21, rue d'Artois, F-75008 PARIS
http : //www.cigre.org

CIGRE US National Committee
2023 Grid of the Future Symposium

A Framework for Assessing Marginal Cost Pricing Through a System Reliability Lens – Is it Back to the Future?

Hala BALLOUZ

Electric Power Engineers, LLC

USA

SUMMARY

In recent years, our ability to ensure that electric energy delivery is at high levels of reliability and resilience has been a specific concern of the public, utilities, and regulatory entities based on the increasing challenges experienced by electric power systems. The challenge of the pursuit of the double-edged goal of lower cost and higher performance has been amplified by an ever-increasing focus on electricity as the primary energy source. As such, some have questioned electric grids' ability to support the transition to meet clean energy goals while also remaining reliable and resilient.

The push for clean energy and electrification is impacted by two primary trends related to grid operations: 1) the uptake of clean, intermittent technologies, such as wind and solar, as well as other alternative electricity production /storage resources like batteries and hydrogen; and 2) a broad adoption of digital tools to manage intermittent resources as well as fluctuating loads with the proliferation of electric vehicles, solar rooftop, and other decentralized energy solutions.

In addition, debates are ongoing about the relative benefits of centralized versus decentralized energy sources and what roles they play in the energy transition, with related discussions about the potential of demand-side management solutions as well as virtual power plants and how that all will be integrated in balancing bulk power systems.

The industry's clean energy transformation, including the integration of significant amounts of intermittent renewable resources, has coincided with the compounding of significant weather events, and the transition from a largely regulated, vertically integrated business model, in many areas of the country, to a segmented structure encompassing generation, transmission, distribution, and retail energy. Central to this deregulated, alternative structure is a competitive bidding system, "marginal cost pricing;" dictating the development and operation of the electricity production system. This paper reveals that, while these competitive markets based on marginal cost pricing have contributed to reducing wholesale power cost pricing in some cases, such models have not incentivized adequate investment in resources and infrastructure to support the reliability and resilience needs of our grid, especially with the additional challenges inherent in the clean energy transition.

The paper identifies flaws in the marginal cost pricing approach and assumptions that have led to the failure of the electricity market in adequately supporting reliability and resilience. These flaws might be best characterized as the "missing money" problem in an imperfect competitive market. There is money missing from the "marginal price" that would otherwise ensure adequate reliability.

Furthermore, the paper discusses the new challenges introduced by the energy transition, both in terms of technology and structure. It calls for a re-evaluation, hinting at a "Back to the Future" approach by retrieving some attributes from the regulated and/or integrated markets to be reflected in new service products for bulk power supply and delivery practices that can move to correct some of the lack of reliability improvement. It emphasizes the need to integrate lessons from our past to create a better future, proposing a framework that creates guardrails to define reliability and resilience and enable elements that balance cost, reliability, and resilience with the latest technological innovations.

KEYWORDS

Marginal cost pricing, competitive markets, Grid Reliability, Resiliency, Resource Allocation, Dynamic Competitive Equilibrium, deregulation

hballouz@epeconsulting.com

Introduction

Today, the electricity industry, along with the various electric grids, is "skating on thin ice" from an operational reliability perspective, having significantly stretched its capacity limits in many areas of the country. This over-extension is causing significant system stress. The shift from regulated integrated systems to deregulated competitive markets in certain regions has given rise to electric systems that are stretched thin in capacity with a lack of robust consumption management. [1,2] This results in highly vulnerable configurations from a reliability perspective. Thus, competitive electricity markets seem less adept at fostering improvements in reliability and resilience, yet generally effective at reducing wholesale energy prices under normal operational states [3].

Marginal cost pricing was introduced in the competitive markets as a mechanism to replace the day-to-day decision-making of identifying the lowest cost solution to meet the power demand. The concept of locational marginal pricing in wholesale energy markets is based on reflecting grid congestion in the bulk transmission system, often preventing the supply from reaching its intended consumer – end-user or distributor/local power company. Yet, the move to heavily deregulated competitive market practices on systems initially designed for vertically integrated markets using traditional tools has led to capacity-lean, energy-oriented systems. In essence, this has resulted in a transition from fully integrated systems (Generation, Transmission, Distribution and Consumption (G,T,D & C)) built with ample capacity to handle disruptions to systems that now struggle with major operational issues with capacity shortfalls, particularly under stressed conditions.

This brings us to a fundamental question about the efficacy of the competitive market frameworks and their use of marginal cost pricing to maintain reliability: It does offer lower wholesale prices, but seemingly at the expense of not incentivizing investments in long term reliability. Our competitive market frameworks have not met our expectations of reliability and resilience, resulting in many grids where underinvestment in infrastructure, technology and the resource mix have left us skating on thin ice rather than delivering "thick," reliable systems. These investment shortfalls are characterized as having a "missing money" problem [4]; there is money missing from the "marginal price" that would ensure adequate reliability.

Our goal in this paper is to frame the challenges of marginal cost pricing in the current energy market structure and examine how to achieve high reliability; addressing the "missing money" problem with new product alternatives.

An Electricity Industry in Transition

The electric industry has experienced upheaval for over three decades. The stages can be conveniently characterized as being in one of four categories: 1) Vertically Integrated; 2) Deregulated; 3) Today's Challenges – a technology revolution across the full spectrum of resources and infrastructure while dealing with extreme weather events along with a major shift to gas fuel generation and renewable resources; and 4) the future challenge of electrification. Not all parts of the world, or for that matter North America, have embraced deregulation, yet are certainly experiencing #3, today's challenges, and will experience the future challenge of #4, electrification of transportation and other sectors.

