load zone.

Monthly forward energy—traded 48 to 1 month ahead, differentiated by day type (weekday, weekend), hour, and load zone.

Hourly forward energy—traded 720 to 8 hours ahead, differentiated by load zone.

Renewable Energy Certificates (RECs)—traded in each forward energy auction with sufficient granularity to let market participants manage jurisdictional renewable requirements.

Trading Methodology
Flow trading (Budish et al. 2023) allows market participants to adjust their portfolio positions efficiently as information changes over time.

Auctions occur every hour for all products. The bidding window starts one minute after the hour and lasts until the hour’s end. During the bidding window, market participants may adjust their orders. The orders at the end of the bidding window (on the hour) are final. Only these final orders are used in the clearing optimization.

Each order is a piecewise linear decreasing demand curve, represented by two or more quantity-price pairs. The order also specifies the linear combination of products for which the demand curve applies. Price is in $/MWh. Quantities are in MW. Thus, quantity represents the rate of trade—the quantity in MWh that trades over one hour.

Market participants either upload their orders or enter them directly into the auction platform. Changes to orders are allowed until the bidding window closes on the hour.

Orders that are not valid are rejected. The auction platform allows the user to see the user’s accepted orders and adjust as needed.

---

3 Since scarcity occurs in real-time, it is intuitive to think that the forward energy market should be a derivative of real-time energy. This would be true in markets without a day-ahead market. However, the day-ahead market provides the opportunity to hedge real-time price risk. Thus, there is no benefit to tying the forward energy product directly to real-time. The sequence of forward trading opportunities is linked to the real-time market. The day-ahead market is directly connected to real-time, and the further forward markets are tied to real-time indirectly through the day-ahead market.
On the hour, the auction platform processes the final orders and determines the prices and quantities that maximize as-bid social welfare. Mathematically, the form of preference expression guarantees unique quantities (Budish et al. 2023, Theorem 1). Prices exist but may not be unique (Budish et al. 2023, Theorem 2). However, unique clearing prices result with an intuitive tie-breaking rule: If a product has multiple clearing prices, the price closest to the prior clearing price is selected. (The use of clearing prices is a cornerstone of economics. For their use in electricity markets, see Kahn et al. 2001 and Watkiss et al. 2023.)

The prices are published in the first minute of the hour. The auction platform revises each participant's position based on the quantities implied by the prices. Each participant can view and download prices and the revised position anytime during the bidding window.

This process repeats every hour.

Orders persist until changed or canceled. Thus, if the user wants to maintain the same preferences because nothing has changed, the user does nothing. The same orders will continue to be processed every hour until the user desires a change.

Settlement and collateral
Every hour, the auction platform updates the settlement for each participant. The user is warned if the user's excess collateral falls below a warning trigger. If the user's excess collateral falls below zero, future trades that would increase the participant's collateral requirement are not allowed. For every order, the portion that would shift the user to a less balanced position is eliminated.

Collateral requirements are based on a to-be-developed optimization. This approach will maximize market stability while minimizing unnecessary capital commitments from participants. The key inputs in determining collateral are 1) the participant's current position, 2) the participant's expected load, and 3) the participant's expected energy production. Each participant reports 2 and 3 to the system operator. Excessive deviations between estimated and realized load and production increase the participant's collateral requirements. The system operator maintains 1. When the system operator evaluates reported production during net peak load conditions, 3 defines the participant's capacity value, or accredited capacity.

Once per week, the accumulated settlement over the prior 7 x 24 hours is reported to each market participant. The system operator automatically transfers money to or from the participant's bank account consistent with the weekly accumulation. If the transfer fails, the user has 24 hours to resolve the issue. After 24 hours, any payment due is taken from the user's collateral account, and the user is prohibited from further trade that would put the participant in a less balanced position.

Transparency
The auction platform publicly posts prices within one minute of the hour. Each participant also learns their revised position. The platform lets users view and download prices, the user's current position, and the most recent trade rates. The slope of aggregate net demand for each product is reported as an indication of liquidity.

Bilateral contracts outside the ISO market must be reported to the system operator. Participants itemize the various bilateral contracts that have changed since the last report. The reports are used for the purpose of collateral requirements.
Participants report the capabilities of each of their resources. The system operator uses these reports in calculating a participant's position.

*Mandatory purchase obligations for load-serving entities*
Load-serving entities have a mandatory schedule of minimum purchase obligations based on the LSE's expected load during the period in question. This obligation increases linearly from 0 percent at 48 months ahead to 100 percent at one month ahead.

Each LSE estimates its expected load for each product and reports and updates these estimates as they change over time. The system operator compares the estimates with realizations. Errors that exceed a specified tolerance may increase obligations and collateral requirements. The change in obligations and collateral is an increasing, convex function in the error size intended to motivate users to report accurate, unbiased estimates of their expected load. Errors during net peak load hours are weighted most heavily.

*Mandatory sale obligations for major generators*
Generators with a market share of accredited capacity exceeding 15 percent have an obligation to sell an increasing amount of their accredited capacity as energy from 0 to 100 percent from 48 to 1 month ahead.

Aside from the mandatory obligations for load-serving entities and major generators, participation in the Forward Energy Market is voluntary.

*Market oversight*
The system operator conducts the Forward Energy Market to ensure smooth operation, enforce these rules, and intervene in market disruptions or anomalies. The system operator provides monthly, quarterly, and annual reports on the Forward Energy Market to its board and regulators (Electricity Authority 2023 is an example). The independent market monitor also studies and discusses the market's performance in regular state-of-the-market reports (see Potomac Economics and Monitoring Analytics for examples). The system operator and the market monitor report to the regulator, the legislature, and the executive branch, who provide ultimate oversight.

*Reliability backstop*
If the resiliency assessment identifies an inadequacy that cannot be resolved from market entry, then the regulator will authorize the system operator to make a timely competitive procurement of the shortfall. The procured emergency reserves, which must be additional resources, would generate energy and operating reserves during reserve shortages when prices are at the price cap and otherwise not be used. The emergency reserve product includes a reliability option. Any deviation between the resource's capacity value and its production during shortage is settled at the price cap. If the emergency reserves are sufficient to eliminate the shortage, they are dispatched in merit order to eliminate the shortage, resources are paid for delivered energy at the real-time price, and the reliability option is not triggered.

**Does the forward energy market limit the system operator's ability to improve the spot market?**
An issue with any longer-term market is the possibility that the term length limits change. Such lock-in is not a feature of the forward energy market. Since the forward energy market is a derivative of the day-ahead market, it automatically inherits any changes to the day-ahead and real-time markets. Market participants understand this when they engage in forward trade.
Changes to the system operator’s markets are common. Many changes occur each year. Today's energy futures do not depend on market rules being held fixed. The same is true if the system operator conducts the forward market.

Indeed, by defining the forward product as energy at a delivery point, the system operator—on instruction from the regulator—could adopt finer demand-side location granularity without changing the forward market rules. The settlement would reflect less demand-side aggregation when the switch is made to the day-ahead energy product.

The system operator should focus on making the spot market as good as possible. The forward energy market helps in that endeavor. It is complementary to the spot market, leveraging the system operator’s crucial work to improve the spot market. As discussed below, one helpful fix is extending the forward market to an intraday rolling settlement. Such an extension is consistent with the forward energy market.

How can a market tied to the spot market work in a world with many zero-marginal-cost renewables? It is precisely this close tie to the spot market that makes the market work. It is reasonable for people to be concerned that investment in generation is challenging in a market that bounces from negative prices to shortage prices in a day. However, long periods of negative prices are not equilibrium predictions based on detailed investment analysis through the energy transition. (For equilibrium predictions, see Cramton et al. 2022.) Instead, they are snapshots of resource scenarios inconsistent with investment incentives.

The reality is that tying the forward energy market to the spot market is sound and consistent with our foundational understanding of markets. As part of this research, we will present an equilibrium-based analysis of the forward energy market at different points in the energy transition.

Is reliability reduced because the forward energy obligation does not include a reserve margin? The incentive for demanders to forward purchase is the reality that they will be paying real-time prices for any quantity not purchased in advance. Suppliers' motivation to sell forward is that their last opportunity to utilize their capacity is in the real-time market. Thus, the incentives are derived from real-time market expectations for suppliers and demanders.

These incentives should be strong enough for the suppliers and demanders to take actions to cover the load even in extreme weather. The resiliency in the market depends on these incentives. The regulator should raise the price cap if the incentives are too weak. The price cap needs to be set to achieve reliability and resiliency goals.

The forward energy market may increase reserve margins in practice. With participants in more balanced positions, there is less quantity trading in shortage events, allowing a higher price cap and stronger performance incentives. The higher price cap encourages entry. Improved risk management also stimulates entry.

Can you provide an example of preference expression and market clearing? To fix ideas, consider an example with three market participants (Ann, Lucy, and George), two products (peak and off-peak), and two times (today and tomorrow). Our participants submit bids in today's forward market to hedge tomorrow's prices. Deviations from today’s position will be realized tomorrow and settled at tomorrow's prices.
Ann is an arbitrageur. She participates in the market to exploit her expert understanding of prices. Her strategy is classic: buy low and sell high; do not drift far from a zero position.

George is a generator with a portfolio of solar, wind, and gas combined cycle generating units. George also has battery resources. George pursues a target strategy designed to maximize profit and limit risk.

Lucy is a load-serving entity. She has a portfolio of consumers whom she is obligated to serve. She participates in the market to maximize profit and limit risk.

Although George is a natural seller and Lucy is a natural buyer, both recognize that it is helpful, like Ann, to participate in buying and selling depending on prices and other circumstances. Thus, all market participants express net demand curves that involve selling or buying depending on price. Participants express quantity as a flow, the rate of trade over a one-hour window (MWh/hour or MW).

First, suppose there is a single forward product, tomorrow’s peak energy. Each demand curve is expressed as a vector of quantity-price pairs.

<table>
<thead>
<tr>
<th>Quantity MW</th>
<th>Price ($/MWh)</th>
<th>Ann</th>
<th>George</th>
<th>Lucy</th>
</tr>
</thead>
<tbody>
<tr>
<td>-60</td>
<td>120</td>
<td>70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-50</td>
<td>100</td>
<td>66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-40</td>
<td>90</td>
<td>62</td>
<td>120</td>
<td></td>
</tr>
<tr>
<td>-30</td>
<td>80</td>
<td>60</td>
<td>92</td>
<td></td>
</tr>
<tr>
<td>-20</td>
<td>70</td>
<td>58</td>
<td>86</td>
<td></td>
</tr>
<tr>
<td>-10</td>
<td>64</td>
<td>56</td>
<td>80</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>60</td>
<td>54</td>
<td>76</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>54</td>
<td>52</td>
<td>72</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>50</td>
<td>48</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>44</td>
<td>46</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>40</td>
<td>42</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>30</td>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>20</td>
<td>14</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Table 1: Piecewise linear net demands (MW) as a function of price ($/MWh)*

Figure 1 shows the net demand curve for each participant.
Ann expects tomorrow’s peak energy to be about $60/MWh. She wants to buy when the forward price is less than $60 and sell when it is higher than $60. To protect herself from going too long or short, Ann bids a net demand curve that becomes steeper when the absolute value of quantity is larger. This shape is a risk management element in all three curves: convex for negative amounts (buying) and concave for positive quantities (selling). It also mitigates adverse selection and moral hazard. For example, George may know that he will take one of his large units offline during the peak period, creating an unexpected price rise in real-time. Ann protects herself from such events by requiring a larger price discount to accept a larger forward position.

