Volume 19 • Number 1 • January/February 2021

IEEE power&energy magazine

for electric power professionals

Choices in Tools and Design

Zero-Marginal-Cost Electricity Markets

Electric Light Pioneer

Charles F. Brush

Stay Connected

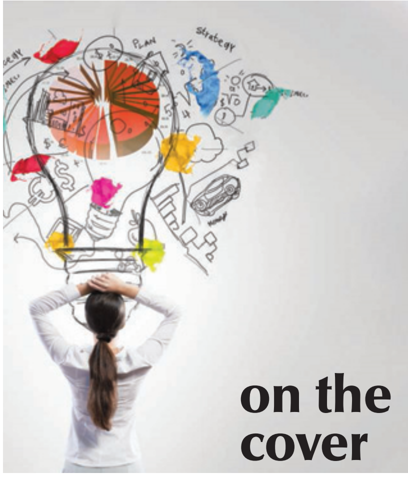
Get Involved With Chapters

IEEE PES Power & Energy Society®

IEEE

IEEE
power&energy magazine
for electric power professionals
Volume 19 • Number 1 • January/February 2021
www.ieee.org/power

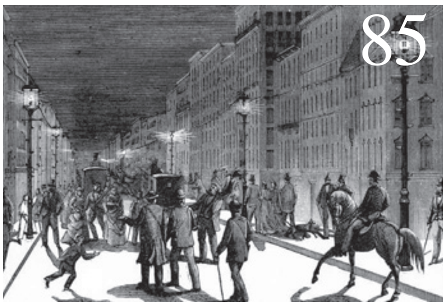
contents



©SHUTTERSTOCK.COM/SERGEY NIVENS

©SHUTTERSTOCK.COM/KHAKIMULLIN ALEKSANDR

features



columns & departments

IEEE POWER & ENERGY SOCIETY (PES)

The IEEE Power & Energy Society is an organization of IEEE members whose principal interest is the advancement of the science and practice of electric power generation, transmission, distribution, and utilization. All members of the IEEE are eligible for membership in the Society. **Mission Statement:** To be the leading provider of scientific and engineering information on electric power and energy for the betterment of society, and the preferred professional development source for our members.

Digital Object Identifier 10.1109/MPE.2020.3033365

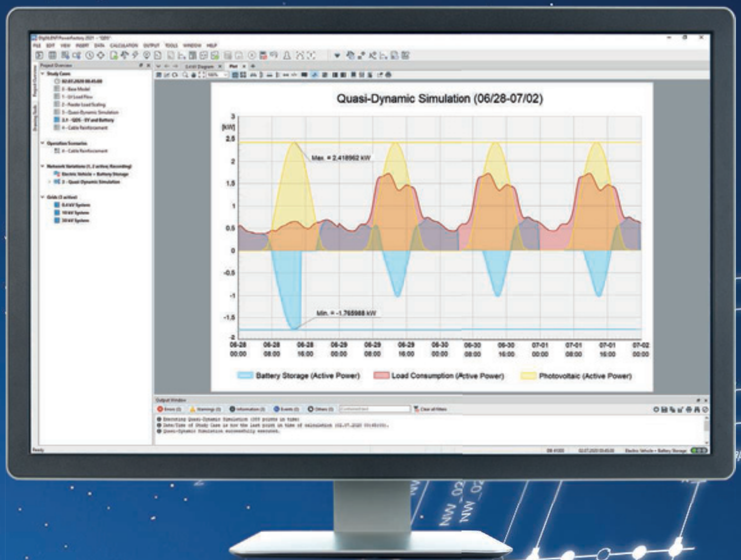
IEEE power&energy magazine

for electric power professionals

Digital Object Identifier 10.1109/MPE.2020.3033366

IEEE Power & Energy Magazine
IEEE Power & Energy Magazine (ISSN 1540-7977) (IPEMCF) is published bimonthly by the Institute of Electrical and Electronics Engineers, Inc. Headquarters: 3 Park Avenue, 17th Floor, New York, NY 10016-5997 USA. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society, or its members. IEEE Operations Center (for orders, subscriptions, address changes): 445 Hoes Lane, Piscataway, NJ 08854 USA. Telephone: +1 732 981 0060, +1 800 678 4333. Individual copies: IEEE members US$20.00 (first copy only), nonmembers US$89 per copy. Subscription Rates: Society members included with membership dues. Subscription rates available upon request. Periodicals postage paid at New York, NY, and at additional mailing offices. Postmaster: Send address changes to IEEE Power & Energy Magazine, IEEE Operations Center, 445 Hoes Lane, Piscataway, NJ 08854 USA. Canadian GST #125634188

Printed in U.S.A.


Quasi-Dynamic Simulation (06/28-07/02)
Battery Storage (Active Power)
Load Consumption (Active Power)
Photovoltaic (Active Power)

Steve Widergren

a new team for 2021

many thank yous and welcomes

THIS ISSUE OF *IEEE POWER & Energy Magazine* marks 18 years since founding Editor-in-Chief Mel Olken launched this periodical. In his words in that first issue, "Our new magazine is 'for electric power professionals' in response to the reality of our industry today and as one means of meeting the goal of the IEEE Power & Energy Society (PES) to reach out to the nontechnical side of the electric power industry and integrate these two critical areas. Feature articles will focus on advanced concepts, technologies, and practices associated with all aspects of electric power from a technical perspective in synergy with nontechnical areas, such as business, environmental, and social concerns."


©SHUTTERSTOCK.COM/KEANGS SEKSAN

Through his leadership and the subsequent leadership of immediate past Editor-in-Chief Mike Henderson, the course of *IEEE Power & Energy Magazine* has remained true to these words. As this millennium's version of the Roaring '20s takes off, the mission of this magazine persists in its objective of clear, plain, and broadly accessible writing on complex subjects that, when told properly, will interest and educate our readers. As your latest editor-in-chief of this magazine, I embrace and commit myself to this mission.

On the Shoulders of Giants

To be associated with the title of editor-in-chief is an honor. However, the position is not about individual prestige and validation; it is about service and commitment to leading a process that regularly produces high-quality magazines for our power industry colleagues. The responsibility that goes with this position gave me pause before I accepted the role. Since Mike Henderson became editor-in-chief and invited me to the editorial board, he has been a mentor. I watched the way he guided the production of the magazine with passion, humor, and good writing. His love for what the magazine stands for and the unique niche it caters to is apparent in the devotion he exhibited to meet the relentless two-month publication deadlines and the quality expectations of you, our readers. I am privileged to continue to have access to his guidance, and his labor of love for this magazine still drives the editorial organization onward.

I want to also recognize John Paserba for jumping into the breach as interim editor-in-chief and managing the past two issues of the magazine with great competence. His commitment to IEEE and this magazine allowed me

Digital Object Identifier 10.1109/MPE.2020.3033367
Date of current version: 6 January 2021

necessary breathing room, and he has been instrumental in my education. The magazine is also fortunate to have him remain as the associate editor for the "History" column. For all of these things, he ascends to volunteering sainthood in my book.

There is another important contributor deserving of recognition. For several years, Robert Henderson has provided editorial support to this magazine with thorough reviews of its content and editorial changes to improve its readability. It is hard work with little glory, but it is so important to the quality of our product. Robert contributed his skills to this, his final, issue. I am grateful to him for his efforts and the reduction in anxiety he provided to me, as I knew that many language details would be captured and properly addressed. Thank you, Robert.

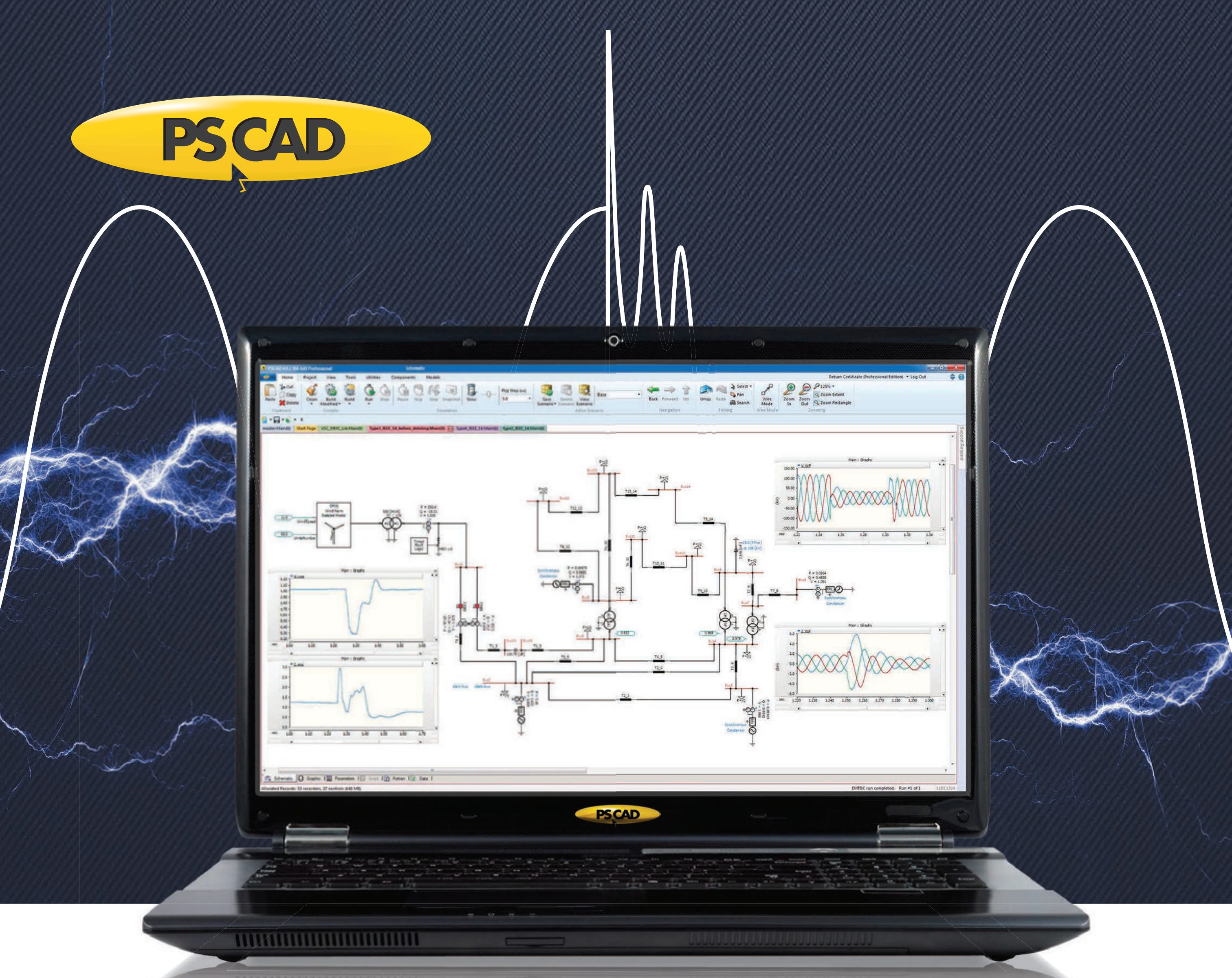
The Team

Along with John, I am joined by three other associate editors and an assistant editor. As assistant editor, Susan O'Bryan provides language editing skills and manages the process and schedule for the creation and submission of articles to the magazine. She works with Geri Krolin-Taylor and IEEE Publishing on the handoff of content and reviews draft publication material to ensure a quality product. As associate editor of submissions, Jianhui Wang fields and manages the process for reviewing unsolicited articles for the magazine.

Also, Antonio Conejo, Ning Lu, and Barry Mather are associate editors for issues. They provide technical review for our feature articles to ensure the content adheres to the magazine's standards of integrity. I am deeply grateful to Antonio for performing the technical editorial review for this issue. He joins me in this column to summarize the fine contributions of the guest editors and authors (see "About This Issue").

Leader's Corner

In the "Leader's Corner" column, Mazana Armstrong shares the challenges posed to our members this past year by the pandemic. In her role as vice president of Chapters, she experienced the sudden decline in events that removed opportunities for members to connect and interact. But she also witnessed the resilience originating from the grassroots aspects of the PES organization: our professional and student Chapters. Thanks to the many champions of engagement in our membership, meetings and conferences were redesigned, switching to virtual events with healthy attendance. From this trial has come fresh, innovative



ideas—a wonderful example of the potency and reward of volunteering.

Book Review

Our "Book Review" column provides insights into a recent publication. Christopher Lee shares his thoughts on *Electric Power Principles: Sources, Conversion, Distribution and Use*, second edition, written by James L. Kirtley, and its place as valuable material for teaching undergraduate electrical engineering students.

PES Elections

In October 2020, IEEE Fellow and current Editor-in-Chief of *IEEE Transactions on Smart Grid* Claudio Cañizares was elected as the 2021 Division VII director-elect. In this position, he will serve on the IEEE Board of Directors in 2022–2023 representing PES. He will succeed Miriam Sanders, who is serving as the 2020–2021 Division VII director. Prof. Cañizares vied for the seat along with Lalit K. Goel. Congratulations to Prof. Cañizares on his new role, and sincere thanks to Prof. Goel for running and providing PES members a choice for their vote.

History

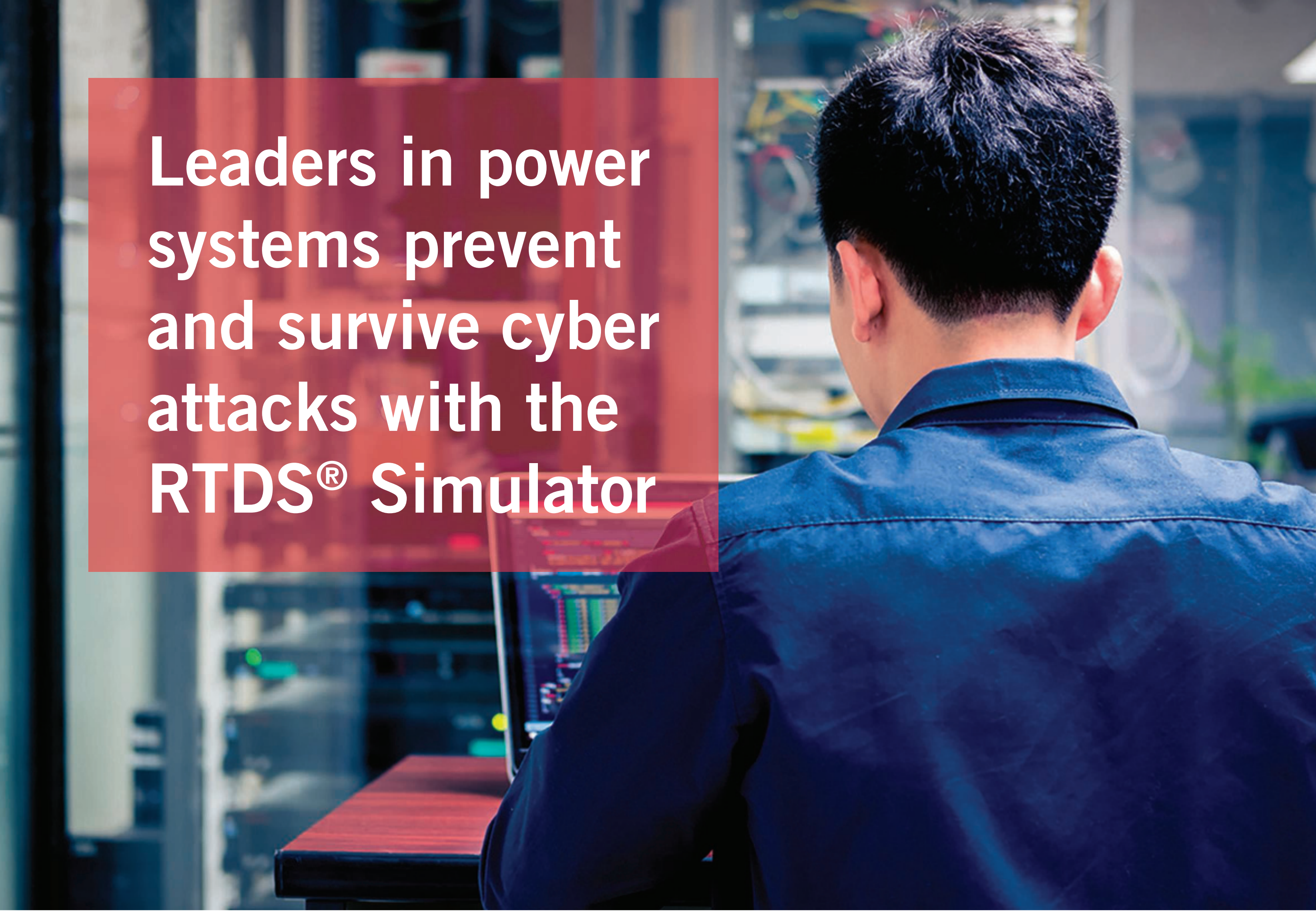
Associate Editor John Paserba brings us an article written by Adam Allerhand about Charles Francis Brush (17 March 1849–15 June 1929): an American engineer, inventor, entrepreneur, and philanthropist. This issue's "History" column explores Brush's contributions to dynamos, arc lighting, and central stations; addresses his competition at the time; and describes some of his work after the Thomson–Houston Electric Co. took control of the Brush Electric Co. in 1889.

Talk to Me

As the new guy on point to bring you *IEEE Power & Energy Magazine*, I am respectful of and aligned with its heritage but well aware that our world and the power and energy landscape is changing. The editorial staff's vulnerability to missteps can be outmatched only by an endless desire to improve the product for our readers. While the magazine receives letters to the editor and good word-of-mouth feedback, the true pulse of our readers remains somewhat veiled and a matter of conjecture to me. As I become more proficient in this role, I will be contemplating how to measure the magazine's success in fulfilling its niche for our audience. You can help. I encourage you to let me know your thoughts or ideas, likes, and dislikes with a message to pem-eic@ieee.org.

In Closing

When a colleague learned of my appointment as editor-in-chief, he noted that I shared a hidden job requirement with my editor-in-chief predecessors at



IT'S NOT JUST WEATHER OR WILDLIFE. PREPARE FOR CYBER INCIDENTS THAT CAN DISRUPT SERVICE AND DAMAGE EQUIPMENT.

Worldwide, the RTDS Simulator is a crucial component of cybersecurity testbeds, offering users the safest, most controllable way to simulate attacks on the power system.

Connect a real-time model of the grid to a communication network simulator and quantify the impacts of a cyber attack. Anticipate the behaviour of real protection, control, and measurement equipment through closed-loop testing. Results are authentic. Service is uninterrupted. This provides you with a realistic, flexible, and contained environment for the validation of energy system security technologies.

Learn more about cybersecurity project opportunities at rtds.com/applications/cybersecurity

RTDS Technologies 25 YEARS

YOUR WORLD IN REAL TIME.
RTDS.COM

In This Issue

Electricity markets were adopted in the late 1990s or early 2000s. At that time, renewable production based on solar or wind energy (weather-dependent resources) was a rarity. Today, renewable production is not only common but also dominant in some jurisdictions, such as Texas and California in the United States as well as Denmark and Spain in Europe. However, the design fundamentals of most electricity markets worldwide have remained essentially unchanged. It is true that diverse instruments have been added to markets that operate systems with high renewable penetration, but the core design of virtually all markets remains that of the early 2000s.

This issue of *IEEE Power & Energy Magazine* on the zero-marginal-cost electricity markets addresses the changes required in market design and market instruments to accommodate an increasing share of renewable production that has an exceptionally low marginal cost. These changes are most important to ensure well-functioning electricity markets in the future. This issue is closely related to the January/February 2019 issue on conversations on design—wholesale electricity markets.

Decarbonization, with important ramifications for the power sector, is a direct consequence of increasing the integration of weather-dependent renewable sources, such as wind or solar power. Thus, in this context, decarbonization is tantamount to renewable integration on the generation side along with changes on the demand side (load flexibility, electrification, and so on). Another consideration is that, for the time being, electricity cannot be stored in great quantities. This impacts the design and operation of electricity markets.

Within the framework for renewable-dominated systems, this issue addresses several relevant questions whose answers are not yet fully clear to the power community.

- Are the economic principles used to design current markets still valid if the generation fleet becomes renewable dominated? If most of the available generating sources have a marginal cost close to zero, can electricity markets function properly? Can a stable flow of revenues be ensured for producers?
- Which are the most effective mechanisms to guarantee investment cost recovery in a renewable-dominated power system? What is the role of capacity payments or capacity markets in ensuring cost recovery? Should capacity markets be adopted, or should producers rely solely on energy markets?
- Which are the most effective financial instruments to materialize investment in renewable generation in both developing and developed regions? Can long-term contracts be adopted without altering the liquidity of short-term markets? How are the roles of electricity financial markets organized?
- What can we learn from market experiences in hydro-dominated systems, such as those in Norway and Brazil? Can we translate the market experiences gained in hydro-dominated systems to renewable-dominated ones?
- Which are the most insightful experiences worldwide of high renewable penetration in power systems? Which are the lessons from Europe and other parts of the world, e.g., New Zealand?

Under the deft leadership of seasoned Guest Editors Luiz Barroso and Hugh Rudnick, a diverse mix of contributors to this issue, coming from industry, government, and academia, addresses these important questions and others, providing insightful analysis, observations, and recommendations.

Antonio Conejo

IEEE Power & Energy Magazine: we each spent formative years of our careers at American Electric Power Co. (AEP), a North American regional utility and power engineering firm whose accomplishments were served in no small measure by the prowess of Philip Sporn, its chief engineer and then chief executive officer, with a career that spanned the 1920s through the 1960s. Sporn was an immigrant who left his mark on our profession in engineering and business in the service of societal advancement. He was an Edison Medal winner and a member of the National Academy of Engineering, and he was deeply revered by those who worked at AEP. We are the progeny of Sporn, someone whose professional and civic calling speaks to the aim of this magazine for blending technology, business, and social progress in an open conversation with our power and energy community.

Lastly, in late September, another immigrant to the United States, National Academy of Engineering member and IEEE Fellow Dr. Eugene Litvinov died. Dr. Litvinov contributed to the advancement of power system analysis, electric market system development, and the application of information technology to many power engineering challenges while at ISO New England. He was a serious man who loved to laugh and developed admiration and friendship with all who worked with him. May he and all colleagues we have recently lost be remembered and rewarded in peace.

Mazana Armstrong

unity in adversity

staying connected during challenging times

THE YEAR 2020 WAS A DIFFICULT one, and we are glad to put it behind us. We feared for our loved ones, we feared for our jobs, and we feared life would never be the same again. This is our new normal, and as hard as that is to accept, we must keep moving forward. In this month's "Leader's Corner" column, I want to reflect on the positive aspects I have seen and experienced this past year in my professional career and as a member of and volunteer for the IEEE Power & Energy Society (PES). This exercise has given me hope that we can survive the pandemic together.

As the PES vice president for Chapters, my role is to oversee the operation of the Society's professional and Student Chapters around the world. We currently have nearly 260 Chapters and 450 Student Chapters in more than 150 countries. PES Chapters provide the means for our Society to connect with members locally. Our Chapters are run by volunteers who are just like the people you work and live with in your community. They organize events to support the professional development of PES members in the community and provide them with networking opportunities. With the start of the pandemic, all local Chapter activities across the world were paused, but only for a brief period. Our volunteers regrouped and soon started organizing virtual events. Conferences and technical meetings were transitioned to virtual events at low to no cost to our members. Distinguished lectures became webinars. Our "engines" were restarted and ran at a high speed in this new virtual world. As PES, we decided to not sit back but instead do just the opposite. We moved forward with strength and support for each other at the time we needed it the most.

Digital Object Identifier 10.1109/MPE.2020.3033368
Date of current version: 6 January 2021

2020 PES Student Congress

Let me tell you a story about the 2020 PES Student Congress. This biannual event was started in 2014, and our Society funds approximately 100 Student Chapters to send their representatives to this training and networking event. The event supports the growth of student membership and ensures that our Society continues to flourish for years to come. These young students will carry the torch lit by the inventors of electric power systems and founders of IEEE/AIEE more than 130 years ago. The fourth Student Congress was scheduled to be held August 2020 in Montréal in conjunction with the IEEE PES General Meeting. When it became clear that all IEEE conferences and events were being switched to virtual events, our organizing committee for the Student Congress was faced with an enormous challenge to create a first-ever virtual congress in an extremely short time and with many uncertainties, such as the number of potential attendees.

> Our organizing team of student volunteers from Sri Lanka, Canada, Tunisia, and Brazil united to pull off an outstanding event.

Our organizing team of student volunteers from Sri Lanka, Canada, Tunisia, and Brazil united to pull off an outstanding event. The virtual congress was free of charge, with registrations exceeding 2,000 student participants from across the world. The sessions included online competitions, debates, soft-skills training, multicultural events, leader forums, and training on how to run successful Student Chapters. The event was broadcast via social media outlets, and it was a huge success.

To make this happen, our student volunteer leaders donated hundreds of hours of their own time on weekends and evenings for one reason only: the positive difference it would make for our members around the world. That truly mattered to them. This volunteering spirit is just one of the examples where I have witnessed the strength of our Society and the support and benefits that emerge when we work as a team.

Become Involved in the PES

Students are the future of our Society, and this is emphasized in the PES's five-year strategic plan. We are currently seeing less participation in our Student Chapters at universities in North America, and fewer students are transitioning to PES membership upon graduation. All of us can do something to change that. No matter where you live in the world, please consider offering a free webinar. Talk to the students at your local university about your job and your career. Tell them why you find a career in the field of power and energy so worthwhile, and explain how they can benefit from being involved with our Society. If you can help create a new PES Student Chapter or revive an inactive one at your local university, please do so.

And if you work with a young colleague just out of school, encourage him or her to join our PES community for a more successful and fulfilling career, just as you and I have experienced. Our Society's foundation is in the development of standards that have enabled us to build and operate arguably the most complex engineered systems known to humankind. Our technical committees are the perfect place for young professionals to grow and mature into technical leaders of the future.

I also encourage you to get in touch with a professional PES Chapter near you. It is very easy to find your nearest Chapter at https://www.ieee-pes.org/pes-communities/chapters/chapter-locator. Check out Chapter activities and consider giving a presentation about your professional interests. I challenge all PES Chapters, both student and professional, to organize at least one virtual event per month. You can help reach that goal by both offering to present and participate virtually in presentations given by others.

If you are still wondering what is in it for you, active involvement with PES Chapters will help you build your professional network, learn new skills, stay up to date professionally, develop leadership skills, and be recognized for your accomplishments through PES Chapter awards. And along the way, you will also create friendships that will stay with you for life. There is absolutely nothing to lose and everything to gain.

I wish that you all stay safe and connected during these challenging times. You are welcome to reach out to me with questions or ideas at mazana.armstrong@ieee.org.

Luiz Barroso and Hugh Rudnick

the growth of renewables

zero-marginal-cost electricity markets

DECARBONIZATION GOALS HAVE created a technological revolution that has enabled renewables—in particular, solar and wind generation—to be in the center of most electricity markets all over the world. Renewables are inherently characterized by high production variability combined with limited predictability and controllability, which have created significant flexibility challenges for power systems planning and operations all around the world.

Renewables also produce electricity at an almost zero marginal cost. In the presence of high shares of renewables, restructured electricity markets based on setting locational marginal prices will also be challenged, as electricity prices can fall to zero or even negative values if abundant renewable generation is observed. Conversely, prices can rise quickly and may reach high figures if renewables are not producing and scarcity takes place. These effects can be exacerbated in the presence of transmission bottlenecks and high levels of distributed generation.

In the abundant presence of (almost) zero-marginal-cost resources, important questions have recently been posed on the power system economics side about the future of current electricity market designs and companies:

- ✔ How will remuneration of generation assets take place?
- ✔ How will financing arrangements be structured?
- ✔ Will revenue uncertainty compromise generation adequacy in the face of new capacity needs due to the decommissioning of existing fossil resources combined with load growth?
- ✔ Can scarcity pricing, revealed through variability in short-term price signals for generation services, create adequate incentives for long-term investments?
- ✔ Are long-term reliability and energy contract markets the way to go, or do capacity markets suffice to create efficient incentives for investments?
- ✔ How can demand-side services and demand response be brought into the market as resources while balancing political considerations and consumers' aversion to risk in electricity pricing?

With increasing decarbonization goals in power systems all over the globe, these are examples of relevant questions to be discussed by policy makers and stakeholders.

The objective of this issue of *IEEE Power & Energy Magazine* is to debate these zero-marginal-cost futures and discuss relevant topics related to generation adequacy and wholesale markets, focusing on conceptual and practical discussions. The issue features authors with wide experience from both academia and industry who focus on the regulatory and market challenges ahead.

Carlos Batlle, Pablo Rodilla, and Paolo Mastropietro open this issue with the conceptual problem statement. They discuss how, more than three decades since the first power markets were implemented in the 1980s, key aspects of the market structure need to be revisited: the interplay between regulation and market forces as well as the relations between transaction characteristics and contractual and other governance structures.

However, this is not the only matter that needs to be properly addressed. The authors argue that short-term pricing mechanisms will continue to be instrumental in guiding optimal operation and investment decisions but also need to be properly coordinated with regulatory-driven, long-term markets. They discuss which mechanisms should be designed, how this should happen, and what market rule modifications are needed to allow for efficient interaction between short-term market prices and long-term complementary signals.

The second article, by Frank Wolak, also discusses conceptual issues, presenting a proposal for a market design in a zero-marginal-cost intermittent renewable future. Wolak suggests key improvements to the design of an efficient short-term wholesale market and posits a long-term resource adequacy mechanism for a system with a large share of zero-marginal-cost intermittent renewables. He argues that his conceptual proposal ensures long-term resource adequacy at a reasonable cost for final consumers while also allowing for the short-term wholesale

Digital Object Identifier 10.1109/MPE.2020.3033369
Date of current version: 6 January 2021

volatility necessary to finance investments in storage and other load-shifting technologies that will be required to manage a large share of renewables.

We then move to the first of three articles on practical experiences. Erik Ela, Andrew Mills, Eric Gimon, Mike Hogan, Nicole Bouchez, Anthony Giacomoni, Hok Ng, Jim Gonzalez, and Mike DeSocio discuss potential pathways of electricity market designs without fuel costs in the United States and Canada. They navigate through some key challenges and efforts to improve market designs today before describing potential options for future designs of electricity markets with these characteristics. This includes arrangements to incentivize investment in and operation of the future supply fleet.

A team of 12 European authors led by Goran Strbac reviews European policy initiatives to address market design challenges. They discuss some European Union (EU) efforts in five selected areas: the missing money problem, the integration of renewables in energy and ancillary services markets, carbon markets, the value of distributed flexibility, cross-border market integration, and the coordination of emerging local energy markets. Open issues and innovative designs are identified to enable a cost-effective and secure decarbonized European electricity system.

Luiz Barroso, Francisco D. Muñoz, Bernardo Bezerra, Hugh Rudnick, and Gabriel Cunha show that some hydroelectric-dominated countries in Latin America (in particular, Brazil) have operated with a lot of generation with zero marginal cost for decades and still managed to incentivize investment in new generation capacity. However, in those settings, long-term markets for financial energy contracts with sufficient liquidity are essential to secure generation financing due to the high volatility of spot prices. The authors discuss the role of long-term markets that have been implemented in the region for decades and potential improvements.

Our last article comes from a team of lenders. Tonci Bakovic, Debabrata Chattopadhyay, Fernando Cubillos, and Marcelino Madrigal present a view from development banking practitioners on the challenges to generation financing in markets with a high penetration of renewables. Financing is essential to attract capital to expand or retain generation capacity. Focusing on developing countries, they argue that policy makers must address both market and institutional design. Supporting the arguments of the previous articles, they claim in particular that long-term contracts to stabilize revenues will continue to be a key design feature to attract competitive generation investments, but more flexible arrangements—such as liquid markets for short- to medium-term contracts that can be rolled over—can be alternatives to avoid the technology lock-up brought by the super-long-term commitments that have emerged in many developing countries.

Financing is essential to attract capital to expand or retain generation capacity.

Finally, Alberto Pototschnig, an economist who served as the first director of the EU Agency for the Cooperation of Energy Regulators and is currently with the Florence School of Regulation, brings his experience to a debate in the "In My View" column. He focuses his discussion on reliability options as a preferred mechanism to address residual adequacy concerns raised by the high penetration of renewables. Focusing on the EU, he argues that reliability options should be designed such that strike prices do not interfere with the functioning of the energy market—under normal or even tight conditions—and that penalties should be imposed to reinforce the incentives for adequacy resources to be available at times of scarcity.

Electricity markets will be dramatically changed in the decades to come, and the discussion of how best to adapt market designs for high shares of renewables will be ongoing. The consequences of the zero-marginal-cost electricity industry go well beyond the wholesale markets discussed in this issue, and we are sure IEEE will continue contributing to these discussions.

We would like to thank the authors for the time, dedication, and articles provided, which shed light on the key topics related to this very relevant discussion. We thank *IEEE Power & Energy Magazine* for providing us with the opportunity to reflect on and analyze such challenging matters, which have taken us and the authors out of our comfort zones to consider many new ideas. A special thank you goes to Editor-in-Chief Steve Widergren, Associate Editor Antonio Conejo, and immediate past Editor-in-Chief Michael Henderson for continuing to provide the conditions for *IEEE Power & Energy Magazine* to remain an IEEE flagship publication.

Digital Object Identifier 10.1109/MPE.2020.3043895

Markets for Efficient Decarbonization

By Carlos Batlle, Pablo Rodilla, and Paolo Mastropietro

IN THE 1980S, SOME SEMINAL WORKS SUCH AS *MARKETS FOR POWER: AN ANALYSIS of Electric Utility Deregulation* by Joskow and Schmalensee and *Spot Pricing of Electricity* by Schweppe et al. set the foundations for electric power system restructuring toward a fully liberalized, marginal price-based market environment in which generators and end users trade. Even then, it was clear that the task was not going to be easy, but an increasingly and always significant portion of the industry and academic community thought that it was at least possible. More than three decades after the first power markets were implemented, the entire "power sector community" continues to discuss the suitability of relying on short-term market prices as an efficient signal to drive investments, especially in the current context in which not only traditional generators but also end users can decide to invest in energy supply resources, and an increasing amount of new generation investments have zero or close to zero variable costs, which some people see as a threat to the short-term market paradigm.

In this article, we discuss the reasons behind the widespread implementation of regulatory-driven, long-term market mechanisms, such as capacity mechanisms or auctions for renewables, storage, or demand response. As a consequence of the former, we argue that in the current (and even more in the upcoming) context, rather than discussing the suitability of relying on any sort of regulatory-driven market or pseudomarket mechanism that provides long-term signals, research efforts should focus on which mechanisms should be introduced and how they should be designed and what modifications of market rules are needed to allow for efficient interaction between short-term market prices and long-term complementary signals.

Digital Object Identifier 10.1109/MPE.2020.3033393
Date of current version: 6 January 2021

1540-7977/21©2021IEEE

Revisiting Market Regulation and Design

The New Market Paradigm Under the New Conditions

The key motivation behind the liberalization of power sectors and the introduction of wholesale market mechanisms was to transfer the responsibility of investment decisions from public administrations to market agents, expecting that competition in generation would spur efficiency and drive wholesale energy prices down. Moreover, competition would remove the responsibility of planning from governments as single decision makers whose errors have massive consequences and leave capacity expansion decisions to the market agents, thus shifting the consequences of planning errors from the consumers to those who make the decisions.

The principal idea was to "let the market" decide what, when, and how to invest in new generation resources. However, reality has contradicted this original premise. There is a growing consensus among power sector regulators worldwide that markets need regulatory mechanisms that protect and complement the expansion processes.

Since the introduction of markets, significant progress has been made on the wholesale competition front. Short-term energy and ancillary services markets, built upon mid-20th century economic dispatch-optimization modeling tools, have worked reasonably well under the designs that characterized power system liberalization during the 1990s and early 2000s. However, the context and perspectives for the future have dramatically changed. Now the complexity and uncertainty of the capacity expansion problem are much larger because of the following:

1) On the supply side, the diversity of technological alternatives is unprecedented.
2) On the demand side, the load for conventional uses is decreasing while, at the same time, it is expected that the whole economy is going to be gradually electrified at an uncertain level and pace.
3) For the first time, end users are starting to have the means not only to respond to short-term price signals but also make their own investment decisions.

These developments raise questions about whether current market designs can be improved to provide short- and long-term price signals able to support an efficient and comprehensive system-wide operation and expansion of the power system, leading to an optimal portfolio of utility scale and distributed generation, storage, and demand-response resources consistent with public policy goals. In this context, governments worldwide have reversed the original expressed intention of letting the market determine investment decisions. Subsidies for renewable energy sources (RESs) have guided a significant part of the investments in generation in the last decade, many competitive bidding tenders and capacity remuneration mechanisms are being implemented, and even direct and centralized key energy-planning decisions are being made.

> Research efforts should focus on what modifications of market rules are needed to allow for efficient interaction between short-term market prices and long-term complementary signals.

Two paradigmatic examples among the many others that could be mentioned include the contract to build Hinkley Point C, signed by the U.K. government, which is currently considering a regulated asset based (RAB) model to finance additional nuclear plants, and the AB 2514 and AB 2868 energy storage mandates in California. Next, we develop an in-depth review of the factors that characterize a power sector context totally different than the one reigning at the time most market-driven restructuring processes started.