Vertically Integrated

In a regulated system, there's a consistent standard: the consumer-level price. A regulatory body, typically a utility commission or local city council, provides the guardrails that set the price to the consumer through a rate structure, aiming to ensure reasonable prices while allowing private utilities a fair return on investment in their designated territory while meeting specific reliability criteria. This arrangement, often termed the utility compact, provides utilities service territory exclusivity in exchange for capped returns on investment and commitments to meet supply requirements (or serving the public interest in the case of not-for-profit utilities). Historically, this has meant local utilities were "vertically integrated," overseeing and designating the infrastructure of production and transmission, including the distribution of electricity to consumers, rolling up all these costs in the price that the consumer paid.

Two crucial points [5] arise:

1. As a vertically integrated utility, there's centralized planning from production to consumption. This centralized control simplifies design and swift execution for reliability requirements.

2. However, concerns prevail that such an integrated structure stifles wholesale competition and introduces inefficiencies that may increase consumer price for electricity. These worries largely prompted the industry's pivot towards deregulation and competitive markets.

Deregulation and Competitive Markets

"Deregulation" emerged as a counter to the norm of highly regulated utilities. In the 1980s, countries like Chile began to deregulate their wholesale electricity sectors. The U.S. followed suit in the 1990s, and this transition continues. Deregulation remains somewhat experimental in the U.S., with some states showing signs of a willingness to reconsider the core programs that have been put in place at the retail level, feeding into questions about market mechanisms at the wholesale level. In recent years, the impetus of these concerns has been related to reliability, although questions persist as to whether lower wholesale prices are being conveyed as lower consumer prices – this latter issue is beyond the scope of this paper. Most regions that have deregulated at the wholesale level have employed marginal cost (competitive) pricing for wholesale energy markets as a foundational element. This process involves an auction-style setup, where generating utilities bid a specified unit of power/energy at a specific price. The lowest bid/bidders secure the contract to supply the total projected demand for energy on an hourly basis. The market operator selects the necessary number of bids to fulfill the total power needs and pays all suppliers the same price, which corresponds to the highest accepted price in the stack of bidders to meet the load demand. Consequently, this means that many of the bidders receive remuneration at a much higher value than their bid price.

The aim of competition is to delineate free-market operations from natural monopolies and open wholesale markets to competitive bidding. Historically, low-voltage power delivery (power distribution) has remained monopolistic for the most part, with local power companies imposing fees atop costs and bundling all costs under a rate structure conveyed to the end-use consumer. Now, end-use consumers in several states have the choice to purchase electricity from any supplier, although the purist form of this model only exists in Texas. Transmission lines, offering alternative paths from generators to distributors or even possibly direct served end-users, have not been viewed as monopolistic assets. Power generators might bid for transmission capacity in certain markets to deliver their electricity to their customers. The Independent System Operator/Regional Transmission Operator (ISO/RTO) overseas the transmission and generation execution and market dispatch operation of wholesale actions, while the Federal Energy Regulatory Commission (FERC) oversees the ISOs/RTOs and determines appropriate “market-based” rates, terms and conditions.

The emergent Distribution System Operator (DSO) concept may redefine local distribution. As Distributed Energy Resources (DERs) gain traction, local power producers could feed electricity into the local grid for another consumer or into the bulk power system, potentially forming "microgrids." These developments further challenge established norms.

In summation [5], deregulation sought to:

- Introduce competition via deregulation at the wholesale power level.
- Unbundle the once vertically integrated system and segregate generation, transmission, distribution, and retail energy consumption into separate entities and/or competitive markets.
- Set up ISOs/RTOs for overseeing the transmission grid and dispatch of bulk central generating resources.
- Encourage cost-effective wholesale solutions and ensure location-specific reliability through ISOs/RTOs using marginal cost pricing, and capacity markets.
- Enable end-user choices at the distribution level, which has not gained traction in most of the country yet.

Today’s Challenge: Technology and System Revolution

In the contemporary landscape, several key electric industry changes (referenced in Table 1) are challenging both reliability and resilience. While the desire of the electric industry is always to keep electricity prices from rising or perhaps to reduce prices, such desire should not come at the expense of reliability. Significant efforts are underway to transition from fossil-fuel solutions to renewables and other non-fossil alternatives. This shift necessitates a profound rethinking of the electricity sector's operations. Changes span from the introduction of new energy sources to reconfigured networks. The debate between decentralized and centralized approaches is ongoing, with many advocating for a comprehensive digital solution to enable a decentralized system. As the momentum builds toward low-carbon energy solutions, there's a marked increase in electricity usage, positioning electrification as the cornerstone of an overall (clean) energy strategy. This underscores the importance of both consumer cost and reliability in our evolving energy paradigm.