George also expects tomorrow’s peak energy price to be about $60/MWh. However, as a natural seller, George is willing to sell at prices a few dollars below $60. At higher prices, George is happy to sell forward an even larger portion of his expected production. At prices well below $60, George is glad to buy ahead, knowing that the opportunity to sell tomorrow should reap profits. Notably, George does not offer his expected solar and wind production at marginal cost, which is negative $30/MWh, because of the investment tax credit George earns on his renewable production. His real-time offer tomorrow will be -30/MWh, but in the forward market, his offer must reflect the opportunity cost of selling the production tomorrow, which is about $60/MWh. For George, like Ann, the forward market is about arbitrage and risk management. His offers reflect opportunity cost, not marginal cost.

Lucy anticipates that tomorrow’s peak energy price will be about $60/MWh. However, she is obligated to purchase her realized load tomorrow. She recognizes the possibility of adverse supply shocks that could send the peak price to levels as high as $5000/MWh. Thus, she adds a significant risk premium to her bids. She wants to buy a large share of her expected load unless the forward price is high. This preference is
why Lucy’s net demand curve is significantly above Ann’s and George’s curves. In other respects, her curve is similar: convex for negative quantities and concave for positive amounts.

All the curves are required to be piecewise, linear, and decreasing. This language gives market participants enormous flexibility in expressing demand. The participant can approximate any continuous, decreasing demand. It is reasonable to assume a participant's true preferences take this form. An essential advantage of this form is that it implies unique prices and quantities, except in unlikely instances of no trade.

To find the clearing price, we add the individual demands in the quantity dimension, which yields the aggregate demand curve in Figure 2, which focuses on the segment of aggregate demand that includes the clearing price. The clearing price is where aggregate net demand is 0, a price of $61.09. The price is unique because the aggregate demand is continuous and strictly decreasing.

![Aggregate net demand curve, tomorrow peak](image)

*Figure 2: Aggregate net demand curve around the clearing price for tomorrow’s peak energy*

Finally, we determine the quantities by evaluating each participant’s net demand at the clearing price, as shown in Figure 3. Lucy buys at a rate of 38.2 MW; George sells at 35.5 MW; Ann sells at 2.7 MW. The net demand is zero, as required by market clearing.
Now, consider two products, peak and off-peak. How does the bid expression and market clearing generalize? With two products, our participants can bid on one or both products individually or on any linear combination of the two products.

Ann bids on the two products individually, and George and Lucy bid on a linear combination of peak and off-peak, consistent with their objectives. Again, each order is a vector of quantity-price pairs.

*Figure 3: Clearing quantities for each participant are uniquely determined from the clearing price*
The resulting demand curves are shown in Figure 4. Each order is a piecewise linear decreasing curve. The participants have the flexibility to submit as many orders as they want. The orders can be for individual products or any linear combination of products.
Ann’s off-peak order is her peak order shifted down by $30. Ann expects an off-peak clearing price of about $30. She wants to buy peak and off-peak energy at prices below $60 and $30; she wants to sell peak and off-peak energy at higher prices.

Lucy bids for a 60-40 split of peak and off-peak. Her expected off-peak demand is 80 percent of her peak demand. Thus, the 60-40 split is consistent with her anticipated demand. She buys more peak energy because the hedging benefit is most significant during peak hours. She knows she will have to pay a risk premium for the extra peak-energy purchase, but she is happy to do so to limit her risk.

George bids for a 50-50 split of peak and off-peak energy, slightly less than the 55-45 split he expects in real-time. He offers less peak energy because of the greater risk it entails. He knows shortages are more apt to occur in the peak hours. He expects to earn a risk premium on the peak energy he sells ahead.

The products clear product-by-product. The number of prices is equal to the number of products. Market clearing involves finding two prices, peak and off-peak energy, that simultaneously balance supply and demand given the collection of orders. The clearing prices of $62.43 and $29.06 for peak and off-peak energy are displayed in Figure 5. These prices imply that Ann sells 6.07 of peak energy and buys 1.57 off-peak. George sells -16.85 each of peak and off-peak energy. Ann buys 22.92 of peak and 15.82 off-peak.

The beauty of the flow trading approach is illustrated in this example. The participants have enormous flexibility in expressing preferences. Then, given a collection of piecewise linear, decreasing demand curves, the system operator finds unique market clearing prices and quantities by solving a linear system.
Larger problems with more products and more participants are solved similarly. The linear system to be solved is larger, but the computational needs are similar. Indeed, as discussed later, computation times tend to increase linearly with the number of products and orders.

**How can market participants manage so many products?**
The flow trading approach allows much finer product granularity than the status quo. The ability to trade thousands of products makes the market more complete. Trading thousands of products is feasible because participants place persistent *portfolio orders* that induce a smooth trade flow among all products. Look at the forward prices to understand how this would work. There are monthly prices for each zone and type of day (weekday or weekend), as illustrated in Figure 6.

![Figure 6: Monthly forward prices for Houston weekday energy ($/MWh) with 24 x 48 = 1152 prices](image)

There are many prices, but they are readily understood by the eye and analyzed with computer modeling. The example above illustrates seasonal and time-of-day price impacts and greater volatility of prices as we get closer to delivery. The prices are updated hourly when the market clears. We made the example realistic using data from ERCOT’s Houston load zone for the year starting in 2019.4

From the system operator’s perspective, each forward energy product is “just another product.” These products are settled according to the prices and quantities as calculated by the system operator according to the market rules.

To settle the forward energy product, the system operator needs the unique prices and quantities determined in the clearing of the forward energy market each hour, the day-ahead quantities and prices determined in the day-ahead market, and the real-time prices and quantities determined in the real-time market. (In the rolling settlement extension, the system operator will also need the prices and quantities for each intraday optimization.) Settlement of these products involves a few matrix multiplications and additions. This computation can be accomplished in less than one second regardless of the number of forward energy products.

---

4 The monthly price process is based on a Brownian Bridge. The average monthly hourly prices give the start of the Brownian Bridge for 2018. The idea is that 2018’s prices strongly shape expectations for 2019’s prices. The four-year long-run average gives the end of the Brownian Bridge. For instance, in 2018, expected prices in 2022 equal the average of the monthly hourly prices in 2015-18.
The hourly prices are shown in Figure 7. Again, to add realism, we tuned the illustration to ERCOT’s Houston load zone, showing forward prices on 11 August 2022.5

Participants trade an hourly forward product up to 30 days ahead.

A simple and effective flow trading strategy is trade-to-target. Participants state their target and the rate at which they want to move toward the target. For example, a load-serving entity might set its target to its obligation, increasing linearly from zero to its expected demand as we move from 48 to 1 month ahead. With flow trading, the participant specifies the rate it desires to make this adjustment as a function of price. A participant wants to buy more quickly when the price is low and sell more quickly when the price is high. Impatience is expressed as a linear net demand curve for each product. The participant's urgency to trade also depends on the size of the adjustment and the closeness to the spot market. Trading faster is preferred when larger adjustments are needed closer to real-time.

How is it possible to trade so many products efficiently?

At a minimum, we envision 48 months of forwards for each of 24 hours of the day for each of two day-types (weekday and weekend), plus 30 days of hourly forwards, resulting in $(48 \times 24 \times 2) + (30 \times 24) = 3,024$ products. The inclusion of RECs to manage jurisdictional requirements may double this number. One final dimension is location. For most markets, a good starting point is load zones. For example, ERCOT has

---

5 The price process follows different Brownian bridges for weekends and weekdays, respectively, to account for weekly seasonality. For weekdays or weekends, the start of the Brownian Bridge is calculated as follows: average hourly prices of the previous month + (average hourly prices of the same month last year – average hourly prices of the previous month last year). For instance, for the matrix starting on 12 August 2022, the weekday bridge begins at the average hourly weekday prices from 13 July to 20 August 2022 plus the difference between the average hourly weekday price from 22 August-19 September 2021 and the average hourly weekday prices from 21 July-20 August 2021. The end of the Brownian bridge is again the four-year long-run average hourly price for weekdays or weekends in the 30 days ahead.
six load zones. With six zones, the system operator would settle 18,144 forward products for hourly energy against the day-ahead hourly price.

Although it may be hard to imagine trading so many products, the flow trading technology makes this easy by exploiting the power of convex optimization. Complaining about too many products would be like complaining about a spreadsheet with too many rows. From a computational complexity viewpoint or a user experience viewpoint, there is little difference between 10 products or 10,000. Bid entry and optimization are readily managed with information technology. Budish et al. (2023) demonstrate how computation times vary with the number of orders and assets; see Figure 8. The computation to find unique prices and quantities can be done on a single server in about 100 seconds, allowing clearing every hour. The Appendix includes sample code for a simple flow trading implementation.

Figure 8: Computation time for challenging cases by the number of orders and assets (Budish et al. 2023)

How difficult is participating in the forward market?
Thanks to a simple and effective method of preference expression, participation is straightforward. A participant's strategy depends on three essential inputs: risk attitude, capital cost, and expected hourly net demand.

The simplest way to represent risk attitude and capital cost comes from standard financial modeling. Two scalar parameters can represent the participant’s utility function. Capital cost is the participant's discount rate or time value of money. Typical discount rates are between 6 and 12 percent. Risk attitude determines the concavity of the utility function. It is standard to assume constant absolute risk aversion. A risk-neutral participant would have a risk parameter of 0, implying linear utility. Risk-averse participants have a risk parameter greater than 0. Larger risk parameters indicate greater risk aversion.

The final input is the participant's hourly expected net demand. For a pure financial participant, net demand is zero for all hours. For others, it is a complex technical calculation that requires good knowledge of one's customers for buyers and one's generating portfolio for sellers. Hybrid participants who own generating resources and serve consumers must estimate their anticipated supply and demand. The need to estimate net demand is not unique to forward energy markets. Participants in any electricity market
must estimate net demand to decide how to participate. Thus, this difficult input is needed regardless of the market's design.

We are developing a participant tool to help regulators, system operators, and market participants understand how the forward energy market works and the ease of participation. Our goal is to show that there few implementation risks. Fear of unforeseen risks often stands as a barrier preventing innovation in markets.

The core of the tool is easy. The user specifies the risk attitude and capital cost parameters. The user then uploads her expected net demand as a comma-separated-value (.csv) file. For an arbitrageur, this is a matrix of zeros; no upload is needed in this case since the zero matrix is the default. The arbitrageur's target position is 0 for all products. For a buyer and seller (or hybrid), the trader's expected net demand (demand minus supply) defines the target position in each hour. The target\(t(m,h) = t(m) \times \text{expected net demand in hour } h\), where \(t(m)\)—the trader's target percentage—increases linearly from 0 percent to 100 percent from 48 to one month ahead. During the current month (month 0), the target stays at 100 percent of expected net demand, and the hourly (rather than monthly) products are used for adjustments. Thus, \(t(m) = 1 - (m-1)/47\) for \(m \in \{1,\ldots,48\}\), \(t(m) = 0\) for \(m < 1\) and \(t(m) = 1\) for \(m > 48\). Load-serving entities have an obligation to procure at least an amount \(t(m)\) less a small amount of slack. This allows the buyer to target \(t(m)\) and satisfy the buyer's obligation.