Old and New Factors Conditioning Electric Power Markets

The Change in the Framing Conditions

Many instrumental framing conditions have changed since electricity markets were originally designed and implemented.

First, politics have changed dramatically. The 1980s and early 1990s were times when economic liberalism was much more "popular" among policy makers. Also, power system liberalization was seen as a helpful tool to support other policy decisions, such as the closure of the mining industry in the United Kingdom.

Currently, the situation is totally different since economic liberalism has lost momentum in a good number of jurisdictions and, more importantly, higher-order objectives (that is, decarbonization via electrification of the economy) are assumed to require tighter control of the investment decisions in the power sector.

At the same time, there was an underlying perception that the degrees of freedom in the decision-making process to plan the future expansion of power systems were rather limited. At the start of the century, nearly all of the new investment alternatives in generation consisted of gas-fired and coal plants with nuclear and hydro in some places. Many expected the expansion of power systems to be led almost exclusively by combined cycle gas turbines (CCGTs). The actual competition was, in fact, assumed to be built along two complementary dimensions. CCGT manufacturers were on one end and gas suppliers on the other via the different contract formats, such as the take-or-pay contracts that prevailed at the start of the liberalization in the European Union (EU) context. The risks that investors were supposed to face were also perceived as largely controllable. Since investing in CCGTs seemed like the only option, leaving these decisions to the market was perceived to be worth a try. At the same time, politicians did not perceive that they were losing much control of one of the key levers they had been using to gear economic policy.

Currently, the number of variables of the capacity expansion problem has skyrocketed. Now, besides the traditional generation technologies subject to new risk factors, such as carbon prices and environmental constraints, there is a diverse portfolio of new alternatives all linked to extremely uncertain learning curves: renewable energy and demand response sources, new nuclear technologies, carbon capture and sequestration, different sorts of storage, hydrogen, and so on. It is not just that there is a diversity of alternatives but that these new choices are qualitatively different from conventional technologies.

Gas-fired generation plants had other relevant expected advantages. There was a consensus that CCGTs turned down the marginal cost of electric energy supply over the long term. In the EU case, in the second half of the 1990s, the levelized cost of energy assumed for CCGTs was in the range of US$40/MWh, well below the energy component embedded in end-user retail rates at the time. CCGTs' higher relative efficiency, lower capital costs, and short building times (it was expected that they could be installed in two or three years), particularly in a context of high-interest rates, were perceived not to have any technological competitor. In addition, CCGTs were supposed to be almost fully scalable and could be sited almost anywhere, so there was no significant need for transmission expansion.

Some of these key advantages, such as building times, did not fully materialize. Relevant uncertainties, such as environmental regulations, have been larger than expected. Scalability has not been significant either, and due to greenhouse gas limitations (for example, carbon taxes), prices will not be as low as originally assumed. Besides, building transmission is increasingly more difficult for a number of well-known reasons.

In a good number of power systems (mainly in the EU case) where liberalization was implemented, there was a significant overcapacity as, until that time, policy makers took care to maintain reliability in a very risk-averse manner. In others, capacity mechanisms, in the form of capacity payments or capacity markets, were implemented in a large number of jurisdictions almost from the very beginning. Only some European countries did not implement capacity mechanisms, but it should not be forgotten that in such jurisdictions, the key incumbents were still publicly owned (such as EDF in France, Enel in Italy, and Vattenfall in Sweden).

Finally, compared with the current state of affairs and financial situation, the willingness of equity to fund investments in generation was much better than it is now. This led, for instance,

> End-user retail rate making also requires a modification from pricing energy to pricing capacity as well as equitably designing fixed charges needed to avoid inefficient self-supply.

to the initial euphoria to invest in CCGTs in the United Kingdom (the so-called "dash for gas"), Italy, and Spain.

At the same time, unfortunately, some problems that were supposed to be healed as markets matured persist after three decades. It is not just that many of these factors have not turned upside down, but it is that many other new and critical ones have arisen. Next, we review some of these issues.

The Persistent "Original Sin": the Inaction of the Demand

One of the basic factors for a market to function is obviously that both supply and demand participate properly in the market. Fundamental to maximizing the net social benefit in both the short and long term, demand plays an active role in the market by reflecting, through its offers, the true value that an asset has in each moment. In economic terms, that means the demand purchase offers reflect an asset's true utility function at all times.

In times of generation shortage, if demand does not actively declare its actual utility function in the market, the price could theoretically rise to levels well above what most consumers would actually be willing to pay. For this and other reasons, since the first electricity markets were implemented, regulators designed artificial price limits. Conversely, in most cases, these caps have been well below the levels that demand would be willing to pay. Thus, price caps not only affect the optimal programming in the short term but also negatively condition investment decision-making processes as they lead to underinvestment. As Joskow has repeatedly argued in his publications, short-term prices are distorted not only by price caps but also by many out-of-market actions by an independent system operator to reduce demand administratively, such as voltage reductions and non-price rationing of demand (rolling blackouts).

In the long term, electricity prices, like those of many other assets or services, are subject to a growing number of uncertainties. These especially affect investors in generation resources due to the capital-intensive nature of these assets. Given this reality, theoretically, it is expected that the demand seeks to sign some type of long-term hedge, which in turn would allow investors to mitigate their risk aversion. The fact that this does not happen is the main cause that provides regulators with a justification to intervene in the market, imposing on the demand the obligation to contract a long-term guarantee, or buying such a guarantee on its behalf. These regulatory "solutions" include not only the so-called capacity remuneration mechanisms (CRM, the name commonly accepted in the EU context) but also other centrally designed long-term mechanisms, such as the RES support mechanisms developed worldwide, auctions for long-term energy contracts implemented in South America, the Hinkley Point C contract signed by the U.K. government, or California's storage mandate (mentioned previously).

Price Caps and the Missing Money Problem

The main theoretical argument that originally justified the need to design capacity mechanisms (payments/markets) was that the existence of price caps and other out-of-market actions previously mentioned led to the well-known "missing money problem." Under scarcity conditions, the market price should be equal to the price the demand is willing to pay for not being interrupted. According to its original formulation, the problem arises when investments made under the expectation to benefit from high prices when the system is tight cannot actually receive this income because the regulator decides to set a maximum price limit.

Although the missing money problem has not ceased to be pointed out as key to the need to implement a CRM, experience does not seem to prove that it has really been instrumental. In fact, the largest investment period in electricity generation occurred in the late 1990s and early 2000s, just when administratively defined maximum prices were far more restrictive. For instance, despite the existence of price caps, according to the data provided by the U.K. Department of Energy and Climate Change in 2005 (Digest of U.K. Energy Statistics), the installed capacity in 1993, two years after the market started, was close to 65 GW and 10 years later it amounted to 75 GW (around 25 GW of new CCGTs were installed in that period). For example, similar phenomena took place in Spain and Italy after the market was implemented in the late 1990s. A similar effect took place in Texas (US$1,000/MWh price cap), where the reserve margin increased during the first decade of market functioning.

Paradoxically, the investment slowdown took place later, when these maximum prices in most markets had been gradually growing. For example, at present, the top price in the European day-ahead market is €3,000/MWh (in the intraday market it is ±€9,999 and there is no price cap in the subsequent market segments; in this case, prices have been freely above €10,000/MWh a good number of times). In the day-ahead energy markets of other continents, high maximum prices have been designed, such as the Electric Reliability Council of Texas in the United States, where the maximum price is US$9,000/MWh, or the National Electricity Market in the southeast of Australia, where it is AUD$14,500/MWh (approximately US$10,000/MWh). This does not mean that the missing money problem is not an issue anymore, but as we later discuss, it is far from being among the main reasons for the short-term market failure to attract sufficient investment.

The Political Will and the Investment Side

The liberalization of traditionally regulated activities was meant to further efficiency in the operation and planning of power systems, mainly based on the proper allocation of risk-management responsibilities between regulators and stakeholders. It could be expected that market liberalization of generation and retail could indirectly even improve the efficiency of the transmission and distribution businesses. On one side, for example, merchant generation investors could put pressure on transmission planners to expand the network in a more efficient way, while on the other side, more active end users could reveal different values of quality of service, useful to support the planning task of distributors. But there were three main dimensions in which market liberalization was supposed to significantly improve the overall system performance: 1) operation (increasing the efficiency of the economic dispatch, which often had lacked transparency and, in some cases, was heavily influenced by diverse political interests); 2) capacity expansion planning, assuming that market agents would make wiser investment decisions, or at least, that they would bear the costs of wrong ones; and 3) eventually, retail as a way to engage end users in the decision-making process at all levels, as it nowadays is starting to take place by taking advantage of the development of distributed energy resources and the Internet of Things.

However, in such a capital-intensive industry, the key dimension in which competition was expected to bring significant gains was on the generation investment side. Economic dispatch could certainly be improved but, comparatively, no great savings could be expected, and it was, and still is, unclear which additional benefits retail market liberalization could entail compared to those derived from the implementation of well-designed, sufficient, and not politically interfered regulated end user rates.

Until investment decisions were left to market forces and since planning errors committed in regulated environments were paid for by customer tariffs, utilities had weak incentives to make efficient decisions. Examples of erroneous-planning investment decisions under traditional regulation were not difficult to find. Some examples are the nuclear development plan in Spain in the 1980s, later followed by an expensive moratorium imposed by the government (25% rate surcharge for 25 years), and the large hydro projects undertaken in Latin American countries that overran their budgets and drastically increased state debt. At the same time, as previously mentioned, publicly owned utilities were unable to stop burning expensive and inefficient national fuels (such as the autochthonous British or Spanish deep-mine coal).

The alleged key objective of market deregulation was to take investment decision making out of the hands of governments and give it to market investors. This is one of the main motivations of the World Bank and other development banks in certain developing countries. But at the same time, some other relevant governments wanted to be relieved of the responsibility of guiding capacity expansion. As previously mentioned, the British government in the early 1990s was a good example.

However, this construct ended with the advent of initiatives to decarbonize the economy and the planet. The electric power sector is a key lever to achieve this aim. Policy makers have decided that different rules are needed to pursue a higher-order decarbonization objective.

Take, for instance, the case in the United States, which includes trans-state short-term energy markets complemented with 1) capacity markets that treat various technologies differently as the methodologies to allocate capacity credits respond to diverse criteria, 2) renewable portfolio standards at the state level fixing different quotas for various technologies and time terms, 3) federal production and investment tax credits, 4) net metering and net billing state policies rewarding distributed generation in diverging ways, 5) energy storage mandates in force in seven states as of June 2020, 6) regulated programs to support nuclear investments (for example, the Clean Energy Standard implemented in New York), and 7) different out-of-market demand response programs.

In the European continent, things are not very different. For instance, in Great Britain: 1) a capacity market was implemented but later suspended in 2019 following the decision of the General Court of the EU; 2) renewables are supported through feed-in tariffs, a contracts for difference scheme, and a tax regulation mechanism; 3) at the time of this writing, the carbon price is made up of the EU emissions trading system price and the carbon price support rate, the latter price set by the British government and largely uncertain after Britain's exit from the EU; and 4) a RAB model for nuclear is currently open and under intense discussion. As these two examples show, mechanisms are not only uncoordinated as they depend on different institutions at different jurisdictional levels but also are extremely subject to unexpected changes.

In addition, due to the push on RES and CO_2 reduction objectives and the increase of gas prices in some jurisdictions (certainly not the case in the United States after the advent of shale gas), governments started to realize that retail prices were not going to necessarily decrease as initially expected. Thus, the initial expectation that political interference was not going to be a factor cannot be held anymore, and also the uncertainty linked to the investment decision-making process is now significantly higher than at the time the liberalization processes started.

As initially mentioned in the enumeration of factors, there are also some other less relevant factors that are turning liberalization into something more complex. When we expected mainly gas-fired plants, transmission planning was not perceived as an instrumental problem to allow for competition. In principle, gas plants could rely on the development of the gas transmission network, but siting was not critical. As a result, transmission planning developed by independent system operators or transmission system operators did not necessarily have to be a key conditioning factor for competitive investors. But renewables change the whole thing. Take, as examples, the onshore wind resources in the U.S. Midwest, the offshore wind resources in the North Sea, and the availability of solar photovoltaics (PVs) in northern Africa. These locations have outstanding renewable potential, but they are far from the main load centers. All of these resources need a previous decision from transmission planners, who have to decide which locations make more economic sense, and therefore critically condition the competition for access. Also, the necessary new transmission links often involve different state administrations, complicating the siting process and leading to unexpected delays that are outside the control of generation investors.

Expansion Planning, Long-Term Supply Guarantee, and Higher-Order Objectives

The main objective of capacity expansion planning has changed dramatically. Originally, the objective was to supply demand growth while guaranteeing certain reliability levels. Since several technologies could achieve this aim, the main consideration was cost-effectiveness. Again, CCGTs were in the best position. Currently, objectives are more ambitious since decarbonizing is a key priority. The challenge is not just that needed technologies happen to be more expensive and complex than the original ones. At this stage, there are no technologies available to go beyond approximately 80% decarbonization. Technological innovation is needed with appropriate market incentives. The promotion of wind and solar has been at least somewhat effective as costs have been reduced to very reasonable levels. Additional development is now needed for energy storage or carbon capture and sequestration, and these technologies appear to be even less mature than wind and solar PVs were a decade ago. We can expect cost reductions on 4-h lithium-ion batteries, but to go beyond 80% decarbonization we need affordable large-capacity batteries or some new technology that has yet to be envisioned. Again, it is not realistic to expect that short-term market prices will provide sufficient incentive for these technological developments.

When the liberalization of the European electrical systems was developed during the 1990s and the first half of the 2000s, it was an incentive for electric companies, which, perceiving the existence of new competitors, launched themselves with determination to undertake new investments. In the Spanish case, between 2002 and 2008, about 22 GW of gas-combined cycle plants were brought online, which represented 32% of the current total installed capacity.

However, as Figures 1 and 2 show, things did not turn out as expected. A good part of the new investments in generation was made at the start of the first decade of this century when demand growth was very significant. In the middle of the decade, this growth began to dissipate. Subsequently, the severe economic crisis that affected the economy of the EU resulted in a dramatic correction of the expected increase in demand that, in fact, not only stopped growing as expected but significantly decreased. From 1995 to 2005, electricity consumption in the EU increased more than 20%, which, projected, would have led to consumption in 2015 of about 4,000 TWh. Between 2005 and 2011, it decreased by about 3%, fewer than 1,000 TWh out of the 4,000 TWh that had been expected by the most optimistic forecasts. This factor in itself was more than enough to disrupt the sector, increasing future risk aversion when facing new investments.

In addition to the aforementioned decrease in demand, technology advancements and regulatory encouragement have boosted the penetration of renewable sources (see Figure 2). This situation has affected the profitability of merchant plants, including those installed after the start of the market. Load factors and prices have been lower than expected and thus income.

These factors led to the utilization factor of the new investments reaching significantly lower levels than investors could expect in the lowest of their original forecasts. For

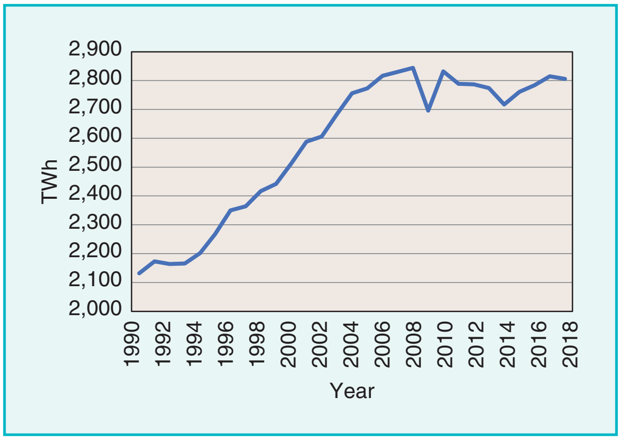


figure 1. Net electricity generation, EU-28, 1990–2018 (data from ec.europa.eu/eurostat).

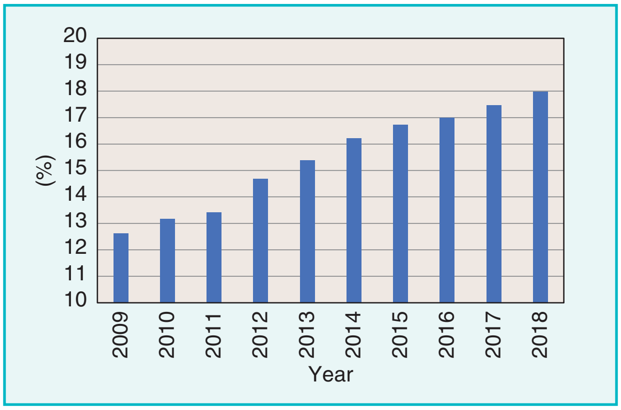


figure 2. Renewable generation quota in the EU-28, 2005–2018 (data from ec.europa.eu/eurostat).

example, between 2010 and 2014 in Italy, the production of the new combined cycle units was reduced by almost half; the load factor, which was 44% in 2010, was 26% in 2014 (see Figure 3).

The Temporal Lumpiness of Solar and Storage

There is an inherent disadvantage to investments in solar PVs and particularly in storage. The learning curve of both technologies keeps improving, so costs keep falling. Contrary to conventional generation, these technologies are fully scalable and can be planned and installed in a few months.

If a solar PVs investor makes a business case analysis today in the face of energy market prices, he or she could reasonably conclude that investing in a solar farm could be in the money. But at the same time, one could quickly conclude that, at the current speed of cost reduction of solar PVs, in a year the investment very likely will be out of money. Thus, if investors cannot find some way to capitalize/hedge the value of their investment, the rational decision is to wait. This is particularly acute in the case of storage. Storage is supposed to be financially justified by its ability to capture off-peak/peak spreads. But it is well known that a modest amount of storage quickly reduces the spread, so the deployment of storage negates its own use case.

The Perennial Market Incompleteness

In principle, the above-mentioned lumpiness should not be an issue in a perfectly functioning and complete market when a fully informed and elastic demand side would be willing to enter into long-term contracts with solar and storage providers, allowing them to proceed with the investment and locking in the benefits their installations would bring to the system. But even optimistically assuming that electricity end users can soon become elastic in the short term (thanks to digitalization, the Internet of Things, and so on), it is far from clear that demand will ever be willing to enter into long-term commitments. This market incompleteness is at the core of the global relevance not only of capacity markets (even in the EU, where they were originally perceived as an anathema; see Figure 4) but also of long-term auctions for renewables.

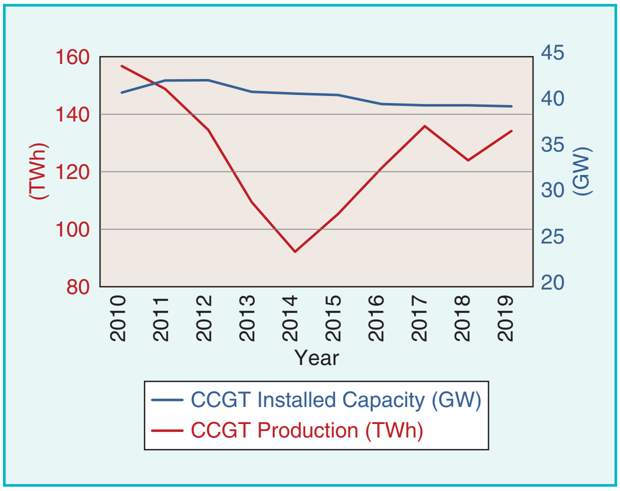


figure 3. CCGT capacity and production in Italy.

In the particular case of RESs, it is increasingly argued that, due to the extreme improvement of the learning curves, there is no longer a need to design any sort of support mechanism, not even auctions for long-term contracts. This argument, in our view, does not properly take into consideration two still relevant key factors: vertical integration (generation and retail) and the need to organize access, thus allowing for the coordinated expansion of the transmission network. Vertical integration, a factor in the majority of systems where retail business has been liberalized, can become an entry barrier for new investors in RESs as they have larger difficulties finding counterparties to sign the power purchase agreements required to properly finance their projects. At the same time, long-term auctions are also a helpful tool to guide future transmission network needs. For example, in the Spanish power system where the peak load is currently lower than 50 GW, the transmission system operator has applications for connecting up to 80 GW of new RESs. While direct subsidies might be no longer needed, keeping some sort of longer-term tendering is still advisable. For the aforementioned reasons as well as some others, it should not be expected that RESs are going to be deployed efficiently if investors do not have access to long-term contracts of any kind.

In summary, a core criterion for designing a proper decision-making procedure is that risks should be allocated among those best prepared to manage them. In the early 1990s, the main risks that investors had to face were linked to construction, fuel contracting (take-or-pay, tolling, and so on), operation and maintenance (O&M), and competitors' decisions. These risks could reasonably be expected to be managed by the private sector. Currently, business investors face significant technological uncertainties and changing public policies and regulations. As market agents struggle with these risks, we argue that policy makers and regulators need to become involved with shaping the marketplace.

Conclusions and Recommendations

There is little doubt that the questions addressed in the seminal works of the 1980s need to be revisited: the interplay between regulation (centralized planning) and market forces must be reconsidered, and the relations among transaction characteristics and contractual and other governance structures need to evolve with the times. In addition, short-term pricing mechanisms will continue to be instrumental in guiding optimal operation and investment decisions and will need to be capable of involving power system agents all the way down to the last kilometer as well as to properly coordinate with regulatory-driven, long-term signals. These

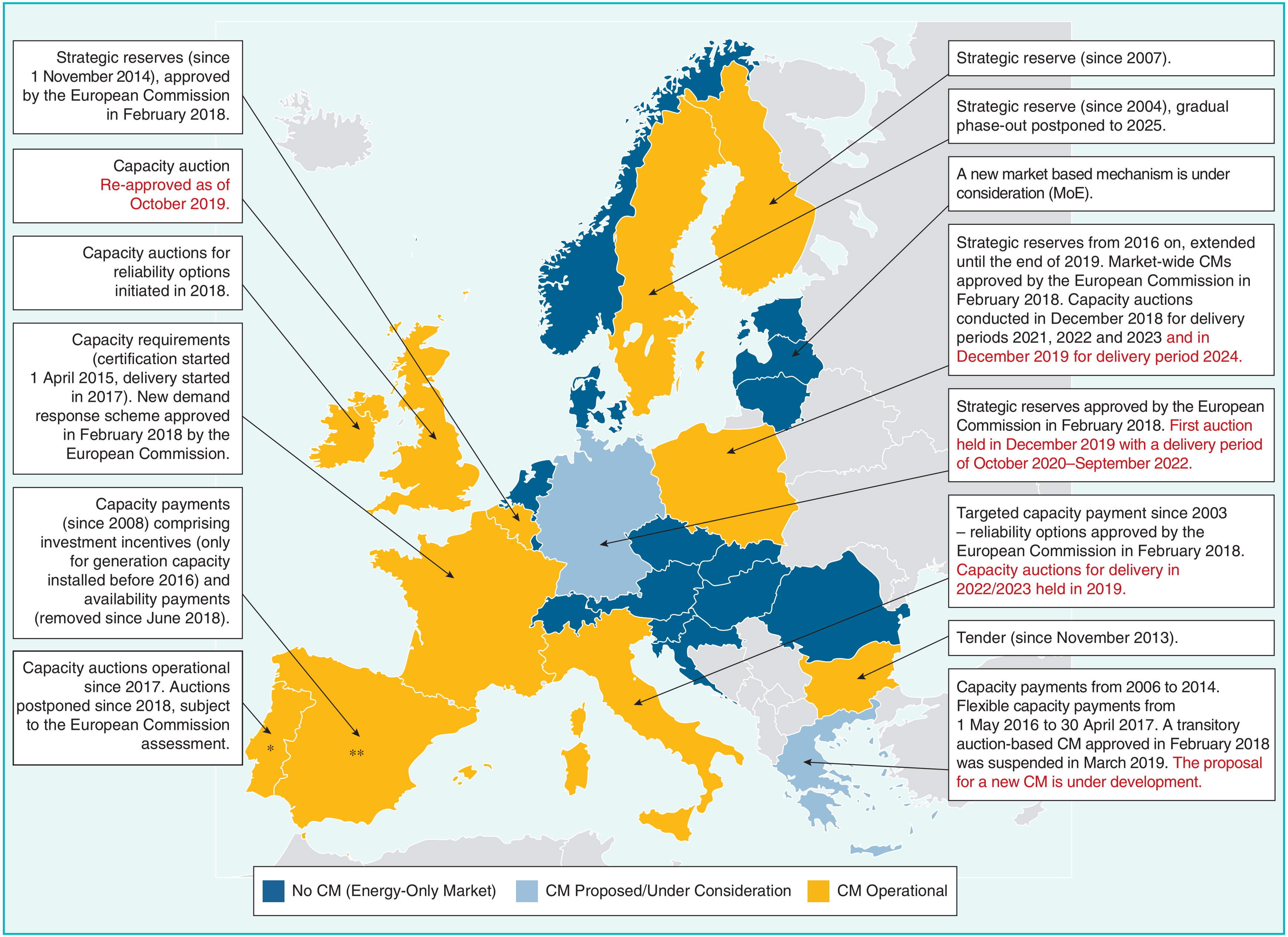


figure 4. The CRMs in Europe in 2019. The red color highlights the most recent changes that occurred in the last year. (Source: acer.europa.eu.)

regulatory challenges can be categorized along three major interrelated lines.

Redesign of Short-Term Market Mechanisms

The variability and low marginal operating costs associated with intermittent generation at scale lead to significant changes in energy market price dynamics. In this context, storage and eventually demand response appear as important tools to maximize the value of RES generation. The short-term market design needs to allow for the efficient and comprehensive integration of such technologies to reward flexibility and accurate forecasting while enhancing liquidity and transparency. Specific issues include, for instance, operating reserve demand curves or intraday and very short-term pricing mechanisms. Auction rules in these mechanisms should be further developed to allow for a level playing field among all technological alternatives, accounting for their new operational constraints without segmenting the market. For better integration of small-scale resources, new market design solutions are required for the participation of independent flexibility aggregators, while at the same time guaranteeing efficient interplay with retailers.

Minimizing Market Distortions

With recent renewed interest in integrated resource planning to achieve decarbonization, it is important to explore ways to reduce the distortions of the regulatory design mechanisms to guide new investments. Examples include explicit and implicit technology subsidies, CRMs, mandated long-term contracts, and new RAB models applied to certain technologies. This effort requires, for instance, a shift in focus from pricing energy to pricing capacity as well as replacing administratively determined subsidy levels with competitive price discovery methods. Additionally, as the expansion of state regulation continues, more fundamental changes are likely to be required in the institutions that determine entry and exit decisions.

Tools for Efficient Behavior of End Users: Market Participation and Retail Rates

Electricity end users have an unprecedented degree of choice. Well-designed price signals are instrumental for the efficient deployment and management of distributed energy resources, electric vehicles, smart appliances, and energy management systems. Yet electricity users invariably face rates and other incentives that offer them little guidance on how their myriad choices affect the cost of electricity provision. Significant developments are still pending, primarily of two types.

1) New market processes need to be developed to properly involve end users in current wholesale market mechanisms (from capacity to regulation markets). End users need to be efficiently exposed to granular price signals (for example, transitioning from the widespread use of plain volumetric tariffs and the current socialization of imbalance costs), while at the same time avoiding inefficient arbitrage opportunities with dynamic retail rates. Significant efforts are also needed to find ways to coordinate the role of aggregators and retailers to avoid undesired cross subsidies.
2) End-user retail rate making also requires a modification from pricing energy to pricing capacity as well as equitably designing fixed charges needed to avoid inefficient self-supply.

These regulatory developments should be accompanied by a redesign of the current institutional and governance framework. Revitalized institutions are necessary to independently inform integrated resource planning as well as collect fixed charges to pay for short-term products and services provided by energy and network suppliers.

For Further Reading

"RAB model for nuclear consultation on a RAB model for new nuclear projects," Department for Business, Energy & Industrial Strategy, United Kingdom Government, London, 2019. [Online]. Available: https://www.gov.uk/government/consultations/regulated-asset-base-rab-model-for-nuclear

"Final report of the sector inquiry on capacity mechanisms report from the commission," European Commission, Brussels, Belgium, SWD(2016) 385 final, Nov. 30, 2016. [Online]. Available: https://ec.europa.eu/energy/sites/ener/files/documents/com2016752.en_.pdf

I. Herrero, P. Rodilla, and C. Batlle, "Evolving bidding formats and pricing schemes in U.S. and Europe day-ahead electricity markets," Massachusetts Institute of Technology, Cambridge, MIT-CEEPR WP-2020-010, June 2020.

W. W. Hogan, "Electricity scarcity pricing through operating reserves," *Econ. Energy Environ. Pol.*, vol. 2, no. 2, pp. 65-86, 2013.

"Renewable energy auctions: Status and trends beyond price," International Renewable Energy Agency, Abu Dhabi, 2019. [Online]. Available: https://www.irena.org/publications/2019/Dec/Renewable-energy-auctions-Status-and-trends-beyond-price

K. Neuhoff, N. Ritter, and S. Schwenen, "Bidding structures and trading arrangements for flexibility across EU power markets," in *Proc. Future Power Market Workshop 2015*, Ofgem, London, June 26, 2015. [Online]. Available: http://hdl.handle.net/10419/111922

Biographies

Carlos Batlle is with the MIT Energy Initiative, Cambridge, Massachusetts, USA, and the Florence School of Regulation, Florence, Italy.

Pablo Rodilla is with Institute for Research in Technology, Comillas Pontifical University, Madrid, Spain.

Paolo Mastropietro is with Institute for Research in Technology, Comillas Pontifical University, Madrid, Spain.

©SHUTTERSTOCK.COM/IURII MOTOV

By Frank A. Wolak

Market Design in an Intermittent Renewable Future

Cost Recovery With Zero-Marginal-Cost Resources

Digital Object Identifier 10.1109/MPE.2020.3033395
Date of current version: 6 January 2021

THE BASIC FEATURES OF AN EFFICIENT SHORT-TERM wholesale market design do not necessarily need to change to accommodate a significantly larger share of zero-marginal-cost, intermittent renewable energy from wind and solar resources. A large share of controllable zero-marginal-cost generation does not create any additional market design challenge relative to a market with a large share of controllable positive marginal cost generation. Regardless of the technology, generation unit owners must recover their fixed costs from sales of energy, ancillary services, and long-term resource adequacy products.

A larger variance in the hourly amount of energy produced by intermittent resources is the primary market design challenge associated with a zero-marginal-cost renewable future. The past 10 years in California have demonstrated that, as the amount of wind and solar generation capacity increases, the variance in hourly energy produced by these resources does too. This increase in supply uncertainty also increases short-term price

1540-7977/21©2021IEEE

volatility, which can finance investments in storage and other technologies that allow consumers to shift their withdrawals of grid-supplied energy away from periods when little wind and solar energy is being produced.

An increased risk of large intermittent energy shortfalls and short-term price volatility implies a greater need for risk management activities. Greater short-term intermittent energy supply risk is likely to require accounting for more transmission and generation operating constraints in the day-ahead and real-time energy markets as well as purchasing more operating reserves and creating additional ancillary service products. Because controllable generation units are likely to have to start and stop more frequently to make up for unexpected renewable energy shortfalls, there will be a greater need to develop short-term pricing approaches that recover the associated start-up and minimum load costs.

The potential for sustained periods of low intermittent energy production creates both a medium and long-term energy supply risk that requires a new long-term resource adequacy mechanism. The traditional capacity-based approach is unlikely to be the least-cost mechanism for ensuring that the future demand for energy is met. In a zero-marginal-cost, intermittent future, wind and solar resources must hedge their energy supply risk with controllable generation resources to maintain long-term resource adequacy. Cross hedging between these technologies accomplishes two goals: First, it can provide the revenue stream necessary for fixed cost recovery by controllable generation units. Second, it ensures that there is sufficient controllable generation capacity to meet demand under all foreseeable future system states with a high degree of confidence.

The remainder of this article first describes the key features of an efficient short-term wholesale market design: a multisettlement locational marginal pricing (LMP) market with an automatic local market power mitigation (LMPM) mechanism, which is the standard market design for all short-term markets in the United States. This section concludes with a discussion of the modifications to this basic design that are likely to be necessary to accommodate a larger share of intermittent renewables.

The second half of the article describes a new long-term resource adequacy mechanism for the efficient short-term market design for an electricity supply industry with a large share of zero-marginal-cost, intermittent renewables. I first explain why a wholesale electricity market requires a long-term resource adequacy mechanism. Then, I describe a mandated, standardized long-term contract approach to long-term resource adequacy that provides strong incentives for intermittent renewable resource owners to hedge their energy supply risk with controllable generation resource owners. I argue that this mechanism ensures long-term resource adequacy at a reasonable cost for final consumers while also allowing the short-term wholesale price volatility that can finance investments in storage and other load-shifting technologies necessary to manage a large share of intermittent renewable resources.

Short-Term Market Design

More than 25 years of international experience with wholesale electricity market design has identified four crucial features of efficient short-term market design. First is the extent to which the market mechanism used to set dispatch levels and locational prices is consistent with how the grid and generation units operate. Second is a financially binding day-ahead market that prices all transmission and generation unit operating constraints expected to be relevant in real time. The third is an automatic LMPM mechanism that limits the ability of a supplier to influence the price it receives when it possesses a substantial ability to exercise market power. The fourth feature is retail market policies that foster active participation of the final demand in the wholesale market.

The early U.S. wholesale market designs in the PJM Interconnection, ISO New England, California, and Texas employed simplified versions of the transmission network configuration and generation unit operating constraints. Similar market designs currently exist throughout Europe and the rest of the world. They set a single market-clearing price for an hour or half-hour for an entire control area or large geographic regions, even though in real time there are often generation units with offer prices below this market-clearing price not producing electricity. Likewise, there are units with offer prices above this market-clearing price producing electricity. This outcome occurs because of the location of demand and available generation units within the region, and the configuration of the transmission network prevents some of these low-offer price units from producing electricity and requires some of the high-offer price units to supply electricity.

This approach to short-term market design provides incentives for suppliers to take actions to exploit the fact that "in real time physics wins," rather than offering their resources into the day-ahead market in a manner that minimizes the cost of meeting demand at all locations in the grid in real time. Instead, suppliers take actions in the simplified day-ahead market that allow them to profit from knowing they will be needed (or not needed) in real time because of transmission and generation unit operating constraints.

Locational Marginal Pricing

Starting with PJM in 1998 and ending with Texas in late 2010, all U.S. wholesale markets adopted a multisettlement LMP market design that cooptimizes the procurement of energy and ancillary services and includes an automatic LMPM mechanism built into the market software. This design has a day-ahead financial market that satisfies the locational demands for energy and each ancillary service simultaneously for all 24 h of the following day. A real-time market then operates using the same network model as the day-ahead market adjusted to real-time system conditions. Deviations from purchases and sales in the day-ahead market are cleared using these real-time prices. Both of these markets price all

> Using short-term pricing to manage the real-time supply and demand balance in a wholesale electricity market is limited by a finite upper bound on a supplier's offer price and/or a price cap.

relevant transmission network and other relevant operating constraints on generation units. As I discuss later, this market design can foster active participation of final demand in the wholesale market.

Only generation unit output levels that are physically feasible will be accepted in both the day-ahead and real-time markets. Prices for the same hour vary depending on whether the location is in a generation-deficient or generation-rich region of the transmission network. The locational marginal or nodal price at a given location is the increase in the minimized value of the "as-offered costs" of serving the locational demands for energy and all ancillary services as a result of a one-unit increase in the amount of energy withdrawn at that location in the transmission network. The price of each ancillary service is equal to the increase in the optimized value of the objective function as a result of a one-unit increase in the demand for that ancillary service.

The recent experience of many European countries with significant wind and solar resources indicates that the cost of making the final schedules that emerge from their zonal markets physically feasible is likely to get even larger as the amount of intermittent renewable generation capacity increases. According to the European Network of Transmission System Operators for Electricity, in 2017 these costs were more than €1 billion in Germany, more than €400 million in Great Britain, more than €80 million in Spain, and approximately €50 million in Italy.

Multisettlement LMP Market

A multisettlement LMP market has at least a day-ahead forward market and a real-time market, each of which employs the same market-clearing mechanism. The day-ahead market typically allows generation unit owners to submit three-part offers to supply energy: start-up costs, minimum load costs, and an energy offer curve. These are used to compute hourly generation schedules, ancillary service quantities, and LMPs for energy and ancillary services for all 24 h of the following day. A generation unit will not be accepted to supply energy in the day-ahead market unless the combination of its offered start-up costs, minimum load costs, and energy production costs are part of the least as-offered-cost solution to serving the hourly locational demands for all 24 h of the following day.

The energy schedules that emerge from the day-ahead market do not require a generation unit to produce the energy sold or a load to consume the energy purchased in the day-ahead market at a given location. Any production shortfall relative to a day-ahead generation schedule must be purchased from the real-time market at that location. Any production greater than a generation unit's day-ahead schedule is sold at the real-time price at that location. Any additional consumption beyond a load's day-ahead energy schedule is paid for at the real-time price at that location, and the surplus of a day-ahead schedule relative to actual consumption is sold at the real-time price at that location.