Table 1: Evolving Industry Changes [5]
*** Electric systems with undersized capacity margins**
*** Climate/weather extreme changes**
*** Aging infrastructure in some cases**
***New resource technologies both centralized and distributed with some at zero marginal cost**
***Shift in traditional resource mix and increase in intermittent resources**
***Changing & increasing consumer demand**
***Electrification including vehicles and heat pumps**
*** T&D seams long dissolved**
*** Siloed G, T & D planning**
*** Increased use of digital technology**

A paramount change in this context is the potential replacement of our current petroleum-based transportation fuel with electric powered vehicles. The reliability of our total energy supply and delivery system is increasingly at risk should we rely on inadequate supply and delivery infrastructure and assets. Connecting and integrating various energy sources and technology through a bulk energy system that was not designed and built for these intended uses puts electric grids at risk. Eventual access to large-scale and long-term energy storage to facilitate operations will alleviate some of this concern, and help our current energy markets, but these solutions are not yet widely available.

Electrification

Major advancements are in process to expand electricity usage. In particular, the goal of ***electrification*** is to convert to electricity those end-use applications that do not use electricity today as their primary energy source, such as transportation vehicles and heating elements. Further, those end-use applications already using electricity are seeking more efficient outcomes. To this end, reliability is paramount to the future electric grid. Now, more than ever, electricity supply needs to be impeccable.

What Is Marginal Cost Pricing

Marginal-cost pricing, within economic theory, refers to the method of setting the price of a product to match the additional cost incurred to produce the last or marginal unit. The marginal cost pricing methodology is used as a mechanism to satisfy market competition needs. Using this approach, producers charge for each product unit the exact amount that accounts for the materials (mainly fuel) and direct labor costs, *excluding fixed overheads*. While businesses might price close to the marginal cost during periods of low demand, this strategy is not viable in the long run as it does not cover sustained operational costs.

During the mid-20th century, advocates of perfect competition—a situation where firms produced almost identical products at identical prices—were in favor of the efficiency that marginal cost pricing introduced. However, some economists, including the likes of Ronald Coase [6], championed the market's capacity to define prices. They appreciated the signaling quality of market pricing, which communicates valuable information about the products to both buyers and sellers. They also noted that mandating sellers to adhere strictly to marginal cost pricing might jeopardize their ability to meet their fixed costs.

Marginal cost pricing and the “merit order” in the wholesale electricity market

In the context of electricity markets, marginal pricing is the method wherein a wholesale electricity price paid to a supplier is established by the variable cost of the most expensive production unit required to meet demand. Typically, such prices are determined in short-term wholesale markets, like the day-ahead market. This pricing mechanism is frequently represented using the "merit order curve," a graphical construct showcasing the marginal power generation costs from all suppliers. Notably, in such systems, all selected generators receive the same price for their production corresponding to the highest marginal cost selected.

Challenges Facing Marginal Cost Pricing Markets

Upon review of operational markets such as ERCOT, PJM, ISO New England, and MISO, a consensus emerges: the current competitive market designs fall short of providing the reliability required for today's demands. The shifting generation mix, emergence of new technologies, policy alterations, aging infrastructure, weather conditions and capacity constraints have collectively contributed to the vulnerability of electric grids and have altered the paradigm for marginal cost pricing assumptions, imposing additional challenges on this mechanism. For example, these market constructs are not designed for resources with zero fuel costs, such as renewable energy. This vulnerability is further accentuated by the merging of gas supply and electric systems as now there is a loss of diversity and certainty of fuels for the electric grid, in turn making the electric system heavily linked to the status of the gas system, for example. Historically, there has been good separation of the gas and electric systems, thus isolating the consequential impacts of one system on the other. The addition of electric vehicles, once they replace gasoline-fueled vehicles, will remove an additional element of diversity by directly linking transportation to the electric supply.

In addition, it is crucial to highlight the evolving challenge of weather extremes. The influence of weather on system stability is becoming increasingly pronounced. From bitterly cold temperatures that hinder equipment functionality and fuel transport to extreme heat that burdens and restricts equipment capabilities – both extremes pose significant operating challenges. The incorporation of weather-dependent power generation methods complicates system operations—something the competitive markets were not initially designed to accommodate. Furthermore, the broad impact of certain weather events can lead neighboring electric systems to experience similar challenges, thereby depleting their marginal assets and impeding resource-sharing efforts.

A recent paper [7] offered some insight as to trends in reliability expectations under deregulation. The paper uses NERC's ten-year annual Long-Term Reliability Assessment (LTRA) report's forecasts for ***anticipated reserve margin*** based on plans effective as a barometer of the ability of the various NERC reliability assessment areas (RAAs) to have adequate supply to meet forecasted load. Further, data from the EIA is used to correlate RSAs with states plus the District of Columbia with those areas where deregulated electric grids have been established. Consequently, six of the eighteen NERC RAAs house the deregulated entities. **For the reference period of evaluation, the 2013 LTRA report spanning 2013-2023 reveals that four of the six RAAs have multiple years where they fail to meet NERC reference reserve margin criteria**. (See Table 2.)