Figure 9: Energy adjustment to reach target position (MWh)

Figure 9 shows the intra-month adjustments for each hour of the next thirty days a participant needs to reach its target position. This is based on a simple model for a load-serving entity with an average load of 4 GWs. I model the energy adjustments by taking the sum of two random processes: 1) A daily shift with mean 0 and standard deviation \(250/(\text{days ahead})\). The daily shift affects all hours within a specific day. 2) An hourly shift with mean 0 and standard deviation \(150/(\text{days ahead})\). The hourly shift differs for each hour. To smooth the random process, I applied convolution to create moving averages across hours of a particular day and across days. This avoids large jumps between adjacent hours within a single day and between two adjacent days. The energy adjustments decline rapidly as "days ahead" increases because the standard deviations of the random processes shrink with "days ahead." This is why the need to trade
grows as we get closer to day-ahead, as shown by the darker colors, green and red, as we move toward day-ahead.

**Figure 10: Flow trade rate at prior energy price (MW)**

Figure 10 shows how the target can be reached with a flow trade rate, assuming a flow trade rate as (energy adjustment)/(8 × days ahead). Flow trade rates are small if the days ahead are large. There are many hours to trade when we are far from day-ahead. This is why the flow trade rates are so low many days from day-ahead. Even close to day-ahead, the required flows are only a handful of Megawatt-hours, which is small for a 4 Gigawatt service provider. The quantity traded is never zero but always small, reducing risk and adverse price impact.

A more advanced version of the tool lets a seller specify the characteristics of her generation portfolio and compute her expected aggregate marginal cost curve in each hour, recognizing the uncertainty about whether the resource will be committed in the day-ahead market. The advanced tool would let buyers express jurisdictional renewable requirements and include these constraints in the trading strategy.

The output lets the user visualize outcomes and how outcomes vary with variations around their specified risk attitude and capital cost. It also helps users determine the increment gain or loss from adding customers (increasing demand) or adding generators (increasing supply). This incremental calculation is essential in pricing and investment decisions. How much should an LSE charge a customer with a particular load shape? Should a supplier build a new combined cycle unit, a battery farm, a solar farm, or a hybrid?

**With so many products, how can liquidity be assured?**

Traditional markets manage liquidity by limiting the number of products. Take wheat. Wheat trading involves many grades and classifications, which vary by country and the organization responsible for grading. The United States Department of Agriculture categorizes wheat into eight classes based on kernel hardness, color, and planting season. Within these classes, wheat is further graded on a scale from 1 through 5 based on additional attributes like test weight, defects, and moisture content. Thus, there are 40 wheat products traded in the US.
As the US nodal systems prove, modern markets, like electricity, can trade products with much richer granularity. Liquidity is managed by allowing near-perfect substitution among products that are near-perfect substitutes. The nodal spot markets and the forward energy market have this feature.

The forward energy market, like the nodal market, has a high level of transparency, robust pricing, and low transaction costs, which favors liquidity. The forward energy market has three further advantages. First, preferences are convex. Market participants enter piecewise linear net demand curves, which yield a quadratic objective in the clearing optimization. Second, because the forward market is conducted well before the day-ahead market, the market participants have time to adjust positions as uncertainty resolves. Third, the frequent batch auction approach allows participants to make thousands of minor adjustments over months. Slow trading enables participants to minimize adverse price impact, improving the market's competitiveness and increasing liquidity.

Flow trading promises smooth movement of forward prices. What might disrupt these price dynamics? Gradual trade and gradual resolution of uncertainty imply that forward prices typically move smoothly. Even discrete events, such as the entry of a new 1 GW combined cycle unit, can impact prices smoothly. First, the event is apt to be anticipated by the market. The probability of approval gradually increases to one. Second, the resource owner, the party most impacted by the approval, is motivated to minimize price impact by trading slowly.

The most disruptive event would be a significant and sustained shock in the natural gas price, as happened in Europe with Russia's invasion of Ukraine in February 2022. Since gas units typically are on the margin, the spot electricity price is set by the marginal gas unit at its heat rate times the price of natural gas. The invasion is a systemic event affecting energy prices every hour of every day for an extended period. The implication for the forward energy market would be a gradual increase in forward prices as the probability of invasion rose. Then, with a discrete price jump in the hour, the likelihood of invasion jumped to one. The discrete change in forward energy prices would require market participants to adjust their target positions. Still, trade would occur smoothly and orderly since each participant is motivated to trade-to-target slowly to mitigate adverse price impact.

Why should the system operator conduct the market?
There are two main reasons: to minimize transaction costs and to adopt an efficient and transparent market.

Regarding transaction costs, the system operator's net cost of conducting the market would be near zero. Most system operators spend a sizeable sum on conducting and maintaining a capacity market (Patterson and Reiter 2016, Christie 2023). The staff dedicated to navigating the stakeholder process and implementing the many changes could be shifted to other tasks. The forward energy market would require some management but would be less of a burden because it rests on a simple foundational structure and is not a vehicle for distributing wealth. Even for systems like ERCOT without a capacity market, the forward energy market would provide valuable information about resource adequacy and improved investment incentives. The net cost would again be near zero.

Only the system operator is motivated to adopt an efficient and transparent firm energy market. ICE and CME would be the apparent alternatives to run the market in the US. After all, they already have a futures energy market for peak and off-peak monthly energy. However, ICE and CME are strongly motivated to continue with their continuous trading methodology rather than flow trading. Their current method makes
their data valuable to high-frequency traders. Their revenue model effectively creates an inefficiency, which they capture through the sale of low-latency data and colocation services. There is a direct conflict between the objectives of efficiency and transparency and ICE and CME revenues. The system operator has no conflict; the system operator specializes in operating efficient and transparent markets. Efficiency and transparency are common themes in the system operators' mission statements.

A comparison between current ISO markets and ICE and CME futures markets is telling. For the ISO markets, anyone with Internet access can freely get the market data by time and location in an organized form. By contrast, the historical electricity and gas futures data, which is far inferior in volume and granularity, costs $37,500 from ICE and $21,000 from CME (CME offers a 50 percent academic discount; ICE has no discount for academic research). These high costs are a prominent and unnecessary entry barrier. The marginal cost of data production is zero. Moreover, the only costly activity in data production is that borne by the market participants in expressing their preferences. Why should ICE and CME own and charge monopoly prices for data produced by market participants?

The system operator already has relationships and data from each market participant. Having the system operator conduct the market would leverage the system operator's market information. Information about the participants' resources and day-ahead positions is essential to the conduct of the market. Likewise, the settlement is trivial for the system operator, as it would add only a single step to the system operator's existing settlement process. Nothing else would change.

Through knowledge of positions, the system operator can establish highly optimized collateral requirements that maximize market resiliency to systemic events with minimal collateral. The collateral requirement would depend on deviations from balanced positions. Markets fail when counterparties become unreliable. Optimized collateral is critical to minimizing this vulnerability at the least cost.

Since the system operator provides a simple, transparent, and friction-free platform for forward contracting, market participants will likely use the platform to adjust positions. If so, then the accounting is trivial. The market platform produces the required information. Trades outside the market platform would primarily be large long-term bilateral transactions, such as multi-year power purchase agreements. These would be reported electronically to the system operator.

System operator conduct would facilitate the inclusion of REC products, allowing market participants to manage climate goals or jurisdiction-specific requirements flexibly. REC trading is a significant issue, especially in multi-jurisdictional markets such as PJM, ISO-NE, SPP, and MISO.

There are critical mandatory elements. Introducing mandatory features into a private exchange or bilateral trade has serious problems. By contrast, the system operator has a well-defined stakeholder process for

---

6 CME Chairman of the Board Emeritus, Leo Melamed, was so tied to continuous trading that following Eric Budish's presentation of "The High-Frequency Trading Arms Race: Frequent Batch Auctions as a Market Design Response" (Budish et al. 2015) explaining the problems of continuous trading, he said to the audience, that if US markets adopted discrete time, that he would take continuous markets and all of his technology to North Korea, so that he could, quote, "show them how they could become the center of finance in about three weeks." (Budish 2017) All markets conducted by electricity system operators use periodic batch auctions to find clearing prices and quantities. Batch auctions are also crucial in financial markets, such as at market opening or closing. From an exchange's viewpoint, continuous trading has the advantage of creating the demand for monopoly-priced low-latency data and co-location services, the arms of the HFT arms race.
adopting and modifying market rules. Managing the stakeholder process with the regulator’s leadership is a core system operator task. The mandatory elements are a) participant obligation to report positions (anticipated load - anticipated production + sales - purchases), b) an obligation to purchase an increasing fraction of net load, increasing from 0 to 100 percent as we move from 48 to 1 month ahead, c) an obligation of dominant incumbents (more than 15 percent market share) to sell an increasing fraction of capacity value, increasing from 0 to 100 percent as we move from 48 to 1 month forward. The mandatory elements are essential to maximize the effectiveness of the market, and these are much more challenging to do with private exchanges.

Finally, what is the advantage of the alternative, not having the system operator conduct the market? A typical reason is to avoid a “monopoly exchange.” Competition is typically desirable to foster innovation, but not in this case. The system operator can efficiently conduct the forward energy market with transparent rules. There is no benefit to competition in clearing and settling the market. We need innovation among the market participants in investing and operating diverse resources. We want market rules that foster innovation among participants. The forward market is best done centrally, like the complementary day-ahead and real-time markets. This encourages competition where it matters: competitive market participants making investment and operating decisions.

The forward energy market in no way restricts bilateral trade or futures trading on ICE and CME. If market participants get value from forward trade outside the system operator’s market, they are free to use alternatives. Competition between the forward energy market and alternatives is desirable. It promotes innovation and reduces consumer rates.

Isn’t the system operator taking on more counterparty risk in conducting the forward energy market? One often thinks of risk as a zero-sum game. Suppose the system operator conducts the forward energy market. Doesn’t this transfer the risk of forward trading from the private market to the ISO market, which ratepayers bear? The answer is that efficient and transparent forward trade reduces counterparty risk and lowers the cost to ratepayers. Vibrant forward trade puts market participants in more balanced positions, reducing risk, market power, and system cost.

To confirm this, one can look at the costly defaults in electricity markets since 2000. In the 2000-2001 California electricity crisis, the utilities entered a long scarcity period caused by drought (low hydro production) with a large short position (Borenstein 2002). The utilities required rescue by the state, costing about $40 billion (California State Auditor 2001). In the February 2021 Texas crisis, the market participants were in much more balanced positions, and defaults were rare despite a real-time energy value of over $50 billion in four days (Cramton 2022). In Britain’s crisis of 2021-2022, poorly hedged suppliers defaulted, costing consumers more than $10 billion (Waddams 2023).

In the forward energy market, imbalanced positions are known, and the associated risk is priced and mitigated through higher collateral. Total system risk is reduced.

Being in a balanced position eliminates price risk and market power, but what about physical risk? Physical risk is something that generators should be concerned about. If your 1GW combined cycle unit goes down minutes before real-time, the real-time price will spike, and you will be a large buyer at this high price. To mitigate this risk, you will want to only sell some of your expected production in the day-ahead market and sell the remainder in real-time. Thus, you want to enter the real-time market as a net seller during normal times.
The forward market helps you manage your sale of expected production to maximize a combination of profit and risk avoidance. Not all risk can be eliminated. The best we can do with physical risk is leave it to those who can do something about it. That would be the generators. If the generators want to mitigate the risk of unplanned outages, they can:

- Maintain their resources.
- Buy better resources.
- Buy a diversified portfolio of resources.
- Set a lower target position.