Mitigating Local Market Power

The configuration of the transmission network, the level and location of demand, and the level of output of other generation units can create system conditions in which almost any generation unit or group of generation units has a significant ability to exercise unilateral market power. The constrained-on generation problem is an example of this phenomenon. The unit's owner knows that it must be accepted to supply energy regardless of its offer price. Without an LMPM mechanism, there may be no limit to the offer price the unit owner could submit and have accepted to supply energy. During the first summer of the California market, when there was no formal LMPM mechanism, suppliers submitted extremely high offers for energy and ancillary services when these system conditions arose. This logic is why market power-mitigation mechanisms typically used in Europe and other industrialized regions and initially employed in the United States, which designate in advance the offers of certain generation units for mitigation for an entire year, miss many instances of the exercise of substantial unilateral market power.

An automated LMPM mechanism built into the market software that relies on actual system conditions to determine whether any supplier has a substantial ability and incentive to exercise unilateral market power is likely to be significantly more effective. This regulator-approved administrative procedure determines 1) when a supplier has an ability to exercise local market power worthy of mitigation, 2) the value of the supplier's mitigated offer price, and 3) the price the mitigated supplier is paid. It is increasingly clear to regulators around the world, particularly those that operate markets with a finite amount of transmission capacity and significant intermittent renewable generation capacity, that an automatic LMPM mechanism is a necessary feature of any short-term market design. Because these LMPM mechanisms are built into the market software of all U.S.

markets and automatically mitigate the offers of suppliers deemed to have a substantial ability to exercise unilateral market power, they are effective at preventing the exercise of significant local market power with little disruption to the operation of the short-term market.

Benefits of a Multisettlement LMP Market

A multisettlement LMP market design can facilitate the active participation of final consumers in the wholesale market and reduce both the input fuel and total variable cost of producing the same amount of thermal energy relative to the multisettlement zonal market design. The presence of an automatic LMPM mechanism and make-whole payments that guarantee start-up, minimum load, and energy cost recovery for the day for all generation units committed to operating in the day-ahead market reduces the incentive for suppliers to exercise unilateral market power. An expected profit-maximizing supplier with no ability to exercise unilateral market power will submit an offer price equal to its marginal cost because make-whole payments ensure recovery of its start-up, minimum load, and energy costs.

Because day-ahead purchases are firm financial commitments, a retailer can sell energy purchased in the day-ahead market at the real-time price by consuming less than its day-ahead energy schedule. This eliminates the need for the regulator to set an administrative baseline relative to which a retailer sells demands reductions. The day-ahead market also allows retailers and large consumers to submit price-sensitive bid curves into the day-ahead market to reduce the market-clearing price and the quantity of energy they purchase in the day-ahead market.

Modifications for Large-Scale Intermittent Renewables Deployment

A multisettlement LMP market design is capable of managing a generation mix with a significant share of intermittent renewables. However, some modifications are likely to be needed as the share of intermittent renewable resources increases. Additional operating constraints will need to be incorporated into the day-ahead and real-time market models for reliable system operation with an increased quantity of intermittent renewables.

Introducing additional ancillary services to accommodate a larger share of intermittent renewable energy may also be needed. For example, California introduced a fast-ramping ancillary service product that compensates controllable generation units not supplying energy during certain hours of the day in order to have sufficient unloaded capacity to meet the rapid increase in net demand (the difference between system demand and renewable generation) in the early evening, when the state's solar resources stop producing. Because controllable resources are likely to have to start and stop more frequently as the share of intermittent resources increases, implementations of convex hull pricing and other market-clearing mechanisms that limit the magnitude of make-whole payments will need to be developed.

Resource Adequacy With Significant Intermittent Renewables

Why do wholesale electricity markets require a regulatory mandate to ensure long-term resource adequacy? Electricity is essential to modern life, but so are many other goods and services. Consumers want cars, but there is no regulatory mandate that ensures enough automobile assembly plants to produce them. They want point-to-point air travel, but there is no regulatory mandate to ensure enough airplanes to accomplish this. Many goods are produced using high-fixed-cost, low-marginal-cost technologies, similar to electricity supply. Nevertheless, these firms recover their cost of production, including a return on the capital invested, by selling their output at a market-determined price.

So, what is different about electricity that requires a long-term resource adequacy mechanism? The regulatory history of the electricity supply industry and the legacy technology for metering electricity consumption results in what I call a *reliability externality*.

The Reliability Externality

Different from the case of wholesale electricity, in the market for automobiles and air travel there is no regulatory prohibition on the short-term price rising to the level necessary to clear the market. Airlines adjust the prices for seats on a flight over time in an attempt to ensure that the number of customers traveling on that flight equals the number of seats flying. This ability to use price to allocate the available seats is also what allows the airline to recover its total production costs.

Using short-term pricing to manage the real-time supply and demand balance in a wholesale electricity market is limited by a finite upper bound on a supplier's offer price and/or a price cap that limits the maximum market-clearing price. Although offer and price caps can limit the ability of suppliers to exercise unilateral market power in the short-term energy market, they also reduce the revenues suppliers can receive during scarcity conditions. This is often referred to as the *missing money* problem for generation unit owners. However, this missing money problem is only a symptom of the existence of the "reliability externality."

This externality exists because offer caps limit the cost to electricity retailers of failing to hedge their purchases from the short-term market. Specifically, if the retailer or large consumer knows the price cap on the short-term market is US$250/MWh, then it is unlikely to be willing to pay more than that for electricity in any earlier forward market. This creates the possibility that real-time system conditions can occur where the amount of electricity demanded at or below the offer cap is less than the amount suppliers are willing to offer at or below the offer cap. This outcome implies that the system operator must be forced to either abandon the market mechanism or curtail load until the available supply offered at or below the offer cap equals the reduced level of demand, as occurred several times in California between January 2001 and April 2001 and most recently on 14–15 August 2020.

Because random curtailments of supply, also known as "rolling blackouts," are used to make demand equal to the available supply at or below the offer cap under these system conditions, this mechanism creates a "reliability externality" because no retailer bears the full cost of failing to procure adequate amounts of energy in advance of the delivery. A retailer that has purchased sufficient supply in the forward market to meet its actual demand is equally likely to be randomly curtailed as another retailer of the same size that has not procured adequate energy in the forward market. For this reason, all retailers have an incentive to underprocure their expected energy needs in the forward market.

The lower the offer cap, the greater the likelihood that the retailer will delay its electricity purchases to the short-term market. Delaying more purchases to the short-term market increases the likelihood of insufficient supply in the short-term market at or below the offer cap. Because retailers do not bear the full cost of failing to procure sufficient energy in the forward market to meet their future demand, there is a missing market for long-term contracts for long enough delivery horizons into the future to allow new generation units to be financed and constructed to serve demand under all future conditions in the short-term market. Therefore, a regulator-mandated, long-term resource-adequacy mechanism is necessary to replace this missing market.

Some form of regulatory intervention is necessary to internalize the resulting reliability externality unless the regulator is willing to eliminate or substantially increase the offer cap so that the short-term price can be used to equate available supply to demand under all possible future system conditions. This approach is taken by the Electricity Reliability Council of Texas, which has a US$9,000/MWh offer cap, and the National Electricity Market in Australia, which has an AUD$15,000 per MWh offer cap. If customers do not have interval meters that can record their consumption on an hourly basis, then they have a very limited ability to benefit from shifting their consumption away from high-priced hours. All that can be recorded for these customers is their total consumption between two successive meter readings so they can only be billed based on an average wholesale price during the billing cycle. Therefore, raising or having no offer cap on the short-term market would not be advisable in a region where few customers have interval meters. Even in regions with interval meters, there would be substantial political backlash from charging hourly wholesale prices that cause real-time demand to equal available supply under all possible future system conditions.

Currently, the most popular approach to addressing this reliability externality is a capacity payment mechanism that assigns a firm capacity value to each generation unit based on the amount of energy it can provide under stressed system conditions. Sufficient firm capacity procurement obligations are then assigned to retailers to ensure that annual system demand peaks can be met.

Capacity-based approaches to long-term resource adequacy rely on the credibility of the firm capacity measures assigned to generation units. This is a relatively straightforward process for thermal units. The nameplate capacity of the generation unit times its annual availability factor is a reasonable estimate of the amount of energy the unit can provide under stressed system conditions. In the case of hydroelectric facilities, this process is less straightforward. The typical approach uses percentiles of the distribution of past hydrological conditions for that generation unit to determine its firm capacity value.

Assigning a firm capacity value to a wind or solar generation unit is extremely challenging for several reasons. First, these units only produce when the underlying resource is available. If stressed system conditions occur when the sun is not shining or the wind is not blowing, these units should be assigned little, if any, firm capacity value. Second, because there is a high degree of contemporaneous correlation between the energy produced by solar and wind facilities within the same region, the usual approach to determining the firm capacity of a wind or solar unit assigns a smaller value to that unit as the total megawatts of wind or solar capacity in the region increases. These facts imply that a capacity-based, long-term resource-adequacy mechanism is poorly suited to a zero-marginal-cost, intermittent renewable feature.

Supplier Incentives With Fixed-Price Forward Contract Obligations for Energy

The standardized fixed-price forward contract (SFPFC) approach to long-term resource adequacy recognizes that a supplier with the ability to serve demand at a reasonable price may not do so if it can exercise unilateral market power. A supplier with the ability to exercise unilateral market power with a fixed-price forward contract obligation finds it expected profit maximizing to minimize the cost of supplying this forward contract quantity of energy. The SFPFC long-term resource adequacy mechanism takes advantage of this incentive by requiring retailers to hold hourly fixed-price forward contract obligations for energy that sum to the hourly value of system demand. This implies that all suppliers find it expected profit maximizing to minimize the cost of meeting their hourly fixed-price forward contract obligations, the sum of which equals the hourly system demand for all hours of the year.

To understand the logic behind the SFPFC mechanism, consider the example of a supplier who owns 150 MW of generation capacity who has sold 100 MWh in a fixed-forward contract for US$25/MWh for a certain hour of the day. This supplier has two options for fulfilling this forward contract: 1) produce the 100 MWh energy from its own units at its marginal cost of US$20/MWh or 2) buy this energy from the short-term market at the prevailing market-clearing price. The supplier will receive US$2,500 from the buyer of the contract for the 100 MWh sold, regardless of how it is supplied. This means that the supplier maximizes the profits it earns from this fixed-price forward contract sale by minimizing the cost of supplying the 100 MWh of energy.

To ensure that the least-cost "make versus buy" decision for this 100 MWh is made, the supplier should offer 100 MWh in the short-term market at its marginal cost of US$20/MWh. This offer price for 100 MWh ensures that if it is cheaper to produce the energy from its generation units (the market price is at or above US$20/MWh), the supplier's offer to produce the energy will be accepted in the short-term market. If it is cheaper to purchase the energy from the short-term market (the market price is below US$20/MW), the supplier's offer will not be accepted and the supplier will purchase the 100 MWh from the short-term market at a price below US$20/MWh.

This example demonstrates that the SFPFC approach to long-term resource adequacy makes it expected profit maximizing for each seller to minimize the cost of supplying the quantity of energy sold in this forward contract each hour of the delivery period. By the logic of the previous example, each supplier will find it in its unilateral interest to submit an offer price into the short-term market equal to its marginal cost for its hourly SFPFC quantity of energy, in order to make the efficient "make versus buy" decision for fulfilling this obligation.

Also, because all suppliers know that the sum of the values of the hourly SFPFC obligations for all suppliers is equal to the system demand, each firm knows that its competitors have substantial fixed-price forward contract obligations for that hour. This implies that all suppliers know that they have limited opportunities to raise the price they receive for short-term market sales beyond their hourly SFPFC quantity. For the previous example, the supplier who owns 150 MWs of generation capacity has a strong incentive to submit an offer price close to its marginal cost to supply any energy beyond the 100 MWh of SFPFC energy it is capable of producing. Therefore, attempts by any supplier to raise prices in the short-term market by withholding output beyond its SFPFC quantity are likely to be unsuccessful because of the aggressiveness of the offers into the short-term market by its competitors with hourly SFPFC obligations.

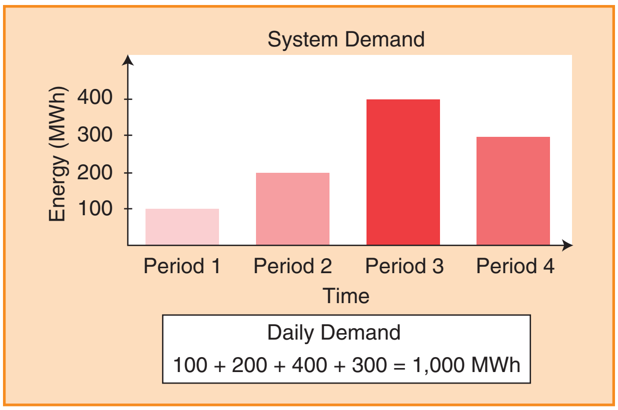


figure 1. Hourly system demands.

The SFPFC Approach to Resource Adequacy

This long-term resource adequacy mechanism requires all electricity retailers to hold SFPFCs for energy for fractions of realized system demand at various horizons to delivery. For example, retailers, in total, must hold SFPFCs that cover 100% of realized system demand in the current year, 95% of realized system demand one year in advance of delivery, 90% two years in advance of delivery, 87% three years in advance of delivery, and 85% four years in advance of delivery. The fractions of system demand and the number of years in advance that the SFPFCs must be purchased are parameters set by the regulator to ensure long-term resource adequacy. In the case of a multisettlement LMP market, the SFPFCs would clear against the quantity-weighted average of the hourly locational prices at all load withdrawal nodes.

SFPFCs are shaped to the hourly system demand within the delivery period of the contract. Figure 1 contains a sample pattern of the system demand for a 4-h delivery horizon. The total demand for the 4-h is 1,000 MWh, and the four hourly demands are 100, 200, 400, and 300 MWh. Therefore, a supplier that sells 300 MWh of SFPFC energy has the hourly system, demand-shaped forward contract obligations of 30 MWh in hour one, 60 MWh in hour two, 120 MWh in hour three, and 90 MWh in hour four as shown for Firm 1 in Figure 2. The hourly forward contract obligations for Firm 2 that sold 200 MWh SFPFC energy and Firm 3 that sold 500 MWh of SFPFC energy are also illustrated in Figure 2. These SFPFC obligations are also allocated across the 4 h according to the same four hourly shares of total system demand. This ensures that the sum of the hourly values of the forward contract obligations for the three suppliers is equal to the hourly value of the system demand. Taking the example of hour three, Firm 1's obligation is 120 MWh, Firm 2's is 80 MWh, and Firm 3's is 200 MWh. These three values sum to 400 MWh, which is equal to the value of system demand in hour three, shown in Figure 1.

These SFPFCs are allocated to retailers based on their share of system demand during the month. Suppose that the four retailers in Figure 3 consume 1/10, 2/10, 3/10, and 4/10, respectively, of the total energy consumed during the month. This means that Retailer 1 is allocated 100 MWh of the 1,000 MWh SFPFC obligations for the 4 h, Retailer 2 is

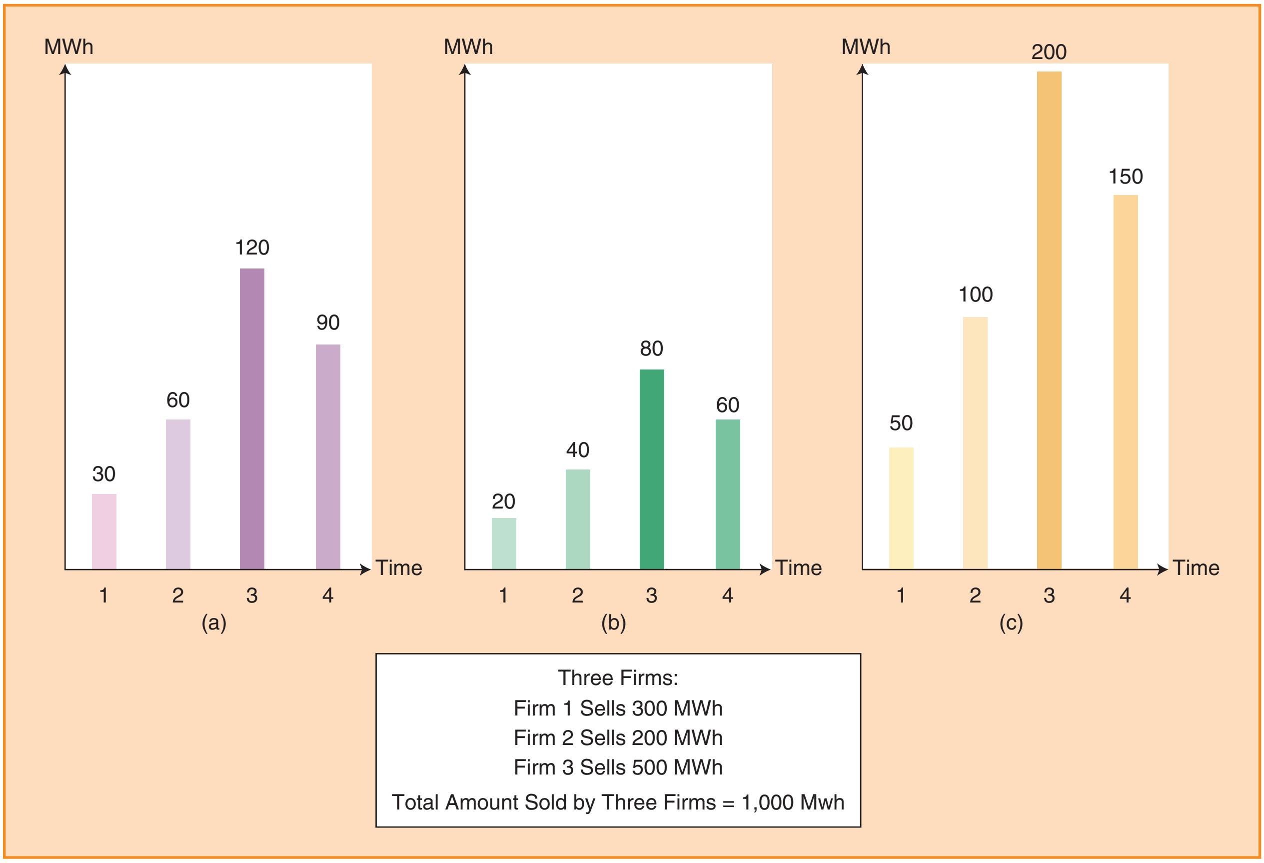


figure 2. The hourly forward contract quantities for three suppliers. The forward contract obligation per period for (a) Firm 1, (b) Firm 2, and (c) Firm 3.

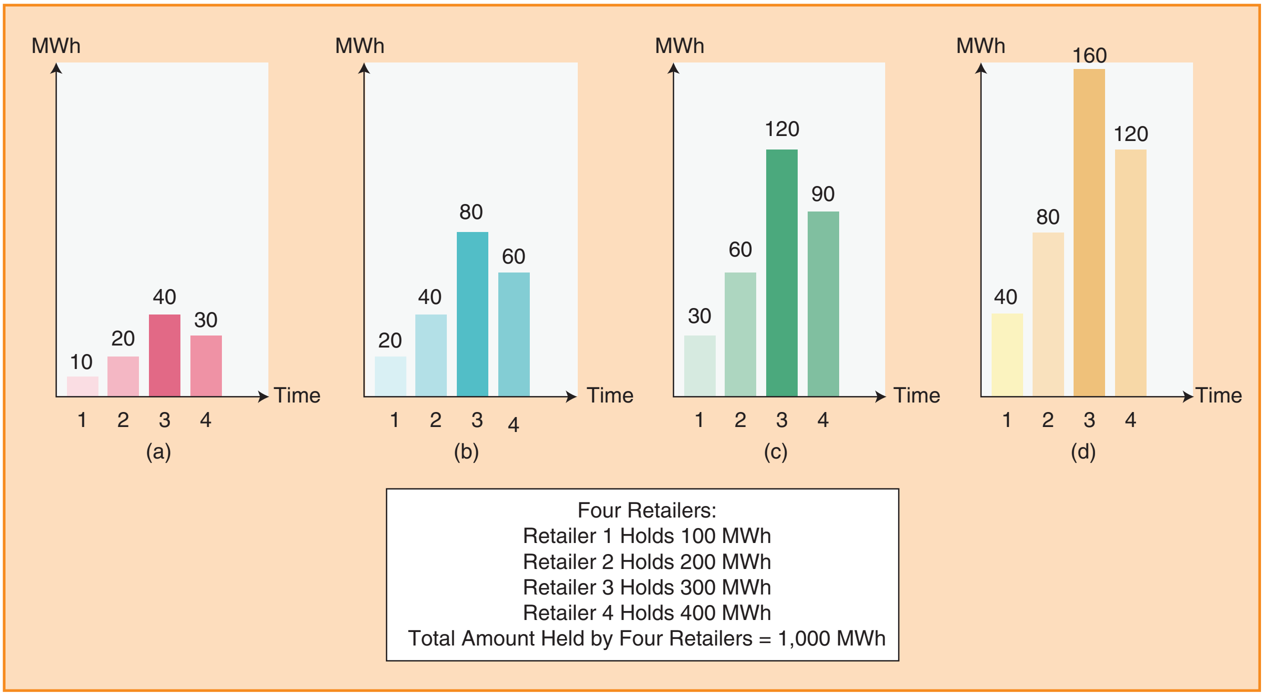


figure 3. The hourly forward contract quantities for four retailers. The forward contract obligation per period for (a) Retailer 1, (b) Retailer 2, (c) Retailer 3, and (d) Retailer 4.

allocated 200 MWh, Retailer 3 is allocated 300 MWh, and Retailer 4 is allocated 400 MWh. The obligations of each retailer are then allocated to the individual hours using the same hourly system demand shares used to allocate the SFPFC energy sales of suppliers to the 4 h. This allocation process implies Retailer 1 holds 10 MWh in hour one, 20 MWh in hour two, 40 MWh in hour three, and 30 MWh in hour four. Repeating this same allocation process for the other three retailers yields the remaining three hourly allocations displayed in Figure 3. Similar to the case of the suppliers, the sum of allocations across the four retailers for each hour equals the total hourly system demand. For period 3, Retailer 1's holding is 40 MWh, Retailer 2's is 80 MWh, Retailer 3's is 120 MWh, and Retailer 4's is 160 MWh. The sum of these four magnitudes is equal to 400 MWh, which is the system demand in hour three.

Mechanics of the Standardized Forward Contract Procurement Process

The SFPFCs would be purchased through auctions several years in advance of delivery to allow new entrants to compete to supply this energy. Because the aggregate hourly values of these SFPFC obligations are allocated to retailers based on their actual share of system demand during the month, this mechanism can easily accommodate retail competition. If one retailer loses the load and another gains it during the month, the share of the aggregate hourly value of SFPFCs allocated to the first retailer falls and the share allocated to the second retailer rises.

The wholesale market operator would run the auctions with oversight by the relevant regulator. One advantage of the design of the SFPFC products is that a simple auction mechanism can be used to purchase each annual product. A multiround auction could be run where suppliers submit the total amount of annual SFPFC energy they would like to sell for a given delivery period at the price for the current round. At each round of the auction, the price would decrease until the amount suppliers are willing to sell at that price is less than or equal to the aggregate amount of SFPFC energy demanded.

The wholesale market operator would also run a clearinghouse to manage the counterparty risk associated with these contracts. All U.S. wholesale market operators currently do this for all participants in their energy and ancillary services markets. In several U.S. markets, the market operator also provides counterparty risk management services for long-term financial transmission rights, which is not significantly different from performing this function for SFPFCs.

SFPFC auctions would be run on an annual basis for deliveries, starting two, three, and four years in the future. In a steady state, auctions for incremental amounts of each annual contract would also be needed so that the aggregate share of demand covered by each annual SFPFC could increase over time. The eventual 100% coverage of demand occurs through a final true-up auction that takes place after the realized values for hourly demand for the delivery period are known.

Consider the following two examples of how the true-up auction would work. Assume for simplicity, the monthly load shares of the four retailers remain unchanged. Suppose that the initial 1,000 MWh SFPFC in the previous example sold at US$50/MWh. However, suppose that the actual demand turned out to be 10% higher in every period as depicted in Figure 4, and the additional 100 MWh purchased in the true-up auction sold at US$80/MWh. If each firm sold 10% more SFPFC energy in the true-up auction, this would yield the hourly obligations for each supplier indicated in Figure 5. The hourly obligations for the four retailers are presented in Figure 6. These would clear against the average cost of purchases from the original auction and true-up auction of US$52.73. If the realized hourly demands are 10% lower as demonstrated in Figure 7, the true-up auction would buy back 100 MWh of SFPFC energy. If all suppliers bought back 10% of their initial sales at US$20/MWh, the resulting hourly obligations would be those in Figure 8. The 10% smaller hourly obligations of the four retailers are provided in Figure 9, and these would clear against the average cost of the initial auction purchase minus the revenues from the true-up auction sales for the required 900 MWh of the obligations of US$53.33.

As depicted in Figures 6 and 9, each purchase or sale of the same annual SFPFC product is allocated to retailers according to their load shares during the delivery month. If three different size purchases are made for the same annual SFPFC product at different prices, then each retailer is allocated

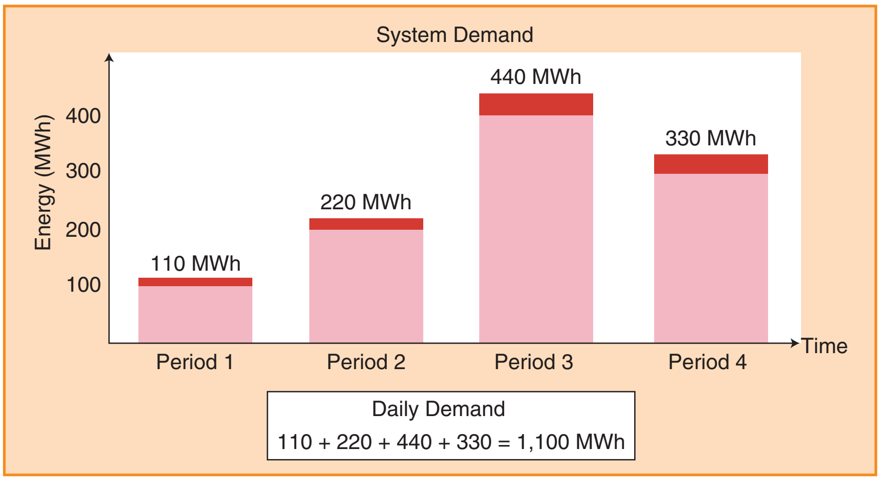


figure 4. Hourly system demands (10% higher).

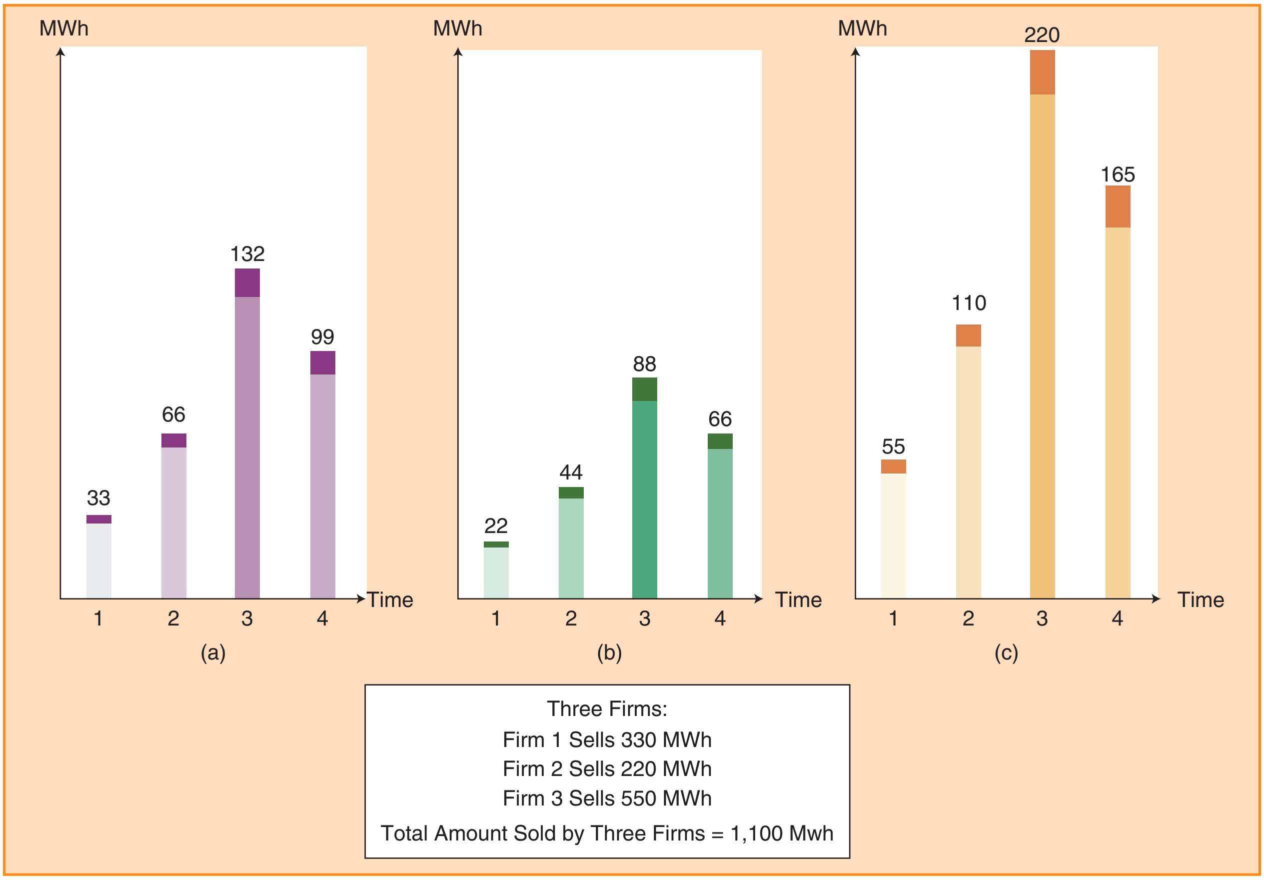


figure 5. The hourly forward contract quantities for three suppliers (10% higher). The forward contract obligation per period for (a) Firm 1, (b) Firm 2, and (c) Firm 3.

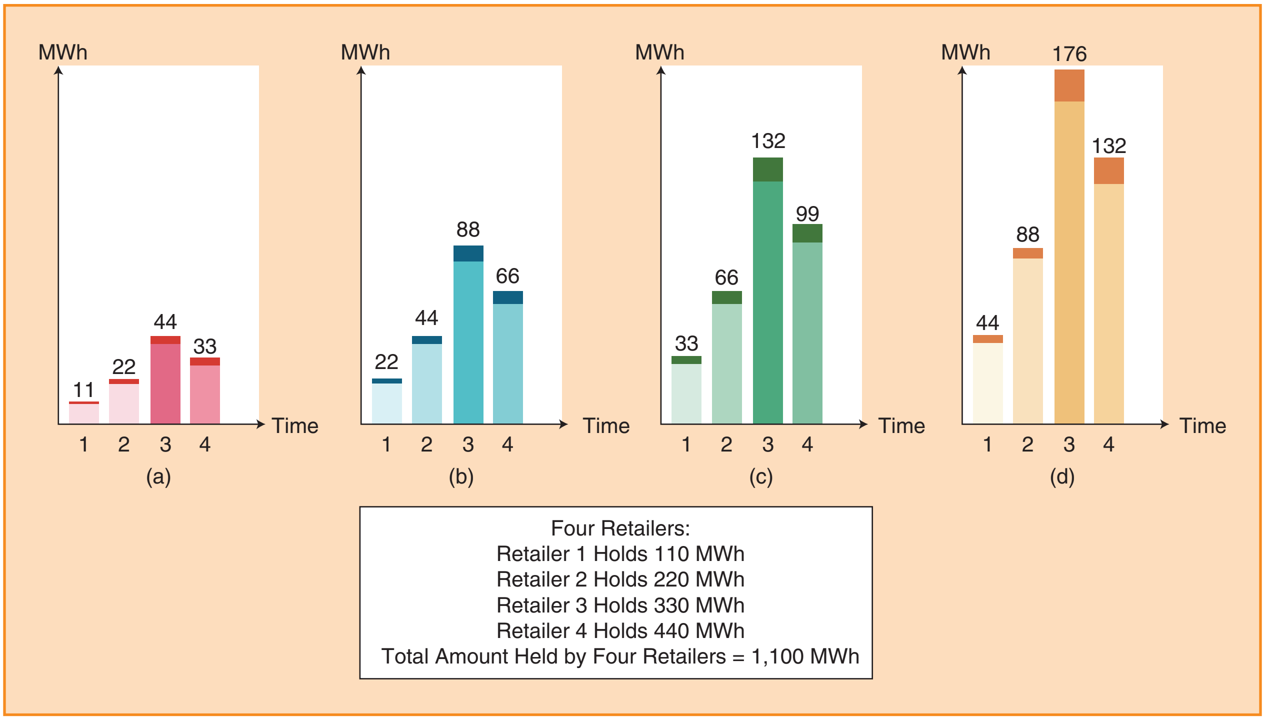


figure 6. The hourly forward contract quantities for four retailers (10% higher). The forward contract obligation per period for (a) Retailer 1, (b) Retailer 2, (c) Retailer 3, and (d) Retailer 4.

its load share for the month of these three purchases. This ensures a level playing field for retailers with respect to their long-term resource-adequacy obligation. All retailers face the same average price for the long-term resource-adequacy obligation associated with their realized demand for the month.

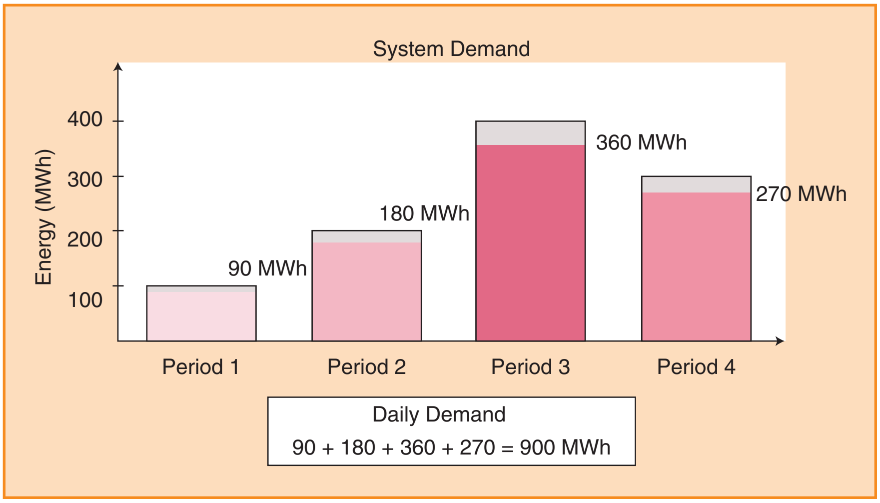


figure 7. Hourly system demands (10% lower).

The advance purchase fractions of the final demand are the regulator's security blanket to ensure that system demand can be met for all hours of the year and for all possible future system conditions. If the regulator is worried that not enough resources will be available in time to satisfy this requirement, it can increase the share of the final demand that it purchases in each annual SFPFC auction. As shown previously, if too much SFPFC energy is purchased in an annual auction, it can be sold back to generation unit owners in a later auction or the final true-up auction.

Cross hedging between controllable generation units and intermittent renewable resources under this mechanism can be enforced by tying the amount of SFPFC energy a generation unit owner can sell on an annual basis to the value of their firm energy. The system operator would assign firm energy values for each generation unit using a mechanism similar to what is

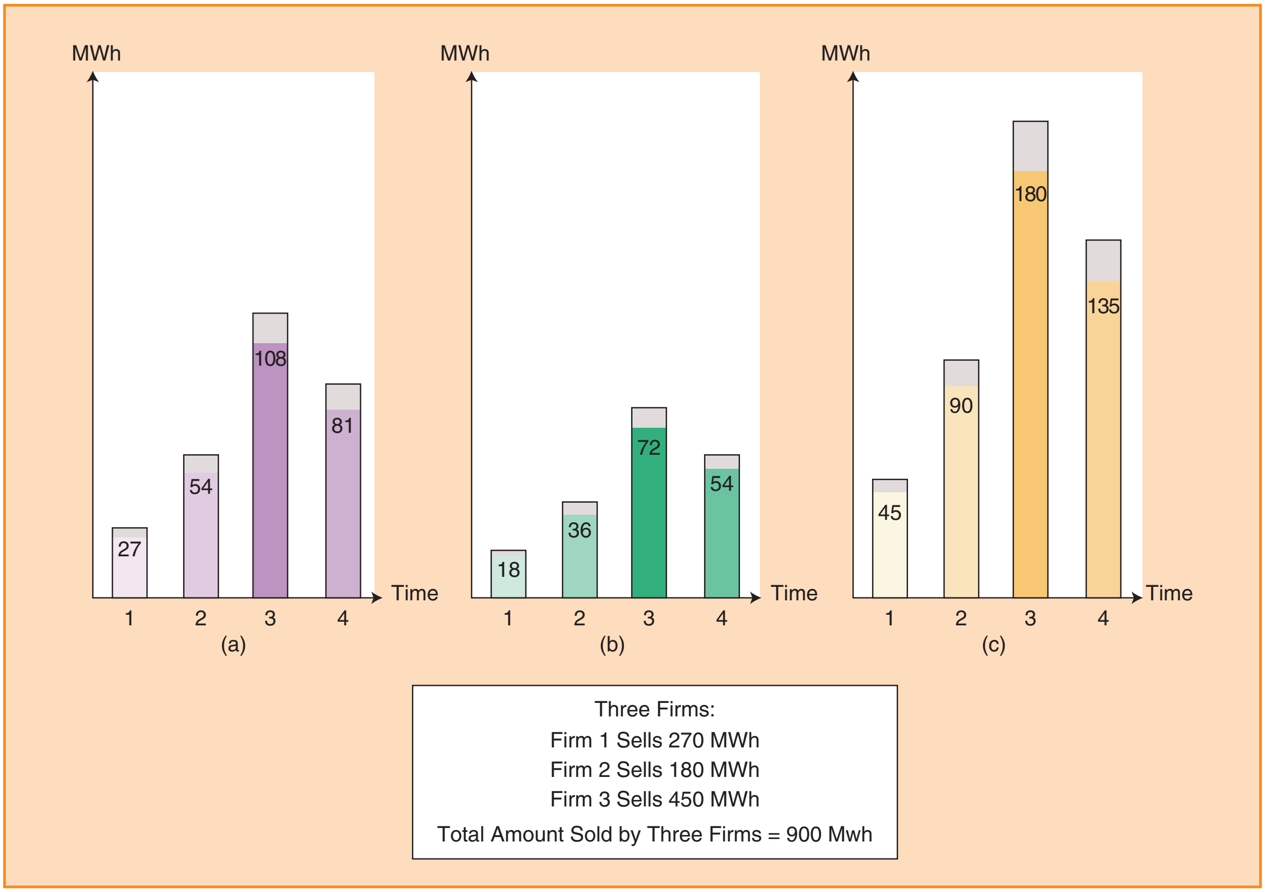


figure 8. The hourly forward contract quantities for three suppliers (10% lower). The forward contract obligation per period for (a) Firm 1, (b) Firm 2, and (c) Firm 3.