Table 2 NERC Forecasted Reserve Margins for the Six Assessment Areas that Include the Deregulated 15 States and the District of Columbia [7]

Region	RRM	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
TRE- ERCOT[1] ****	13.75% ****	13.74% ****	11.59% ****	10.34% ****	10.46% ****	9.34% ****	7.36% ****	6.44% ****	5.91% ****	5.11% ****	4.43% ****
MISO[1] ****	14.2%	18.28%	12.13% ****	7.00% ****	6.29 % ****	5.54% ****	4.86% ****	5.65% ****	4.90% ****	4.16% ****	4.43% ****
NPCC-New England (ISO-NE)[1] ****	13.85%	29.02%	24.70%	22.39%	16.12%	15.36%	14.64%	14.02%	13.30% ****	12.69% ****	12.07% ****
NPCC-New York (NYISO)[1] ****	17.0%	22.71%	21,22%	19.75%	18.85%	17.89%	16.85% ****	15.77% ****	14.73% ****	13.83% ****	13.03% ****
PJM Interconn. **	15.9%	30.86%	24.93%	23.40%	22.67%	21.39%	21.07%	19.66%	18.30%	17.11%	15.93% **
WECC- NORW	17.48%	33.94%	N/A	N/A	N/A	33.54%	N/A	N/A	N/A	N/A	24.27%
Total: 18 U.S. AA[2]	14.94%	27.52%	N/A	N/A	N/A	20.24%	N/A	N/A	N/A	N/A	15.38%

[1] The four asterisks identify the four NERC U.S. Assessment Areas failing to meet NERC Reference Reserve Margin Standards, shown by year, from 2014-2023.When forecast reserve margins.

[2]NERC forecast reserve margin totals, from 2014-2023, include all 18 NERC U.S. Assessment Areas (AA),

At least three of these markets continue to face struggles as cited later in this paper and recorded by NERC. Such a revelation speaks to the fact that deregulated markets are largely and significantly failing to incentivize the necessary infrastructure and resources to keep an adequate reliability outlook based on planned supply and forecasted demand. Further, when all eighteen RAAs are aggregated, it shows that anticipated reserve margin is satisfied for the entire US grid; this indicates that regulated entities are sufficiently large to offset those RAAs where compliance is not achieved. It is of note that regulated markets have also experienced challenges, but they have largely been successful using known outcomes since those markets have the traditional vertically integrated control over the designation of resources, infrastructure and management of load.

Energy-Only vs. Capacity Markets

An "energy-only" market compensates for power once it has been produced. In contrast, a "capacity" market rewards the mere readiness or availability of capacity for power generation. Energy markets have proven the ability to lower wholesale energy prices through competition. However, to guarantee a consistent supply, energy-only markets introduce various “flexibility mechanisms or alternative products.” These modifications are needed to make up for the shortfall of the competitive market to provide capacity for reliability. Conversely, a Capacity Forward Market is a long-term wholesale electricity platform aimed at ensuring resource adequacy for reliability requirements. This market setup is crafted to stimulate economic investments in both supply and demand capacity resources. These auxiliary offerings are sometimes referred to as “product alternatives.”

Table 3 [5] provides a summary of system deficiencies that have been observed at various ISOs. These energy and capacity market designs must demonstrate that capacity incentives will not merely generate profits in specific situations without genuinely encouraging the right development of generation capacity and infrastructure to maintain reliability. In ERCOT's case, the prevailing market structure, characterized as predominantly marginal/surge pricing (an alternative product offering), has not/ does not foster long-term investment in vital resources and infrastructure to uphold reliability and resilience. Further, related to capacity markets, FERC Commissioner Christie [8] pointed out, "The real challenge at FERC is determining whether these multi-state capacity markets, including New England's, can fulfill our expectations of delivering reliable power at affordable prices." He further speculated that devising solutions could take up to two years. Concerning the New England ISO, Commissioner Allison Clements [8] reiterated her focus on the region's reliability issues, “The recent weather events in Texas, the Midwest, and the Tennessee Valley highlighted the differing demand management strategies by ERCOT and TVA”. TVA's ability to efficiently reduce load through collaborations with local power entities – a "virtual" vertical integration – stands out. While TVA can still make improvements, it clearly had a robust flexibility plan to manage unforeseen weather events.

	Storm	System	Market Description	Insights	Unplanned
2011	Arctic Winter Blizzard	ERCOT	Energy Only Market	Near blackout: unplanned outages, resource adequacy issues, and operational challenges	29,700 MW
2014	Polar Vortex	ERCOT & Eastern Interc.	Energy & Forward Capacity Market	ISO-NE - system continues to operate on a knife’s edge for extreme winter weather events	19,500 MW
2018	Arctic Cold Front	MISO	Forward Capacity Market	Sharp energy price increases with expensive natural gas & transmission congestion	15,800 MW
2021	Winter Storm URI	ERCOT, SPP & MISO	Energy & Forward Capacity Market	Compounded with uncertainty in Wind availability	61,300 MW
2022	Winter Storm ELLIOT	TVA & PJM	Varied: Regulated and Unregulated	Compounded with current and upcoming power plant closures (52 GW)	90,500 MW

Table 3: Electric System Shortfalls [8]

The contrast between ERCOT's ineffective load shedding plan, which caused significant supply/load balance issues, versus TVA's prompt load shed actions with limited general load interruptions to managed outcomes, is stark.

Observations made by NERC & FERC on its September 21,2023, report, addressing the latest winter storm events, observed striking commonalities amongst all five events where the majority of the causes lied in the following categories that failed in planning for such events:

- Significant levels of incremental unplanned electric generating unit losses with top causes found to be mechanical/electrical related to freezing.
- Significant Fuel Shortages – namely natural gas production and outages resulting in significant unavailability of electric generating units.
- Short-range forecasts of peak electricity demand significantly less than actual demand for some Balancing Authorities in event area.