The forward energy market reduces risk by enabling market participants to take positions at known prices consistent with their anticipated needs or capacities. Production uncertainty varies by resource type, location, and other factors. Suppliers account for this production uncertainty in their forward trading strategies. A wind resource may choose to sell ahead a quantity close to its capacity value during net peak hours. The capacity value is production conditional on a shortage event. Then the wind resource will not have to buy energy at the shortage price. Since the mandatory obligation is on demanders, not suppliers, the wind resource can pursue this strategy to avoid downside risk.

**If the purchase of forward energy is mandatory, who establishes the requirement?**

Load-serving entities have mandatory purchase obligations. The system operator determines these obligations with inputs from each LSE. The LSE estimates its expected load by hour and load zone over the forward period. Of course, there is uncertainty in this forecast, which gradually resolves as we move closer to the day ahead. This gradual resolution of uncertainty is another compelling reason it makes sense for the obligations to increase progressively from 0 percent 48 months ahead to 100 percent one month ahead. (The other reason is to mitigate market power.) Thus, there is little harm if demand 48 months ahead is mismeasured. The requirement is zero; purchase is entirely voluntary. Only with time, as uncertainty resolves, does the requirement become significant. The LSEs are motivated to closely tied forecasts to reality since, ultimately, demand is realized. Persistent bias would be observed and corrected. Persistent bias is penalized with higher collateral requirements and purchase obligations.

The advantage of consistent obligations among service providers is that it helps coordinate purchases. Both sides of the market are needed to trade. There is often a coordination problem in markets. Buyers trade when sellers do, but when? The obligation on service providers mitigates the coordination problem, mitigates market power, and promotes resource adequacy.

**How do load-serving entities deal with customers switching with retail choice?**

Service providers typically have many thousands or millions of customers. Most of the customers are small. This allows service providers to understand and manage the risk around switching. For larger customers, longer-term contracts can help manage churn.

Another way to manage switching is by having less of it. A common problem with retail choice is companies exploiting behavioral biases to make money. The primary bias is "inattention." Customers have learned over many decades to be inattentive to utility bills. Many service providers offer a below-market price for a short time, such as six months or a year. The consumer adopts, and then the provider doubles the price at the contract end, invoking the "automatic renewal." Unfortunately, most regulators do not protect customers from this practice, so it is common.
Some markets with retail choice and capacity markets have complex provisions for LSE obligations to move with the customer automatically. One can do that to reduce further switching risk, but I do not recommend it. The problem is poor rate plan regulation that induces switching among the more attentive customers.

A better practice is for the regulator to have a default contract for the consumer with fair terms that do not exploit inattention and other behavioral biases. Instead, the default contract is a simple time-of-use contract. Service providers would be required to offer the default contract, competing on price and reputation.

The best service providers would offer innovative contracts that take advantage of the customer’s low-carbon technologies, such as electric vehicles, batteries, and solar panels. The contract would track the real-time wholesale price but would employ hedging. The customer would see and feel the real-time energy price on the margin, but nearly all her electricity would be purchased at much more stable forward prices. The contract would automatically manage the customer’s low-carbon technologies to create maximum value for the customer, charging the EV when the price is low, discharging when the price is high, anticipating the customer’s needs, and the production of the customer’s solar panels. Such a contract would create substantial value for the consumer and system, motivating investment in low-carbon technologies. The forward energy market and retail choice are complements in this consumer engagement. In the Appendix, I explain how consumer engagement improves resiliency.

By focusing solely on energy, are you missing other essential services?

As the share of renewables grows, the system operator may need other services for operating reliability. For example, thermal generation retirements may create an inertia shortage. The system operator identifies and introduces the required service as a product in the day-ahead and real-time optimizations. It is natural for the set and quantity of ancillary services to change with the resource structure. I do not anticipate a need to include these ancillary services in forward products before the day-ahead apart from reserves, discussed later.

What about oversupply? Are there times when renewable production exceeds load?

Oversupply is best addressed with negative prices. There is no need for arbitrary rationing. Price will do the job. Price will also lead to the correct response. Show me a market with persistent negative prices, and someone will start mining Bitcoin there.

How does the forward energy market differ from a capacity market?

A state-of-the-art capacity market has excellent performance incentives (Cramton et al. 2013), making it like the forward energy market. A forward energy market always settles against the day-ahead market, providing perfect performance incentives. If a day-ahead position differs from a forward position, the difference settles at the day-ahead price. There are no exceptions. Likewise, the best capacity markets are a financial option to deliver energy during reserve shortages. Again, there are no excuses. The sale of capacity comes with an obligation to provide energy during reserve shortages. Real-time deviations from the obligation settle at the real-time price. Both markets provide a financial hedge to electricity demanders. Both markets have a similar mantra, "Buy enough in advance."

A difference between the markets is the hedge’s form. The forward energy market provides complete price coverage. Demanders are hedged regardless of the day-ahead price, and the day-ahead market provides an efficient hedge of the real-time price. By contrast, a capacity market defines the energy obligation as an option: only during a reserve shortage does the capacity supplier have an obligation to deliver energy.
at the shortage price. (There is an obligation to offer day-ahead, but the offer may be at the price cap and not lead to energy delivery.) Thus, the capacity market only provides price coverage at the shortage price. Variations of the market offer broader price coverage by triggering the obligation at a lower scarcity price, say $200/MWh rather than $5,000/MWh. A lower strike price makes sense in markets where the scarcity events tend to be of long duration, such as Colombia, where El Nino can cause low hydro production for months (Cramton and Stoft 2007). In the Eastern US, where shortages are of short duration, typically 5-120 minutes, a high trigger price is used to minimize the role of the capacity market in energy contracting.

The market designer’s desire to minimize the capacity market's role in energy contracting stems from a fundamental capacity market problem: market power. Capacity markets use big-event auctions to procure capacity. Typically, a single auction procures 100 percent of the required capacity three years ahead. Since some market participants are prominent, the scope for exercising market power is considerable. Thus, the capacity market must have market rules to mitigate the exercise of market power. These rules limit the behavior of participants in potentially undesirable ways. Market power mitigation is controversial and imperfect.

By contrast, the forward energy market mitigates market power by allowing market participants to trade forward energy slowly. Market power is effectively eliminated, and no controversial rules limit behavior. The equilibrium behavior is to trade slowly and not exercise market power.

The second essential difference between the markets is the role of accreditation, which defines how much generators can sell in a capacity market. The best capacity markets use state-of-the-art forecasting methods to determine a resource’s capacity value, which is its ability to deliver energy during reserve shortages. Electricity spot markets often have missing money—the profits in the spot market are insufficient to cover the long-run average cost. The capacity market in equilibrium will restore the missing money, which implies that the money paid to generators from selling capacity will exceed the financial cost of the option. Every resource wants to sell as much capacity as possible. Accreditation has a direct impact on generator revenues.

The stakeholder process about capacity values becomes a debate about how market rules allocate money. In such discussions, money trumps efficiency. The market participants care more about getting more money than making the market more efficient. The result is an unproductive stakeholder process. Every system operator with a capacity market has experienced this problem.

A capacity market is like training wheels on a bike. The training wheels do not necessarily help the child learn the critical aspect of biking: balance. Yet, the stakeholder process results in a bike with training wheels and other bells and whistles (Figure 11, left). The child is better off with a far simpler balance bike that focuses on the art of balance (Figure 11, right). Aagaard and Kleit (2022) discuss these challenges.
By contrast, the forward energy market limits accreditation's role to optimizing collateral. A resource's capacity value will determine the physical hedge the resource brings to the generator's resource portfolio. The greater the physical hedge, the lower the collateral without introducing counterparty risk. This use of capacity values is second order. It does not directly lead to money transfers to the generator. It has a modest impact on the generator's collateral requirement.

The forward energy market motivates the proper response to missing money. If the market is missing money, there will be too little generator investment. Resource adequacy assessments will identify this shortfall, and the regulator will restore the missing money by raising the price cap or shifting the operating reserve demand curve upward. The higher price cap and demand curve will increase spot market revenues to encourage efficient investment. Alternatively, the system operator will redouble efforts to reduce entry barriers if that is what is limiting investment.

We can think of investment dynamics as an optimal control problem. The control is the price cap. Raising the price cap increases revenues to generators—holding resources fixed. More revenues increase the incentives for investment, raising the reserve margin. The investment only occurs if the payments cover costs. There is no missing money. Thus, the long-run equilibrium is that forward prices equal the expectation of real-time prices plus a risk premium. There is no missing money. Further, long-run cost is minimized because the participants have the information and incentives to make efficient operating and investment decisions. Should things get out of equilibrium, the market monitor, the system operator, and the regulator have the forward prices to recognize potential trouble and make an adjustment.

Significant sources of missing money are capacity market flaws and unpriced operator decisions, primarily reliability unit commitments. The forward energy market eliminates the capacity market flaws. It mitigates missing money from reliability unit commitments with the rolling settlement reform of the two-settlement system. Rolling settlement sets prices for what is unpriced today. It also injects convex arbitrage bids into the optimization, which mitigates the non-convexities, another source of missing money.
The primary role of accreditation is resource adequacy assessments, a technical modeling exercise all markets do. A further advantage of the forward energy market is that it provides transparent and reliable information about resource adequacy. The forward energy prices are an outstanding measure of resource adequacy, and far-in-advance price information gives market participants and regulators ample time to respond.

Finally, the role of scarcity pricing—the operating reserve demand curve—vanishes as demanders provide sufficient flexibility so that reserve shortages do not occur. The primary administrative device becomes irrelevant. Given the rapid adoption of electric vehicles, we should expect the day of negligible reserve shortages to arrive within a decade or two. Outages will be at the distribution level, not system-wide.

The forward energy market promises a gradual transition to a perfectly competitive market. The capacity market does not. As long as stakeholders are engaged in a fight over the allocation of money, there will be no tendency for the market to become more efficient. With the forward energy market, the stakeholders are not fighting over money and can focus on efficiency. The result is a bicycle ruthlessly optimized for the required task (Figure 12).

![Figure 12: A road bike optimized for maximum efficiency and agility on varied terrain; a triathlon bike optimized for maximum efficiency on fairly flat terrain without drafting](image)

No single bicycle is best for all races. Likewise, no single market design is best for all settings. The road bike on the left is optimized for maximum efficiency on varied terrain while drafting a quarter meter behind another rider. The triathlon bike on the right is optimized for maximum efficiency on fairly flat terrain without drafting, which is illegal in triathlon. The triathlon bike puts the rider in a much more aerodynamic position to minimize wind resistance but handles poorly and affords poor visibility.

The most significant differences in electricity markets stem from the different supply and demand characteristics. The best market designs tune the parameters of the market to the setting. However, all the best market designs are grounded in the same fundamental principles, like the bicycle designs above. There are more similarities than differences. Indeed, the best market designs reoptimize the market parameters dynamically as circumstances change. For example, operating reserve requirements may differ by season or as the share of intermittent resources grows. A cyclist has multiple bikes: an ultralight bike
for the mountain stages, a triathlon bike for the time trials, and an aero road bike for the other stages. A market with dynamic parameters can motivate investments in resources that can profit from the changing setting. For example, the operator of a state-of-the-art combined cycle unit can configure it differently based on its best use in the market. In periods where the unit runs constantly, the operator tunes it for maximum efficiency to produce the most energy per fuel input. When the unit provides flexibility, it can be adjusted to operate like several gas peakers with rapid ramping and lower startup costs. The best markets encourage such resources to bid multiple configurations into the market and let the system operator optimize the unit's configuration daily.