The advance purchase fractions of the final demand are the regulator's security blanket to ensure that system demand can be met for all hours of the year and for all possible future system conditions.

currently used to compute firm capacity values. Multiplying a unit's megawatts of firm capacity by the number of hours in the year would yield the unit's firm energy value, which is the upper bound on the amount of SFPFC energy the unit owner could sell in all auctions for an annual compliance period. Because the firm capacity of a generation unit is defined as the amount of energy it can produce under stressed system conditions, this limitation on the annual sales of firm energy implies that intermittent wind and solar resources would sell much less SFPFC energy than the total megawatt hours they expect to produce in a typical year, and controllable generation unit owners would sell significantly more SFPFC energy than the total megawatt hours they expect to produce in an typical year.

In most years, a controllable resource owner would be producing energy in a small number of hours of the year but earning the difference between the price at which the energy was sold in the SFPFC auction and the hourly short-term market price times the hourly value of its SFPFC energy obligation for all of the hours that it does not produce energy. Owners of intermittent renewables would typically produce more than their SFPFC obligation in energy and sell the additional energy at the short-term price. In years with a low renewable output near their SFPRC obligations, controllable resource owners would produce close to the hourly value of their SFPFC energy obligation, thus making average short-term prices significantly higher. However, aggregate retail demand would be shielded from these high short-term prices because of their SFPFC holdings.

Advantages of the SFPFC Approach to Long-Term Resource Adequacy

This mechanism has many advantages relative to a capacity-based approach. There is no regulator-mandated aggregate capacity requirement. Generation unit owners are allowed to decide both the total megawatts and the mix of technologies to meet their SFPFC energy obligations. There is also no prohibition on generation unit owners or retailers engaging in other hedging arrangements outside of this mechanism. Specifically, a retailer could enter into a bilateral contract for energy with a generation unit owner or another retailer to manage the short-term price and quantity risk associated with

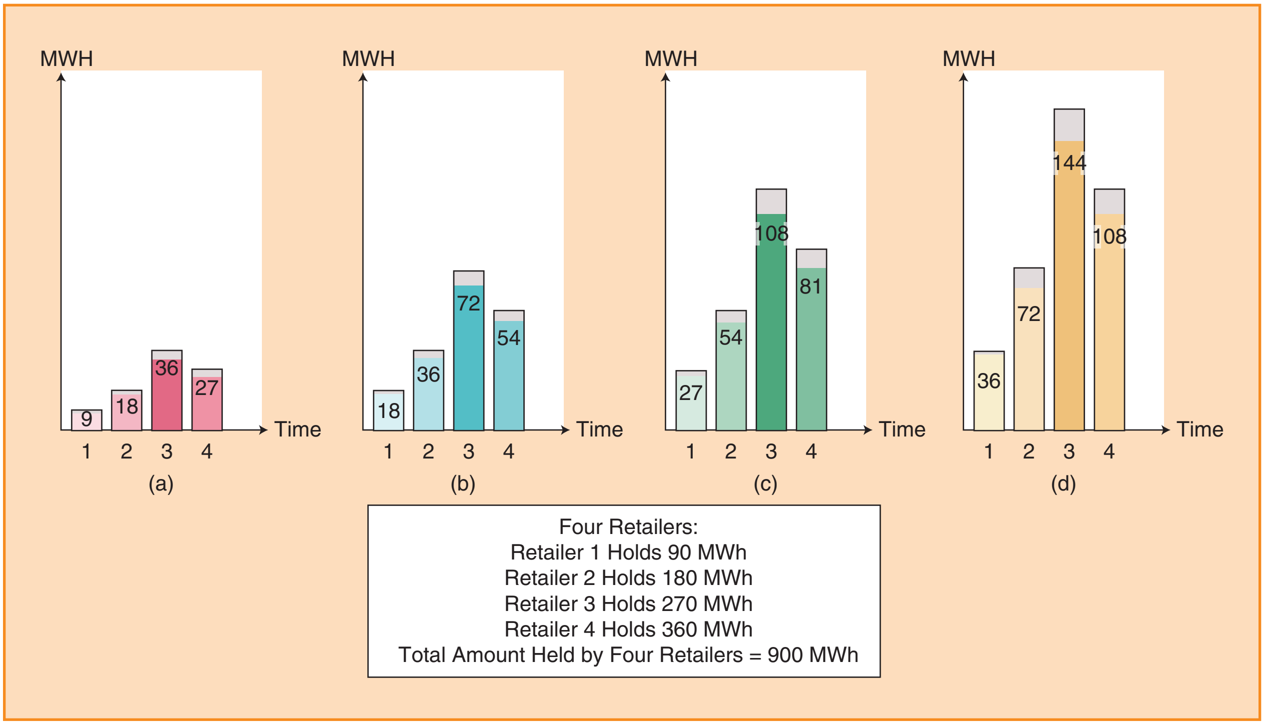


figure 9. The hourly forward contract quantities for four retailers (10% lower). The forward contract obligation per period for (a) Retailer 1, (b) Retailer 2, (c) Retailer 3, and (d) Retailer 4.

> This mechanism provides a nudge to market participants to develop a liquid market for these bilateral contract arrangements at horizons to a delivery similar to the SFPFC products.

the difference between their actual hourly load shape and the hourly values of their retail load obligation. This mechanism provides a nudge to market participants to develop a liquid market for these bilateral contract arrangements at horizons to a delivery similar to the SFPFC products. Instead of starting from the baseline of a no fixed-price, forward contract coverage of system demand by retailers, this mechanism starts with 100% coverage of system demand, which retailers can unwind at their own risk.

For the regulated retail customers, the purchase prices of SFPFCs can be used to set the wholesale price implicit in the regulated retail price over the time horizon that the forward contract clears. This would provide retailers with a strong incentive to reduce their average wholesale energy procurement costs below this price through bilateral hedging arrangements, storage investments, or demand response efforts.

There are several reasons why this mechanism should be a more cost-effective approach to long-term resource adequacy than a capacity-based mechanism in a zero-marginal-cost, intermittent future. First, the sale of SFPFC energy starting delivery two or more years in the future provides a revenue stream that will significantly increase investor confidence in recovering the cost of any investment in the new generation capacity. Second, because retailers are protected from high short-term prices by total hourly SFPFC holdings equal to system demand, the offer cap on the short-term market can be raised to increase the incentive for all suppliers to produce as much energy as possible during stressed system conditions. Third, the possibility of higher short-term price spikes can finance investments in storage and load-shifting technologies and encourage active participation of final demand in the wholesale market, further enhancing system reliability in a market with significant intermittent renewable resources.

If SFPFC energy is sold for delivery in four years based on a proposed generation unit, the regulator should require the construction of the new unit to begin within a prespecified number of months after the signing date of the contract or require the posting of a substantially larger amount of collateral in the clearinghouse with the market operator. Otherwise, the amount of SFPFC energy that this proposed unit sold would be automatically liquidated in a subsequent SFPFC auction, and a financial penalty would be imposed on the developer. Other completion milestones would have to be met at future dates to ensure the unit can provide the amount of firm energy that it committed to provide in the SFPFC contract sold. If any of these milestones were not met, the contract would be liquidated.

Final Comments

There is no perfect wholesale market design. There are only better wholesale market designs, and what constitutes a better design depends on many factors specific to the region. The long-term resource adequacy mechanism should be coordinated with short-term market design. Although there is general agreement on the key features of a best-practice, short-term market design, many details must be adjusted to reflect local conditions. For this reason, wholesale market design is a process of continuous learning, adaption, and, hopefully, improvement. The standardized energy contracting approach to long-term resource adequacy described in this article is an example of this process. While it has many features likely to make it significantly better suited to a zero-marginal-cost, intermittent-renewables electricity-supply industry, there are many details of this basic mechanism that should be adapted to reflect local conditions.

For Further Reading

C. Graf, F. Quaglia, and F.A. Wolak, “Simplified electricity market models with significant intermittent renewable capacity: Evidence from Italy,” Stanford Univ., CA, Jan. 2020. [Online]. Available: https://ngi.stanford.edu/sites/g/files/sbiybj14406/f/GrafQuagliaWolak_SimplifiedElectricityMarketModelsRenewables%281%29.pdf

P. R. Gribik, W. W. Hogan, and S. L. Pope, “Market-Clearing Electricity Prices and Energy Uplift,” Harvard Electricity Policy Group, Cambridge, MA, Dec. 31, 2007.

W. W. Hogan, “Electricity market restructuring: Reforms of reforms,” *J. Regulatory Econ.*, vol. 21, no. 1, pp. 103–132, 2002. doi: 10.1023/A:1013682825693.

F. C. Schweppe, M. C. Caramanis, R. D. Tabors, and R. E. Bohn, *Spot Pricing of Electricity*. Springer Science & Business Media, 2013.

F. A. Wolak, “Wholesale electricity market design,” in *Handbook on the Economics of Electricity Markets*, J. M. Glanchant, P. L. Joskow, and M. Pollitt, Eds. Cambridge, U.K.: Cambridge Univ. Press, 2020.

Biography

Frank A. Wolak is with Stanford University, California, USA.

Electricity Market of the Future

©SHUTTERSTOCK.COM/KHAKIMULLIN ALEKSANDR

Potential North American Designs Without Fuel Costs

By Erik Ela, Andrew Mills, Eric Gimon, Mike Hogan, Nicole Bouchez, Anthony Giacomoni, Hok Ng, Jim Gonzalez, and Mike DeSocio

Digital Object Identifier 10.1109/MPE.2020.3033396
Date of current version: 6 January 2021

ELECTRICITY MARKETS IN THE UNITED STATES and Canada have evolved since their inception in the late 1990s and early 2000s. Not all states and provinces moved toward restructured organized electricity markets, but rather those that have belonged to markets operated by independent system operators (ISOs) and regional transmission organizations, with designs developed through stakeholder processes and approved through state, provincial, or federal agencies, such as the Federal Energy Regulatory Commission (FERC).

Areas in the western United States are also beginning to join organized markets. Differences in design exist due to regional characteristics and stakeholder processes, but most continue to converge to a common set of design features: locational prices based on marginal costs, bid-based security-constrained economic dispatch, and day-ahead and real-time auctions for energy co-optimized with ancillary services for common grid services. A question that often comes up is whether these market designs are sufficient for systems dominated by resources lacking fuel costs and possessing other unique characteristics or whether substantial changes may be necessary to ensure economic efficiency and reliability.

States, utilities, and companies have introduced mandates or goals to supply 100% of energy by renewable resources or nonemitting resources. (See Figure 1.) As of early May 2020, 16 states have adopted 100% clean/renewable mandates or targets, and more have adopted less-stringent goals. Finally, many of the organized markets are already experiencing high levels of instantaneous amounts of variable renewable energy (VRE), such as wind and solar. These experiences demonstrate that studying power systems with 100% zero-fuel-cost supply is not an academic exercise. Efficiently designed electricity markets can enable solutions to meet these goals while providing affordable and reliable electricity to consumers.

In this article, the authors discuss some key challenges and potential options for designing electricity markets when the supply fleet lacks fuel costs. This includes the transition to meet these goals as well as the designs incentivizing the investment in and operation of the future supply fleet. Before describing potential future designs, it is important to highlight current efforts to overcome challenges and improve market designs.

Key Questions Facing Market Designers

With decarbonization goals, the future supply fleet may look quite different from the current one. It may consist of substantial amounts of VRE and hydropower, other enabling technologies like short-term or seasonal electric storage, greater levels of responsive demand, and local resources of numerous technology types (either on the distribution system or customer sited). It may also consist of other low-carbon resources like nuclear power and some remaining efficient thermal plants. Except for some remaining fuel-burning technologies, the future and current supply fleets will have something in common: variable operating costs that are not dependent on fuel costs.

At the core of any future scenario is VRE. VRE has several unique characteristics that are important to consider, given the quantities of VRE that may be present in these scenarios. VRE production depends on the weather, meaning that the available energy changes across time and cannot be predicted with perfect accuracy. VRE also has other unique technical characteristics, such as its inverter-based interface. Finally, because VRE depends on the weather for production, it has essentially zero variable costs, with most of its costs tied to capital. Each of these characteristics may influence future electricity market outcomes in different ways.

Similar to other commodities, wholesale electricity prices indicate when supply is limited and demand reduction is most valuable (high-price periods) or when supply is abundant and increased demand can be met with little additional cost (low-price periods). These wholesale pricing signals provide a coordinating role across various decisions, both for short-run operational decisions and long-run investment and retirement decisions. In recent years, the dominant driver of annual changes in average wholesale electricity prices has been natural gas prices, as natural gas generators have been the predominant price-setting technology. Over the past decade, the boom in U.S. shale gas production has driven prices well below their historical averages.

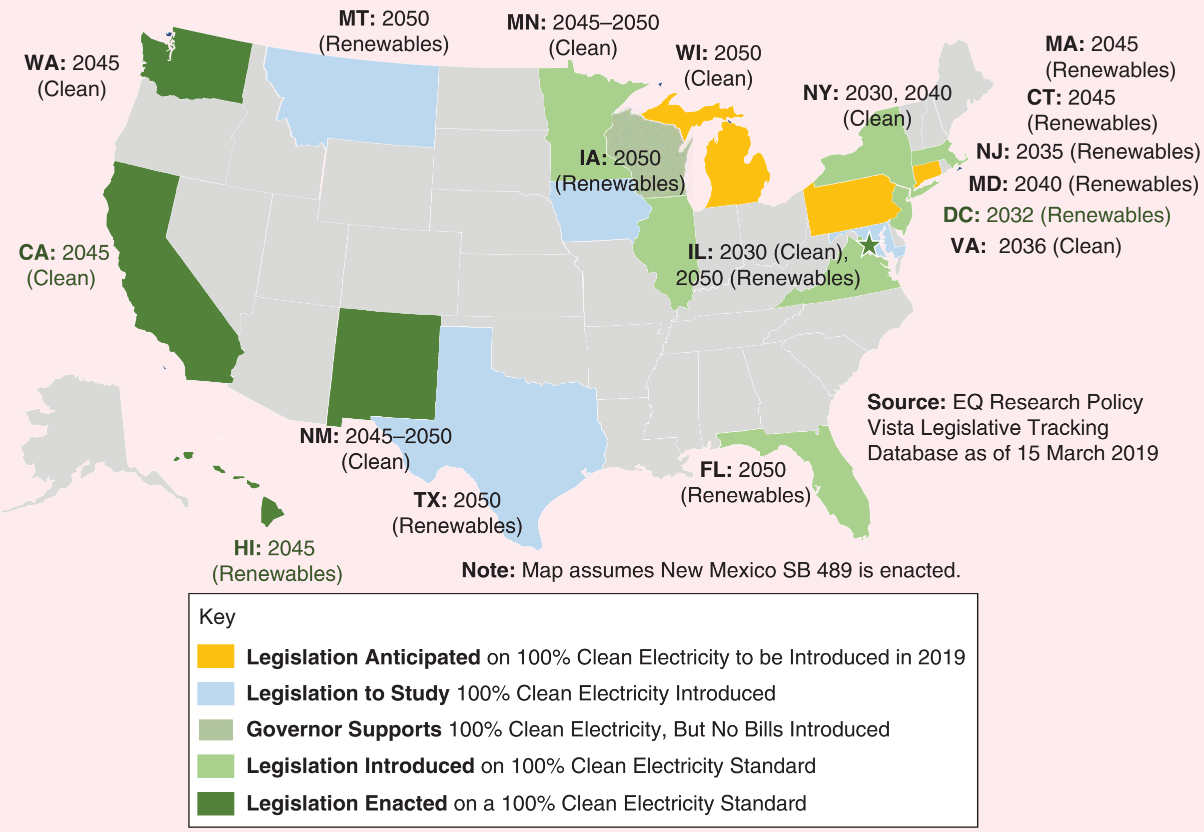
Growth in VRE is starting to have noticeable effects on wholesale electricity prices. VRE's lack of fuel costs pushes the supply curve out during periods of high VRE production. Without corresponding growth in demand or the retirement of surplus capacity, this results in the merit-order effect, that is, lower electricity prices. It also can lead to more variable prices across time and space as well as impacts on the prices of ancillary services, depending on conditions. However, the true impact on prices is not always simple to understand or predict.

The impacts of increased solar production on price patterns are obvious in the California ISO (CAISO), where solar produced more than 18% of annual demand in 2019. This has contributed to lower prices during midday, particularly in spring, but also pushes high-priced periods into the early evening after sunset. Thermal resources that are decommitted during midday may find it more difficult to supply energy after sunset because of commitment constraints. These temporal patterns and variability effects of prices can incentivize increased flexibility from both the supply and demand sides. The springtime supply abundance can also impact the ability to provide downward reserves from resources that are required to be online and generating above a minimum level. Thus, reductions in energy prices can simultaneously occur with increases in downward reserve prices (Figure 2).

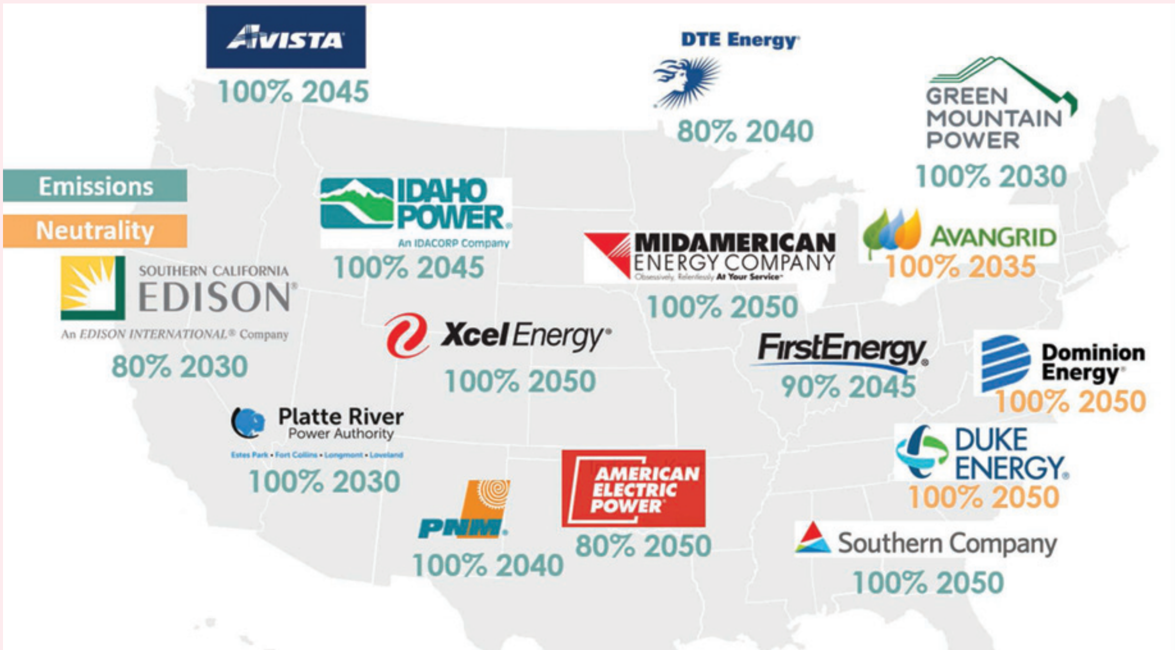
Lawrence Berkeley National Labs has performed several simulations of market prices in futures with higher VRE penetrations for various U.S. regions, which show similar trends as the historic declining price impacts. Higher VRE levels were observed to lower average energy prices, increase price variability, increase the frequency of zero-energy prices, and increase prices for ancillary services. A variety of other studies have shown a range of wholesale price impacts from VRE, using a variety of different assumptions affecting the results (Table 1). The range in values demonstrates the difficulty in trying to predict this impact.

Although these simulations show a reduction in average energy prices due to increased VRE, this may not necessarily be the case on future systems approaching 100% renewable energy. Several assumptions in these studies may not always hold in practice. It is not clear that wholesale prices will simply decline, as observed in studies. This may depend on many factors, such as

- ✔ the market structure, including compensation and investment incentives beyond energy markets (e.g., capacity markets)
- ✔ exogenous planning reserve margins
- ✔ outside policies influencing investment
- ✔ responsiveness of demand to price
- ✔ the existence and settings of administrative shortage pricing
- ✔ VRE locations and the correlation of production



(a)



(b)

figure 1. U.S. goals for (a) states and (b) utilities (representative examples). (Continued)

- ✔ the cost-effectiveness and ability of enabling technologies that mitigate temporal supply and demand concentrations.

Other questions have piqued the interest of market designers. If electricity markets do not incentivize resources for supplying emission-free energy, how would the resource mix transition to the scenarios being discussed? If the transition is supported outside of electricity markets, how will this impact optimal solutions? If the variability of VRE increases the need for flexibility, will the markets incentivize those attributes? How important will future ancillary service markets be? Will VRE variability increase price volatility, or will enabling resources take advantage and reduce the variability? Will VRE forecast errors cause greater uncertainty of prices and divergence between day-ahead and real-time markets? Will unit commitment procedures be necessary, or, if not, how must the market clearing models be enhanced? How will transmission flows and congestion price hedging be impacted if increased variability is present? How will growing amounts of small resources, either residentially owned or located on distribution systems, compete in wholesale markets? Finally, will wholesale and retail designs enable consumers to react to prices in meaningful ways? These are some of the many questions that market designers must consider when evaluating how markets may evolve to allow for an economic, reliable, and environmentally responsible electric power system.

What Does a Future Market Design Need to Do?

Today, marginal cost pricing provides several benefits. It incentivizes resources to use low-cost fuels and improve heat-rate efficiency to generate more energy per unit of fuel consumed. It also provides rents for resources that are inframarginal. Locational pricing motivates suppliers to build in areas with the highest value. Ancillary service co-optimization prompts resources to provide services that are most valuable to the grid at the least cost. Lastly, designs such as shortage pricing primarily incentivize resources to provide energy and services at critical time periods. Many of these attributes will remain important in the future system, but some may be less significant. For example, there may not be fuel to procure

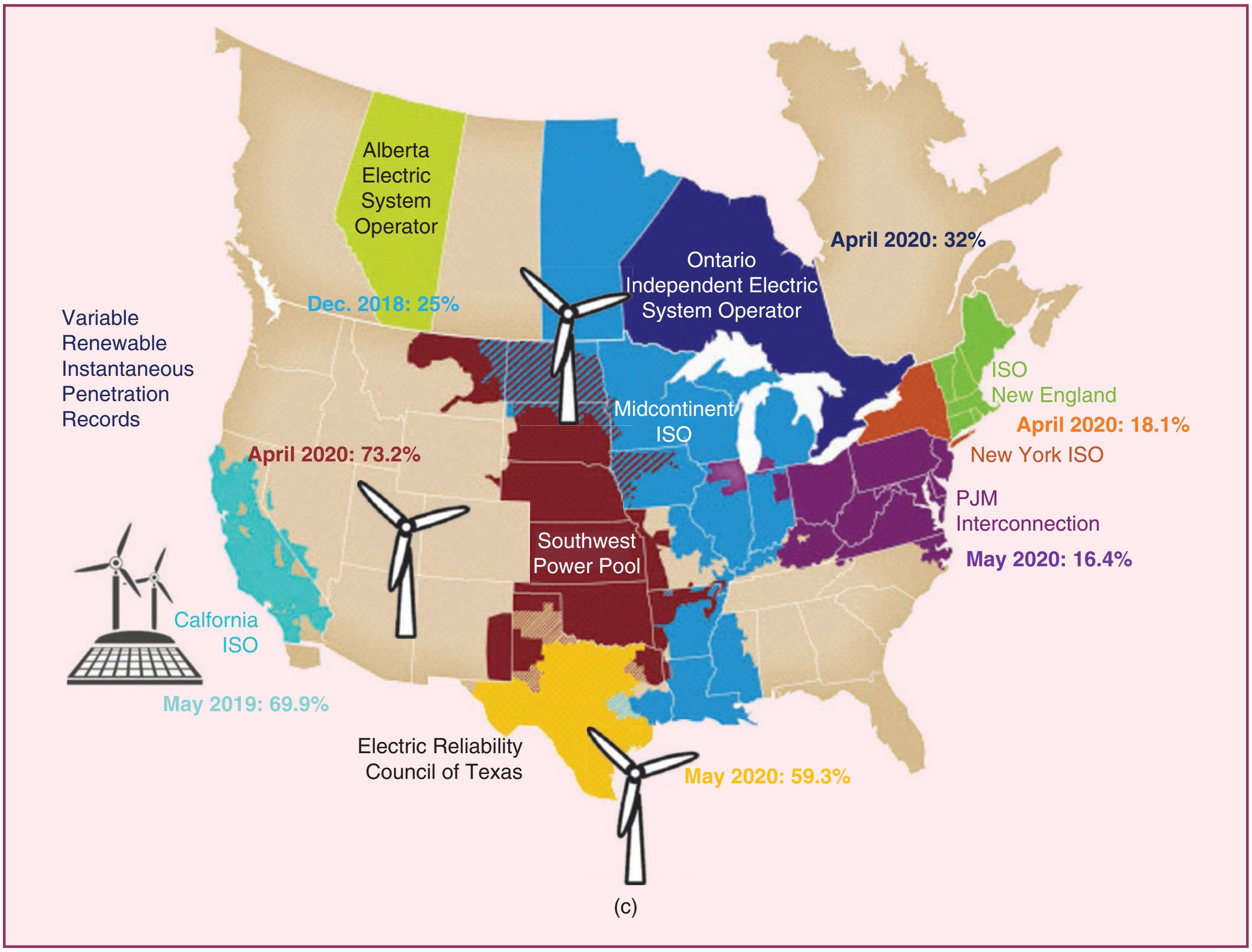


figure 1. (Continued) U.S. goals for (c) The most recent instantaneous VRE penetration records. Data are accurate as of September 2020. (Source: Electric Power Research Institute and Energy Systems Integration Group; used with permission.)

nor heat rates to make efficient. Unit commitment costs may be negligible, and installed capacity may not be the primary attribute signifying supply adequacy. Many participants may bid into the market based not on fuel costs but on opportunity costs. For example, the opportunity cost of demand-side resources is based on forgoing or shifting consumption, and the opportunity cost of energy-limited resources is based on the potential lost profit if energy produced cannot be sold later due to lack of energy stored. Moreover, price signals may be needed to incentivize attributes or behaviors that are abundant today but may become crucial and in short supply in the future.

At the onset of electricity sector restructuring, market designers considered what markets should be signaling. This is as important now as it was then, and while the resource mix is changing dramatically, the principles are mostly unchanged. A research team led by Energy Innovation, an energy policy research firm, recently evaluated possible options for future electricity market designs. The team established 10 key principles for wholesale electricity markets, which ensure economic efficiency, reliability, and technology neutrality. The Energy Systems Integration Group (ESIG) also held a workshop "toward 100%," with six important tracks on key challenges. One of those tracks was on future markets, where the participants discussed key challenges and potential strategies. Numerous ISOs and regional transmission organizations in North America have also authored studies and reports looking at future resource mixes with very high VRE, including New York Independent System Operator (NYISO), Southwest Power Pool (SPP), and others.

The Energy Innovation team established that wholesale electricity markets should do the following:

1) accommodate rapid decarbonization, providing opportunities for the participation of zero-carbon resources

figure 2. Trends in the (a) net load, (b) energy prices, and (c) downward regulation reserve prices in the CAISO, which reflect the impact of increased solar.

2) support grid reliability so that the incremental costs of reliability do not exceed 1) the amount customers would knowingly be willing to pay or 2) incremental benefits
3) promote short-run efficiency through the optimized dispatch of the lowest-cost resource mix
4) facilitate demand-side participation and grid flexibility
5) promote long-run efficiency, including efficient, competitive entry into and exit from the market, under conditions of significant uncertainty
6) minimize the exercise of market power and manipulation
7) minimize the potential for distortions and interventions that would prevent or limit markets' ability to achieve efficient outcomes, consistent with the public interest
8) enable the adequate financing of resources needed to deliver cost-effective reliability based on the efficient allocation of risk (i.e., those that can best mitigate risk should bear it), preventing customers from bearing the cost of poor investment decisions
9) be capable of integrating new technology as needs evolve, adapting as technology changes
10) have designs that are readily and realistically implementable.

Three broad philosophies stem from these principles. First, real-time prices should indicate reliability needs and incremental changes in supply and demand in the most granular way possible. Second, a market must transform physical system risk shared by all into fiscal risk shared out proportionally (no free riders/no market manipulation). Energy markets should aim to be able to manage as many situations as possible by raising or lowering prices. Finally, the market should be investable. The market must provide sufficient revenue to attract investment in assets that improve reliability or economic efficiency and promote the orderly retirement of costly, inefficient resources that are no longer needed.

The markets track of the ESIG workshop developed 20 questions requiring further examination. The track discussed key challenges and explored two exercises to see whether the solutions differ: designing markets for 1) a system that is 100% renewable and whose market could be designed from scratch or 2) a system still in the process of transitioning to 100% renewable. Topics ranged from how and what reliability services need to be incentivized to how an optimal resource mix can be attained. Price-responsive demand was a key enabler in all the discussions. With it, the group found the challenges easier to address, but without it, the challenges were difficult to overcome. Given the lack of fuel costs, the following behaviors and attributes, which may need signals to incentivise them, were highlighted:

- ✔ reducing fixed, capital, and operations and maintenance costs
- ✔ locating resources where they provide value and with the least overall cost, including infrastructure
- ✔ locating resources where they can provide the most energy without severely impacting reliability
- ✔ reducing the negative effects of forecast errors
- ✔ providing the most important reliability services at times when they are most needed
- ✔ transferring energy from times of ample supply to periods where supply is needed
- ✔ consuming energy at times when the cost to do so is acceptable and reducing consumption when it is not.

table 1. The change in average wholesale electricity price and the VRE penetration increase for several recent studies in the United States.

Study	Market Region	Change in Price (US$/MWh) per Percentage Increase in VRE Penetration
Brancucci Martinez-Anido et al. (2016)	ISO New England	$−0.15
Deetjen et al. (2016)	ERCOT	$−0.25
EnerNex (2010)	Eastern Interconnection	$−0.45
Fagan et al. (2012)	Midcontinent ISO	$−0.28
GE Energy (2014)	PJM	$−0.50
LCG Consulting (2016)	ERCOT	$−0.52
Levin and Botterud (2015)	ERCOT	$−0.41
Mills and Wiser (2012)—Solar	CAISO	$−0.13
Mills and Wiser (2012)—Wind	CAISO	$−0.10
New England States Commission on Electricity (2017)	ISO New England	$−0.80
New York ISO (2010)	NYISO	$−0.45

ERCOT: Electric Reliability Council of Texas.

Market Operators Are Adapting Now to Prepare for the Transition

In the United States and Canada, many organized markets are already observing high VRE levels. In the early morning of 27 April 2020, the SPP reached more than 73% of its instantaneous power provided by VRE. Many regions also have the challenge of market designs that must harmonize with policy decisions made outside of the market. Thus, market operators across the continent have been facilitating design changes to enable new technologies to participate, ensure reliability in the face of emerging challenges, and provide signals that lead to optimal operation and investment of the supply fleet.

95% Zero-Carbon Energy Production in Ontario

In the span of 10 years, Ontario transformed its supply mix to produe 95% of its energy carbon free by phasing out coal resources from 2005 through 2014. The existing capacity from nuclear and hydro largely remained, while coal was replaced by natural gas, wind, and solar. The electricity market operated fundamentally in the same manner throughout this transformation using bid-based economic dispatch. Transmission-connected VRE was required to be dispatchable and bid into the market. As more VRE capacity connected to the system, electricity prices trended lower, demonstrating the merit-order effect, as discussed previously (see Figure 3). Recent prices have been, on average, one-third of the levels cleared from the market before the start of phasing out coal.

This transition has not resulted in binary market pricing outcomes of either zero dollars when VRE is marginal or a large spike when natural gas resources are. Ontario has a unique combination of zero-carbon resources in its hydro fleet that, while having no fuel costs, does have other imposed variable costs due to its dependency on water conditions. These prices range from negative values, representing costs incurred with nonproduction, to hundreds of dollars per megawatt, indicating limited water availability and opportunity costs. This is in comparison to natural gas fuel costs ranging in the tens of dollars to produce a megawatt hour of energy. The hydro portion of the supply curve has been preserved but shifted by the addition of VRE. For Ontario, the market design in place has been sufficient to efficiently dispatch the new supply mix. The variability and uncertainty of VRE resulted in greater price volatility, demonstrating the increased need for system flexibility and other products.

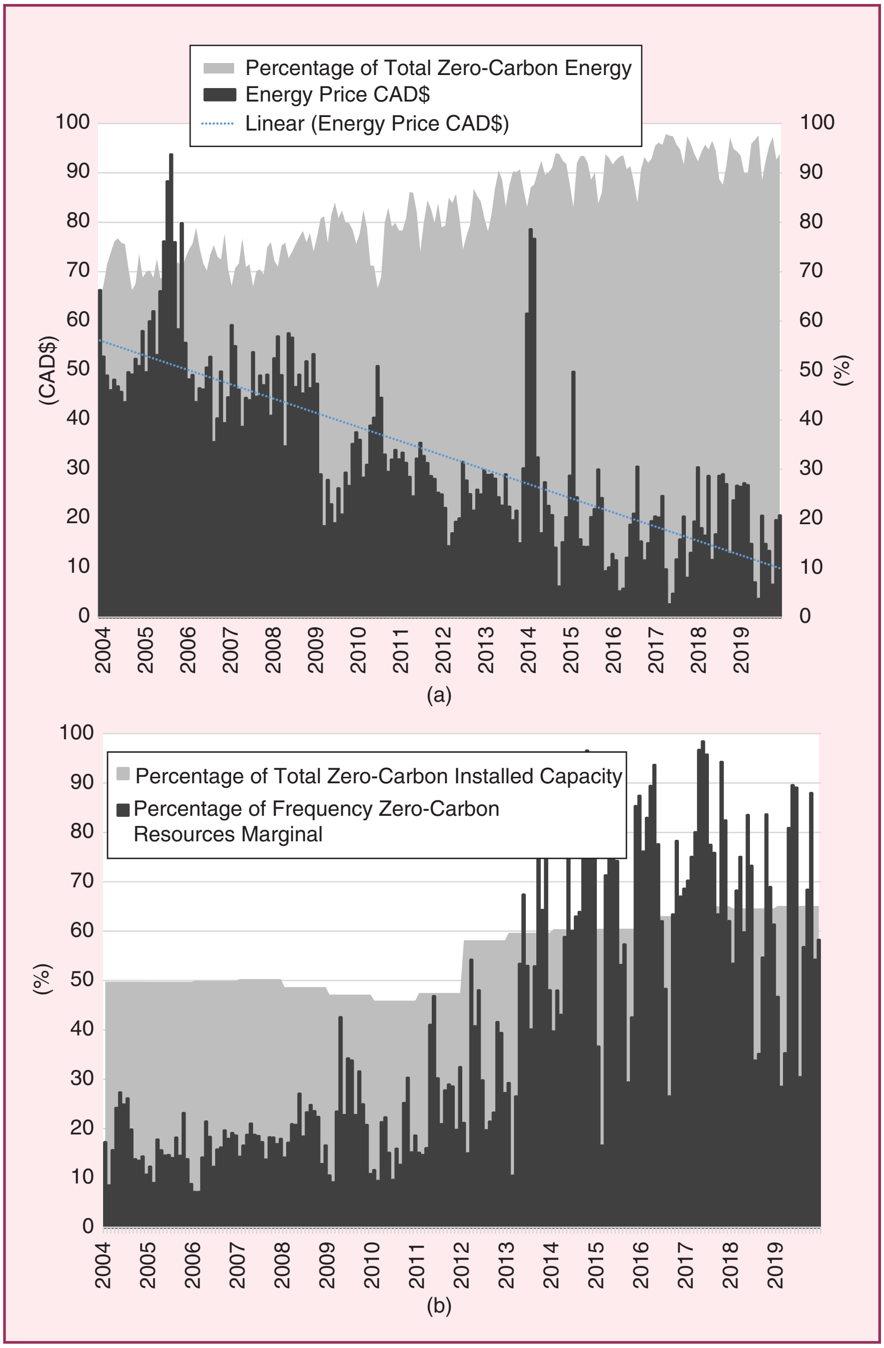


figure 3. (a) The merit-order effect of lowered average energy prices as carbon-free resources with VRE have increased. (b) How increased VRE resulted in more volatile energy prices, as VRE output is dependent on weather conditions that can change rapidly.