A close examination of Texas and the ERCOT system (see Table 4) provides some good insight on the Energy Only Market. Further, it is an excellent example of a competitive market heavily designed using real time energy price signals.

The Texas Alternative Products	The Challenge	Lessons Learned
Goal: Incentivize investment by generation companies and improve system reliability	The Price Cap approach revealed the limitations of a flawed marginal pricing' heavy market environment.	Major Flaw: Time delay in market response to price signals
Implementation of market price caps		Price Cap band-aids worsen the energy pricing signal flaws
In 2011, the Public Utility Commission of Texas raised the caps from $3,000/MWh to $9,000/MWh.		We must also learn from PJM & other Capacity Market shortcomings.
In 2023, the Public Utility Commission unanimously approved the Performance Credit Mechanism (PCM)	Many skeptics question the effectiveness of the PCM, which resembles a capacity market with the latter facing challenges	Alternative reward systems are needed for long-term investment certainty for reliable energy systems

Table 4: ERCOT and the Texas System [5]

Shortfalls of Marginal Cost Pricing [10]

Marginal cost pricing often falls short in promoting long-term investments [1,2,3]. For the bulk power system, installing generators, power lines, and other necessary equipment can take years, while it may require months for smaller assets. For instance, introducing an operational transmission line or a power plant can consume up to a decade in California. In Texas, even though the timeline is comparatively shorter, it's generally not under three years. Deregulation aimed to foster a competitive wholesale market while simultaneously prompting the development of more efficient assets, but that has not occurred in favor of reliability and resilience.

The marginal cost pricing, deregulated structure, functioned adequately for a period when our grids possessed sufficient capacity to ensure reliability. As time progressed, with short term energy prices serving as the principal commodity steering these markets and pricing acting as the dispatch settlement metric, the systems gravitated towards meager reserve margins and asset maintenance. This situation was observed in various U.S. regions, resulting in grids that were precariously balanced. New stakeholders have been introduced that employ intricate processes to share information and synchronize activities, currently the advanced technology, however transparency to envisage or manage the grid as an integrated entity is lacking. Today, the most striking chasm lies between transmission and distribution at both the grid and resource levels, leaving our systems devoid of the unified metrics essential for maintaining comprehensive reliability in planning and operations.

In light of the 2021 power system crisis [11,12] experienced by ERCOT as well as SPP and parts of MISO, certain conclusions appear clear (though the causes and implications warrant detailed scrutiny):

- The market blueprint was ill-equipped to spur stakeholders into investing in system capacity, maintenance, winterization, or reliability. This negligence culminated in extensive vulnerabilities across the grid, impacting generation capacity, transmission and distribution networks, contingency reserves, and beyond.

- Numerous stakeholders, largely load entities, grappled with exorbitant electricity bills, leading to unpaid dues to grid operators and utility companies, consequently spiraling into multiple bankruptcies, while some other entities benefited from the market surge pricing to make very large profits. This scenario starkly illustrates that surge pricing failed to address the inherent deficiencies of a market sculpted around the marginal cost of power.

Operating reserves are not planning reserves.

A nuanced, but paramount, flaw in contemporary pricing models pertains to the way balancing authorities delineate operating reserves. This discrepancy became glaringly apparent during the ERCOT debacle [11,12]. During periods when thousands faced service disruption (but operating reserves were deemed adequate), the market-wide locational market price (LMP) decreased to $1,200/MWh. Identifying the misalignment between the value of lost load (VOLL)-based scarcity pricing mechanism and actual market scenarios, the Public Utility Commission of Texas embarked on the “unprecedented step” of directing ERCOT to pin market prices at $9,000/MWh. Their rationale was clear: if consumer loads were being curtailed, the price should rest at $9,000/MWh. Adhering to this mandate, ERCOT maintained this price point until it declared the cessation of the power emergency on February 19, 2021 [13]. Unfortunately, major disruptions continued, and it’s evident that customers facing service disruptions effectively provided ancillary services to the grid, ensuring adequate operating reserves. These customers, not the generation companies, should have been compensated during this crisis with the high prices that generators were paid, which demonstrates yet one more flawed outcome of current market structures.

Marginal Cost Pricing in a World with Imperfect Competition

There are some well-documented shortfalls in imperfect electricity markets. In part, some issues are a result of roles that electricity providers are tasked with protecting. These distortions to a fully competitive market have created unintended consequences to what was envisioned to be a very progressive step.

In 2017, the Department of Energy’s National Renewable Energy Lab (NREL) wrote a report [15] on marginal cost pricing when applied to the imperfect electricity market. The essence of this report was a treatise on the shortcomings of marginal cost pricing or “auction-based competitive wholesale markets” when deployed in the real world. Moreover, the report recognizes early on that electricity markets are not “perfectly” competitive because they support certain public obligations and embrace other externalities peculiar to the public electricity business. Several paragraphs have been extracted from the Executive Summary of the NREL report that summarize some key features.

The U.S. electric industry changed from the regulated vertically integrated structure to a deregulated competitive one with auction-based competitive wholesale markets. The intent was to provide a reliable supply of power at the lowest reasonable cost to the consumer with transparency while mitigating market power. All intent is to prevent a market actor from influencing the market price or excluding competition. This necessitates market structures and operating rules that ensure revenue sufficiency for all participating generators, which are needed for resource adequacy purposes. Wholesale electricity markets employ marginal-cost pricing to provide cost-effective dispatch such that generators are compensated for their operational costs.