I see how energy price risk is addressed, but what about risk from reserves and unanticipated demand? It may be desirable to encourage forward trade of reserves beyond day-ahead. This is easily accomplished. The best markets co-optimize energy and reserves starting day-ahead. The forward reserve product is defined with the same locational granularity as the day-ahead reserves, typically system-wide, although finer granularity is possible.

One sensible approach would trade forward reserves for every hour up to 30 days ahead. To coordinate trade, service providers would have an increasing obligation from 0 to 100 percent of their expected load share as we move from 30 to 1 day ahead. Based on the service providers’ reports, the expected shares are identical to those in the forward energy market. This approach would provide complete reserve price coverage yet adds little complexity. It accommodates to locational reserves, such as by load zone, since the gradual trading addresses market power.

One final concern is price coverage for hours where realized demand exceeds anticipated demand because of extreme weather. Usually, this is fine; the service provider buys more in the spot market at reasonable prices. However, the price could be extreme at times of net peak load. The service provider can mitigate this risk by buying extra forward energy for hours of likely net peak load. The service provider expects to be a seller in the day-ahead market on most net-peak hours but is protected in the event of a demand spike. The service provider can also mitigate this price risk by encouraging some of its retail customers to be price-responsive. An EV owner would welcome dynamic prices. The service provider can create value and limit risk by optimizing the charging of the customer’s vehicle. Optimization of discharging can provide further value where vehicle-to-grid is possible.

These risk mitigation measures may sufficiently address unanticipated demand at the net peak. However, an alternative is to offer a financial option for energy when the price is above a trigger, say $1,000/MWh. Whenever the real-time energy price exceeds $1,000, the service provider receives energy for the option quantity at the $1,000 price.

To coordinate trade in this net-peak option, the service provider could have an obligation to purchase a quantity based on the provider’s share of the net peak load. The obligation would increase from 0 percent to 100 percent of the reserve margin as we move from 48 to 1 month forward. The regulator would set the reserve margin, say 15 percent, to facilitate risk management.

This financial product avoids the flaws of the capacity market. It represents a means for service providers to mitigate net peak price risk by purchasing an energy option. Generators can reduce this risk by building resources that perform well at the net peak. Service providers can reduce their need to buy net-peak options by encouraging price-responsive demand. A well-constructed reliability option may let market participants better manage risk. Modeling can shed light on the benefits of a net-peak option.
If flow trading is so great, why has it not been adopted in financial markets?

Financial markets suffer from the same limitations of the stakeholder process as electricity markets. The most dominant stakeholders lobby the regulators to adopt market rules that favor them. For example, high-frequency traders dominate the technical committee advising the Commodity Futures Trading Commission. It is no wonder that the CFTC is slow to adopt reforms that limit the profits enjoyed by traders with speed advantages, especially since the other influential stakeholders, the exchanges, make most of their money selling tools—data and collocation services—to high-frequency traders. There is little tendency for the market to adopt efficiency-enhancing reforms. Regulators are risk-averse and easily scared that a reform may have unintended consequences. Such is the tyranny of the status quo (Budish et al. 2015, 2021).

Indeed, the market design challenges in financial markets are worse than in electricity markets. Important financial stakeholders were entrenched when information technology made the reforms discussed here possible. Thus, these stakeholders provided immediate resistance to change. A norm of transparency did not exist and indeed was prevented by legislation from the early 1900s, for example, that prevents the disclosure of bidding in Treasury markets—even after 100 years. By contrast, electricity markets were new, and the importance of transparency was well-understood by early market designers.

How does your proposal motivate flexibility?

Efficient forward prices reward those providing flexibility. Market participants can easily see and enjoy the value of flexibility. Storage technologies, such as batteries, are especially adept at providing value. With excellent price information, storage devices are trivially optimized with a linear program (Crampes and Trochet 2019). Similarly, demanders are encouraged to be price-responsive. Transparent and efficient prices will motivate the demand-side innovation essential to consumer engagement. Consumer engagement breeds resiliency (Bobbio et al. 2023).

How would participation in the market work for resources with variable renewables, such as wind?

To manage risk, a wind resource should be hesitant to bid more than its capacity value far in advance. The resource’s capacity value is its expected production conditional on a reserve shortage. Only as uncertainty is resolved should the resource gradually shift toward its expected production. Capacity value is far less than its anticipated production since wind will produce little during shortages. The resource does not want to have sold expected production at a low forward price and then be forced to buy at $5,000 in a day-ahead shortage. This means that wind resources will be sellers in the days leading up to day-ahead much of the time. Waiting to sell close to day-ahead is no problem. The resource has no obligation to sell earlier, and the system operator will expect their performance to be poor during shortages.

The forward energy market works well for variable renewables, even when renewables dominate. The renewables have the flexibility, price, and production expectations to optimize positions throughout the 48 months.

Is a forward energy market pro-competitive?

Current markets in the US and Europe lack liquid forward markets. The price information could be better. Poor pricing and limited liquidity create challenges for small market participants and opportunities for dominant market participants, perpetuating an undesirable market structure. Prominent market participants can better manage risk and optimize the scheduling of their resources. Self-scheduling is a viable option for participants with large resource portfolios, whereas small market participants need the system operator to optimize their resources centrally. Thus, a lack of price transparency harms small
market participants. Dominant participants can benefit from opaque markets since they can do large bilaterals at more favorable terms. Transparent prices level the playing field.

**Do all contracts have to be disclosed? Are price and quantity disclosed? Can the contracts be options?** The disclosure requirement is like the mandatory disclosure of resource plans in the day-ahead and real-time markets, which is essential for system operation. Forward position transparency helps system operators establish optimized collateral requirements and assess market power. The disclosure can be price and quantity or quantities only.

Translation is needed if the trade is inconsistent with the forward energy products. This translation is an essential issue for legacy contracts. Once the market is established, there should be no reason for parties to transact outside the forward energy market and no reason to trade products inconsistent with forward energy products.

The translation is not perfect for options, but it should be sufficient for the primary use, collateral optimization. For example, an option to buy at $100/MWh is equivalent to purchasing and selling the same quantity in forward energy since the collateral concern primarily arises from exceptionally high prices.

**Can the market address the states’ preferences to set renewables portfolios and clean energy standards?** States’ standards can be included by defining renewable electricity certificates (RECs) as a product. Qualifying generators produce both energy and RECs. The RECs are defined based on the location of the generator. Buyers can then purchase energy and RECs consistent with their state’s renewables portfolio and clean energy standards. The standards can vary by load zone and hour, providing states with a rich way to establish standards and giving service providers enormous flexibility in managing energy and climate obligations.

**Can this approach reduce the need for reliability must-run contracts?** Reliability must-run contracts are a sign that something is wrong with the market. Entry needs to be faster to keep up with the system’s needs. There are three primary causes: 1) interconnection entry barriers, 2) poor price signals, and 3) inadequate transmission. The forward energy market mitigates 2. However, the system operator, market monitor, and regulator should investigate the root cause of the RMR need and attempt to address it. The changes required may be improvements in the interconnection process, an increase in the price cap, or better transmission planning with faster execution.

Although the forward energy market cannot directly mitigate entry barriers and transmission shortcomings, it does so indirectly. The improved price information helps the planning process determine where and when new resources—supply, demand, and transmission—are needed. Thus, better price information leads to a better planning process, which allows a more timely and effective response to system needs. In this way, the forward energy market makes RMR contracts less likely and of shorter duration.

**How will the benefits of the forward energy market change through the energy transition?** The importance of price transparency grows with the rise of intermittent renewables. Shocks in supply due to weather will increase, creating a greater demand for flexibility both in amplitude and duration. While an efficient day-ahead market can suffice for the efficient operation of short-duration batteries, long-duration storage requires efficient price information days, weeks, and months ahead. The efficiency gains from the forward energy market will grow throughout the energy transition.
How do you see the market evolving?
As the energy transition progresses, it will make sense for the granularity of the market, especially for location, to become finer. Today’s load zones are an excellent place to start, but congestion must be more finely managed as more renewables, batteries, and price-responsive demand arrive. Including finer load zones is easy, and computing prices will remain feasible. Indeed, a nodal forward energy market may be possible and is worthy of exploration. The forward energy market would then replace the firm transmission rights (FTR) market as the congestion pricing would be internalized. In a nodal forward market implementation, the network constraints must be fully represented as in the FTR market. Including the network constraints would imply a much more challenging optimization. As a result, the clearing frequency would likely need to be daily rather than hourly for monthly products.

To facilitate an easy transition to finer locational granularity, we can define the product in terms of the wholesale energy price at the point of customer interconnection. Today, this location is the load zone, but at some point, pricing on the demand side will become more granular. The locational granularity can transition simultaneously without changing the contract terms.

Is four years in advance enough?
Beginning obligations 48 months ahead is a suggestion. It is longer than we see with current capacity markets, which tend to be three years ahead. It is possible and desirable to have forwards start ten years or even twenty years ahead. Doing so may reduce capital costs for generators.

Even without far-forward products, the price information embedded in the forward energy market would provide investors with much greater assurance of revenues than the exiting capacity and spot energy markets.

As with locational granularity, it makes sense to start four years ahead and then increase the window over time as participants gain experience with the market. Nonetheless, I would have no problem if the market started with seven- or ten-year forwards to reduce capital costs.

Are there complementary reforms to improve the spot market?
Extending the forward prices to intraday trade with a rolling settlement is an important complementary reform. The market’s two-settlement system needs improvements to accommodate rapid innovation and incentivize efficient investment. Past spot markets were served well by a financially binding day-ahead market to optimally commit fossil-generating resources for the next day, followed by a real-time market for physical operation and the settlement of deviations from day-ahead plans. Today, the challenges of intermittent generation require more robust intraday pricing and settlement to manage incentives throughout the day. Essential flexibility providers, both supply and demand, need improved price signals to optimize their use and system value.
The rolling settlement extends the forward prices intraday for improved incentives and efficiency throughout the day. Figure 13 illustrates the forward prices for the Houston load zone on 26 August 2023. The day-ahead prices are in the left column. ERCOT reports these prices at 4 pm (16.00 on 25 August). The forward prices evolve until real-time, the diagonal of the matrix. In this illustration, the day-ahead and real-time prices are the actual realizations for 26 August. The other prices are a geometric Brownian bridge in log prices that moves from day-ahead to real-time.

In PJM, this matrix would include 24 hourly energy prices for each day-ahead forward, as well as 24 five-minute prices for forwards within two hours, except for the last ten minutes. The last two five-minute periods are for system dispatch; trades are locked, and the real-time market is physical. The forward intraday prices would be updated hourly for products more than two hours ahead and every five minutes for the five-minute products less than two hours ahead.

The rolling settlement has the same core idea as the forward energy market: efficient and transparent periodic clearing. The difference is that beginning day-ahead, non-convex commitment decisions are introduced, complicating the optimization, as is done today. What is new is that the prices from the intraday optimization are used to make the quantity adjustments financially binding. These additional settlements improve incentives and give participants essential information to provide optimal flexibility.

The rolling settlement prices allow participants to address non-convexities on their own. This independence adds convex bids, reduces the importance of non-convexities, and reduces computational challenges.

The rolling settlement also reduces the need for scarcity pricing to smooth the jump to the price cap. The prices throughout the day are smoothed and evolve with the probability of a real-time reserve shortage. Participants have many auctions with which to make smooth adjustments to minimize physical risk.