Some Regions Are Evaluating Putting a Price on Carbon

There is a strong potential synergy between wholesale electricity markets and renewable technology targets. Applying a price to carbon dioxide emissions in wholesale electricity

Leakage refers to a situation in which there are shifts in generation and emissions from resources subject to a carbon price to higher-emitting resources that are not.

markets would help send efficient price signals to market participants about the value of clean energy resources and align electric systems with decarbonization goals. It may also accelerate the transition to a clean energy future by directly incentivizing new entry of low-carbon resources in locations where they would displace the most carbon dioxide emissions.

In New York, there has been interest in pricing carbon dioxide emissions in addition to the state's current participation in the Regional Greenhouse Gas Initiative (RGGI), a cooperative effort among 10 Northeastern states to cap and reduce power sector carbon dioxide emissions. NYISO, in conjunction with its stakeholders, developed a design where the state sets a social cost of carbon as a price per ton of carbon dioxide emitted based on state goals and the environmental impact. The emitting generators pay for the carbon dioxide they release into the atmosphere. Participants receive economic incentives to invest in low-carbon technologies, and existing participants receive incentives to reduce their carbon emissions. The revenue collected from emitting resources is then returned to wholesale customers. The design also addresses emissions-leakage concerns with neighboring states. *Leakage* refers to a situation in which there are shifts in generation and emissions from resources subject to a carbon price to higher-emitting resources that are not. Leakage can hinder emissions-reduction policies when energy from higher-emitting resources outside of the carbon pricing region displaces efficient, lower-emitting resources within the region. Also, unmitigated leakage can potentially impact investment decisions and consumer costs throughout the system. The way the NYISO design alleviates these concerns is by charging a price at New York's electrical border that does not reflect the carbon price so that a cheaper, higher-emitting supply would not gain market share.

An analysis of the proposal has shown that it would

- ✔ reduce the consumer cost of reaching the state's goal of 100% carbon-free emissions by 2040
- ✔ help grow investment and innovation in clean energy generation
- ✔ promote innovation and efficiency in fossil fuel technology
- ✔ improve public health by encouraging retirement of the highest-emitting generators.

The design has proceeded through the NYISO stakeholder process and now awaits support from New York State. If supported by the state and approved by stakeholders, the NYISO Board of Directors, and FERC, carbon pricing in New York would be implemented (see Figure 4).

PJM Interconnection, which operates in 13 states and the District of Columbia, has a unique challenge related to the diverse range of emission policy initiatives across the region it serves. Three states currently participate in RGGI, while two others have taken steps to join (see Figure 5). Eleven states have renewable portfolio standards or goals employing renewable energy credits, and four states have or are investigating providing subsidies to a broader subset of zero-emitting generation.

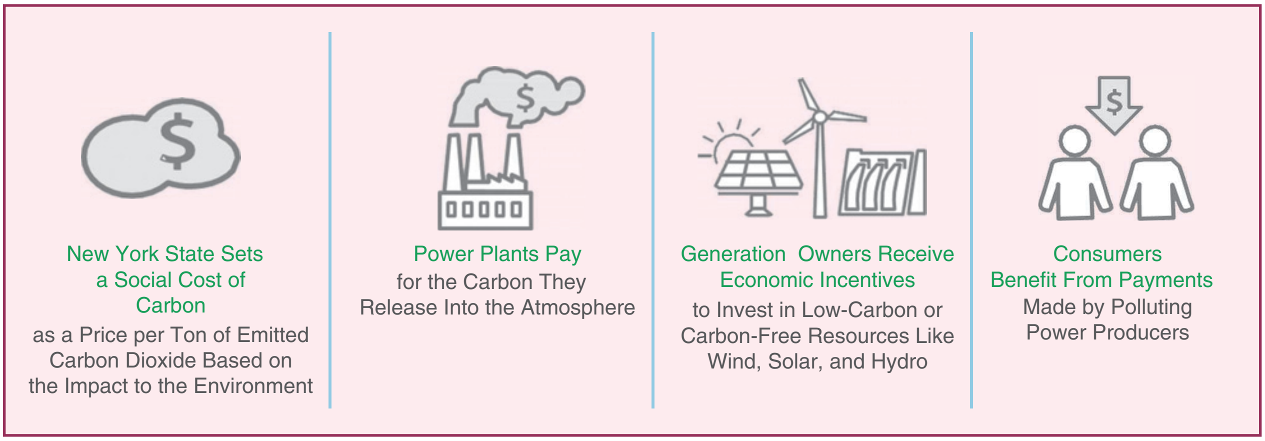


figure 4. An overview of NYISO carbon pricing design.

As a result of participation in RGGI, generation in several states is subject to a carbon price. When implementing a carbon price on a subregional basis, the carbon price can have an impact on both emissions levels and energy prices throughout the system through leakage, as described previously. In July 2019, the Carbon Pricing Senior Task Force was formed as part of the PJM stakeholder process to investigate leakage-mitigation approaches. Both one- and two-way border adjustment approaches were explored to mitigate leakage between the states that participate in RGGI and those that do not. A one-way border adjustment approach adjusts the price of transfers into a subregion subject to carbon pricing to account for the carbon price, while a two-way border adjustment approach also adjusts the price of transfers out of a subregion subject to carbon pricing to remove the impact of the carbon price.

The Evolving Challenge of Determining Which Reliability Services Are Essential

Customers measure the reliability of their electric service simply by whether electricity is available when they need it to be. Grid operators fulfill this need using several types of services or products that procure attributes to support the delivery of energy, thus supporting electric reliability. North American markets have several common design features for these reliability services. Nearly all areas are either currently or are planning to co-optimize ancillary services with energy production, use a cascading hierarchy to assign the highest quality services with the highest prices, and use shortage pricing when there is an insufficient supply of services. The names and existence of different services vary across different regions, and the emphasis on different types of services is evolving (see Figure 6).

A few of the services in Figure 6, secondary and tertiary contingency reserve and regulating reserve, are specific products in all the market regions in North America. Others, like flexibility reserve and primary contingency reserve, are products for a subset of regions. Still others, like inertia, do not have specific products in any region. The reasons for the differences include how soon the need for a certain product has arisen, the existing requirements for other products, and stakeholder prioritization processes.

Resource adequacy refers to having sufficient capacity installed to meet long-term reliability targets. With a traditional generating fleet, if there is enough installed capacity to serve the peak demand, there should also be sufficient capacity for all other times. However, as the fleet moves toward more VRE and enabling technologies, the task of ensuring resource adequacy changes markedly due to the inherent uncertainty and temporal nature of these technologies.

As more VRE is integrated, the quantity and type of services procured to maintain reliability may change to account for its variability and uncertainty. Regions have implemented or proposed changes to 1) increase the reserve requirement to account for needs beyond contingencies, 2) improve the locational scheduling of reserve so that it can be delivered to where it is needed, 3) address shortage pricing to reflect the importance of different services, and 4) utilize demand curves such that the market places value on procuring more reserve than the minimum requirement. In a growing number of market regions, new products have been introduced (the flexibility/following reserve in Figure 6), highlighting the key difference between the characteristics of services related to addressing variability and uncertainty from those that are needed to address contingencies.

Another key area of evolution in reliability service markets is the mix of resources that participate. In regions with large levels of VRE, operators have shifted to separating products into up (increase supply, reduce demand) and down (decrease supply, increase demand) services, creating opportunities for nonconventional resources to supply the service while still ensuring reliability. To benefit the reliable operation of the system, reserve services must also be deliverable and not awarded to resources unable to respond due to transmission congestion. Much of North America has been making significant changes to allow for the participation of electric storage, demand response resources, and even resources located on the distribution system within all the

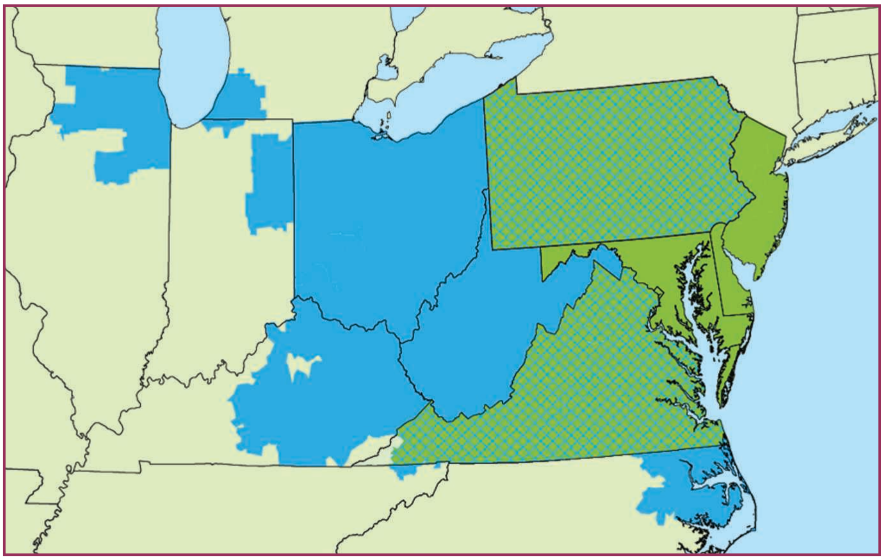


figure 5. PJM Interconnection's footprint (shaded in green, checkered green, and blue) with the states that currently participate in RGGI in green, the states that have taken steps to join in checkered green, and the regions that do not participate in blue.

different reliability services in a cost-effective and high-performing manner.

Characteristics such as the ability to maintain nominal voltages, respond to frequency excursions, and ensure stability are all necessary for grid reliability. Historically, generators have inherently provided these services, but as more inverter-based resources are integrated, these attributes are becoming more important. Different technologies may provide the attributes in different ways, making it essential when designing markets to incentivize the attribute provided to the grid and not how the specific technology provides it today. As an example, sufficient synchronous inertia is required to maintain stability, and inertia markets are under discussion as a future possibility. Although not the exact same thing, future systems with extremely fast controls coming from inverter-based resources can replicate some of the support that inertia provides. Research has even been conducted on ways that grid-forming inverter technology can work without any synchronous inertia. As the resource mix continues to evolve, it will be important to understand the types and amounts of these attributes and the corresponding products necessary to support the reliable operation of the grid, regardless of technology.

What Will the Future Market or Regulatory Structure Look Like?

Market structures differ not only across the globe but also within North America. For example, of the nine organized electricity markets, four have centralized capacity markets (one voluntary), one is transitioning to a centralized capacity market, two have bilateral resource adequacy requirements, and two others have no resource adequacy requirement. It is impossible to predict what the future structure will look like, and it is possible that different regions will differ in regulatory practices, carbon/renewable goals, and stakeholder and consumer opinions, among other features. That said, researchers and practitioners have started looking at a few structures and market designs that can meet some of the principles discussed previously.

There are generally three schools of thought regarding the future electricity market structure. First, existing market designs will function just as well as they do today, or with minor incremental changes. Second, substantial changes are required for markets to function properly, given the future resource mix. Third, markets should be eliminated or minimized in favor of a return to central planning, vertical integration, and cost-of-service pricing. There are several variations of the actual market design across each option, and readers are encouraged to review the reports referenced in the "For Further Reading" section. Some of the options that have been proposed in previous studies are briefly discussed, some by the authors, without claiming any one option is superior to another.

The market design philosophy that underpins North American wholesale energy markets—marginal cost pricing using bid-based, security-constrained economic dispatch

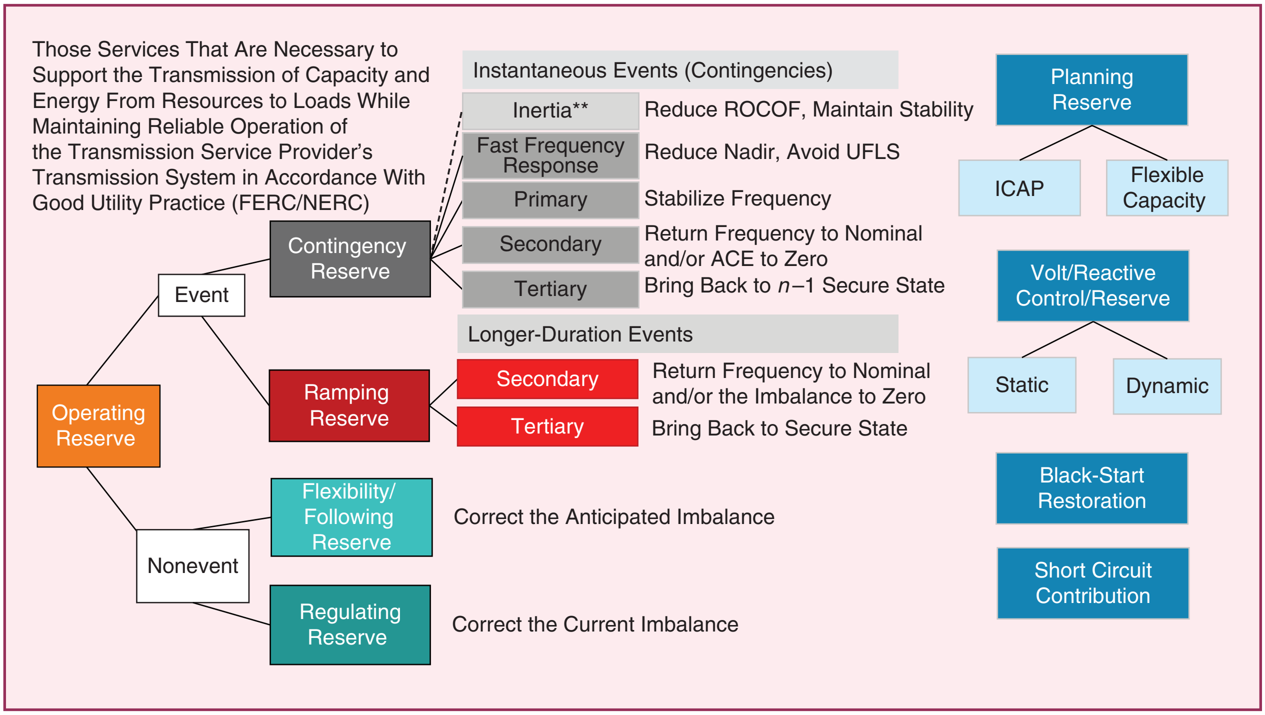


figure 6. An example set and categorization of reliability services. NERC: North American Electric Reliability Corporation; ROCOF: rate of change of frequency; ULFS: under-frequency load shielding; ICAP: installed capacity. (Source: Electric Power Research Institute; used with permission.)

with locational pricing—was conceived to deliver reliability efficiently, regardless of the mix of resources or their short-run production cost profile. That market design, well applied in practice, may be relied upon to perform those tasks in a low-carbon power system. Proposers of this option suggest that this structure, with economic dispatch as its bedrock, along with active decentralized forward procurements between buyers and sellers, could lead to an adequate supply mix with short-term signals to ensure that the system maintains reliability. Some proposers suggest additional measures (e.g., setting minimum financial standards for retail service providers) to support needed liquidity in bilateral trade in long-term options, through which generators and load-serving entities can mitigate their risks in the spot market, while others do not think that is necessary. Otherwise, the proposers suggest that energy market pricing focused on operational needs, combined with endogenizing the value of carbon dioxide emissions, could provide sufficient revenue to meet operational reliability and investment needs. This type of structure can also improve the participation of responsive demand that could respond by consuming less when there is insufficient VRE or provide opportunities for storage that might sell energy at the opportunity cost of being unable to sell later during a critical, high-price period.

A second option discussed is to pair the existing energy market with some type of organized forward market. The key difference in these proposals, compared to the previous ones, is the notion that energy markets alone may not sustain efficient investment. This is due to the possibility that short-term energy prices may have greater volatility and may not average to long-term marginal costs as well as the uncertain prospects of capital recovery of infrequently used assets. The proposers suggest that the energy market alone could get investment right, but that we must also consider the risks that this may not happen. This may be particularly challenging for the set of enabling resources providing flexibility and additional reliability services during low VRE production. The quantity of the enabling resources must be large enough to buy energy when VRE

table 2. Corneli et al. provide a common set of considerations for long-term market design, but each has their own proposals, with the key differences shown.

Key Features	Configuration Market (Corneli)	Long-Term Energy Market (Pierpont)	Firm Market (Gimon)
• How is a long-term market portfolio selected?	Bid-based, region-wide system co-optimization model	Through exogenous guidance from policy makers and system planners	Bid-based, region-wide system co-optimization model
• What is the objective function of the long-term market?	Minimize the expected cost of meeting reliability requirements across a wide variety of possible weather-, load-, and resource-availability scenarios	Minimize the cost of meeting a share of total load, specified by policy makers, from the eligible resources that choose to bid	Minimize the cost of producing a significant share of total energy through a "default dispatch," which short-term markets take as a baseline for real-time operation
• What products are bought in the long-term market?	Capabilities to perform as needed to meet objective functions	Annual energy output, subject to shape; location; resource type; and guidance from policy makers	Long-term energy schedules
• How is fixed cost recovery carried out for selected resources?	Resources selected are eligible for fixed-cost recovery through a variety of means, including power purchase agreements, tolls, regulated tariffs, and clearing prices as worked out through additional design work	Long-term power purchase agreements for energy, which may be either pay-as-bid or uniform market clearing price	Pay-as-bid long-term power purchase agreements
• Is participation mandatory?	No	No	Participation is presumed but not required
• How often is the long-term market conducted, and how much does it purchase?	Periodically, e.g., once every three to five years	Annually	Periodically to cover incremental amounts of needed resources
• Does the long-term market drive rapid decarbonization and how?	Where co-optimized clean energy resources are cheapest, the market will naturally select decarbonizing choices but will otherwise reflect carbon prices and efficient policies	Presumably, both through clean energy resources becoming increasingly competitive and through policies	Where co-optimized clean energy resources are the cheapest, the market will naturally select decarbonizing choices but will otherwise reflect carbon prices and efficient policies

production is excessive (to raise prices) but small enough not to consistently exceed energy needs during scarcity conditions. Stochastic simulation tools could be used to determine the optimal set of resources with the needed attributes while being able to support investment where short-term prices may be too uncertain to make those decisions. Although these options are reminiscent of existing forward-capacity markets and transmission planning processes, proposers suggest voluntary participation and a focus on the incentivizing attributes needed in the future resource mix while primarily relying on short-term energy markets. Three possible options for long-term forward markets are shown in Table 2.

Other options are possible. A recent set of awards was provided through the U.S. Department of Energy's Advanced Research Program, which were aimed at evolving system operations and electricity market operations to a more risk-driven paradigm. The projects will propose and develop new operating and market designs that evaluate and structure performance into market incentives, establish transparent risk-assessment methods, leverage existing approaches to quantify and mitigate risk, and identify how resource performance assessment can create new business opportunities to mitigate risk. It is expected that the market and market clearing algorithms will capture uncertainty, allocate the cost of uncertainty to those who cause it, and reward those who mitigate it.

Another option is moving back toward a more regulated system. If the benefits of competition from these future power systems are not realized and monopolies of power supply and reliability services are seen as inevitable, a regulated system may be a feasible option. That does not make things simpler; the way that the system is planned and operated would continue to be just as complex. The decisions, whether made by one entity or multiple parties, should use the same engineering and economic principles for this future resource fleet, with poor decisions still resulting in inefficient or unreliable outcomes.

Conclusions

Electricity markets have always been complex due to their unique physics of electricity supply and delivery. That will continue regardless of the future grid. There is no crystal ball foretelling how best to achieve a system that emits no carbon and how to get there cost-effectively. During this transition, innovations may cause paradigm shifts that require rethinking. Although regions across North America are seeing substantial levels of VRE, conversations about what market structure and design may be most appropriate for each region are just beginning. Further work is needed to evaluate the different options and how they may work across different jurisdictions. To supply the energy and services for this future system, engineering and economic principles are needed to provide the foundation for evaluating which options are best to support a system that is reliable, economically efficient, and allows the needed resources an opportunity to recover their costs and be rewarded for effective innovation.

For Further Reading

S. Aggarwal et al., "Wholesale electricity market design for rapid decarbonization," Energy Innovation, San Francisco, CA, June 2019. [Online]. Available: https://energyinnovation.org/publication/wholesale-electricity-market-design-for-rapid-decarbonization/

A. D. Mills et al., *Impact of Wind, Solar, and Other Factors on Wholesale Power Prices: An Historical Analysis—2008 Through 2017.* Berkeley, CA: Lawrence Berkeley National Laboratory, Nov. 2019. [Online]. Available: https://emp.lbl.gov/publications/impact-wind-solar-and-other-factors

J. Tribulski and S. Coyne, "Carbon Pricing Senior Task Force," PJM.com, 2020. [Online]. Available: https://www.pjm.com/committees-and-groups/task-forces/cpstf.aspx

M. Hogan, "Follow the missing money: Ensuring reliability at least cost to consumers in the transition to a low-carbon power system," *Electricity J.*, vol. 30, no. 1, pp. 55–61, Jan./Feb. 2017. doi: 10.1016/j.tej.2016.12.006. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1040619016302512?via%3Dihub

"Reliability and market considerations for a grid in transition," New York Independent System Operator Rensselaer, Dec. 2019. [Online]. Available: https://www.nyiso.com/documents/20142/9869531/Reliability%20and%20Market%20Considerations%20for%20a%20Grid%20in%20Transition%20-%2020191220%20Final.pdf/7846db9c-9113-a85c-8abf-1a0ffe971967

"Toward 100% renewable energy pathways: Key research needs," Energy Systems Integration Group, Oct. 2019. [Online]. Available: https://www.esig.energy/download/toward-100-renewable-energy-pathways-key-research-needs

Biographies

Erik Ela is with the Electric Power Research Institute, Palo Alto, California, USA.

Andrew Mills is with Lawrence Berkeley National Laboratory, Berkeley, California, USA.

Eric Gimon is with Energy Innovation, San Francisco, California, USA.

Mike Hogan is with Regulatory Assistance Project, Montpelier, Vermont, USA.

Nicole Bouchez is with the New York Independent System Operator, Rensselaer, New York, USA.

Anthony Giacomoni is with PJM Interconnection, Audubon, Pennsylvania, USA.

Hok Ng is with the Independent Electric System Operator, Toronto, Ontario, Canada.

Jim Gonzalez is with Southwest Power Pool, Little Rock, Arkansas, USA.

Mike DeSocio is with the New York Independent System Operator, Rensselaer, New York, USA.


©SHUTTERSTOCK.COM/FOTOIDEE

Decarbonization of Electricity Systems in Europe

Market Design Challenges

By Goran Strbac, Dimitrios Papadaskalopoulos, Nikolaos Chrysanthopoulos, Ana Estanqueiro, Hugo Algarvio, Fernando Lopes, Laurens de Vries, Germán Morales-España, Jos Sijm, Ricardo Hernandez-Serna, Juha Kiviluoma, and Niina Helistö

Digital Object Identifier 10.1109/MPE.2020.3033397
Date of current version: 6 January 2021

DRIVEN BY CLIMATE CHANGE concerns, Europe has taken significant initiatives toward the decarbonization of its energy system. The European Commission (EC) has set targets for 2030 to achieve at least 40% reduction in greenhouse gas emissions with respect to the 1990 baseline level and cover at least 32% of the total energy consumption in the European Union (EU) through renewable energy sources, predominantly wind and solar generation. However, these technologies are inherently characterized by high variability, limited predictability and controllability, and lack of inertia, significantly increasing the balancing requirements of the system with respect to historical levels. The flexibility burden is currently

1540-7977/21©2021IEEE

> The decarbonization of end-demand segments (e.g., heating, transport, and industry) will also require significant investment, and appropriate energy policy initiatives will need to be developed.

carried by flexible fossil-fueled conventional generators (mainly gas), which are required to produce significantly less energy (as low operating cost and CO_2-free renewable and nuclear generation are prioritized in the merit order) and operate part loaded with frequent startup and shut-down cycles, with devastating effects on their cost efficiency.

Furthermore, the decarbonization agenda is also envisaged to affect the demand side, mainly through the electrification of segments of the transport, heating, and cooling sectors that are currently heavily reliant on fossil fuels. However, this electrification is expected to yield a disproportionately higher increase in peak electricity demand levels than the associated increase in the overall electrical energy consumption due to the temporal patterns in the usage of vehicles and heating/cooling appliances. This implies that capital-intensive investments in new generation capacity and network reinforcements will be required, and this new infrastructure will be significantly underutilized. Considering these challenges, as the decarbonization initiatives further develop, the utilization of the generation and network infrastructure is constantly reducing, and the total electricity system costs are dramatically increasing.

Beyond the technical challenges associated with increasing balancing requirements and peak demand levels driven by the decarbonization of the European energy system, there are growing challenges associated with the design of electricity markets. The key market challenges include a) the "merit-order effect" of renewable generation and the resulting "missing money" problem faced by the generation side; b) the integration of variable renewable generation in energy and ancillary services markets; c) the design of effective carbon emissions markets; d) the capture of the full system value of distributed flexibility in energy and balancing markets; and e) the geographical integration of different market segments, including the development of a harmonized pan-European market and the coordination of emerging local energy markets. This article aims at providing evidence of these challenges in the European setting, reviewing European policy initiatives to address them, and identifying open issues toward developing innovative electricity market designs to enable the cost-effective and secure development of a highly decarbonized European electricity system.

The Need for a Radical Change of Electricity Market Design

Beyond setting ambitious carbon reduction targets, the EC remains committed to a deregulated electricity market paradigm, according to which the investment and operation of the generation, demand, and energy storage components are driven by competitive markets encapsulating profit-driven market participants. This implies that both the large-scale integration of low-carbon generation as well as the realization of the system benefits of flexibility resources will have a significant impact on the current market dynamics and require a fundamentally new market design.

The most fundamental feature of this new market design lies in shifting the focus from the operation timescale and the short-run marginal cost of the system toward the investment timescale and the necessary capital investments to support the decarbonization agenda, as qualitatively illustrated in Figure 1. Under a large-scale integration of renewables, the short-run marginal cost and consequently the prices in the energy market will be massively reduced due to the very low (nearly zero) marginal production costs of these resources; this is widely known as the merit-order effect of renewables. On the other hand, the value and prices of ancillary services will be increased by an order of magnitude, mainly due to the higher balancing requirements driven by the variability of renewables.

Despite the controversy around capacity markets, their size is also expected to grow in the next decade due to the increasing need to remunerate and recover the investment costs of conventional generation,

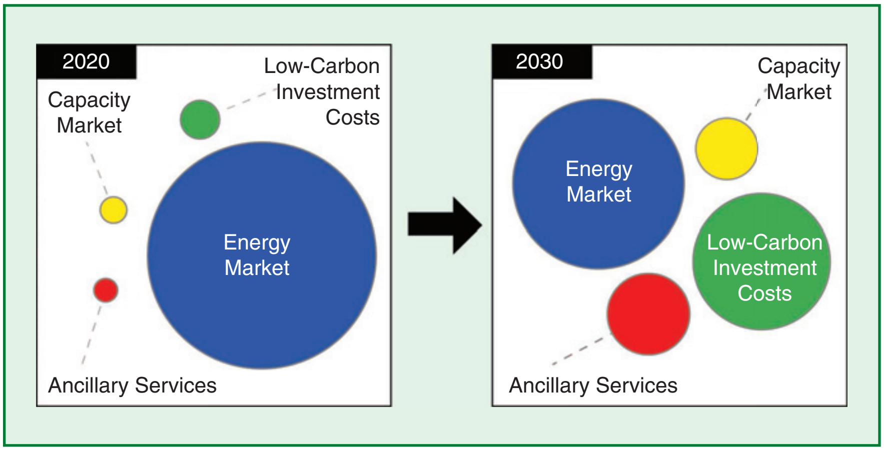


figure 1. A qualitative illustration of market evolution.

preserving the security of supply. More importantly, a large increase in the total low-carbon generation investment costs is expected (despite the reducing unit costs of renewables) to achieve the ambitious carbon reduction targets. Furthermore, the decarbonization of end-demand segments (e.g., heating, transport, and industry) will also require significant investment, and appropriate energy policy initiatives will need to be developed.

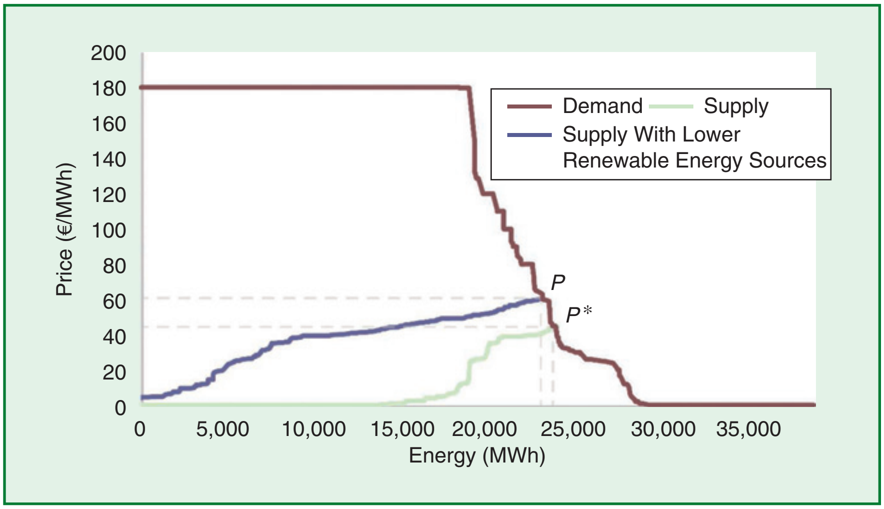


figure 2. An illustration of the merit-order effect on the Portuguese day-ahead market prices.

Merit-Order Effect in the European Electricity System

Various studies have recently investigated the merit-order effect of renewable generation in European markets. Two representative examples are presented in this article. The first one focused on the Portuguese day-ahead prices and was conducted by the National Laboratory of Energy and Geology. The time period of the study ranged from 1 January to 30 June 2016 (a total of 4,368 h). The simulations considered data extracted from the Iberian (Spanish and Portuguese) electricity market and were performed with the agent-based simulation tool, MATREM (for Multi-Agent Trading in Energy Markets). The results indicate that a wind penetration of 28.1% in the Portuguese system yielded an average price reduction of about 17 €/MWh during the first half of 2016. The highest reduction in the study period, about 25 €/MWh, was observed in January, which was a particularly windy month. The merit-order effect for one time period (2 January 2016, 7 p.m.) is illustrated in Figure 2; a larger penetration of renewables shifts the supply curve to the right, thereby reducing the energy prices from P to P*.

The second study focused on the Northern European system (including the Nordic countries, the Baltic countries, Poland, and Germany) and was conducted by the VTT Technical Research Centre of Finland through a combined investment and operation modeling approach linking the Balmorel and WILMAR-JMM models. Specifically, a sensitivity analysis on the share of variable renewable energy (VRE) resources was performed, while the portfolio of conventional generation technologies was optimized considering two different time horizons (2030 and 2050). Figure 3 presents the results of this analysis, where (a) and (b) correspond to a 40% and 60% share of VRE, respectively, while both include the current VRE share in the region (22%) for reference. The merit-order effect is evident in both graphs: increasing the VRE share from the current level to 40% and (especially) to 60% reduces the energy prices substantially. Interestingly enough, the prices in 2050 are higher than in 2030, especially in the 60% VRE share case; this is due to the fact that a large part of the existing baseload thermal generation capacity, although remaining in the system until 2030, is expected to retire before 2050.

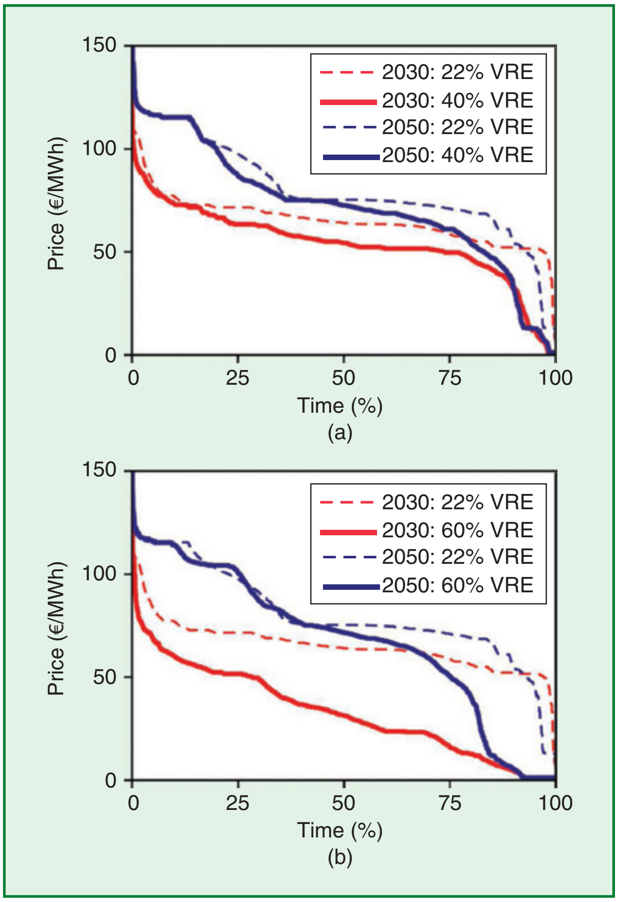


figure 3. Price duration curves in the Northern European system for a 22% VRE share [dashed lines in both (a) and (b)], a 40% VRE share [solid lines in (a)], and a 60% VRE share [solid lines in (b)].

In this context, maintaining the current market design, which focuses on the trading arrangements for energy as a basic commodity, risks creating a scenario in which the large generation and storage players are unable to recover their investment costs and thus are motivated to leave the market. This critical market challenge is usually referred to as the revenue insufficiency or missing money problem and entails the dangers of compromising the security of supply (considering the potential market exit of conventional generators) and/or compromising the carbon reduction targets (considering the potential exit of low-carbon generators).

In the case of conventional generators, recent European market design initiatives have contributed to addressing the missing money problem. First, the design of balancing markets is continuously refined through the introduction of additional balancing products, the harmonization of procurement and activation processes among different countries, and the gradual shift toward a joint energy and reserve market clearing process. These policy changes are expected to enhance the cost reflectivity of balancing markets and increase the associated revenues of balancing providers. Second, following the U.S. paradigm, some European countries (e.g., the United Kingdom, Ireland, Italy, and Poland) have started implementing capacity markets, remunerating participants that can contribute to the required adequacy levels in a cost-efficient fashion through competitive auctions; however, capacity remuneration mechanisms remain controversial and have been characterized as market distortive measures.

Finally, the concept of scarcity pricing has been recently highlighted as a means to resolve the missing money problem: during periods of high demand and scarce supply, the energy price is set at the marginal benefit of the demand side, which is often estimated as the value of lost load. Considering the very high value of this marginal benefit, the activation of scarcity prices during a limited number of periods per year can theoretically secure sufficient revenues for generators to recover their investment costs. In this context, the EC has recently recognized [in Regulation (EU) 2019/943] scarcity pricing as a key feature of the future low-carbon electricity market, with Belgium being the first European country that has decided to implement such a mechanism, which is scheduled to start in late 2021.

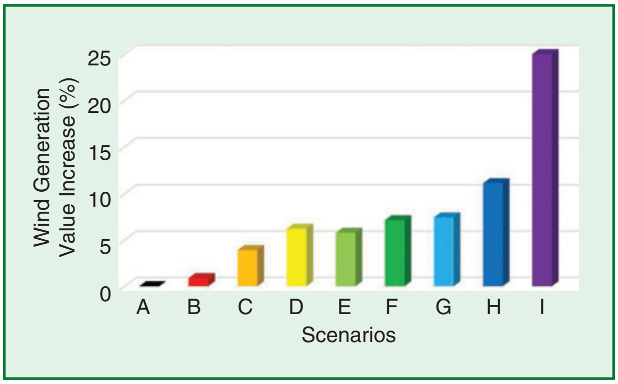


figure 4. An increase of wind generation value in the Iberian market for different scenarios.

Market Participation of Renewable Generation

Not only conventional generators but also renewables face significant challenges in the emerging market environment. First of all, most European countries, with the support of the EC policy framework, are gradually abandoning out-of-market incentive mechanisms (such as feed-in tariffs, green certificates, and long-term contracts for differences) that had been introduced in the 1990s to provide the initial push for investments in renewable generation, on the grounds of fully integrating renewables in the deregulated market environment. Second, given the long gate closure times applied in many European markets, the renewable generators' bids are typically based on 12–36 h ahead forecasts in the day-ahead market, entailing significant forecast errors due to the stochastic nature of renewables' output. As a result, the deviations between forecasted and actual output need to be compensated in intraday and balancing markets, with the latter involving the payment of substantial penalties, which compromise the renewables' market profitability. Finally, variable renewable generation is not generally qualified for participation in capacity markets, considering its inherent inability to provide firm power.