However, marginal-cost pricing alone cannot guarantee cost recovery outside of perfect competition, and current electricity market structures have at least six attributes that preclude them from functioning as perfectly competitive markets. These attributes are shown pictorially in the graphic of Figure 1.

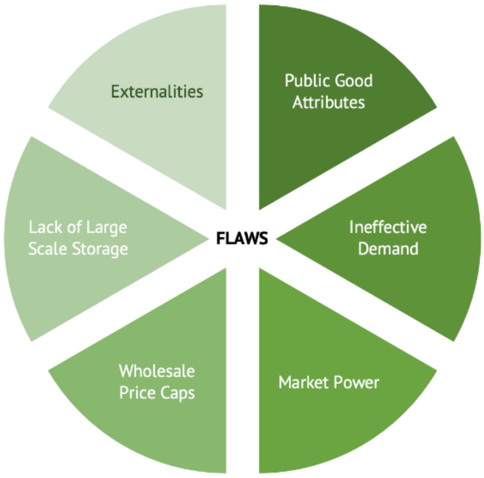

Figure 1: Flaws in Marginal Cost Pricing Model [15]

Some of these attributes—namely externalities and public good attributes—are classic sources of market failure but exist in today's electricity wholesale market construct. The ineffective demand curve feature is arguably the most impactful contributor to market failure and refers to the inability of consumers to express their actual demand for electricity. It results from the "demand-side flaws" of demand inelasticity and the system operator's inability to control the real-time flow of power to specific customers. Other attributes—primarily market power, lack of large-scale storage, and wholesale price caps—have compounding causes and effects that amplify underlying market failures. For example, market power exists in part because of significant barriers to entry that yield monopolistic tendencies. Electricity markets are regulated to minimize this market power, but the resulting regulatorily-imposed average-price based retail rate structures contribute to the observed inelastic demand mentioned above.

Similarly, the wholesale price caps that are implemented to restrict market power may also prevent prices from reaching levels needed to ensure adequate revenues for generators. In addition, imperfect information related to the planning and operation of the power system, such as uncertainty in load growth and future economic and policy factors, further amplifies these current market structure failures.

Until (and unless) these contributors to market failure are ameliorated, some form of corrective action(s) will continue to be necessary to improve market efficiency so that prices can correctly reflect the needed level of system reliability. Many of these options necessarily involve a contract with some form of administrative or out-of-market actions, such as scarcity pricing, capacity payments, bilateral or other out-of-market contracts, or some hybrid combination. A key focus with these options is to create a connection between the wholesale electricity market and long-term reliability/loss-of load expectation targets, which are inherently disconnected in the native, unaltered markets as previously described.

The addition of low marginal cost resources, such as generators fueled by low-cost natural gas and near-zero marginal cost wind and solar generators, can further exacerbate revenue sufficiency and resource adequacy concerns caused by these underlying market failure contributors. These low marginal cost resources effectively suppress energy prices and reduce the capacity factors of conventional generators through the merit-order.

Additionally, the uncertainty and intermittency of variable generation resources, such as wind and solar, requires more system flexibility. This can be achieved by a wide range of supply-side and demand-side options for physical flexibility. This also necessitates institutional flexibility options, including new market designs, greater regional coordination, and fair cost allocation.

Future research is needed to assess optimal market designs as alternative products that are technology neutral, robust to generator fleet composition, and politically/socially acceptable, while ensuring revenue sufficiency of power system assets needed for reliability. Various modeling tools are needed for this effort to span multiple time horizons, including planning and operational decision time-frames.

An additional study [16] indicates that fully deregulated and competitive markets are not providing lower costs of electricity as the players end up marking up cost to play and the increased costs offset any marginal cost gain. The money is not sufficiently going to resources contributing to reliability attributes of our resources and systems, but rather to making more money in an imperfect competitive market environment.

Further, the study [16] provides additional insight into the impact of deregulation on wholesale electricity prices and reliability. The work effort does focus on advancing the theoretical attributes of marginal cost pricing methodology and its shortcomings in the practical application to wholesale electricity markets. It is of note that market-based prices do provide incentives for profit-maximizing firms to reduce costs for power production, but firms that have ***market power*** also have an incentive to increase markups. Market power has also been identified in the NREL report [15] as one of the flaws in applying marginal cost pricing techniques. When cost efficiencies are outweighed by an increase in markups, market-based prices can result in higher values than regulated rates. Without efforts to protect and strengthen competition, such as regulatory oversight and antitrust enforcement, markets may yield lower consumer welfare. When it is difficult to limit market power, consumers may prefer a regulated monopoly to markets, and regulators face a tradeoff between efficiencies in production and higher prices.

The study [16] does explore this tradeoff in the context of the deregulation of the U.S. electricity sector, which began in the 1990s. Deregulation efforts included the introduction of market-based wholesale prices and restructuring measures to introduce competition into the upstream generation market and the downstream retail market. Over 20 years later, it largely remains indeterminant as to the precise consequences of many of these efforts. But, much more evidence exists about the impacts on prices. Contrary to the objectives of deregulation, this study shows that prices increased in deregulated markets, despite a modest reduction in marginal and average variable costs. Thus, the ***increase in markups dominated the efficiency gains***, indicating the widespread exercise of market power. These findings show that deregulation does not necessarily lead to lower prices for consumers. The above cited research provided some insight into one of the known flaws (market power) in current competitive markets. It reveals some reduction in wholesale power costs, but little of that conveyed to the consumer. Moreover, it shows how market power is a deterrent to achieving the goals using the marginal cost pricing technique. The study shows substantial price increases for consumers in deregulated states relative to consumers in regulated states; deregulation can lead to unintended higher prices and markups when market-based equilibrium prices are above marginal costs.