Many opportunities to make financially binding trades implies more granular trade, more balanced positions, and less market power. Real-time automated market power mitigation becomes less important.

A circuit breaker that reduces the price cap to $1000/MWh after a systemic shortage causes a sustained emergency is the one remaining essential element to reduce physical risk. Long durations at the price cap can lead to large wealth transfers that do little to improve incentives but magnify physical supplier risk.
Dropping the price cap to $1000/MWh after 10 hours at the price cap over two weeks is a reasonable circuit breaker. However, the natural gas market must have a coordinated policy to ensure that as much gas as possible is available to generators. Otherwise, the circuit breaker may aggravate the shortage.

How would the adoption of a forward energy market proceed in different markets?
The seven US electricity markets have distinct challenges and histories. Their market rules reflect these differences. Still, each market addresses resource adequacy, as shown in Table 1. ERCOT stands alone in its absence of a capacity requirement, relying on the voluntary response of market participants to anticipated scarcity pricing. The others have a capacity requirement. Each load-serving entity must buy sufficient capacity to cover its expected demand. The capacity requirement is met through bilateral trade or in a capacity auction conducted by the system operator.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PJM</td>
<td>PJM Interconnection</td>
<td>Capacity market</td>
<td>Three-year forward purchase in capacity auction</td>
<td>$5,500</td>
<td>Operating reserve demand curve</td>
</tr>
<tr>
<td>ISO-NE</td>
<td>ISO New England</td>
<td>Capacity market</td>
<td>Three-year forward purchase in capacity auction</td>
<td>$3,500</td>
<td>Pay-for-performance</td>
</tr>
<tr>
<td>NYISO</td>
<td>New York Independent System Operator</td>
<td>Capacity market</td>
<td>Capacity auctions six and one month ahead</td>
<td>$4,000</td>
<td>Operating reserve demand curve</td>
</tr>
<tr>
<td>CAISO</td>
<td>California Independent System Operator</td>
<td>Capacity mechanism</td>
<td>Bilateral contracting</td>
<td>$2,000</td>
<td>Operating reserve demand curve</td>
</tr>
<tr>
<td>MISO</td>
<td>Midcontinent Independent System Operator</td>
<td>Capacity mechanism</td>
<td>Bilateral contracting</td>
<td>$3,500</td>
<td>Operating reserve demand curve</td>
</tr>
<tr>
<td>SPP</td>
<td>Southwest Power Pool</td>
<td>Capacity mechanism</td>
<td>Bilateral contracting</td>
<td>$2,000</td>
<td>Operating reserve demand curve</td>
</tr>
<tr>
<td>ERCOT</td>
<td>Electricity Reliability Council of Texas</td>
<td>Reserve demand curve</td>
<td>None (voluntary response to spot scarcity pricing)</td>
<td>$5,000</td>
<td>Operating reserve demand curve</td>
</tr>
</tbody>
</table>

Table 1: Comparison of resource adequacy approaches for US system operators

The forward energy market would dramatically enhance participants' flexibility in meeting capacity requirements. Participants can easily engage in efficient and transparent trade to create value and manage risk. Participants would see robust forward prices to improve investment and operating decisions. All participants would enjoy these benefits.

For regulators and system operators, the forward energy market would improve market efficiency, simplicity, transparency, and fairness. Perhaps most importantly, the forward energy market would improve resiliency and reliability. Robust forward prices would amplify incentives to invest in technologies encouraging price response demand, such as electric vehicles and smart homes. Retail providers would be encouraged to offer dynamic rate plans that allow consumers to create and enjoy value by being flexible. These dynamic rates could include automatic hedging via forward purchase so that the dynamic rate
reduces downside consumer risk relative to a fixed rate. Good price information four years ahead would give households, service providers, and industry the information necessary to make these decisions.

The forward energy market would replace the capacity markets in PJM, ISO-NE, and NYISO and the capacity mechanisms in CAISO, MISO, and SPP. This transition would benefit regulators, system operators, and market participants. Regulating and operating the capacity markets and mechanisms has proven to be an enormous challenge over nearly two decades. The primary reason is that the complex details of these capacity approaches directly impact the distribution of money among market participants. Endless debates about these details dominate the stakeholder process.

By contrast, the forward energy market is based on fundamental economic ideas complementary to the core idea of the spot market: maximize as-bid social welfare subject to network and resource constraints. A transparent, efficient, and simple market process results. There is much less for the stakeholders to debate, and the market provides the ammunition to deflect rent-seeking behavior. Regulators and system operators can then focus on much-needed efficiency improvements. It allows market participants to compete on a level playing field and focus on improving their goods and services rather than seeking rents in the stakeholder process.

For the six markets with capacity markets and mechanisms, I anticipate these advantages of the forward energy market will lead to its adoption despite the inertia of the status quo.

One incremental implementation approach is to introduce an hourly 30-day ahead hourly forward energy (and reserve) market, while retaining existing capacity requirements. The incremental implementation would operate as described here but with a weakening of the mandatory purchase obligation so that the obligation starts at 0 percent of anticipated load to avoid a discontinuous purchase obligation. LSE obligations could increase linearly from 0 to 100 percent from 30 to 1 day ahead.

An advantage of this incremental approach is that stakeholders would be able to see the advantages of forward trade and improved pricing before making a more radical departure from existing capacity mechanisms. The value of forward trade is especially valuable close to day-ahead, as is illustrated in the simulations of a 4GW service provider. It is closer to day-ahead where volumes and volatilities are largest.

For ERCOT, the benefit that is apt to carry the day is improved resiliency. The February 2021 crisis was devastating regarding human and dollar costs (Cramton 2022). Since the event, the Texas legislature, the PUC, and ERCOT have made good and bad changes. The intent has been to improve resiliency and reliability. The Contingency Reserve Service discussed earlier is an example of a change not working as intended.

The Texas PUC and ERCOT seek more effective ways to improve resiliency and reliability. The forward energy market is a promising choice, especially since it emphasizes the same core principles guiding ERCOT’s market. The forward energy market is built on the same foundation as the day-ahead and real-time markets, which guide behavior with efficient price signals for commitment day-ahead and operation in real-time. The forward energy market extends that logic to produce transparent and efficient forward prices to guide investment and operating decisions.

Is the forward energy market desirable in places that do not have the US two-settlement system?

The forward energy market makes sense for markets worldwide. European markets have struggled with liquidity in forward energy. Less liquidity hinders competition because smaller market participants need
liquid markets to help manage risk. The forward energy market would resolve this issue. The markets in Britain and Ireland seem especially ripe for this innovation. In Canada, the restructured markets in Alberta and Ontario would benefit from the forward energy market. In South America, the Colombian market could replace its firm energy market (Cramton 2007) with the forward energy market.

The Australian and New Zealand electricity markets would find similar benefits. Like Alberta, both are energy-only markets without a day-ahead market.

New Zealand’s Market Development Advisory Group recently explored improving price discovery. Their report (MDAG 2022, quotes in italics) set out five pre-conditions that need to be satisfied for our ‘energy-only’ arrangements to work well, namely:

(a) Wholesale prices reflecting real supply and demand conditions, including very high prices in times of scarcity;

The forward energy market allows very high prices in times of scarcity because it encourages participants to be nearly balanced near real-time. Balanced positions mitigate market power, improve efficiency, and avoid large wealth transfers, which can cause markets to fail from counterparty default.

(b) Confidence among wholesale buyers and sellers that the high prices make sense (which means confidence in the structure and rules of the market, including the sufficiency of competition);

Again, the balanced positions mean that participants are not motivated to exercise market power. Competitive behavior builds trust in the pricing and markets. It also makes the high prices sustainable since little volume trades at the high prices. The high prices provide ideal incentives for efficient operation and investment yet avoid wealth transfers.

The transparency and efficiency of the forward energy market also build trust. The prices are based on market fundamentals; they are transparent and readily checked. The prices evolve as a continuous process, so there are few surprises other than outages near real-time.

(c) Availability of ‘tools’ for wholesale buyers and sellers to manage their exposure to those spot price risks;

Risk management is the core of the forward energy market. The market provides a platform for market participants to best manage risks, profits, and costs. The participants have robust price information to guide their investment and operating decisions. They can hedge price risk with little or no transaction costs.

(d) General public and political acceptance that volatility and high prices (in times of scarcity) in the wholesale market are, in fact, in the best long-term interest of consumers and that measures to ‘soften the landing for unhedged participants’ can trigger a vicious circle of undermined investment incentives and higher future prices; and

The forward energy market enables near-costless hedging and robust pricing information, making managing risk easy for all market participants. Some physical risk remains, but suppliers are best placed to manage this risk. There is no excuse for a market participant to go unhedged. Indeed, it would be expensive to do so, absent fraud, because the collateral requirements would be high for the poorly hedged participant.

(e) Confidence among consumers/politicians that investment will be timely and competitive.
The forward prices give excellent information about prices four years ahead. Four years is ample time to recognize whether investment is adequate. A reserve margin that is too tight would show up in the resource adequacy assessments, especially in the forward prices. The high forward prices should induce investment. If they do not, then there is some barrier to investment, or the prices must still be higher. The regulator should mitigate the entry barriers and, if necessary, raise the price cap. Both actions will further induce investment.

New Zealand has a market with essential differences from many others. The two most significant differences are: 1) The abundance of hydro resources allows for a more straightforward market design. Non-convexities due to fossil-resource startup costs and other constraints are less important. 2) As a small country, it is harder to maintain competition in infrastructure industries with large scale economies. 1 is a blessing. 2 is a challenge.

Although I cast the forward energy market in the US setting (nodal systems with energy and reserves co-optimized in a financially binding day-ahead market followed by real-time physical dispatch for efficient settlement), the forward energy market would be ideal in the energy-only markets of New Zealand, Australia, and Alberta. It would eliminate the need for the complications of a day-ahead market, an FTR market, and a capacity market. The forward energy market would do all three with much greater simplicity, efficiency, and transparency.

In New Zealand and Australia, the nodal real-time market would be unchanged. All forward products would be derivatives of real-time energy. There would be a single, unified, forward energy market, with an increasing granularity of products as we move closer to real-time: monthly (48 to 1 month ahead) and 30 minutes (intramonth). The locational granularity would be customer delivery points, currently load zones, but as in the US, demand-side locational granularity will become finer through the transition. Finer granularity is automatically accounted for in the product definition.

One significant benefit of transparency is expert market analysis. The periodic reports of the Electricity Authority are an excellent example (Electricity Authority 2023). By contrast, there is no free expert analysis for non-transparent government-run markets, such as US Treasury Auctions. The lack of transparency in US Treasuries confirms George Stigler’s (1971) insight, "...as a rule, regulation is acquired by the industry and is designed and operated primarily for its benefit..." Fortunately, the better electricity markets are the counterexample that proves Stigler's rule.

What are the next steps in this research?

We are performing a full-scale simulation of the market. We will choose one market for a pilot study. ISO-NE, PJM, ERCOT, or SPP would be good choices. The four have a long history of consistent operation, which makes them suitable for a backcast analysis. The PJM and ISO-NE markets appear to have more missing money in their spot markets—a problem that was at least partially addressed with the introduction of stronger performance incentives in their capacity markets. ERCOT would be one of the more straightforward markets, with only six load zones and one jurisdiction. ERCOT has 11.5 years of consistent data with its nodal market and scarcity pricing, enabling a backcast from 2011 to 2023.