Nevertheless, various measures have been recently proposed to address these challenges and enhance the profitability of renewable generators, including both renewables' operational strategies (e.g., advanced forecasting techniques, aggregation strategies) and new market designs (e.g., postponing gate closure times, shortening market resolution, and allowing participation of renewables in balancing markets). In an effort to quantitatively analyze the effects of such strategies, within the European research program IRPWIND, the normalized value of wind generation (calculated as the difference between its overall market revenue minus its imbalance penalties, divided by the overall energy production) in the Iberian market has been quantified through the agent-based simulation tool MATREM for the following set of scenarios (Figure 4):

- ✔ *Scenario A*: In this reference scenario (with respect to which percentage of wind generation value increase is calculated in the remaining scenarios), the wind generators' market bids are based on deterministic (expected) wind power forecasts.
- ✔ *Scenario B*: The wind generators employ a more advanced, probabilistic quantile-based forecast approach.
- ✔ *Scenario C*: Multiple wind generators within a given control area are aggregated and then participate in the market as a single entity (with a certain degree of power controllability) to limit the overall forecast errors.
- ✔ *Scenario D*: The gate closure time of the day-ahead market is postponed by 2 h (from 12 p.m. to 2 p.m. Central European Time) to take advantage of more accurate forecasts.

- ✔ *Scenario E*: Wind generators are allowed to participate in balancing markets, in line with the current market arrangements in certain European countries (e.g., Spain, Germany, Denmark, and the United Kingdom).
- ✔ *Scenario F*: Beyond the existing markets, wind generators participate in two new balancing markets—the renewable power band market and the energy reserve market—that have been proposed by the IRPWIND program. These markets are similar to the secondary and tertiary reserve markets, respectively, with the difference that their temporal resolution is 15 min (instead of 1 h), and the wind generators can submit bids up to 15 min ahead of real time to enable them to reduce their imbalance payments.
- ✔ *Scenario G*: This scenario is a combination of scenarios B and F, with wind generators participating in the two new markets and employing a probabilistic forecast approach.
- ✔ *Scenario H*: This scenario is a combination of scenarios C and F, with wind generators being aggregated and participating in the two new markets.
- ✔ *Scenario I*: Beyond participating in the two new markets and employing a probabilistic forecast approach, wind generators can participate in the formation of a new type of bilateral contracts, short-term energy contracts. Following the logic of the new markets, these short-term energy contracts are formed in a 15-min temporal resolution and enable wind generators to trade their energy imbalances.

Prioritization of Renewable Generation in Merit-Order Dispatch

Another important market design issue around renewable generation lies in its prioritization in the merit-order dispatch, with most European markets accepting its curtailment only when the technical limits of the system are breached, on the grounds that such curtailment always increases the operating costs and the CO_2 emissions of the system. However, this assumption is not always valid, as demonstrated by the following example. This example involves a wind generator (which is assumed to be able to produce 100 MWh across the considered 4-h operating horizon), a biomass unit, and a conventional natural gas unit (Table 1), which need to supply a total demand of 160 MW at hours $t = 1$ and $t = 2$ and 380 MW at hours $t = 3$ and $t = 4$.

When the wind generator is prioritized in the dispatch and is forced to deliver its maximum possible output at all hours, the resulting optimal generation dispatch, system operating costs, and system CO_2 emissions are as presented in Table 2. Although the biomass generator constitutes the cheapest available conventional unit and has the capacity to cover the remaining demand at all hours, its maximum ramp rate limit does not allow it to cover the demand at $t = 3$ (since demand increases from 160 MW at $t = 2$ to 380 MW at $t = 3$, while the maximum ramp rate of the biomass generator is 120 MW/h). Therefore, the more expensive and polluting gas generator needs to be activated at $t = 3$ to cover the remaining 100 MW of the demand.

On the other hand, when the dispatch prioritization of the wind generator is relaxed, the resulting optimal generation dispatch, system operating costs, and system CO_2 emissions are as presented in Table 3. Although the wind generation exhibits lower operating costs and zero CO_2 emissions, its whole output is curtailed at $t = 2$, enabling the biomass generator to reach a higher output at this hour; the wind generator, along with the

table 1. Generators' data.

	Maximum Power (MW)	Maximum Ramp Rate (MW/h)	Marginal Cost (€/MWh)	Marginal CO_2 Emissions (Metric-Tons/MWh)
Wind	100	—	0	0
Biomass	300	120	31	0
Gas	150	100	70	0.32

table 2. The results with wind dispatch prioritization.

	Power $t = 1$ (MW)	Power $t = 2$ (MW)	Power $t = 3$ (MW)	Power $t = 4$ (MW)	Total Output (MWh)	Operating Cost (€)	Total CO_2 Emissions (Metric-Tons)
Wind	100	100	100	100	400	0	0
Biomass	60	60	180	280	580	17,980	0
Gas	0	0	100	0	100	7,000	32
System	160	160	380	380	1,080	24,980	32

table 3. The results without wind dispatch prioritization.

	Power $t = 1$ (MW)	Power $t = 2$ (MW)	Power $t = 3$ (MW)	Power $t = 4$ (MW)	Total Output (MWh)	Operating Cost (€)	Total CO_2 Emissions (Metric-Tons)
Wind	100	0	100	100	300	0	0
Biomass	60	160	280	280	780	24,180	0
Gas	0	0	0	0	0	0	0
System	160	160	380	380	1,080	24,180	0

biomass generator, subsequently provides the required ramping flexibility at $t = 3$. As a result, there is no need to activate the more expensive and pollutive gas generator at $t = 3$, and thus, although the total wind output is reduced by 25% with respect to the scenario with wind dispatch prioritization, the total operating costs and CO_2 emissions are reduced by 3.2% and 100%, respectively.

This simple example has demonstrated that the strict prioritization of renewable generation in the merit-order dispatch is not always the most effective strategy in terms of both operating costs and CO_2 emissions. Although this particular example is driven by the ramping requirements of electricity systems, a recent study conducted by the Netherlands Organisation for Applied Scientific Research has presented numerous examples where renewable generation flexibility constitutes an effective market strategy in reducing both operating costs and CO_2 emissions. By adopting similar smart curtailment strategies, renewable generation can be transformed from the cause of flexibility problems to part of the solution (such as contributing to ramping requirements in the above example), thus lowering the system flexibility dependency on conventional generation.

Carbon Pricing

Another crucial policy instrument toward incorporating the ambitious emissions reduction targets within the deregulated market environment is the introduction of carbon markets, which effectively penalize the production of emissions and incentivize investment in low-carbon technologies. In Europe, such a market mechanism, the EU Emissions Trading System (EU ETS), was established in 2005 and remains the EU's flagship policy toward a market-based reduction of emissions. The EU ETS is based on cap and trade principles, meaning that a maximum (cap) is set on the total amount of emissions that can be produced by the system (which is reduced over time to gradually achieve the carbon reduction targets), and a certain number of EU emissions allowances covering this cap are then auctioned and can subsequently be traded. Participants emitting greenhouse gases need to purchase sufficient allowances, lest they face significant fines. In electricity markets, given that the carbon allowance price is passed on by fossil-fueled generators in the electricity price, the revenues of low-carbon generators are increased, partially addressing their missing money problem.

The effectiveness of the EU ETS has been demonstrated in practice, with the EU estimating that the emissions from sectors covered by the system have been reduced by 21% in 2020 with respect to the 2005 levels. However, certain questions have arisen around the long-term economic efficiency of this mechanism, particularly regarding the variability of the CO_2 allowances price. Although the gradual reduction of the CO_2 cap should theoretically lead to an increasing CO_2 price over time, in practice, this price has been unstable. After the global financial crisis of 2007–2008, the CO_2 price dropped from around 25 €/ton to as low as 5 €/ton in 2013; after many years, the price exceeded the 20 €/ton level in 2018, but if the current COVID-19 crisis causes a sustained reduction of energy demand, the price may decline again.

This CO_2 price variability creates significant uncertainties and risks for both potential investors in low-carbon technologies as well as electricity consumers. The potential of a very low CO_2 price discourages investments in low-carbon generation, while the potential of a very high CO_2 price implies an undesired increase in the consumers' energy bills and their subsequent resistance to emissions reduction policies. Although a market stability reserve has been recently introduced to address this challenge by adjusting the number of auctioned allowances, its effect on CO_2 prices is indirect and thus uncertain.

In this context, new designs for reducing the price risks of the EU ETS have been lately brought forward, including the introduction of CO_2 price floors and price ceilings (i.e., minimum and maximum CO_2 price limits). A price floor has already been implemented in the United Kingdom and has been announced in The Netherlands. In an effort to analyze the impacts of these CO_2 market designs, Delft University of Technology has conducted a study through the agent-based model EMLab that simulates self-interested companies' generation investment decisions in alternative technologies (e.g., coal, gas, nuclear, carbon capture and storage, and renewables). Figure 5 presents key results of this study, including the (a) emerging CO_2 prices and (b) CO_2 emissions in Europe in different years (x-axis) under alternative CO_2 market designs (Original ETS, Price Floor, and Price Floor and Ceiling); these results include median CO_2 prices and emissions as well as 50%/90% envelopes as Monte Carlo simulations have been carried out to capture the uncertainties around the evolution of demand levels and fuel prices.

Under all market designs, the CO_2 price is relatively low, and the CO_2 emissions are relatively high during the early years due to the higher CO_2 cap. After about 10 years, however, the CO_2 cap becomes stricter, and thus CO_2 prices increase significantly, reaching very high values in scenarios with high demand growth. Consequently, with a delay corresponding to investment lead times, investments in low-carbon technologies start emerging, and CO_2 emissions start dropping. After two investment cycles, the market stabilizes, and emissions decline steadily.

Under the current market design with neither price floors nor ceilings (Original ETS in Figure 5), the CO_2 price variability is immense, particularly with respect to the extremely high prices (reaching an extreme value of 500 €/ton) observed after the first 10 years in scenarios with high demand growth. Under a market design with a price floor, the CO_2 price variability is drastically reduced in terms of avoiding both the very low (nearly zero) levels as well as the very high levels observed under the Original ETS design. As a result, the risks associated with low-carbon investments are reduced, such investments emerge sooner, and CO_2 emissions drop faster. Consequently, when the CO_2 cap becomes stricter, part of the required investments has already taken place, and the CO_2 price remains at lower levels.

Finally, under a market design with a carefully considered price ceiling, very high CO_2 prices and subsequently very high consumer energy bills are avoided while the CO_2 emissions reduction targets are not compromised. In conclusion, the introduction of a price floor and a price ceiling in the EU ETS constitutes an effective way to achieve carbon reduction targets with reduced risks for both low-carbon investments and electricity consumers.

The Role and Value of Flexibility

The new flexibility resources, predominantly energy storage and demand-side response (DSR), play a key role in reducing the costs associated with the transition to a low-carbon energy future. An important part of these flexibility resources corresponds to the large-scale technical solutions, such as bulk, long-duration energy storage, that can deal with extreme events of long periods of low wind and solar generation output and DSR from large industrial/commercial consumers that can flexibly schedule some of their processes. However, in the emerging decentralized and digitalized energy paradigm, another very promising part corresponds to small-scale and distributed forms of flexibility sources at the local distribution level, such as residential smart appliances, smart-charging electric vehicles (EVs) with potential vehicle-to-grid capabilities, distributed generation, and distributed energy storage, including heat storage. These resources are owned by small electricity customers who, enabled by the advancements in digital technologies, are gradually transformed from passive electricity consumers to active prosumers, considering their dual ability to flexibly manage their electricity demand and produce electricity through microgeneration. This paradigm change is reflected in the Clean Energy for All Europeans package recently presented by the EC, which highlights the empowerment of energy end users through the active involvement in energy system operation and planning.

According to a comprehensive study conducted by Imperial College London through an advanced whole electricity system model, the potential cost savings brought by the intelligent coordination of flexibility in the Great Britain system are around £3.8 billion/year in a system meeting the Great Britain benchmark emissions target of 100 gCO_2/kWh in 2030 and around £8 billion/year in a system meeting a more ambitious target of 50 gCO_2/kWh, as illustrated in Figure 6. The components of these cost savings include.

- ✔ savings in operating expenses, by the avoided curtailment of zero-cost renewable generation and the more cost-efficient provision of the required balancing services (operating expenditures)
- ✔ savings in capital expenses associated with reinforcing distribution [D capital expenditures (CAPEX)] transmission (T CAPEX), and interconnection assets (I CAPEX) driven by reduced peak demand levels and the cost-effective management of network constraints
- ✔ savings in capital expenses associated with investments in conventional generation [G CAPEX (conventional)], driven by reduced peak demand levels and reduced requirements for generation flexibility
- ✔ savings in capital expenses associated with investments in low-carbon generation [G CAPEX (low-carbon)] while meeting the carbon target, which is the most dominant benefit in the lower carbon emission scenario of 50 gCO_2/kWh due to the high cost of firm low-carbon generation technologies (i.e., carbon capture and storage and nuclear), driven by the much more efficient utilization of lower-cost variable renewable generation.

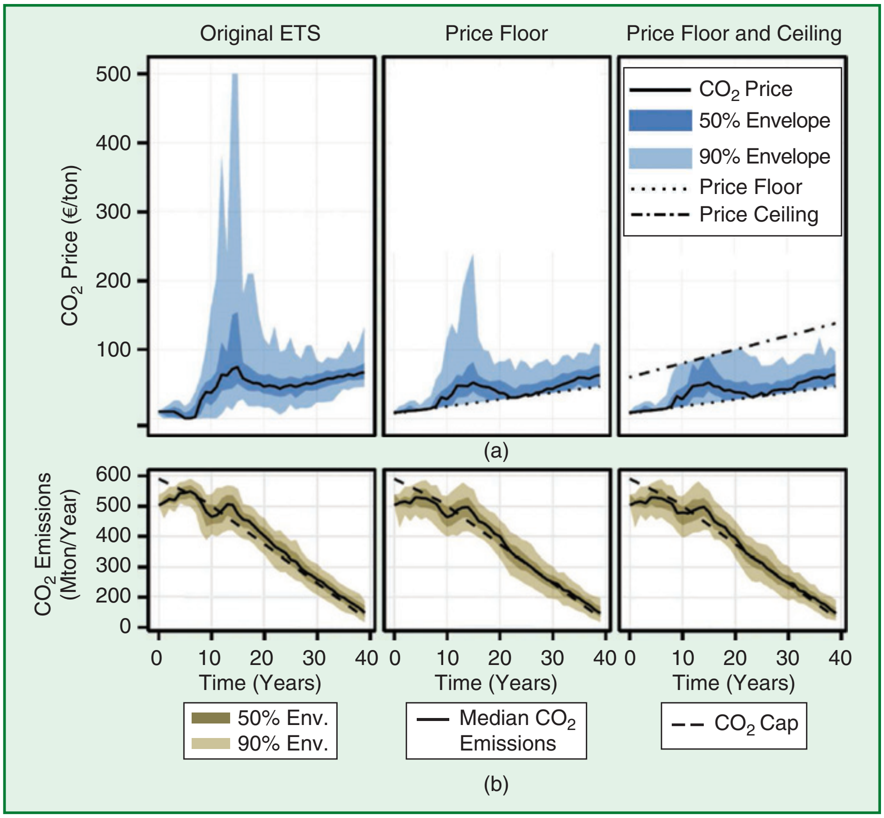


figure 5. The impacts of alternative CO_2 market designs on CO_2 prices and emissions. Env: Envelope; Mton: metric ton. [Courtesy of Jörn Richstein of the German Institute (DIW, Berlin), based on data from Richstein et al.]

The current European market design does not capture the whole spectrum and full extent of the benefits of flexibility resources, thus hindering their further development. In thve energy market segment, most small consumers and prosumers are still facing flat retail tariffs that do not reflect the time-variable value of energy in the system; therefore, they are prevented from activating their flexibility resources to consume energy during periods of abundant renewable generation and/or produce energy during periods of low availability of renewables.

Figure 7 presents the results of a study conducted by Imperial College London, which aimed at quantifying the impacts of domestic demand flexibility (in terms of smart-charging EVs, electric heating with heat storage, and smart wet appliances) on both system operation and the domestic consumers' energy bills in the Great Britain system. Different scenarios have been examined with respect to the percentage of consumers owning the flexibility resources (0, 25, 50, 75, and 100%) and the generation mix (including the current mix in 2020 and the projected mix in 2030). It has been assumed that activation of demand flexibility respects the consumers' service requirements in terms of traveling (for EVs), the indoor temperature (for heating), and the timely completion of the wet appliances' cycles, implying that such demand flexibility does not reduce their overall energy consumption but merely redistributes it in time.

Figure 7(a) demonstrates that demand flexibility can greatly reduce the levels of renewable generation curtailment, especially in the 2030 system with higher renewable integration. Figure 7(b) quantifies the energy bill savings that flexible consumers enjoy with respect to inflexible consumers for the same amount of energy consumed, assuming that these bills are based on fully cost-reflective tariffs capturing the system operation conditions. It can be observed that these savings are very high (especially in the 2030 system, where they even reach a level of around 44%), implying that under a cost-reflective energy pricing framework, as we move toward a lower-carbon system, the implications of the temporal patterns of consumers' demand on their bills become more important than their overall energy consumption. It is also noticed that the savings achieved by flexible consumers are reduced as the percentage of these consumers increases, implying that early adopters of flexibility will enjoy the highest benefits.

Second, the majority of European balancing and capacity markets impose excessively strict limits on the type, minimum size, and minimum temporal availability of the participants. In combination with the lack of regulatory clarity around the role of aggregators in many European countries, the value of distributed flexibility in reducing the system balancing and capacity costs remains largely unexploited. Notable examples include forbidding demand-side resources from accessing certain markets or not participating on a

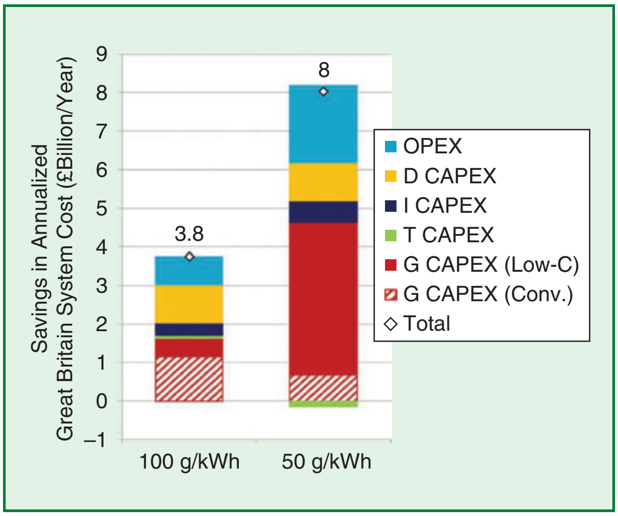


figure 6. The gross system savings of decentralized flexibility in the Great Britain system under different carbon emissions targets. (Positive/negative values indicate that the respective type of expenses is reduced/increased due to the effects of decentralized flexibility.) D: distribution; I: interconnection; G: generation; T: transmission; OPEX: operating expenditures; CAPEX: capital expenditures; Low-C: low-carbon; conv.: conventional.

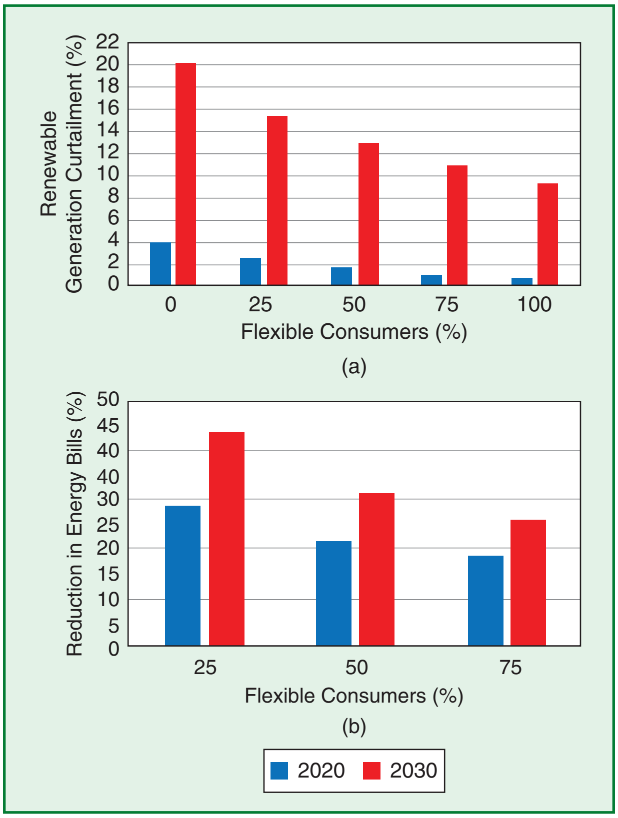


figure 7. (a) The impacts of domestic demand flexibility on renewable generation curtailment and (b) the relative reduction of flexible consumers' energy bills with respect to inflexible consumers in the Great Britain system.

The value of flexibility resources in reducing the low-carbon generation investments required for the achievement of carbon targets is not currently captured by any European market design.

level playing field with large-scale generation (e.g., shorter contract lengths).

Furthermore, European balancing and ancillary services markets generally ignore the time-coupling operating properties of DSR since each ancillary service product is cleared independently. As a result, market outcomes are not fully cost reflective and may overestimate the value of some flexible resources. As an example, a study conducted by Imperial College London has quantified the value of frequency-response service provided by thermostatically controlled loads in the Great Britain system under the independent and simultaneous clearing of frequency response and reserve services. In this example, a case when refrigeration provides the primary frequency control by reducing its consumption will be naturally followed by a load recovery effect (i.e., the demand in a subsequent period will be higher than the level it would follow if the provision of frequency response had not taken place to restore temperature at the desired set-point), implying that the secondary reserve requirements of the system may increase. Therefore, the actual value of the frequency regulation service when accounting for this effect is visibly lower than the one projected by the current independent clearing approach.

Moreover, the location-specific component of the distribution network charges constitutes a very small proportion of the overall charges in most European countries, and the largest amount of network costs is socialized, preventing distributed flexibility resources from taking actions to avoid/defer distribution network reinforcements. Last but not least, the value of flexibility resources in reducing the low-carbon generation investments required for the achievement of carbon targets (which constitutes the most significant value stream in the low-carbon future, as illustrated in Figure 6) is not currently captured by any European market design, to the best of the authors' knowledge, and constitutes a key market design challenge going forward.

Geographical Integration of Electricity Markets: The European-Wide Approach and Local Energy Markets

As previously discussed, a cost-effective transition to the low-carbon energy future involves a combination of large-scale renewable generation and the deployment of small-scale distributed flexibility resources at the local level. In this context, another major policy challenge lies in the introduction of suitable market mechanisms at multiple geographical levels, ranging from the European-wide level to the local community level.

Concerning the former, previous work has demonstrated that a coordinated European-wide approach for the integration of renewable generation can offer very significant benefits compared to a member state-centric approach by taking advantage of the significant geographical diversity of renewable energy resources' availability, including the higher capacity factors of wind generation in Northern Europe and the higher capacity factors of solar generation in Southern Europe. Specifically, if such diversity is combined with a full harmonization and integration of the different countries' electricity markets, the same amount of renewable energy can be produced with 150 GW fewer renewable generation capacity with respect to the member state-centric approach, entailing around €200 billion of savings in capital investments until 2030. Although such a European-wide approach has been outlined in the European Renewable Energy Directive, it has not yet been realized. Furthermore, interconnections to the Middle East and Africa could potentially further increase these benefits by exploiting the high solar generation availability in those regions.

At the other end, despite the massive value of distributed flexibility resources enabled by the digitalized energy paradigm, the effective integration of large numbers of such small resources in electricity markets is extremely challenging due to scalability limitations and privacy concerns raised by the end consumers/prosumers. In this context, local energy markets (LEMs) constitute a new market mechanism attracting continuously increasing interest. LEMs enable the direct trading of energy and flexibility among the end users of a local community, coordinated either in a centralized fashion (e.g., by an independent community manager) or in a fully distributed fashion through emerging peer-to-peer trading architectures.

Beyond addressing scalability and privacy concerns, a LEM promises a number of significant benefits, including a) limiting the energy dependency of active consumers/prosumers on the incumbent electricity retailers and consequently enhancing the competitiveness of the latter; b) avoiding distribution network reinforcements as a result of matching local demand with local generation; c) enhancing the engagement of local end users in system operation by creating a local identity and promoting social cooperation; and d) revitalizing the local economy by shaping opportunities for local investment, creating new jobs at the community level, and promoting self-sufficiency. The EC has recognized these

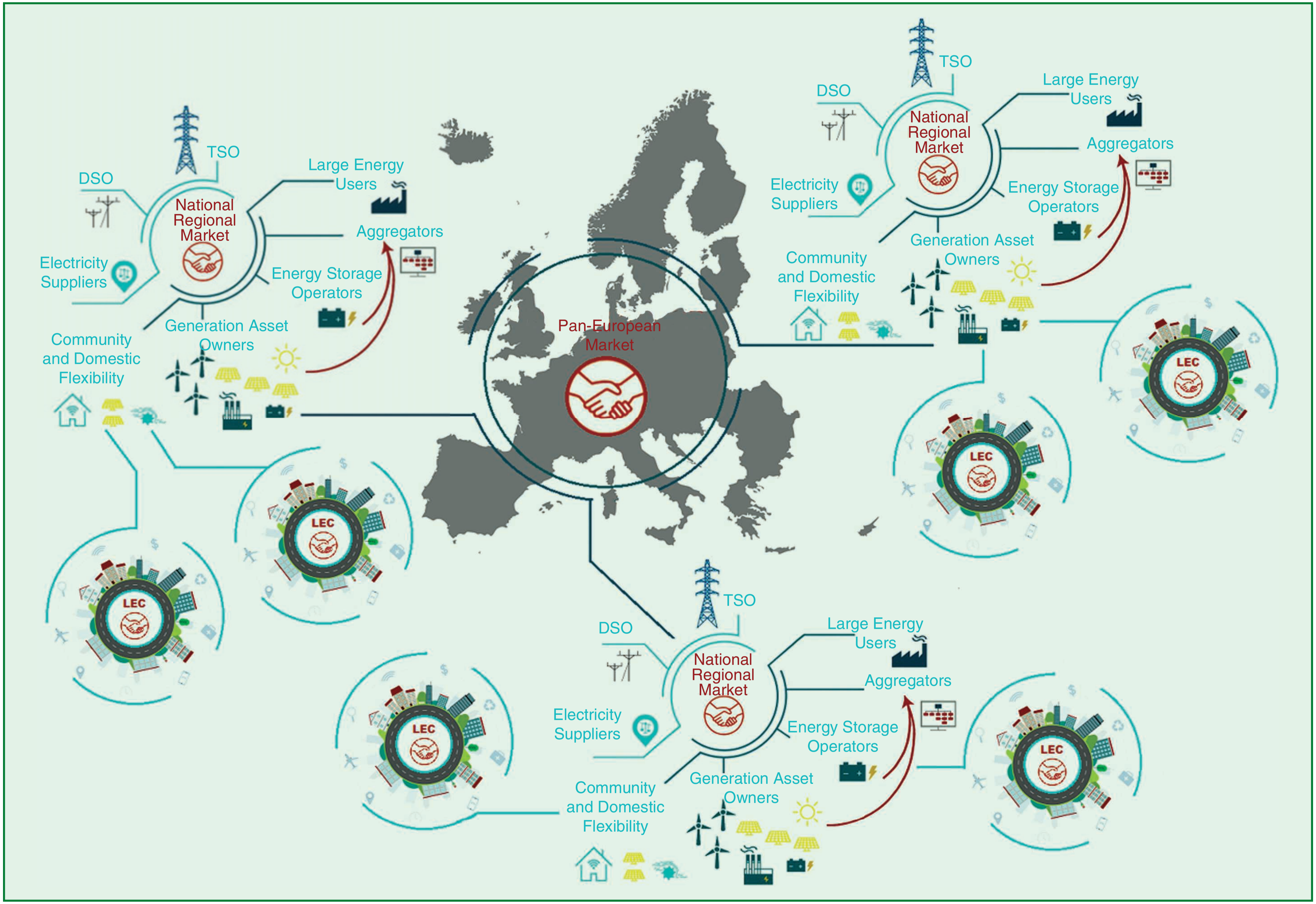


figure 8. The TradeRES vision for an integrated market architecture. TSO: transmission system operator; DSO: distribution systems operator.

> Another major policy challenge lies in the introduction of suitable market mechanisms at multiple geographical levels, ranging from the European-wide level to the local community level.

benefits by establishing and promoting the concept of local energy communities.

The TradeRES Vision

The vision of the recently initiated Horizon 2020 TradeRES project (www.tradeRES.eu) lies in developing and testing innovative electricity market designs that will enable the cost-effective and secure development of a nearly 100% renewable power system and realize the full extent of the system-wide benefits of flexibility resources. Such market designs should be capable of addressing the key challenges identified in this article, including a) the missing money problem faced by both renewable and conventional generators, b) the participation of renewable generation in ancillary services markets and the provision of flexibility, c) the design of effective carbon emissions markets, d) the incorporation of small-scale distributed flexibility in energy and balancing markets, and e) the pan-European harmonization of electricity markets and the economic utilization of cross-border interconnections. Following the previous discussion, one of the main objectives of the project involves the development of an integrated European market architecture that encapsulates the pan-European market, national/regional markets, and local energy communities in a fashion that enables the maximum utilization of available renewable generation and flexibility resources, as reflected in Figure 8.

Acknowledgment

This work has received funding from the EU Horizon 2020 research and innovation program under project TradeRES (grant 864276).

For Further Reading

A. Shakoor et al., "Roadmap for flexibility services to 2030," Report to the Committee on Climate Change, London, U.K., May 2017.

F. Lopes, "MATREM: An agent-based simulation tool for electricity markets," in *Electricity Markets With Increasing Levels of Renewable Generation: Structure, Operation, Agent-based Simulation, and Emerging Designs, Series: Studies in Systems, Decision, and Control*, vol. 144, F. Lopes and H. Coelho, Eds. Cham, Switzerland: Springer-Verlag, 2018, pp. 189–225.

N. Helistö, J. Kiviluoma, and H. Holttinen. "Sensitivity of electricity prices in energy-only markets with large amounts of zero marginal cost generation," in *Proc. 14th Int. Conf. European Energy Market (EEM)*, Dresden, Germany, 2017, pp. 1–6. doi: 10.1109/EEM.2017.7981893.

A. Papavasiliou, Y. Smeers, and G. Bertrand, "An extended analysis on the remuneration of capacity under scarcity conditions," *Econ. Energy Environ. Pol.*, vol. 7, no. 2, pp. 1–19, 2018.

G. Morales-España and J. Sijm, "Simultaneous reduction of emissions and costs by curtailing renewables in optimal operation of power systems," in *Proc. ISGT-Europe2020: IEEE PES Innov. Smart Grid Technol. Europe*, Den Haag, The Netherlands, Oct. 2020, pp. 1–6.

J. C. Richstein, E. J. L. Chappin, and L. J. De Vries, "Cross-border electricity market effects due to price caps in an emission trading system: An agent-based approach," *Energy Pol.*, vol. 71, pp. 139–158, Aug. 2014. doi: 10.1016/j.enpol.2014.03.037.

Biographies

Goran Strbac is with Imperial College London, United Kingdom.

Dimitrios Papadaskalopoulos is with Imperial College London, United Kingdom.

Nikolaos Chrysanthopoulos is with Imperial College London, United Kingdom.

Ana Estanqueiro is with the National Laboratory of Energy and Geology, Amadora, Portugal.

Hugo Algarvio is with the National Laboratory of Energy and Geology, Amadora, Portugal.

Fernando Lopes is with the National Laboratory of Energy and Geology, Amadora, Portugal.

Laurens de Vries is with Delft University of Technology, The Netherlands.

Germán Morales-España is with the Netherlands Organisation for Applied Scientific Research, The Hague, The Netherlands.

Jos Sijm is with the Netherlands Organisation for Applied Scientific Research, The Hague, The Netherlands.

Ricardo Hernandez-Serna is with the Netherlands Organisation for Applied Scientific Research, The Hague, The Netherlands.

Juha Kiviluoma is with the VTT Technical Research Centre, Espoo, Finland.

Niina Helistö is with the VTT Technical Research Centre, Espoo, Finland.

Zero-Marginal-Cost Electricity Market Designs

©SHUTTERSTOCK.COM/LISA KOLBASA

By Luiz Barroso, Francisco D. Muñoz, Bernardo Bezerra, Hugh Rudnick, and Gabriel Cunha

Digital Object Identifier 10.1109/MPE.2020.3033398
Date of current version: 6 January 2021

Lessons Learned From Hydro Systems in Latin America Might Be Applicable for Decarbonization

LARGE REDUCTIONS IN THE COST OF RENEWABLE energy technologies, particularly wind and solar, as well as various instruments used to achieve decarbonization targets (e.g., renewable mandates, renewable auctions, subsidies, and carbon pricing mechanisms) are driving the rapid growth of investments in these generation technologies worldwide.

Despite the overall benefits of producing electricity using renewables instead of relying on fossil fuels, incorporating large amounts of solar and wind generation can be challenging for power systems. Solar irradiance and wind speeds are variable and, to some extent, unpredictable, which can compromise the stability of the power grid. Private investors, electric utilities, and independent system operators (ISOs) are addressing this challenge through a combination of measures that include the geographical diversification of resources, utilization of energy storage, and implementation of demand-response programs.

Another feature of renewables that is making market participants, lenders, policy makers, and regulators concerned is their effect on equilibrium prices of electricity. Most wholesale electricity markets set real-time prices (also referred to as spot prices or locational marginal prices, depending on the implementation) as the marginal cost of producing one incremental unit of electricity at any given instant. Under this paradigm, there are concerns that increasing levels of generation from technologies with near-zero marginal costs such as renewables will inevitably depress spot prices to the point that revenues from the energy spot market will be insufficient to cover the capital costs of merchant generation technologies. This has raised questions about the ability of current electricity markets based on spot pricing to incentivize investments that will deliver efficient and reliable power systems in situations with high shares of renewables.

This article calls upon the experiences of hydro-dominated Latin American electricity markets to highlight the parallels among renewable-driven energy systems of the 21st century and hydro-driven systems that Latin America has been operating for decades (in particular, during the 1990s). Several Latin American countries have found that the liquidity of long-term financial instruments is essential to incentivizing investments in generation capacity. This is particularly important in situations where spot prices are extremely volatile, alternating between periods with high prices, driven by scarcity pricing mechanisms, and extended periods of time with zero prices, when hydro resources are abundant. In many Latin American countries, regulators have imposed minimum mandatory forward contracting requirements to guarantee a minimum level of liquidity of long-term financial instruments, complementing voluntary bilateral markets for these products.

Centralized auctions for long-term contracts are also common in the region. They act as market-based mechanisms to procure electricity and ensure some tariff stability for retail customers when there is no competition in the retail segment. Mandatory long-term products can be simple forward energy contracts (as they are in Chile and Peru), energy bundled with reliability products (in Brazil), or contracts for a standalone reliability product (in Colombia), with energy contracts traded in bilateral markets. The experiences of countries in Latin America dealing with systems with high shares of generation from near-zero-marginal-cost resources can be useful for electricity markets in other parts of the world, particularly if new renewable-dominated systems do not have enough liquidity of long-term financial contracts to hedge risk.

The overall experience of Latin America's long-term markets to attract and retain investors in new generation has been positive. However, there are some issues related to the design of both long- and short-term markets that need to be addressed. For instance, auctions for long-duration contracts facilitate investments and benefit lenders; however, they can introduce inflexibilities to the market, preventing cost reductions in technology from being passed on to consumers. Short-term markets will also need to be improved to accommodate increasing shares of generation from renewables. Some of the needed enhancements will require mirroring features of short-term markets in the United States and Europe, such as increasing the temporal granularity of real-time prices, introducing multisettlement mechanisms (absent in many prominent Latin American markets, such as Brazil and Chile), and allowing emerging technologies and demand-side resources to participate in wholesale markets. In this vein, there could be learning opportunities both for Latin America and electricity markets in the United States and Europe to find the best market design to accommodate increasing shares of generation with zero marginal cost.

The Pricing of Electricity: What Does the Theory Say?

The foundations of electricity pricing were developed in the mid-1980s, with Fred Schweppe in particular having made significant contributions to its underlying theory and practice. Spot pricing is fundamental to the design and operation of electricity markets worldwide and has powerful implications for the efficiency induced by these market-based signals. Under some specific assumptions (such as perfect competition), if the wholesale electricity price in each period and location reflects only the short-run marginal cost of an incremental change in demand (plus the cost of reducing demand when capacity is scarce) without the consideration of capital costs, a spot market can guarantee the efficient operation of generation units in the short term and incentivize the entry and exit of generation units of the right size, with the right technological characteristics, and at the right locations in a transmission network.