A reference to the United Kingdom's experience with marginal cost pricing and carbon free generation is referenced in [14] as "Navigating the Crises in European Energy: Price Inflation, Marginal Cost Pricing, and Principles for Electricity Market Redesign in an Era of Low-Carbon Transition." The reference accentuates the struggles with marginal costs pricing in current markets. Within the report, challenges and corresponding implications for policy are identified. Short-run (in electricity, half-hourly) marginal-cost-pricing means that the most expensive operating sources set the electricity price for most of the time. This forms an inappropriate basis for funding investment in long-run assets with low operating costs. *Fossil fuels set the electricity price for most of the time, at levels which are now much higher than the energy cost of at least half the system* (recent renewables and existing nuclear) – so the wholesale price of electricity is way above the average cost of generating it. This dependence on fossil fuels to set the wholesale price in practice introduces high volatility and uncertainty in the price that non-fossil investors would receive in the market, making it an extremely inefficient basis for funding large-scale renewables. Renewables investment in practice has mostly been funded outside the wholesale market, leading to large technology cost savings which only partially feed through to electricity prices. The result is an increasingly disjointed system with prices to most consumers still mainly set by fossil fuels even as governments accelerate efforts to decarbonize electricity – a combination which itself is unsustainable. Regulators face a very difficult problem.

Last but not least, the concept of the "missing money" problem [4] adds a significant layer to the shortfalls of marginal cost pricing and imperfect competitive markets, addressing the money missing from the "marginal price" that otherwise should have addressed providing adequate reliability.

Developing Alternatives to Marginal Cost Pricing

An alternative concept to wholesale short-run marginal cost pricing has been introduced by one of the authors of this paper, which extends to the grid edge. The concept is called ***dynamic competitive equilibrium (DCE)*** [10]. This concept emphasizes incorporating consumer preferences into a holistic generation and delivery strategy, diverging from the prevailing marginal cost pricing model. A notable insight from the model challenges a standard assumption, highlighting that real-time equilibrium prices often are not linked directly to marginal costs or values. By sidestepping pitfalls commonly associated with traditional pricing models that can severely compromise system reliability, DCE makes a compelling case for a departure from the marginal cost pricing paradigm.

Outlined below are key principles underpinning DCE:

Planning:

Constructing a reliable grid demands a multi-faceted planning approach encompassing:

- Long-term strategies for reliability, spanning multiple years.
- Resource allocation, informed by load forecasts over several days.
- Flexibility with real-time response mechanisms to unpredictable events, such as fluctuations in wind or solar generation or unexpected consumer behavior, necessitating actions within minutes.

Successfully navigating such intricacies demands collaboration among experts and active participation from all stakeholders. To ensure enduring reliability, we need to consider alternatives to the transient nature of marginal cost pricing models.

In many industries, services are procured via meticulously crafted, financeable contracts. Central to this is the proposed creation of a Reliability System Operator (RSO). This entity, functioning as a centralized planner, would craft optimal resource expansion strategies across markets. The RSO would assimilate many roles currently undertaken by RTOs or balancing authorities (BAs), forging contracts with generators and resource aggregators to guarantee cost-effective reliability. Importantly, real-time pricing wouldn't be the foundation for these contracts. A detailed account of the RSO's comprehensive roles can be found in the original paper. [10].

For instance, PJM effectively leverages planning and contract tools. Recognizing potential severe winter generation capacity gaps, it employs regulated forward supply contracts, buttressed by ***stringent penalties*** for delivery lapses during supply crunches. PJM's triennial system mandates utilities to secure capacity to meet cumulative customer demand [8]. Another example is the British capacity market, which orchestrates annual auctions targeting both immediate and long-term capacity needs. Such coordinated strategies address the temporal complexities and contractual needs often sidelined in conventional models that assume generation companies are already equipped to hedge against delivery failures.

Distributed Intelligence:

Drawing inspiration from the Internet – where equilibrium between supply and demand is realized through distributed network control – regulations can be devised for ensuring grid stability and consumer privacy simultaneously.

Embracing the emerging science of ***demand dispatch,*** a key attribute of flexibility, forming a resilient, decentralized, and automated control design could cater to grid operators' needs while safeguarding consumer comfort, privacy, and convenience [10]. Resource aggregators, like Enbala Power Networks, Comverge, CPower, Enel X (and potentially utilities in an expanded role), are pivotal for actualizing demand dispatch blueprints. Consequently, there's an imperative for economic models and regulatory frameworks that elucidate the roles of these aggregators, especially prioritizing grid reliability [5]. Further, policy designers and engineers should collaboratively enhance mandatory appliance standards [5]. This ensures appliances can virtually store energy, bolstering grid reliability without inconveniencing consumers.