Regardless of the pilot market, we begin with a forecast of net load, defined as load minus renewable production. Then, we forecast the natural gas price and the day-ahead price on a forward basis. Market participants use these forward price distributions in their bidding strategies.
We develop highly parameterized trade-to-target strategies for natural sellers and buyers. There is heterogeneity among participants in risk attitude and capital cost. Market participants with more extensive resource portfolios and balance sheets are less risk-averse and have a lower capital cost.

We also include arbitrageurs, purely financial players with a target position 0. These players have a greater risk tolerance and a lower cost of capital. Arbitrage is the first principle of finance: buy low, sell high. Arbitrageurs bring value to the market by improving prices. Arbitrage is profitable when it moves prices toward improved efficiency. We can see this arbitrage benefit in the virtual bidding in day-ahead energy markets (Jha and Wolak 2023, Birge et al. 2016).

The baseline strategy for buyers assumes that the LSEs have target positions increasing from 0 percent to 100 percent from 48 to 1 month ahead. The LSEs deviate from target positions based on a slope parameter—the slope of their net demand for each product.

A generator derives its trade-to-target strategy from its forecast of the day-ahead price, the marginal cost curve, and expected production. Like the buyers, the seller's target increases linearly from 0 percent to 100 percent from 48 to 1 month ahead. The generators deviate from the target position based on the slope of their net demand curves.

The simulation optimizes the market participants' strategy parameters to determine approximate mutual best responses. We can then look at the properties of market equilibrium. We can evaluate risk relative to a but-for world where hedging is limited to the day-ahead market. We can develop optimized collateral requirements that assure resiliency at the least cost.

The simulation provides a detailed proof-of-concept. Stakeholders will see the computational feasibility and market implications of the forward energy market based on an analysis tied to the characteristics of an actual electricity system.

The backcast helps us understand how the market would work today. However, the market must work well throughout the energy transition. To test future performance, we select two points in time, 2040 and 2060, to develop scenarios roughly halfway and at the end of the energy transition. These scenarios will have quite different resource structures with much higher penetration of renewable resources than we see today. The future-year analyses will mimic the backcast study, demonstrating its future performance.

Once the pilot analysis is complete, we will have a template for a similar study in other markets. Since the forward energy market improves electricity markets regardless of design, we intend to perform similar analyses worldwide. Our research is conducted under open access and is freely available.

Many broader research questions can be studied (suggested by Hung-po Chao):

- What experience do we have with successful and unsuccessful forward energy markets, and what lessons can we glean from this experience?
- What is the role of financial transmission rights and transmission network modeling within the framework of the forward energy market?
- What are the implications for inter-ISO market trading and interregional coordination when implementing forward energy markets in different regions?
- What challenges do market participants encounter during the transition to a forward energy market, and can these challenges be effectively managed through a phased-in approach?
Our modeling will be done on a mini cluster of servers selected for this task. The compute cluster includes 288 cores running at 3.7GHz with 1,152GB of DDR5 RAM (4800MT/s), supporting 512-bit advanced vector operations. Data management is handled by a dedicated database server with 36 cores and 768GB of RAM and 10Gb network. The cluster is housed in a secure umd.edu server facility. The cluster is needed to model the day-ahead and real-time markets, which involves solving large nonconvex problems to co-optimize energy and reserves throughout the day.

Closing thoughts
I conclude with a few principles.

Physics trumps economics.

Reliable electricity delivery is the norm in all modern electricity markets despite different approaches to market design. The reason is that all markets recognize the paramount role of physics in electricity delivery. All markets deliver electricity in a way consistent with physics—economics be damned.

What good economics can do is reduce the cost of reliable electricity. Since there are millions of decision-makers in an electricity market, this is best done with prices consistent with physics. Only then is the goal of least-cost delivery achieved. The cost of electricity varies widely by market and is one indicator of market performance. As of December 2022, Germany’s household electricity price was 50.61 per kWh. As of June 2023, a Texan household paid 0.14 per kWh. German electricity is more reliable than Texas electricity, but the price difference reflects an expensive German energy policy: prices ignore physics, and there is zero consumer engagement. A shift to nodal pricing, a forward energy market, and the installation of smart meters would remedy this situation, driving down consumer costs.

Why does Germany not follow this sensible path? Stigler (1971) again provides the answer. The profits of the dominant utilities would fall if regulators allowed a competitive electricity market. The regulator’s chief concern is keeping the lights on, and this is something the dominant utilities do well.

A better market design achieves the same reliability at a lower cost. Good electricity market design will be increasingly important as the world transitions to a sustainable future, necessitating a more significant role for electricity. The pressure on regulators to reform markets will increase.

Good market design minimizes the probability of political intervention.

A significant benefit of good market design is that politicians are less apt to intervene. When the lights are on, politicians view electricity as dull and technical. So long as the cost is reasonable and the lights are on, politicians are unlikely to participate in market design discussions. Their absence is a good thing and a leading explanation for why electricity market designs are as effective and innovative as they are.

Thus, the system operator’s primary goal of reliable electricity at the least cost is consistent with good electricity market design. Good market design keeps the lights on and the cost low and keeps the politicians away.

Ideas are incestuous.

The ideas presented here are familiar. The most important points have been well-understood for decades if not centuries.
The inefficiencies created by imbalanced ownership appear in Myerson and Satterthwaite (1981), Cramton et al. (1987), and Ausubel et al. (2014) and are empirically documented in many studies (Borenstein et al. 2002, Wolfram 1999, and Wolak 2003). Vickrey pricing (1961) can mitigate market power but only with non-anonymous, discriminatory prices that seem unfair to many and are anticompetitive in favoring larger parties. More frequent trade provides a better means to mitigate market power (Black 1971, Coase 1972, Kyle 1985, and Vayanos 1999), especially when dynamic trade is natural to manage risk as circumstances change. Several have proposed long-term energy derivatives as an approach to resource adequacy (Cramton and Stoft 2006, 2007; Cramton et al. 2013; Gimon 2020; Pierpont 2020; Wolak 2021, 2023).

The form of trade matters. Frequent batch auctions can eliminate an arms race for speed (Budish et al. 2015), especially when combined with flow trading (Budish et al. 2023), which brings a practical language of preference expression. Participants can adopt simple trade-to-target strategies, allowing flexible risk management and providing efficient and transparent price signals. These prices summarize the essential information for efficient investment throughout the energy transition.

References
ACCC (2023) "Competition in Retail Electricity Market not Delivering for all Customers," Australian Competition and Consumer Commission Media Release, 15 December.
Bivens, Carrie (2023) "IMM Concerns with the AS Methodology and Recommended Improvements," Potomac Economics.
Bobbio, Emmanuele, Simon Brandkamp, Stephanie Chan, Peter Cramton, David Malec, and Lucy Yu (2023) "Resilient Electricity Requires Consumer Engagement," University of Maryland Working Paper.
Budish, Eric (2017) "Will the Market Fix the Market?" Transcript of AEA/AFA Joint Luncheon Talk, 6 January.


Electricity Authority (2023) "Market Performance Quarterly Review," April-June.


Patterson, Delia and Harvey Reiter (2016) "Chasing the Uncatchable: Why Trying to Fix Mandatory Capacity Markets is Like Trying to Win a Game of Whack-a-Mole," Public Utilities Fortnightly, June.


Appendix

I answer some additional questions here. These questions are less central to understanding the forward energy market.

Can you provide sample source code that implements flow trading?

For a zip file with sample source code, see cramton.umd.edu/electricity and click on Sample Code. The sample code does not include the advanced tricks, decompositions, or preconditioning that are performed
in a production-grade implementation. The sample code is intended to reflect the mathematical logic of flow trading. The code is useful for validating or benchmarking more advanced implementations.

Consumers like flat rates. Why would they tolerate dynamic rates?
Although price-responsive demand is not an essential feature of the forward energy market, I mention it because it is aligned with the forward energy market in that it supports resiliency.

Consumers like low electricity bills. They ask for dynamic rates when it lets them save money. Saving money is why consumers who have invested in low-carbon technologies want to be price-responsive. Electric vehicle (EV) owners top this list. An EV owner usually has her car parked in the garage. The owner is indifferent to when the car charges if she has enough battery for her daily trips.

California has low EV rates in the super-off-peak window from midnight until 6 am. I pay 15.4 cents/kWh at night versus 81.6 cents/kWh at peak. My Tesla understands this and starts charging after midnight. I will do even better in the future by discharging to the grid in the on-peak window. Even now, my nighttime consumption is massive relative to my on-peak consumption (my July on-peak was -72 kWh, including solar production versus 523 kWh super-off-peak). The July savings from shifting my consumption to nighttime was $346.23 (a monthly bill of $21.79 versus $368.02). The scope for consumers benefiting from being price-responsive will grow as the pricing and algorithms improve, and the cost of low-carbon technologies decline. A minority of price-responsive consumers is enough to enhance market resiliency dramatically. The majority can keep flat rates as long as they like.

Consumer adoption of electric vehicles has reached a tipping point. Today, over one-quarter of California automobile purchases are electric vehicles; in Norway, 83.5 percent of new car purchases are EVs. The scope for price-responsive demand is rapidly increasing.

Using household-level data in Britain, we have measured consumers’ willingness to respond to dynamic rates. A one-percent price increase reduces demand by 0.26 percent (Figure 14). Moreover, this response increases by a factor of three for those who own electric vehicles (Figure 15).

Figure 14: Customers on dynamic rates respond to price, Britain 2020-21
Don't dynamic rates expose consumers to unacceptable levels of risk?

No. Well-designed dynamic rates expose the consumer to the wholesale price on the margin yet include a built-in hedge against real-time price volatility. The service provider buys the consumer's expected consumption in the forward market, hedging price volatility. The hedge transforms what would be a downside risk—paying a high price during scarcity events—to an opportunity—consuming less than expected during these events. The consumer then automatically sells her surplus back to the market at the high scarcity price. The behavioral optics of dynamic rates are turned upside down when the hedge is included. Consumers with less income who are more price-sensitive become the winners from the shift to dynamic rates.

One challenge with dynamic rates is that they benefit those with low-carbon technologies the most. Wealthier consumers are most able to invest in low-carbon technologies. Thus, wealthy early adopters initially enjoy the advantages of dynamic rates. Programs subsidizing low-income consumers' adoption of low-carbon technologies can reverse this tendency. Such subsidies are especially desirable when carbon prices are too low, and borrowing rates are regressive.

How much price-responsive demand does one need to improve resiliency substantially?

As a thought experiment, we took the British elasticity of 0.26 percent and asked what fraction of Texans would need to be on dynamic rates to eliminate the large gap in supply during winter storm Uri. The answer is 44 percent. Thus, Texas could have survived the extreme winter storm with an outage if only a significant minority of retail customers were on dynamic rates (Figure 16).
The calculation is crude and invokes extreme extrapolation. The point, however, is valid. A minority share of price-responsive consumers can significantly improve market resiliency.

Over time, electric vehicles and other low-carbon technologies will be adopted, increasing incentives for price-responsive demand. Eventually, we will move to the ideal where consumer engagement is sufficient to eliminate all shortages, and scarcity pricing becomes irrelevant. However, before we get there (20 years?), the market provides near-ideal incentives and robust information to invest and operate resources like batteries.