Figure 1(a) illustrates the supply and demand curves for a hypothetical system with four generation units with different marginal costs. The intersection of supply and demand curves defines both the spot price and the demand level that must be supplied with the available generation units. Note that, for most demand levels, the intersection point occurs at one of the horizontal segments of the supply curve, which means that spot prices coincide with the marginal cost of some generation unit (e.g., the resulting spot price for Demand_2). However, the intersection of supply and demand can also occur at vertical segments of the supply curve when generation capacity is scarce. In these cases, the resulting spot prices are higher than the cost of the most expensive generation unit running in the system (e.g., the resulting spot prices for Demand_1 and Demand_3) and are often referred to as *scarcity prices*. In equilibrium, this set of spot prices, including scarcity prices, allows for all the efficient units to cover both operating and capital costs.

The supply curve in Figure 1(a) is representative of most historical and current systems, with a steeply sloped demand curve (low elasticity) and a supply curve (also called a merit-order curve) formed by a portfolio of units with different marginal costs, such as solar, wind, hydro, nuclear, coal, gas, and/or diesel generation. In those systems, supply sets the spot price most of the time, meaning that prices are equal to the short-run marginal cost of the most expensive unit in operation. In such cases, many of the low-priced generation units recover a large fraction of their investment during times when more expensive generators set the spot price. This is particularly true for renewables, which have extremely low variable costs (typically, only a few dollars per megawatt-hour and linked to wear and tear and other operating costs) and can be considered as virtually equal to zero for all intents and purposes. For example, a wind unit that runs at a time when demand is high and the spot price is set by a diesel generator earns a short-run profit equal to the difference between the marginal costs of the diesel unit and the wind generator. In the relatively mature markets of the United States and Europe, the somewhat predictable behavior of the supply and demand curves has resulted in moderately stable spot prices, which eases the predictability of revenues needed by generators to secure financing with lenders.

In contrast, electricity systems with high renewable shares may have much less variety in the variable cost of generation technologies. Figure 1(b) shows the supply and demand curves for a hypothetical system where all the generators have zero marginal cost, akin to how supply curves would look in a system with a lot of generation from abundant hydro, wind, and solar resources. Note that, in those cases, any time that demand is low enough, spot prices are equal to zero. However, when demand goes up, prices can increase dramatically, allowing all technologies to recover their investment costs. Although all generation units benefit from scarcity prices, their occurrence in the case depicted in Figure 1(a) is particularly important to ensure that peaking

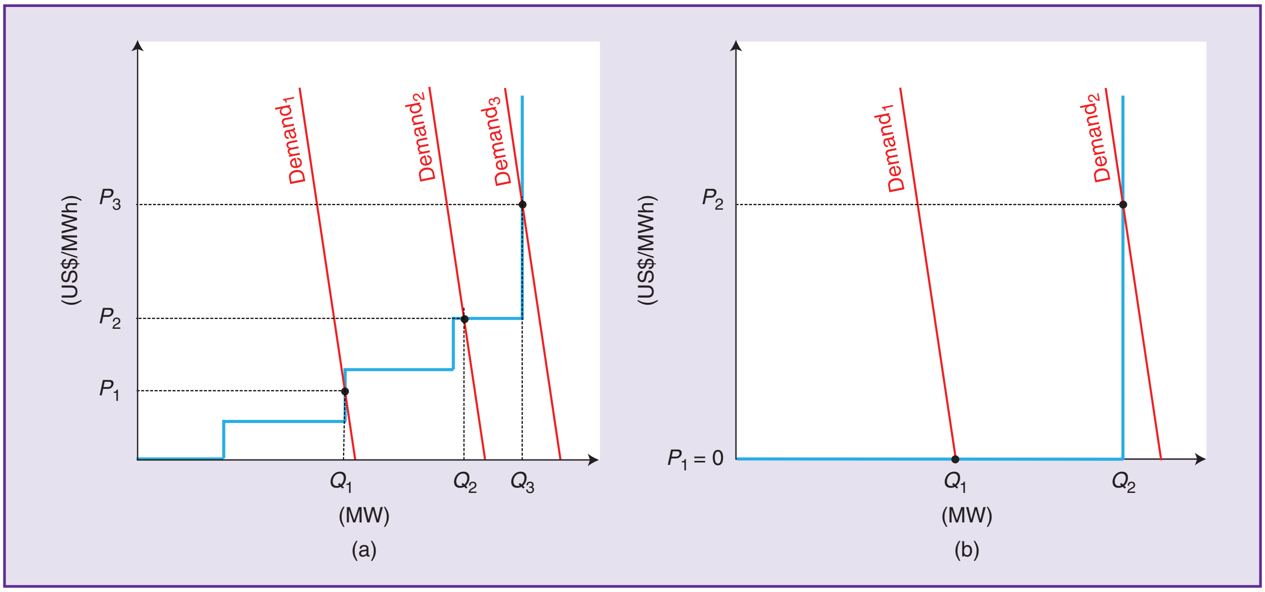


figure 1. The supply and elastic demand curves for a system (a) with a mix of different conventional generation units and (b) where all units have zero marginal cost.

units (e.g., diesel generators) are able to recover their investment costs. Otherwise, incentives to the entry and exit of capacity work in the same manner in both examples. Note that in the situation depicted in Figure 1(b), there is a more extreme discrepancy between peak spot prices (potentially very high scarcity prices) and off-peak prices (virtually zero), which provides incentives for demand-side resources to adjust consumption, achieving a similar effect to increasing or decreasing generation capacity in the long run.

Hence, conceptually, there is nothing that prevents the application of the classic spot-pricing theory to systems with high shares of generation from resources with zero marginal cost. As we show in Figure 1(a) and (b), the only difference is that, in systems with a lot of generation from technologies with zero marginal cost, scarcity pricing becomes the main mechanism to ensure cost recovery in the long run because spot prices are likely to be zero for extended periods of time. If liquid financial markets used to hedge the price-volume risk over different time frames are in place, the optimal capacity-expansion mix is secured (and able to be financed). In these situations, consumers can also define their optimal reliability needs and participation in the market as an active demand response based on private preferences (e.g., risk aversion).

In mature electricity markets following these design principles, spot prices can increase dramatically during scarcity times due to high price caps (as seen in Australia and Texas). In practice, however, the administrative estimate of the value of lost load (VoLL) used to determine price caps is sometimes driven by political instead of technical considerations, which can introduce distortions. A relatively recent development regarding price formation in periods of scarcity (which is addressed in the following section) has been the implementation of sloped operating reserve demand curves (ORDCs) employed in some markets in the United States and Mexico, where price-dependent curves replace vertical demand curves for operating reserves.

In practice, low price caps, illiquid financial markets for long-term contracts, and the lack of demand response can pose real challenges for electricity markets in their purest form, which choose to rely solely on spot pricing—including scarcity pricing—to provide expansion incentives. These challenges are especially pronounced for countries with fast load-growth rates or increasing levels of decommissioning of existing generation capacity, where the lack of new supply may result in shortages. Furthermore, it is likely that these challenges will become even more pronounced with increasing shares of generation from renewables. Not only is there a tendency toward a feast-or-famine situation with regard to equilibrium prices, as illustrated in Figure 1(b), but the technological disruption of renewables has profoundly altered the landscape of expectations for the electricity sector. In particular, there are significant uncertainties regarding the rate at which the cost of renewables will continue to fall and their share in the expansion mix will continue to rise as well as the rate at which additional innovations, such as the emergence of distributed energy resources, demand response, and storage technologies, will be disseminated. This combination of spot price volatility and uncertainty with regard to the future evolution of the system creates an environment that may threaten potential investments and loans, creating an even greater motivation for long-term markets for financial and/or reliability products.

Spot Pricing in Hydro Systems in Latin America: One Form of Scarcity Pricing

Latin America is formed by 16 countries and has a power system with roughly 400 GW of installed capacity, where hydropower accounts for approximately 50% of the generation mix. Load growth rates have historically hovered near 5% per year in a region where energy consumption is roughly 1,500 TWh/year. Figure 2 shows a general depiction of the main wholesale market design elements.

On the pricing side, only Colombia and the Central American regional electricity market have adopted a bid-based scheme for generation dispatch and spot price formation, as represented in shades of red in Figure 2. All the other countries in the region utilize cost-based arrangements, where generators report only their directly attributable marginal production costs (i.e., fuel costs) to build the merit-order curve for the dispatch and pricing of electricity done by the ISO. Water values are used as proxies for marginal production costs for hydro plants, which are calculated by the ISO based on a set of administratively defined assumptions and with the aid of stochastic optimization models. Given the cost-based merit-order curve, spot prices are defined as the cost of the marginal unit needed to meet demand in each settlement period.

Although cost-based markets have some disadvantages compared to bid-based ones, most countries in Latin America have opted for cost-based market designs for the following reasons: 1) to ensure transparency (the dispatch and spot prices are calculated by computer models using well-known algorithms, with the software and system data publicly available to all market participants), 2) to guarantee efficiency in the dispatch of hydro plants in cascade with independent owners and multiple water uses, and 3) to avoid potential issues with market power that could arise in bid-based markets. When electricity markets were first implemented in Latin America, regulators perceived that it was important to tackle these issues to imbue investors with the confidence to invest in new generation capacity, which was the main goal of industry reform in these countries. In addition, regulators were concerned that the cost of implementing a bid-based dispatch and pricing mechanism could be prohibitively high due to the need to set up sophisticated trading platforms and market power-mitigation mechanisms as well as educating state-owned companies to bid rationally into these markets.

On the other hand, one of the main criticisms of the current centralized scheme used to determine water values has to do with the sensitivity of probabilistic simulation models to input parameters, such as the probability distribution of future hydrological conditions. Despite the efforts made by ISOs to ensure the transparency and replicability of results, conflicts (sometimes leading to court cases) occurred frequently because of discrepancies between the assumptions made by the authority and each firm's private view of what should and should not go into the simulation model. This is because assumptions about input parameters affect not only the centralized estimate of the value of water but also dispatch decisions, prices, and revenues for private firms that participate in the market.

As highlighted in Figure 3, there are several interesting commonalities and contrasts between the renewable-dominated systems that may become prominent in future and hydro-dominated systems such as those in Latin America (especially prior to the introduction of large amounts of thermal capacity in the 1990s and 2000s).

The first common feature is related to the volatility of spot prices. Systems with high shares of generation from resources with zero marginal costs commonly face extended periods of time when spot prices are zero, followed by periods of high prices when renewable resources are not available. This happens because these systems are typically designed to ensure that demand can be supplied even in the most adverse weather conditions considered in the simulation model, which, in practice, do not occur frequently. For this reason, it has been common for hydro-dominated systems in Latin America to face excess energy and low spot prices for extended periods of time. Nevertheless, in extremely dry seasons, some demand rationing may occur, and prices can climb up to the price cap. This phenomenon is illustrated in Figure 4, which shows the observed monthly spot prices in the Brazilian Southeast system from January 1993 until August 1997, when an electricity

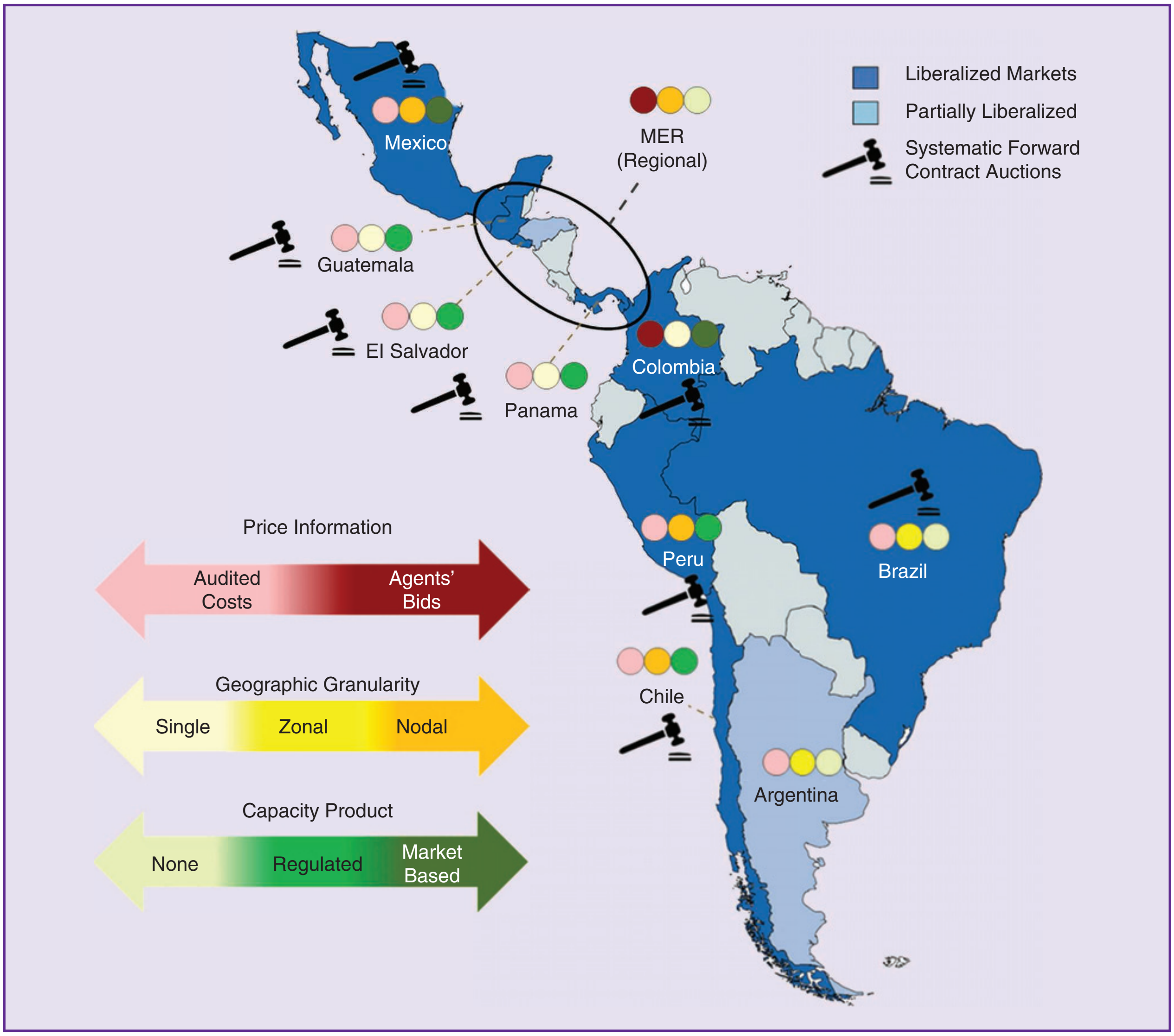


figure 2. The outlook of wholesale market design elements in Latin America. MER: the Central American regional electricity market.

	Hydro-Based Systems	Renewables-Based Systems
Timescale of Main Variability Components	Weekly, Seasonal, and Interannual	From Minutes to Daily Patterns
Effect on Short-Term System Operations	Hydro Provides Reserves	Renewables Require Reserves
What Kind of Support Generation Is Needed?	Resources to Remain On (Off) for Entire Dry (Wet) Seasons	Fast-Acting Resources to Make Up for Sudden Imbalance
Risks of Shortage	Rationing (Energy Constrained)	Blackouts (Capacity Constrained)
Price Formation	Based on Probability and Cost Rationing	Based on Probability and Cost of Sudden Blackouts

figure 3. The parallels between renewable- and hydro-dominated markets.

market reform was discussed in the country.

As we show in Figure 4, the spot price was close to zero in 36 out of the 56 months depicted, and the longest low-price period lasted for nearly two years (21 months). This behavior is similar for other countries in the region with many hydro resources, and it is similar to the price behavior expected in renewable-driven systems. It merits noting, however, that price volatility in hydro-based systems occurs at a different timescale. Energy prices tend to exhibit low volatility of spot prices in the short term, as reservoirs can easily transfer hydro energy from off-peak to peak hours and modulate load supply.

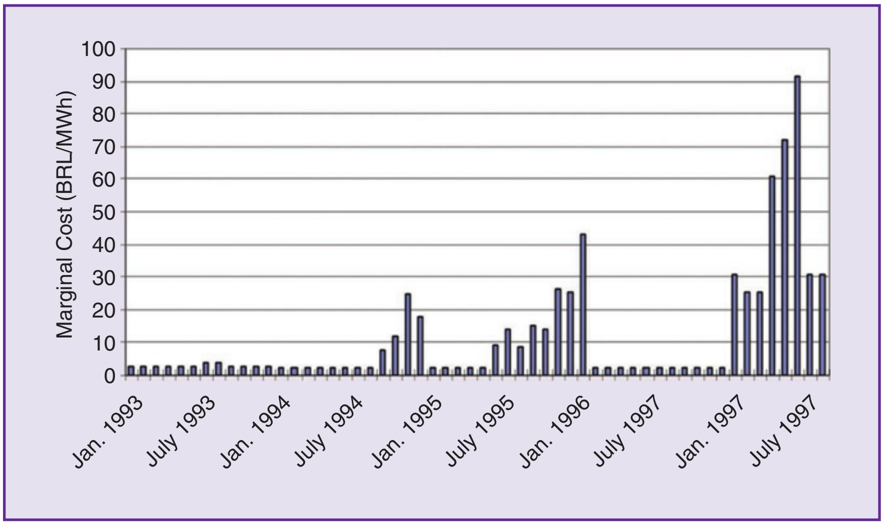


figure 4. The historical monthly spot prices in the Brazilian system (BRL/MWh).

The second commonality has to do with the principle used to estimate the value of water in cost-based markets in Latin America and the idea behind the implementation of sloped ORDCs in U.S. markets. Sloped ORDCs are constructed based on the notion that stochastic fluctuations in the supply-demand balance must be accommodated by dispatching part of the system's reserves, or short-term operational flexibility, committed in the ex-ante market. If a lower amount of reserves is procured, therefore, there is a probability that the system will not be able to respond and that some demand will need to be curtailed. A probabilistic simulation model representing multiple sources of uncertainty can be used to assess this type of risk. By multiplying this probability of shortage as a function of the amount of online reserves by the VoLL, one obtains an estimate for the demand side's marginal willingness to pay for an incremental amount of reserves to avoid a shortage. It is clear from this depiction that increasing the VoLL has a direct influence on the shape of the ORDC, thus leading to a more conservative assignment of resources in the short term, pressuring spot prices upwards and inducing a larger capacity margin at the market equilibrium, leading to a more reliable system in the long term.

A very similar logic applies to the calculation of water values in hydro-dominated systems. Provided that all available generation resources have zero marginal cost, the water value can be approximated by the product of the VoLL (or, more precisely, the cost of rationing) multiplied by a probability of energy shortages, reflecting the opportunity cost of not having water for generating power in the near future. Water values are calculated by simulating the system's operation over several periods using a probabilistic model that considers different scenarios of hydrological conditions (e.g., dry, average, and wet). Figure 5 illustrates how a probabilistic simulation model weights the future opportunity cost of water according to the probabilities of each scenario to define today's marginal value of hydropower. When the system is unable to meet demand in a given dry scenario and stage, the opportunity cost of water is equal to the cost of rationing, as the only alternative to replace a reduction in hydro generation is to curtail demand. In scenarios with intermediate inflows, the value of water is usually equal to the cost of the cheapest thermal plant in the system that could increase its output if hydro generation was reduced. In contrast, the water value is zero in scenarios where dams are overflowing, which is often the case in extremely wet seasons.

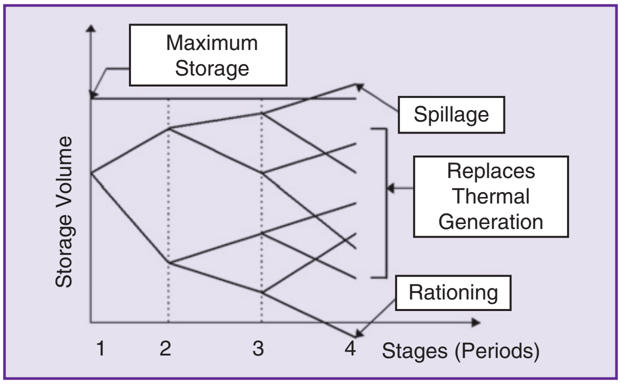


figure 5. The calculation of the opportunity cost of water.

The third parallel we see between these two types of systems is the challenge of incentivizing efficient investments in situations with highly volatile spot prices. As we mentioned in the previous section, theory states that volatile spot markets give generators and consumers incentives to engage in long-term financial contracts to hedge risks. However, in situations where markets for long-term contracts are illiquid or insufficiently mature, it can be difficult for developers to gain access to project finance agreements from lending entities to support new investments in generation capacity. It was because of the high-volatility of spot markets with extended periods with zero prices, the lack of liquidity of long-term contracts, and the pressing need for new generation capacity that many countries in Latin America chose to implement centralized auctions for long-term contracts. Long-term contracts can also provide insurance against both policy uncertainty and some political risk, reducing the risk premium required to justify investments.

Figure 6 shows two examples of actual merit-order curves of the Brazilian system for two historical snapshots in wet and average weeks. Note that the availability of hydro resources can have a large effect on spot prices and incentives for market-based system expansion.

The Need for Hedging Instruments: Centralized Auctions for Long-Term Contracts

Since the initial implementation of Latin American market reforms in the 1990s, several countries have introduced capacity mechanisms to ensure that there are always adequate supply resources to meet demand. Capacity mechanisms are designed to correct potential issues that result

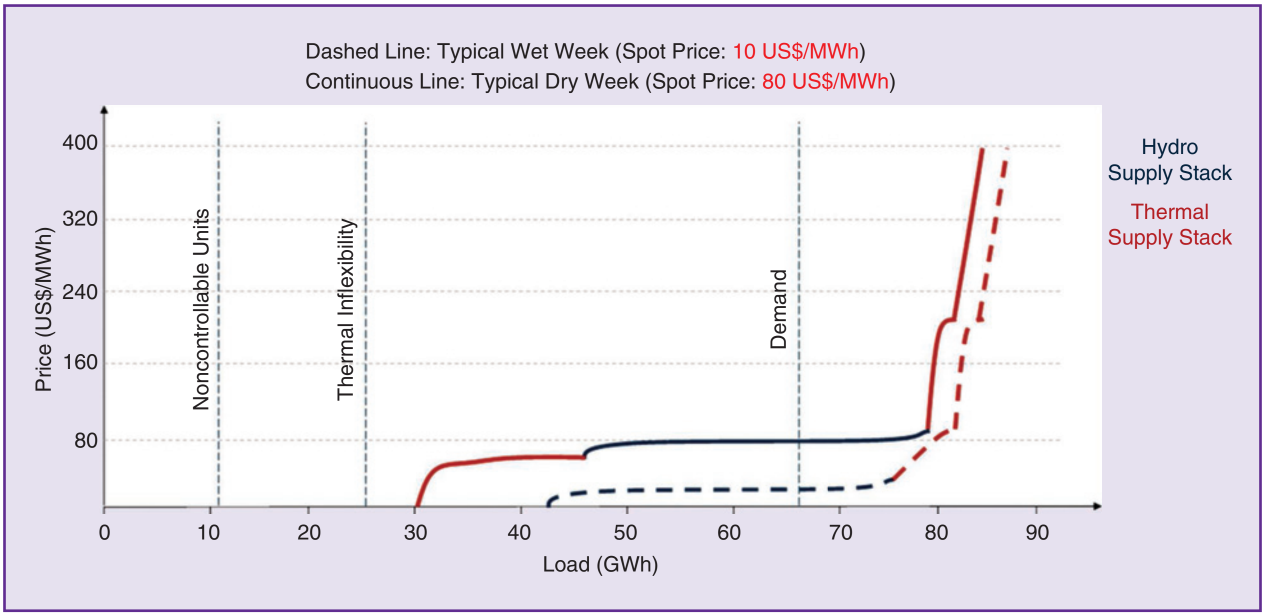


figure 6. Examples of the variance of merit-order curves in Brazil.

Long-term contracts can also provide insurance against both policy uncertainty and some political risk, reducing the risk premium required to justify investments.

from price caps that are set too low (leading to artificially low spot prices on average and, thus, insufficient incentives for system expansion) and also as instruments to stabilize revenues for generators in light of volatile signals from spot markets. They operate just as capacity markets do in the United States, relying on an administrative definition of what constitutes the firm capacity product that drives a component of agents' remuneration.

At the time of the initial market reforms in Latin America in the 1990s, most implementations of capacity mechanisms were in the form of a regulated capacity payment (as seen in Colombia, Chile, and Peru), where the capacity price was determined by the regulator based on an administrative estimate of the cost of new entry. At that time, Brazil was the only country that imposed a forward contracting requirement on load-serving entities and deregulated consumers to mandatorily cover a high percentage of their loads through energy contracts, which had to be negotiated bilaterally and backed by firm energy credits. These initial designs, however, faced numerous challenges in practical implementation during the following decade.

In the case of capacity payments, experience showed that both the administrative definition of capacity price and capacity product could have a large impact on investment incentives for individual firms. In addition, forward contracting requirements alone did not ensure that regulated distribution companies were incentivized to choose least-cost contracts for retail customers, thus leading to self-dealing issues and inflated prices. The previous challenges, combined with the fact that short-term prices did not provide an adequate incentive for generation expansion (given the absence of liquid marketplaces for financial hedging), motivated a second wave of market reforms in Latin America in the 2000s, following the first wave of reforms in the 1990s. This second wave of reforms was focused on improving the mechanisms used to safeguard resource adequacy and led to the introduction of market-based capacity products (or reliability products) in some countries.

Brazil pioneered this new wave of reforms in 2004, introducing a mechanism that put auctions front and center and served as an inspiration to several other countries, such as Chile in 2005 and Colombia and Peru in 2006. These new designs typically combined centralized auctions with a quantity-based mechanism that required a minimum level of contracting for loads. Implementation details vary among countries, particularly regarding the following core elements, which can be used to describe most auction mechanisms for long-term contracts introduced in real-world electricity markets over the years:

- ✔ *Demand-side obligations*: These include, among others, 1) the assignment of responsibility for forecasting the demand several years ahead for the procurement of contracts, 2) a mechanism for assigning the cost of forward contracts to consumers, and 3) rules to specify under which conditions agents can opt out of the standard mechanism to procure their own demand. For example, although the auctions in Colombia involve only the purchase of a reliability product, the auctions in Peru and Chile involve only a forward contract, while Brazil requires contracts for bundles of reliability products and forward contracts.
- ✔ *Supply-side liabilities*: These include, among others, 1) what exactly generators' reliability commitments entail regarding firm supply backing and 2) penalties for noncompliance with contract clauses and specific obligations for energy and/or reliability delivery, often tailored to physical attributes of different generation technologies.
- ✔ *Auction design elements*: These include, among others, the definition of 1) lead times; 2) contract duration; and 3) the eventual technology segmentation of potential suppliers, such as differentiating between existing projects and new projects for contracting purposes.

In practice, the implementation of minimum contract requirements and centralized auctions by Latin American regulators in the 2000s emerged as a practical short-term solution to issues that were primarily related to the lack of investments in past periods, when demand was growing too fast compared to new capacity additions. Consequently, there was a need to take some action to accelerate private investments in new capacity. Furthermore, although it might not have been the primary intent to correct for the market failure that results from incomplete financial markets for risk sharing, there is now robust empirical and theoretical evidence suggesting that mechanisms that introduce this type of financial contract can indeed improve market liquidity and market efficiency.

Long-term electricity auctions are now one of the driving forces for the expansion of the power sector in Latin America. To date, more than 100,000 MW of new generation capacity from all technologies have been contracted and delivered at competitive prices via those auctions. In addition, since the late 2000s, countries all over the world have started using different variants of these auctions as

table 1. Recommendations for improving several design elements in the wholesale market.		
Design Element	**Current Status in Latin America**	**Suggested Improvement**
Wholesale spot price formation	In many cases, spot prices are computed ignoring transmission constraints and without the co-optimization of energy and reserves.	Spot prices should be computed considering all transmission and generation constraints, plus reserves, simultaneously. This approach ensures that all constraints are reflected in spot prices.
Temporal granularity of spot prices	Most countries in the region compute prices at hourly time intervals, with the exception of Peru (every 30 min) and Brazil (three load blocks in weekly prices).	A time granularity of at least 1 h is recommended to allow spot prices to better reflect the physics of the system, which is particularly important for units that impart flexibility. Increasing the frequency of dispatch and settlement intervals also decreases the need to activate reserve products.
Spatial granularity of spot prices	Some countries employ a simplified version of nodal or zonal pricing using merit-order curves to determine the spot price for prespecified pricing zones (Brazil), which can also be a single large pricing zone that includes the whole country (e.g., Colombia).	Countries should implement locational marginal pricing with mechanisms to allow market participants to hedge congestion risks. locational marginal pricing provides efficient signals for the entry and exit of generation units by reflecting information about the incremental value of generation at each location in the transmission network.
Cost- or bid-based arrangements for dispatch and price formation	With the exception of Colombia, all of the short-term electricity markets in Latin America to date have been cost-based markets.	Whenever possible (sufficient political will, human capital, and competition), we recommend bid- instead of cost-based markets. Practical experience indicates that having a central agency that relies on a single view of the future to make decisions may lead to conflicts and legal disputes. Concerns about the coordination of hydro units in cascaded systems, multiple water uses, and market power concerns can be addressed with a combination of property rights and active market monitoring.
Scarcity pricing	Countries with lots of hydro resources have a form of a scarcity-pricing mechanism that reflects the administratively calculated socioeconomic cost of curtailing demand sometime in the future if hydro resources are not available (as assessed by a simulation model). To our best knowledge, Mexico is the only country that has implemented sloped ORDCs.	Although there is not enough demand response for scarcity prices to naturally emerge, we recommend that the VoLL or cost of deficit parameter used in simulation models should be at least equal to the price at which demand would be willing to reduce consumption (in line with the resource-adequacy target of the system). The use of a sloped ORDC can also improve price formation during times when reserves are scarce and prevent abrupt price drop offs.
Ancillary services	In most countries, the provision of ancillary services is mandated by regulations that compensate units only for the directly attributable costs of providing the service (e.g., fuel costs). In most cases, energy and reserve products are not co-optimized.	We recommend migrating to schemes that co-optimize the provision of energy and ancillary services. We also recommend remunerating ancillary services based on uniform price, ensuring that all agents providing the same service are remunerated equally. Mechanisms that compensate only for the directly attributable costs of providing these services are discriminatory and do not provide incentives for the entry of efficient units in the long term.
Multisettlement markets	Most countries use day-ahead scheduling and only one settlement. In both Chile and Colombia, for example, day-ahead prices are not used to settle any transaction; rather, they rely solely on real-time prices.	We recommend implementing day-ahead markets that will allow forward financial commitments to be settled against real-time prices and evaluate the need for additional settlements.
Capacity mechanism	Countries rely on different criteria to define firm energy and firm capacity values, both of which determine remunerations for firms that contribute to the system with these products. Some countries also employ administrative capacity payments without necessarily aiming for an explicit resource-adequacy target.	Countries that choose to rely on capacity mechanisms should pay attention to the definition of firm capacity of renewables and energy storage technologies. We recommend crediting firm capacity based on some reliability metric that treats all resources, including demand-side ones, equally. We also recommend that countries define a resource-adequacy target that is aligned with the administrative estimate of the cost of unsupplied demand used to price scarcity, ensuring consistency between value assessments of additional transmission reinforcements (planned centrally) and the profitability of new generation investments (based on market signals).
Centralized auctions for long-term energy contracts or minimum contracting requirements	Centralized auctions for contracts are used in many countries as a mechanism to ensure that distribution companies will procure power at the least possible cost for retail consumers. Brazil also relies on centralized auctions for contracts, with physical backup as a resource-adequacy mechanism.	Countries should consider reducing the duration of mandated contracts and explore options to introduce more liquidity into long-term financial markets, fostering the participation of financial agents and retail aggregators. Contracts of a long duration (e.g., 15–20 years) can be effective at incentivizing generation investments, reducing risk for generation firms. However, they also prevent customers from benefiting from cost reductions due to technological disruptions in generation technologies, locking customers into prices that might become too high compared to average spot prices. Additionally, we recommend that regulators consider counterparty and price risks as part of the selection criteria used in centralized auctions.

mechanisms to procure power from renewables and support the development of these technologies, thus promoting robust investment markets.

Despite the success of auction mechanisms in Latin America, there is now evidence indicating that some aspects of the original design elements of these auctions could be improved. For instance, some of the first contracts were auctioned for periods that were excessively long, which led to rigid commitments that ended up allocating too much risk to customers. Other design and implementation issues are related to the bundling of energy and reliability in a single product, contract enforcement, and selection criteria used in centralized auctions when contracts incorporate different sources of risk, such as indexation clauses (e.g., fossil fuel price risk, renewable generation profile risk, and spot price risk). Administrative definitions of firm capacity and firm energy could also result in biases against emerging technologies and impair them to compete on equal footing against, for instance, conventional generation technologies.

Improvements Needed in Latin America to Accommodate High Shares of Renewables

Even though current market designs in Latin America have served their purpose, they were not originally tailored to accommodate increasing shares of generation from variable and unpredictable resources in short time intervals. As introduced previously, hydro-dominant systems, although with their own share of challenges, have relatively high short-term flexibility compared to systems with high shares of generation from renewables (e.g., wind and solar photovoltaic). This explains why, in general, most markets in Latin America have rather simple mechanisms to settle imbalances in the short term, while they lack most of the advanced features of more highly developed electricity markets. Going forward, we recommend improving several design elements in the wholesale market as described in Table 1.

Conclusions

There are several aspects of the electricity markets in Latin America that could be improved. Some of the needed enhancements will require mirroring features of short-term markets in the United States and Europe, such as increasing the temporal granularity of real-time prices, opening wholesale markets to demand-side resources as well as emerging technologies, and introducing multisettlement systems. Long-term markets could also be improved by ensuring that contracting requirements and auction mechanisms allow for all technologies to compete on equal footing. Reducing the duration of contracts would also allow consumers to benefit from technological disruptions in generation technologies in the coming decades. It is also possible to introduce more liquidity into markets for long-term contracts by implementing marketplaces for these instruments and opening them to financial agents and retail aggregators.

Nevertheless, the experience in Latin America (Brazil, in particular) shows that many hydro systems in the region have operated with large amounts of generation with zero or near-zero marginal cost for decades and have still managed to incentivize investments in new generation capacity. However, in those situations, long-term markets for sufficiently liquid financial contracts are essential to secure generation financing, allowing investors to reduce their exposure to the high volatility of spot prices. Additionally, some types of long-term contracts can also provide insurance against both policy uncertainty and political risk, which can be large in some countries in the region.

Finally, from our perspective, the experience of some countries in Latin America that have relied on markets for long-term contracts offers some learning opportunities for countries with advanced short-term markets (e.g., the United States and Western Europe). This experience could be useful if the volatility of spot prices due to increasing shares of generation from renewables becomes a barrier to incentivize investments in new generation capacity.

Acknowledgments

This research was supported by CONICYT FONDECYT 1190228, ANID/PIA/ACT192094, ANID/FONDAP/15110019 (SERC-Chile), ANID-Basal Project FB0008, and the Complex Engineering Systems Institute (ANID PIA/APOYO AFB180003).

For Further Reading

G. De Maere d'Aertrycke, A. Ehrenmann, and Y. Smeers, "Investment with incomplete markets for risk: The need for long-term contracts," *Energy Policy*, vol. 105, pp. 571–583, June 2017. doi: 10.1016/j.enpol.2017.01.029.

S. Mocarquer, L. A. Barroso, H. Rudnick, and B. Bezerra, "Balance of power," *IEEE Power Energy Mag.*, vol. 7, no. 4, pp. 26–34, 2009. doi: 10.1109/MPE.2009.933417.

F. D. Munoz, S. Wogrin, S. S. Oren, and B. F. Hobbs, "Economic inefficiencies of cost-based electricity market designs," *Energy J.*, vol. 39, no. 3, pp. 51–68, 2018. doi: 10.5547/01956574.39.3.fmun.

H. Rudnick, L. A. Barroso, R. Moreno, R. Ferreira, and E. Pereira, "Facilitating the integration of renewables in Latin America: The role of hydropower generation and other Energy storage technologies," *IEEE Power Energy Mag.*, vol. 15, no. 5, pp. 68–80, 2017. doi: 10.1109/MPE.2017.2708862.

Biographies

Luiz Barroso is with the Institute for Research in Technology, ICAI, Comillas Pontifical University, Madrid, Spain.

Francisco D. Muñoz is with the Universidad Adolfo Ibáñez, Santiago, Chile.

Bernardo Bezerra is with PSR, Brazil.

Hugh Rudnick is with Pontificia Universidad Católica de Chile, Santiago.

Gabriel Cunha is with PSR, Brazil.

©SHUTTERSTOCK.COM/ANTIKWAR

Renewable Energy Financing and Market Design

A View and Recommendations From Development Banking Practitioners to Developing Countries

By Tonci Bakovic, Debabrata Chattopadhyay, Fernando Cubillos, and Marcelino Madrigal

Digital Object Identifier 10.1109/MPE.2020.3033399
Date of current version: 6 January 2021

DEVELOPING NATIONS WITH HIGH GROWTH RATES of energy demand and, in many cases, aging fossil fuel-based generation capacity have a massive opportunity to rapidly grow variable renewable energy (VRE) capacity. This should ideally be done through a transparent and competitive wholesale market, which is instrumental not only for attracting competitive VRE generation but also for providing the flexible capacity essential for managing the variability of solar/wind and, thus, the security of the power system. However, there are challenges, including the need

 1540-7977/21©2021IEEE

to meet certain minimum preconditions to form a market, such as

- ✔ adopting a cost-reflective end-user tariff
- ✔ unbundling the sector and imposing market share limits to reduce concentration in generation
- ✔ strengthening institutional capacity, especially in the areas of regulation and system operation
- ✔ developing effective rules for electricity grid access and expansion.