Back to the Future

General conclusions in this paper are based on actional outcomes that drive the process to keep electricity prices competitive and fair but ensure that appropriate levels of reliability and resilience are achieved. Multiple studies have been show-cased that demonstrate that marginal cost pricing methodology has provided lower wholesale prices in many applications, but appears to not be providing improvements in supporting the right infrastructure for reliability and resilience. Advances in renewable resources are not being effectively used in the generation mix because marginal cost pricing principles were built with fossil powered generation options in the underlying assumptions, and not renewables with near zero marginal cost price. Advances in clean energy strategies and electrification are adding pressures to grid reliability with rapid retirement of fossil powered generating resources and devices. And, lastly, extreme weather events are testing grid system integrity and have become major tests for grid reliability.

Consequently, there is a strong need to step back and clearly articulate what features must be introduced to and altered in the competitive market setting, starting with defining metrics for reliability and resilience that are fit to address the ecosystem of today. Features from the past have some application in the marketplace; the use of integrated planning activities across transmission, distribution, generation and grid-edge are desired to set the guardrails to reliable load service. Mechanisms to provide the integrated system benefits to the market and stakeholders must be created and applied. New alternative products to marginal cost pricing need to be derived within these guardrails to better utilize all resources and infrastructure to the fullest of their value and address the right reliability and resilience metrics. These alternative products, pricing and other, should embrace the following principles:

- **Define:** Specify the product being bought/offered, that is, **--** energy, capacity, flexibility, operation and maintenance, load shed amount and duration, etc. The competitive market has failed to signal correct product features desired.
- **Incent:** Provide sufficient fiscal incentives via remuneration practices that compensate resource providers for products offered. A full offering price model is required to attract power suppliers and demand response entities with correct power providing features.
- **Perform:** Establish performance measures that have sufficient details to specify clear requirements and mechanism to reward effective contribution to reliability and resilience. Resource providers for example must have details on performance measures such as when, where and under particular operating conditions. The purpose is to guide stakeholders with a clear understanding of expectations and compensation measures to effectively contribute to reliability and resilience.
- **Penalize:** Develop a penalty structure for suppliers with fiscal assessments to offset a failure to deliver performance. Market participants need to provide full service, with requirements to participate in both energy and capacity markets, as well as any other essential products. This requires careful design to ensure that penalties do not introduce risk factors that drive much needed firm resources out of the market.

Acknowledgment

The author thanks Tom Reddoch (Consultant), Joy Ditto (Joy Ditto Consulting, LLC), and Mohamad Ahmed (EPE, LLC) for their thoughtful reviews, comments, and suggestions.

BIBLIOGRAPHY

[1] Kavya Balaraman, “The power grid faced heat waves, record demand and tight conditions in 2022. What happens next?,” Utility Dive, November 2022.

[2] Andy Stone, Bruce Edleston, and Mark Kolesar “Grid Forward Debate: Has Electricity Deregulation Led to Better Community Outcomes? – a podcast,” Kleinman Center for Energy Policy, October 27, 2020.

[3] Judy Chang, Johannes Pfeifenberger and John Tsoukalis, “Potential Benefits of a Regional Wholesale Power Market to North Carolina’s Electricity Customers,” Brattle Group Report, April 2019.

[4] Michael Hogan, “Follow the missing money: Ensuring reliability at least cost to consumers in the transition to a low-carbon power system,” Economic Journal, January/February, 2017, Pps 55-61.

[5] Hala Ballouz, “Marginal Cost Pricing from a Reliability Lens,” presentation @ the IEEE Summer Power Engineering Meeting, Orlando, Florida, July 2023.

[6] Ronald Coase, British Economist, Clifton R. Musser Professor of Economics at the University of Chicago Law School.

[7] Eric L. Prentis, “U.S. Electrical System Reliability: Deregulated Retail Choice States' Evidence and Market Modeling,” Liebnez Center for Economics, 2021.

[8] Ethan Howland, “PJM, MISO, ISO-NE. – The capacity markets are not all right,” Utility Dive, March 17, 2023.

[9] David Patton, Potomac Economics - MISO Independent Market Monitor, “IMM Quarterly Report, June 13, 2023.

[10] Hala Ballouz, Joel Mathias, Sean Meyn, Robert Moye, and Joseph Warrington, “Reliable Power Grid: Long Overdue Alternatives to Surge Pricing,” March 12, 2021.

[11] “Winter Storm Uri 2021 -- The Economic Impact of the Storm,” Controller Office, Texas Government, October 2021.

[12] “The February 2021 Cold Weather Outages in Texas and the South Central United States.” FERC-NERC Staff Report, November 16, 2021.

[13] R. Gold, “Texas overcharged $16 billion for power during freeze,” says Independent Monitor from the Texas PUC, March 2021.

[14] Michael Grubb, “Navigating the crises in European energy: Price Inflation, Marginal Cost Pricing, and principles for electricity market redesign in an era of low-carbon transition,” UCL Institute for Sustainable Resources, Series Navigating the Energy-Climate Crises, Working Paper #3, September 5, 2022.

[15] Michael Milligan, Bethany Frew, Kara Clark, and Aaron Bloom, “Marginal Cost Pricing in a World without Perfect Competition: Implications for Electricity Markets with High Shares of Low Marginal Cost Resources,” NREL Technical Report NREL/TP-6A20-69076 December 2017.

[16] Alexander MacKay and Ignacia Mercadal, “Deregulation, Market Power, and Prices: Evidence from the Electricity Sector,” Working Paper Series, CEEPR WP 2022-008MIT, April 2022.