How does your proposal differ from Frank Wolak's and Australia's mandatory contracting proposals?
My proposal is closest to Wolak (2021, 2023). There are differences in details, which I hope Frank Wolak will view as friendly amendments. Two differences are the trading format and the obligations of service providers. That service providers procure forward energy through big-event auctions. He envisions far fewer products than what I suggest. Moreover, the schedule of service provider obligations ratchets up much more steeply: 85 percent four years ahead, 87 percent three years ahead, 90 percent two years ahead, and 100 percent one year ahead.

By contrast, I propose that trade occurs in frequent batch auctions. Auctions occur every hour, allowing participants to trade slowly and minimize adverse price impact (Black 1971, Kyle 1985, Vayanos 1999). Doing so reduces trading costs. It also eliminates a wasteful arms race for speed (Budish et al. 2015), especially when implemented with flow trading (Budish et al. 2023). I propose a schedule of service provider obligations that is much more gradual, increasing linearly from 0 percent at 48 months to 100 percent at one month. Service providers retain enormous flexibility in their trading strategies. The most
A straightforward approach would be to follow the linear path of obligations, employing a *trade-to-target* strategy.

Trade-to-target would result in a gradual purchase flow. A critical advantage of this approach is that it effectively eliminates any market power issues. Purchases are tiny and continuous. No market participant is making large purchases or sales in any auction. Moreover, each market participant trades each product in each auction with probability one. The market participants are willing to buy more or sell less if the price is low or high. The net demand curve bid will have a negative slope and zero trade at the target. Service providers will rarely be far from their target positions; only minor adjustments are needed at each auction. The implication is that each product will have liquidity. The aggregate net demand curve will be reasonably flat.

Australia has mandatory contracting requirements to ensure service providers are sufficiently hedged. The most significant difference between the forward energy market and the Australian market is that the Australian Electricity Regulator mandates contracting but does not provide a means for satisfying the requirement. The forward energy market provides a simple, transparent, and efficient trading platform to adjust positions consistent with requirements flexibly. This is a significant distinction. Australian service providers, especially small retailers, have had trouble finding counterparties for mandated contracting at reasonable prices (ACCC 2023).

**Are there ways to amplify the positive resiliency benefit of the forward energy market?**

Yes. The best way to enhance the resiliency of the electricity market is to adopt the same reforms in the natural gas market. It would be easy for regulators to adopt a similar transparent and efficient trading mechanism for natural gas using the same principles and methodology. Electricity markets depend on gas markets, as the 2021 Winter Storm Uri demonstrated in Texas. The failure of the gas market was a proximate cause of the Texas crisis. Generators could not get gas, which led to a massive loss of thermal units that lacked fuel. Even those with dual fuel ran out because the delivery of oil was made impossible by failures in the transportation network. The critical infrastructures are linked.

Unfortunately, the gas market has none of the reliability protections we have for electricity. The absence of protections is the result of poor regulation. Gas is essentially unregulated within Texas. Transparency of prices or volumes is poor. The gas regulator is the Railroad Commission, comprised of commissioners who advocate for the gas industry. There is little consideration of consumer interest.

During Winter Storm Uri, the most significant loss was from natural gas production, which declined by about 15 GWs. Coal and nuclear energy sources also had failures. Wind resources performed better than expected in almost all hours during the storm.

As we add renewables, we must rely even more on natural gas to provide energy during shortages, at least until there is an economic alternative for long-duration storage. Therefore, a significant concern is the poor performance of natural gas during a crisis. Preliminary evidence suggests that many gas units failed because of a lack of gas or inadequate gas pressure. Gas supply was a problem. Texas natural gas production dropped **45 percent** when gas demand surged because of the extreme cold. Detailed forensic work will uncover to what extent the shortage was from lack of gas or gas pressure versus freezing at the generator. This distinction is essential. Texas can spend billions winterizing its gas-generation fleet. The grid will remain vulnerable to winter storms unless the gas supply is also winterized.
During the Texas crisis, the gas market had a significant shortage. There was a failure to deliver at any price. Many who bought gas forward and did not receive it got a force majeure excuse rather than compensation for the failed delivery.

Technically, reforming the gas market to require efficient and transparent trade is easy. The same flow trading methodology would provide an efficient and transparent gas market framework. All that is needed is a regulator motivated to maximize social welfare rather than one captured by special interests.

**Do you need to worry about a short squeeze as in other forward commodity markets?**
The forward energy market settles against the day-ahead price rather than the real-time price because the day-ahead market already provides the hedge between day-ahead and real-time. Thus, there is no need to do so with the forward energy market.

However, as in any forward market, a short squeeze is possible. The squeeze would take place in the real-time market. Such a squeeze is possible today. A participant takes a significant imbalanced position in the day-ahead market, causing others to take imbalanced positions day-ahead and then squeeze them in real-time. A dominant supplier has a comparative advantage in executing the squeeze in electricity. A generator buys a large quantity in the day-ahead market, leaving others short, then offers supply at high prices in real-time and strategically withholds with an "unplanned" outage. During periods of scarcity, enhanced market power improves the effectiveness of such strategies.

The Forward Energy Market mitigates this possibility through the transparency of positions. The system operator and market monitor would observe the imbalanced position, which would prompt regulatory action. Moreover, the single-price auction makes a squeeze prohibitively expensive.

Recall Salomon Brothers' famous squeeze in the US Treasury markets in 1990-1991. To be successful, Salomon Brothers needed to hold a considerable position. They acquired majority shares in some Treasury auctions. Although illegal, winning a majority was possible because of the pay-as-bid pricing and large price-tick size at the time. Salomon Brothers could acquire most of the issue and squeeze the short dealers in the subsequent market by bidding one tick above the obvious clearing price. Developing such a significant stake would be prohibitively expensive with single pricing, which we have in electricity.

Hundreds of market participants exist in the forward energy and day-ahead markets. The participants include natural buyers, natural sellers, and arbitragers. Each natural buyer and seller also acts as arbitragers—the arbitrage behavior results in price convergence. The day-ahead price equals the expected real-time price plus a small risk premium (Jha and Wolak 2023).

The market is highly competitive. Therefore, the scope for strategic bidding is limited. Flow trading further mitigates incentives for strategic bidding by incentivizing participants to seek balanced positions to manage risk and limit collateral. With balanced positions, there is no incentive to distort bids.

Market power only arises at the day-ahead market and within the day. Then, market participants can take actions that may result in more significant and favorable price impacts because other participants will not have time to take corrective measures to mitigate this behavior. Intraday market power is best addressed with the rolling settlement discussed later.
Will clearing be network-constrained and require specification of sources and sinks?
Contract clearing will be consistent with the product. In markets without persistent congestion, having a single-zone forward energy market and letting the firm transmission rights (FTR) market handle locational hedging will suffice. In markets with persistent congestion, the products are readily split into zones. In a market with six load zones, there would be about 18,000 products rather than 3,000.

There is a tradeoff in deciding how many locational elements should be hedged in the forward energy market versus the FTR market. Many markets now have three-year ahead FTRs. The FTR market can remain the primary method for congestion hedging.

Indeed, the FTR market remains essential to price transport of generation from one node to a load zone. A generator in zone A, selling zone B forward energy, would need to hedge the congestion cost of delivery in the FTR market. The FTR market design could adopt a similar flow trading methodology with frequent batch auctions to improve the hedge. However, since the FTR clearing optimization is challenging, it likely makes sense to clear the FTR market daily rather than hourly. This change to smooth trading and frequent clearing would improve the FTR market's liquidity and mitigate market power.

Is the forward energy market consistent with strategic reserves?
Strategic reserves, sometimes called emergency or reliability reserves, stand idle until a scarcity event, most commonly a reserve shortage. Only in shortage do they run to mitigate the shortage. The "market" price throughout the shortage is the price cap, say $5000/MWh. The strategic reserve is procured periodically via competitive bid. The winning resources receive the clearing price for the quantity procured. The payment is in $/MW of capacity. The product is a reliability option to deliver 1 MWh of energy for each MW of capacity whenever a reserve shortage occurs. Performance deviations settle at the price cap. The resources cannot participate outside these shortage events in the energy or reserve markets.

The strategic reserves do not harm the market when structured in this way. They are decidedly outside the market. Further, they face the same strong performance incentives of resources inside the market. The strategic reserves aim to provide additional insurance to reduce the probability and duration of shortages. In this way, they can support reliability and resiliency.

There are, however, some problems. When energy prices are high for an extended period, it is tempting for regulators to step in and use the strategic reserves to reduce prices. This pre-shoragte use harms the market because it becomes difficult for generators to guess under what circumstances the strategic reserves are used—a classic moral hazard problem. It is difficult for the regulators to commit to not interfering with the market. A recent example is Germany's strategic reserve. Germany had rules about when the system operator would use the strategic reserve. However, when sustained high energy prices arrived, the regulator discarded the rules and instructed the system operator to use the strategic reserve to reduce market prices. Such behavior creates missing money.

Still, provided regulators can commit to only using the resources during shortages, strategic reserves can provide additional insurance during the transition. The forward energy market works perfectly with these strategic reserves. It is simply a matter of cost. Is buying more strategic reserves or having a higher reserve margin within the market more cost-effective? The answer depends on the setting and is primarily a question of price discrimination. Suppose the price of strategic reserves is low because there are ample near-retirement resources whose best use is to stay idle except in a shortage. In that case, it may be
economical to buy strategic reserves. However, if these resources can create more value in the energy market than in emergency response, then it is likely that welfare will be greater without the strategic reserve.

A second problem with strategic reserves arises when they are adopted without sufficient lead time. Until the expansion of resources is consistent with the increase in the reserve requirement, the strategic reserves are simply a wealth transfer from load to generators. The reason is that strategic reserves only enhance reliability if they expand the set of resources; otherwise, they raise energy prices.

A good example is the ERCOT Contingency Reserve Service introduced in summer 2023. CRS is a version of strategic reserves, procuring reserves only used during reserve shortages. When procured on a short-term basis, this amounts to mandating the withholding of resources until the price is at the price cap. The near-term implication of this required withholding is higher prices for consumers with little improvement in resiliency or reliability. This implication was borne out in summer 2023, as reported by the Independent Market Monitor (Bivens 2023). The additional cost to consumers for the three months of June-August was about $8.5 billion. The cost through November was about $12.5 billion.

Relying solely on strategic reserves for reliability would be astronomically expensive. The forward energy market is a better approach, which provides the forward prices to guide investors and operators to optimize welfare.

However, emergency reserves can provide a reliability backstop if the forward prices suggest an emergency. If everything has been done to encourage entry and shortage is still likely, a reasonable emergency response would be for the regulator to authorize the system operator to make a timely competitive procurement of the shortfall. The procured resources, which must be additional resources, would be used as emergency reserves to generate energy and operating reserves during reserve shortages when prices are at the price cap. This backstop does not distort the market since the resources are only used during shortages; no quantity is taken from market resources.

Are you asking the participants in the forward market to anticipate all risks?

No. Market participants will anticipate foreseeable events. The clearing prices will reflect the market consensus about these risks. A virtue is that the market participants are motivated to estimate risks as best they can and mitigate risks, including the possibility of unforeseen events. The market provides a means to best manage the risk by entering the day-ahead market in a balanced position. Doing so means that little volume trades at extreme prices, allowing the market to maintain efficient prices during scarcity without bankrupting counterparties. To the extent that a party chooses to hold an imbalanced position, the party's collateral requirement increases, guaranteeing the settlement even if outcomes become unfavorable. Some residual risk of nonperformance will remain, but the probability of performance is maximized, recognizing the collateral cost.