Under an energy spot market for power, prices exhibit some degree of volatility, including periods of high and low—if not very low/zero—prices. Such volatility is typical of markets and has its role in creating the necessary signals to induce investment in generation, with both fixed and variable costs, and flexible/peaking capacity, which is coupled with capacity remuneration in some markets to avoid under- or overinvestment.

Attracting investment for low-cost generation traditionally requires access to long-term financing, which, in turn, depends on the predictability of revenue that the market architecture delivers to investors. A careful balance must be achieved between attracting competitive investments by shifting risks away from investors and achieving low-cost outcomes for consumers in the short and long terms.

In the choice of market design, this issue needs to be considered within the broader context of preconditions. If a market is highly concentrated and there is inadequate regulatory oversight, especially of state-owned utilities operating in competitive environments, even the most sophisticated market design is unlikely to contain market power and, hence, likely to fail.

Even if these preconditions are met, price volatility entails risk for buyers as well as sellers. Market participants need to develop a risk appetite and have access to financial products to manage risk. Without a well-developed financial market, it is unlikely that these products would be available. Generators, retailers, and distributors unable to manage their spot price risk resorting to the safety of long-term contracts or power purchase agreements (PPAs), which, depending on their design, could be very inflexible and more advantageous to one of the parties.

While PPAs have been very effective at attracting investment, the extent to which they can be relied upon depends on each country's conditions and, especially, on expected demand growth projections. With the advent of higher shares of renewables moving away from PPA contracts, there is additional uncertainty in spot markets. Also, in many developing countries, pure spot markets have challenges simply because, in many instances, it is not socially acceptable or politically manageable to charge consumers highly variable monthly electricity bills.

In some developing countries, expenditures on energy bills can be costly. Even in higher-middle-income countries in Latin America, such as Mexico and Chile, the poorest quintiles of the population spend 6.9% and 8.3% of their income on energy (see Carvajal et al. in the "For Further Reading" section). Subsidies administered via electricity bills are still widespread in the developing world, which complicates retail price formation alternatives and, with that, the effectiveness of some market design options.

Objectives, institutional capacity, social acceptance of price fluctuations, and the risk appetite of policy makers vary across countries, signifying that market design should adapt to local conditions and balance them to achieve desired outcomes. Clear objectives are especially important in this process, and the tradeoff of design options needs to be weighted properly.

In this article, the experience of development bank practitioners in supporting sector reform and financing private generation assets in deregulated electricity markets of the developing world is described. The article also shows that an increased share of renewables is already demonstrating the need for a redesign of how electricity markets function.

When structuring the project financing of generation assets, such as renewables, lenders assess many associated risks to determine the bankability of the project, such as

- ✔ transmission-curtailment risk
- ✔ balancing cost risk inherent to VRE generation
- ✔ dispatch risk (in countries that have stranded PPAs and must-run thermal plants that bite into the merit stack before zero-variable-cost renewables)
- ✔ contractual risk, including a volume–risk mismatch of contractual commitment and VRE generation
- ✔ energy yield risk, banking on P90, P75, or P50 production (where P*x* means that, with probability *x*, the output of a given power plant will be above a certain level; energy output risk is key in project financing)
- ✔ engineering procurement and construction risk
- ✔ demand risk (when the offtake contract is "pay as demanded," and distributed generation is a threat)
- ✔ payment risk (when the offtake is not in good financial standing)
- ✔ foreign exchange risk (when the contract is not in the currency of the loan, usually U.S. dollars)
- ✔ environmental and social risk, including cumulative impact on the watershed, airshed, and community
- ✔ technological disruption risk.

These are the main risks, and there are others, depending on the country and sponsor of the project, such as reputational risks. Different banks have various ways of perceiving, assessing, and weighing these risks. As such, they may perceive the same project differently. To attract investment in power systems, it is important for market designers to consider that, on top of market risks—the ones addressed by market design—most of the listed risks are also assessed by investors and lenders. More complexity may add to efficiency gains in market outcomes but may make projects riskier, given the overall risk assessment.

A View From Inter-American Development Bank's Experience in Financing Renewable Energy Projects

Latin America is abundant in natural resources, and overall the region's power sector has a larger share of renewables than many other developing regions, especially thanks to the extensive development of hydropower. As of 2018, the electricity mix in Latin America and the Caribbean, in energy production terms, was 47% hydropower, 39% thermal generation, 2% nuclear power, 1% geothermal power generation, 5% wind power, and 1% solar generation, with the rest from other sources. Some countries, such as Uruguay and Chile, already have a larger market penetration of variable power generation, such as wind and solar: 35% in Uruguay and 6.7% in Chile. In 2018, Costa Rica ran with 99% renewable generation, mainly hydro followed by wind power.

In addition to having a rather clean power mix, the region is close to being the first developing area to achieve 100% electricity access. As of 2018, the electricity access rate in Latin America and the Caribbean stood at 97.6%. Haiti and Honduras are the only two countries with access rates well below the others, at 39% and 81%, respectively.

Given the load growth in the region, electricity demand in the period 2012–2018 grew by 88% overall in the region, and, in some subregions, such as Central America, demand grew by as much as 18%, as reported by the Latin America Energy Organization. In addition, many countries in the region increased policy support for renewables, thus growing the need for effective mechanisms to incentivize low-cost generation, such as long-term electricity auctions.

The development of new renewable technologies, such as wind and solar power, has proceeded rapidly in Latin America, primarily due to long-term contracts awarded via auction mechanisms. This system was pioneered in the region by Brazil in 2005 and Uruguay in 2006 and rapidly extended to other countries, such as Chile and Mexico. The use of electricity auctions—a wholesale market design option that consists of competition for the market rather than in the market—has proven quite effective to accelerate the uptake of renewables in Latin American countries with growing demand and clear policy goals.

The auctions are successful because those entities awarded long-term contracts, which usually range from 15 to 30 years in a region's first auction, provide revenue certainty to RE producers, help attract lower-cost financing, and, in tandem with reduced costs of RE, deliver lowest-cost renewables. Revenue comes not only from the contract but also from other important factors, such as the creditworthiness of the contract offtakers and the market structure where they operate, improved by guarantee schemes when necessary.

While auctions continue to be very effective mechanisms to deploy renewables at a rapid pace and for competitive costs, there are signs in some markets that high penetration levels of renewables may introduce risks in the medium term. Auctions in Brazil, Mexico, Argentina, Chile, Peru, and Jamaica, from 2016 to 2019, awarded projects for a total of US$46.8 billion of new investments in RE. Projects commissioned from 2014 to 2019 added 27 GW of RE (mainly wind and solar power), mostly in the three largest markets in the region: Brazil, Mexico, and Chile. These additions represent historical increases in the addition of renewables in the region, as shown in Figure 1.

While long-term contracts at a fixed price are the common design of auctions in Latin America, projects still face

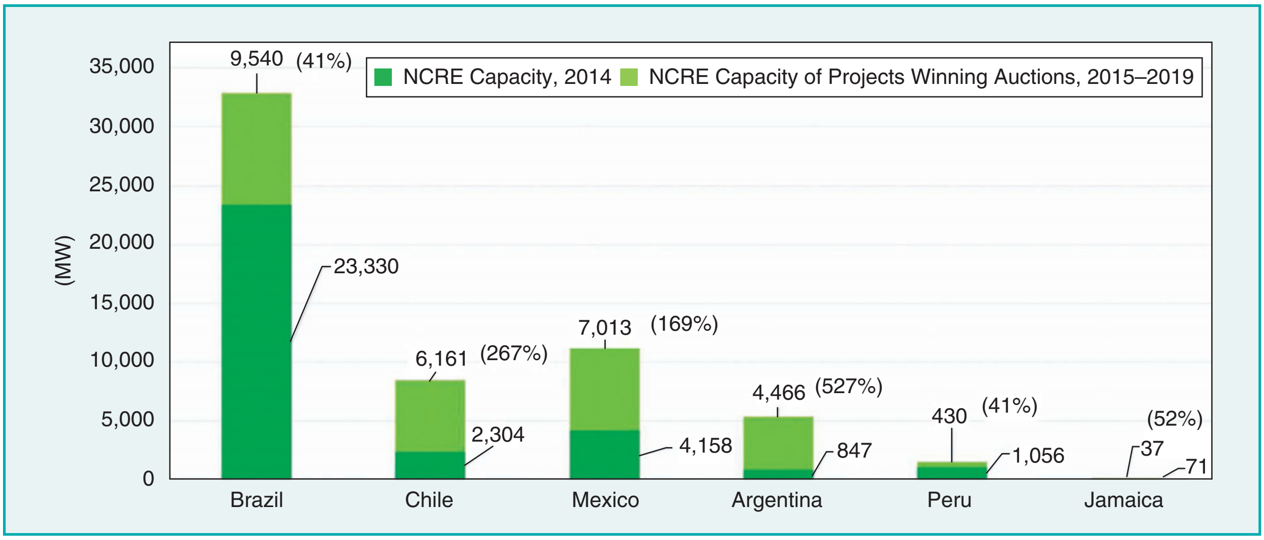


figure 1. The nonconventional RE (NRCE) capacity additions represented by the winning projects in auctions held from 2015 through June 2019 compared to the 2014 renewable capacity in Latin America and the Caribbean. The percentages in parentheses show the increase over the 2014 NCRE capacity, assuming all projects are completed. [Source: Viscidi and Yepez (2019); used with permission.]

A careful balance must be achieved between attracting competitive investments by shifting risks away from investors and achieving low-cost outcomes for consumers in the short and long terms.

other challenges, such as dispatch risks that can result from sudden reductions in demand; increased inflexibility from old fossil fuel contracts or other must-run provisions; delays in transmission deployment; and regulatory rule changes, which are not uncommon. All of these factors can immediately impact both the performance of the contract and any partial merchant revenues from the market. Sudden shocks in demand growth projections have also led to oversupply situations in some markets, such as Uruguay, Peru, and Costa Rica. With these scenarios in mind, the importance of improving the governance and technical quality of auctions seems to be more pressing as the share of renewables increases.

The successful development of RE resources in Latin American countries has been supported by the Inter-American Development Bank (IDB), both through technical assistance in developing the policy and regulatory environment for renewables and through financing projects with a variety of instruments, such as loans, guarantees, and policy support lending. Other results in improving access and reducing emissions are indicated in Figure 2.

From 2016 to 2018, the IDB supported 4,400 MW of new renewable generation in the region, primarily wind and solar power as well as some hydropower. Most of these projects are privately owned and part of the successful auctioning processed in many countries.

Based on extensive experience financing RE projects, the IDB has identified key elements that improve a project's bankability and access to long-term financing. These include an experienced sponsor with strong technical capabilities and financial expertise; a strong business model with a stable revenue stream from a credit-worthy off-taker and competitive price; merchant revenues, if any, projected conservatively for debt-sizing purposes; proven technology and in-depth analysis of output generation based on reliable data; and a stable and predictable regulatory framework.

The deployment of new technologies requires an enabling regulatory framework that ensures revenue for a reasonable timeframe, which could vary for subsequent additions but allows for low-risk/low-cost financing. Technology risks could be mitigated with concessional or blended finance.

The rapid evolution of costs is posing a challenge in different markets. Early PPAs with RE projects are proving to be misaligned with current market prices. A solar photovoltaic (PV) PPA signed in 2016 is almost double the price of a PPA signed in 2020, reflecting the rapid drop in the cost of the technology. Given the increase in efficiency, decreasing prices will continue to be a challenge until prices stabilize.

However, for new technologies, like batteries or other storage solutions, a significant decrease in prices is expected as demand grows and efficiency gains are achieved at the manufacturing and design levels. In these cases, the challenge will be how to cope with the obsolescence risk from both the market designer and investor perspectives. As explained before, to achieve the virtuous process of decreasing costs, low-risk financing needs to be available for such technologies. How can long-term financing be achieved with a potential obsolescence risk? On the other hand, how can regulators prevent assets from being stranded with high prices when new technologies become available?

One possible solution is to structure a buy-back feature in the regulatory scheme. Using this feature, the regulator may include an option to buy the stranded asset at the remaining asset value (nondepreciated value). Thus, a new technology with significantly lower costs could be

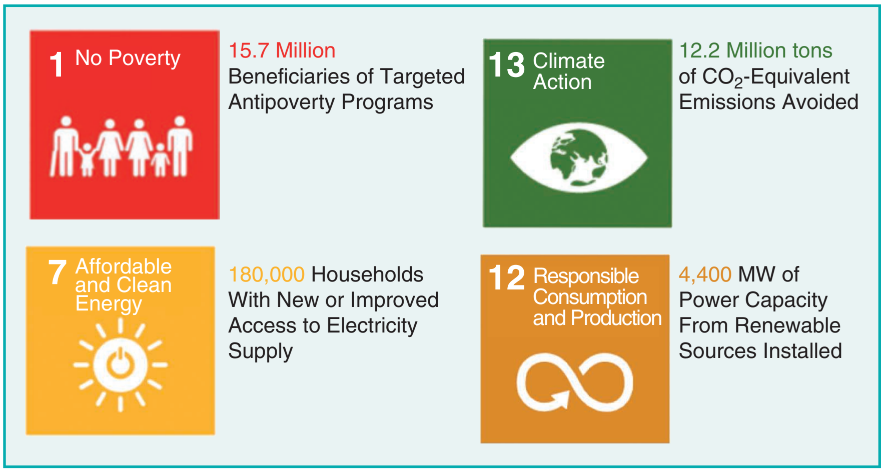


figure 2. Some results of the IDB in financing access to energy (Sustainable Development Goal 7) and cleaner RE supplies (Sustainable Development Goal 12), 2016–2018. [Source: *IDB Development Effectiveness Overview* (2019); used with permission.]

In addition to generation, IFC has invested in 32 distribution companies, which serve close to 60 million customers, and also in six transmission companies.

deployed and replace the stranded asset. The system would benefit from a significantly lower cost of the provided service but would have to repay the replaced asset at the non-depreciated value. The option value should be sufficient to cover the nonamortized value of the loan and equity so the financing risk is duly mitigated. Though proper balancing of the risk borne by developers and consumers is required, these potential innovative market design options should not be ignored if rapid decarbonization and low-cost supplies are the main objectives.

A View From the International Finance Corporation Investing in Renewable Merchant Projects With Decreasing Marginal Costs

The International Finance Corporation (IFC) has financed close to 55,000 MW of private power plants around the world. Nearly 10,000 MW of these investments are in hydroelectric generation, 5,000 MW are in solar PVs, 5,000 MW are in wind, 1,500 MW are in geothermal power, and 350 MW are in concentrated solar power. IFC has also financed several gas-fired power projects, including four large liquid natural gas-to-power projects and two floating-storage regasification unit terminals.

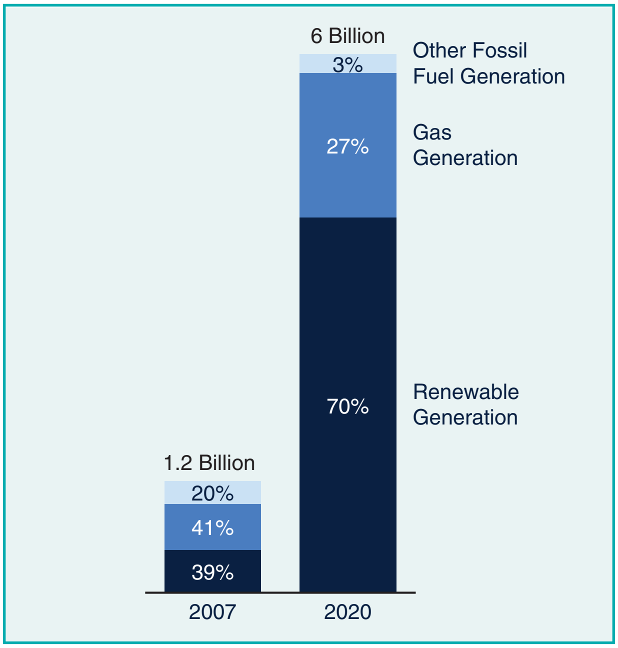


figure 3. IFC's power generation portfolio by subsector for fiscal years 2007–2020, showing the percentage of generation portfolio share. (Source: IFC, www.ifc.org/infrastructure.)

Presently, IFC manages a US$6 billion power generation portfolio (see Figure 3). Since 2007, IFC's power generation business has grown rapidly by fivefold, with renewables representing 75% of this growth. In addition to generation, IFC has invested in 32 distribution companies, which serve close to 60 million customers, and also in six transmission companies.

In most of its transactions, IFC has followed on the reform steps of its sister organization, the World Bank. In the 1990s, IFC successfully invested in the first independent power producer models in several countries and regions, including Mexico (1,700 MW), Pakistan (1,500 MW), the Philippines (1,500 MW), and Sub-Saharan Africa (1,000 MW). In the early 2000s, as the World Bank introduced unbundling and competitive, deregulated wholesale markets, IFC started building a portfolio of 9,400 MW of fully or partially merchant power plants in more than 10 countries around the world (see Table 1).

As Table 1 shows, IFC has a long track record of investing in merchant projects, financing many first-of-its-kind projects across multiple countries. With 9,400 MW financed on a merchant basis, IFC is a leader in investing in merchant projects in developing countries that bet on transparent price signals in deregulated markets. The results of these merchant investments have been largely successful. The next sections comment on the lessons learned from some of these investments and how IFC's appetite for merchant investments is changing with an ever-larger penetration of zero-marginal-cost renewables.

Large Penetration of Renewables and Zero Marginal Cost

In 2013, IFC was among the first to invest in large utility-scale solar PVs in Chile. When IFC made its initial investment, there was only a small, 2-MW solar plant. Today, there are more than 2,200 MW of solar generation in Chile. IFC invested in the first 300 MW. Three of the four projects were merchant, and one had a PPA. Similarly, in Turkey, IFC was the first bank to take merchant risk on 4,000 MW of power generation, with RE projects accounting for half. The other half, invested in, are among the most efficient combined-cycle gas turbines in Europe and are totally flexible since they are not contracted. These plants play an important role in firming up the intermittency of a larger penetration of renewables.

Among the challenges merchant investments faced in these two countries was an ever-larger penetration of renewables, driven in large part by technological advances and the downfall in renewable equipment costs. The additional factors of the collapse in oil prices from US$100/barrel (bl) to US$35/bl, the economic slowdown of China (Chile's largest copper export market), and the recession in Europe (Turkey's largest export market) prompted electricity spot prices to collapse from an average of US$100/MWh to US$40/MWh.

In Chile's case, the situation was exacerbated due to transmission constraints. Renewables were built faster than transmission lines, and the Atacama region, home to Chile's best solar resources, were decoupled from Santiago, the largest demand center. Solar projects experienced zero marginal cost for several years due to delays in transmission construction. This is also a problem in other markets, such as India and China, that have seen large renewable curtailments due to transmission constraints.

With proper due diligence and robust financial structuring, IFC was able to cope with the above shocks or, rather, the size of their impact, technological disruption was perhaps the only aspect that was completely unexpected. This is what IFC calls "first-mover disadvantage." In fewer than five years, solar PV investment costs collapsed from US$2,200/kW to under US$1,000/kW.

Of Chile's four separate electricity market segments, 67% of installed capacity (13 of 20 GW) was in the Central Interconnected System (SIC), serving the country's central zone, which includes Santiago and about 92% of the population. The second-largest system was the Northern Interconnected System (SING), with 20% of installed capacity (4 GW) serving the desert-mining regions in the north. At the time of IFC's investments, up until November 2017, the SIC and SING systems were not interconnected, largely due to social opposition to transmission expansion.

Figure 4(a) shows the spot prices before and Figure 4(b) illustrates the spot prices after the commissioning of the transmission line and its effect on electricity spot prices. At certain hours, this included zero marginal costs in the northern SIC [the yellow lines in Figure 4(a)]. In consequence of the transmission constraints and larger renewable penetration, Investors and lenders in Chile's solar PV projects experienced zero marginal cost in the northern SIC for almost four years.

Lessons From Chile: Technological Disruption and Merchant Investments

The Chile merchant projects were structured based on spot price forecasts from reputable power market consultants, which relied on what were then considered conservative oil price forecasts compared to other estimates in the market. Still, the rapid descent in power prices in the northern SIC system (to zero at some hours in the day for specific network nodes) was not foreseen by anyone in the market. From this experience, several lessons can be drawn.

- ✔ *Cannibalization of spot prices from rapid RE development*: The projections relied heavily on free-market forces, namely, that investors stop investing (and banks stop financing) when faced with prospective overcapacity or transmission constraints. In reality, power producers rushed into the market, encouraged by 1) collapsing solar equipment costs, 2) the availability of relatively inexpensive financing, and 3) specific features of the power market model that made generators indifferent to signing a contract with a distribution company to sell at the node price versus taking full merchant exposure to sell on the spot market. Despite the strong decoupling phenomenon identified (between the two transmission extremes of the SIC), investors kept building solar PV projects (mostly merchant), which rapidly decreased spot prices in the region. IFC's projects were the first 300-MW solar projects in Chile's north, followed by another 2,200 MW.
- ✔ *Transmission curtailment*: Delays in transmission expansion severely impacted node prices. Moreover, renewable projects can be built at a staggering speed compared to the necessary transmission, which increases the risk of a price collapse. Transmission curtailment risk is one of the key issues when financing RE projects around the world. Many provinces in

table 1. IFC merchant or quasi-merchant investments.

Country	Year of First Project	Technology	Number of Projects	Megawatts Installed
Chile	1991	Hydro, solar, thermal, and wind	11	2,017
India	2004	Hydro	1	192
Turkey	2008	Thermal, hydro, and wind	7	4,476
The Philippines	2008	Hydro and thermal	1	1,035
Colombia	2008	Hydro	3	10
Panama	2010	Hydro, wind, and thermal	2	679
Romania	2010	Wind	2	228
Georgia	2011	Hydro	3	267
Peru	2011	Hydro	1	168
Mexico	2013	Solar	3	329
Total			**34**	**9,400**

(Source: IFC, www.ifc.org/infrastructure.)

India and China have also curtailed renewable generation due to transmission issues.

- *Correlation to commodity prices*: Electricity spot prices are heavily affected by commodity price fluctuations. When investing in a high oil price cycle, one should consider the possible effects of a future decline in oil prices.
- *Technological disruption*: Early investors in RE have faced "first-mover disadvantage." The collapse in renewable capital costs has also affected feed-in-tariff schemes in Europe, with many being renegotiated. In all of its new merchant investments, IFC runs a technological disruption scenario. When IFC invested in Chile's first solar projects, the investment cost of PV solar was at US$2,200/kW. Today, PV solar projects can be built for fewer than US$1,000/kW.
- *PPAs are not necessarily the best solution for all problems*: In a multinodal market hampered by transmission constraints, PPAs are not necessarily a better alternative than merchant projects for investors and lenders alike, given the potential for negative contract margins due to transmission constraints. Furthermore, technological disruption is making first-mover PPAs unsustainable for both the producer and off-taker, and those PPAs are quickly becoming out of market, which poses long-term problems for 15–20-year contracts.
- *Capacity payments*: IFC's past success in merchant plant finance, in Chile, was mainly due to the existence of capacity payments. It was the capacity payments that made those projects bankable. (In some of IFC's merchant investments in Chile, close to 30–40% of the revenue came from capacity payments. These capacity payments were guaranteed since they were regulated payments.) With technology disrupting the power sector, capacity payments are at risk, especially for VRE, since they are meant to remunerate plants that provide firm capacity at critical hours only, and as more solar power penetrates the systems, the critical hours keep shifting to the night. A lack of capacity payments means that the structuring approach to merchant plant financing needs to be re-evaluated and refined.

Focus on Asia: The Need for Transition and Exit Strategies

Global coal capacity stands today at just more than 2,000 GW. Almost 70% of this capacity is in Asia, with 1,000 GW in China, 205 GW in India, and another 200 GW in the rest of Asia. There are other significant coal-dominated systems,

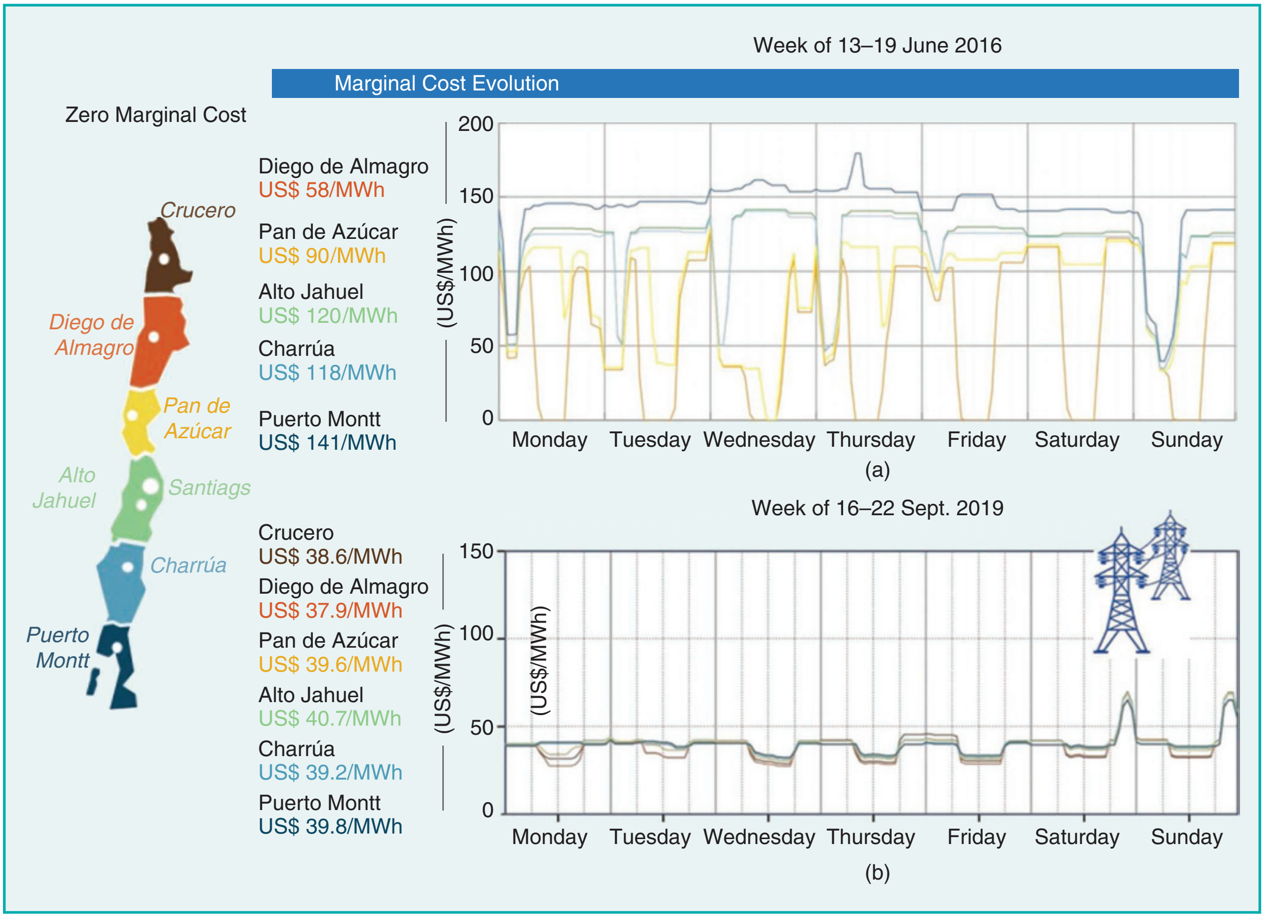


figure 4. Average weekly nodal prices (a) before and (b) after line commissioning. (Source: Antuko Energy.)

> Attracting new investment in competitive-cost generation to supply the demand in a timely manner is a priority objective.

including South Africa (40 GW) and Indonesia (40 GW). Global coal generation increased from ~6,000 to ~10,000 TWh from 2000 to 2015 (and global coal production has increased from 4.6 to 7.3 billion tons over the same period).

However, there is an interesting twist to this growth story in the last five years, as coal generation is widely believed to have peaked around 2014–2015, stabilized since then, and even decreased slightly. The entry of renewables globally, especially feed-in-tariff-driven subsidized renewables in many markets (including China and India), contributed to this stabilization of coal. The average utilization of coal plants globally has dropped from a peak of 66% in 2015 to 51% in 2016, raising concerns around stranded capacity and the rapid retirement of coal assets.

A 2018 report by the Rocky Mountain Institute posits an accelerated phase-out of 200 GW of coal over the next decade and more than 1,200 GW of coal worldwide over the next 30 years (see the "For Further Reading" section). According to the report, 42% of the coal capacity globally could already be operating at a loss—a number that is expected to grow to 56% by 2030.

Coal generation must be reduced substantially to make room for renewables in Asia and parts of Africa. However, this process will be difficult and expensive. There are US$255 billion worth of stranded assets in addition to US$234 billion in decommissioning costs, i.e., a nearly half-trillion-dollar puzzle that needs to be solved to pave the way for RE. To integrate renewables into markets successfully, solutions must be found to difficult and expensive challenges, such as paying for stranded assets, closing mines and plants, undertaking environmental remediation, and developing social programs to compensate/reskill employees.

These solutions would not necessarily lie in the scope of markets, but unless they are addressed somehow, it is difficult to see how the transition from coal to renewables can be accelerated. Many of the long-term contracts that made it difficult for solar/wind to come in or led to significant curtailment (e.g., 20–30% in parts of India and China in recent years) could persist.

There are other challenges in supporting solar and wind that we have alluded to in the preceding discussion, namely, building and, in some cases, upgrading and reorienting transmission and distribution to accommodate variability in solar/wind. These investments can be substantial depending on the level of penetration of solar/wind and will, for the most part, remain in the domain of public finance. Some of the mega transmission projects in India and China have already accounted for US$2–6 billion per project for ±800–1,100-kV ultra-high-voltage (HV) dc and 1,200-kV HVac projects. Integrating the southern, western, and eastern power pools in Africa would also cost tens of billions of dollars over the next decade to ensure the connection of regional scale-efficient hydro and solar/wind.

In general, these upgrades to transmission and distribution are an expensive proposition in most parts of the developing world, from small island nations to very large power systems. Even if these upgrades are undertaken, there remains the issue of flexible generation capacity that needs to accompany solar and wind. While their lower costs make solar and wind attractive propositions to displace expensive thermal energy sources (especially liquid fuel and gas) in some parts of the world, flexible generation—be it storage/pumped-storage hydro, open-cycle gas, imports from another system, or battery storage—can be expensive.

As markets in Europe, parts of the United States, and Australia have demonstrated, although it is possible, to some extent, to attract flexible generation through appropriate capacity and balancing markets, much work remains to be done. As noted before, less mature markets in other parts of the world make renewable integration an even more challenging proposition. Absent adequate transmission/distribution capacity and/or flexible generation, a very large penetration of solar/wind would either not be achieved or, if strongly incentivized with a subsidy, would continue to result in significant curtailment, inefficient dispatch, and grid stability issues.

Potential Solutions: Focus on India

Strategies exist that could accelerate the process of decarbonization, bringing in solar/wind/hydro and other renewables and storage as well as raising the necessary investments. We discuss these in the context of India, which is a significant RE hub poised for rapidly scaling up its renewable base.

The Indian power system has a total installed capacity of 370 GW (as of March 2020), including 205 GW of coal, 35 GW of solar, and 37 GW of wind. There is a renewable policy goal to first get to 175 GW of renewables by 2022, including 100 GW of solar and 60 GW of wind. The medium-term goal is to scale further to 450 GW by 2030. Although the country has made significant strides initially with wind and, more recently, with solar, it has also experienced many of the problems discussed previously.

Institutional capacity is at the core of achieving good results of many market reforms in the power industry of the developing world.

Highly inflexible PPAs force inefficient coal units to run at 55–70% minimum loading, although 62% or ~127 GW of the existing coal capacity has operating costs greater than those of new renewable projects as of 2018. As coal PPAs dominate generation, the day-ahead market (DAM) has extremely low liquidity of 4% after nearly 12 years of operation. There has been no participation of solar/wind in the market. There are close to 50 GW of hydro in the system, but the larger storage units are multipurpose projects with restrictions on their ability to provide ancillary services. There is limited flexibility in the system and no ancillary services market set up to incentivize entry of new capacity.

The interstate transmission system is very robust, with more than 100 GW of capacity already in place, including ±800 kV of HVdc and a 1,200-kV HVac line under construction. However, intrastate transmission capacity has been a bottleneck, and renewable curtailment in some states, especially wind in the south, has been triggered by such constraints. Regional electricity trade has been well below 1% of the annual regional electricity demand consistently for more than a decade, making it one of the least connected regions in the world.

Going forward, here are suggestions for possible solutions to address these issues.

- ✔ A structured coal plant retirement program needs to be developed by expanding on the 25-GW coal plant retirement (by 2027) slated in the current government plan. Coal plants can be repurposed wherever possible to retain or even enhance the dynamic reactive power and inertia services, e.g., through conversion of the generator to a synchronous condenser. There is also the possibility of using part of the site to install RE generation (e.g., solar PVs) and battery storage. The World Bank is currently developing an accelerate coal transition facility to develop these ideas and finance these projects initially in India and South Africa.
- ✔ The average solar and wind contract prices have fallen below US$0.044 per kWh (₹3/kWh) since late 2019, making them lower than the operating cost of many of the existing coal plants. Our analysis of 15-min wholesale spot prices in the DAM over the past five years suggests that solar and wind would have been competitive in the market. It would make sense to bring solar and wind—at least part of the new VRE projects—to bidding in the market to enhance liquidity. It would also be important to develop a market-based mechanism (e.g., contract for differences) to bring some of the thermal capacity currently on long-term PPA into the market. These issues are being actively discussed by the Ministry of Power in India as part of a road map to redesign the future electricity market. India needs a minimum of US$10 billion of new investment in solar and wind annually over the next decade to meet its renewable policy targets by 2030. Active market participation in a liquid DAM and real-time market that started on 3 June 2020 is expected to facilitate such investments to be led by the market mechanism.
- ✔ An ancillary services market, co-optimized with the energy market, is also recognized as a critical component to ensure new-storage hydro, pumped-storage hydro, open-cycle gas turbine, and battery storage can all be developed in a market-oriented way. A World Bank analysis of historic DAM prices suggests that a combination of ancillary services payments at US$2/MWh and arbitrage in the DAM would render a battery electric storage system moderately attractive, and at US$4/MWh, it would be a reasonably strong business case.
- ✔ Finally, the regional dimension of the market is also critically important, as India can catalyze faster development of hydro in Nepal (with an economic potential of 42 GW) and Bhutan (25 GW) than it can import to support solar/wind. In turn, India can also export power to countries like Bangladesh and Sri Lanka that have limited primary energy supplies. Given the complexity of developing large-scale hydro in India, regional hydro projects can be a potent option. One of our analyses indicates that every gigawatt of hydro, even with limited storage, can support the variability of at least 4 GW of solar and wind in India through the provision of spinning reserve. We also find that a combination of cross-border hydro, complemented by battery storage for additional frequency-control ancillary services, can be an important part of the decarbonization strategy in the region. Developing additional cross-border transmission infrastructure rapidly would be a core part of this strategy, and this will require up to US$2 billion in the short to medium term until 2030, followed by additional investments to form a deeply integrated ac network. There is a thriving economic case for these projects supported by multiple studies conducted by the World Bank, U.S. Agency for International Development, and the Asian Development Bank.

Since no market design can solve structural concentration issues, structural changes, such as the separation of functions or legal limits to market share, should be priority activities.

Conclusions and Recommendations From Developing Banking Practitioners

Implementing market reform and the appropriate market design for developing countries needs to be viewed in the overall context of the institutional capacity, market structure, social acceptability of fluctuating prices, and many competing objectives, such as effectiveness in bringing needed supply and the efficiency of the different mechanics. With that said, some general messages and recommendations from more than 30 years of market design in LAC can be drawn, looking forward to grids with ever-growing shares of renewables.

Market Design and Institutional Design Should Be Considered Together

Institutional capacity is at the core of achieving good results of many market reforms in the power industry of the developing world. Market design should always be adapted to local realities and consider the needed institutional arrangements and capacities to execute, supervise, and generate social acceptance of the implementation and outcomes of a market design. Price fluctuations at the retail level tend to be less socially acceptable in developing countries since expenditures in energy can represent a larger share of consumers' total expenditures and also due to other local social and political realities.

Market design should not be separated from the institutional design process and other regulatory aspects, such as the effectiveness of price formation at the retail level. Some market designs, e.g., energy-only markets, would be more difficult to implement in a context where retail price stability is an objective and attracting low-cost investments is a primary objective.

Prioritize Design Options Based on Specific Objectives

Over the past few years, developing countries faced rapid demand growth. Attracting new investment in competitive-cost generation to supply the demand in a timely manner is a priority objective. Some market designs are more effective at achieving such objectives than others, and a clear set of goals should be used to weigh different options. If demand growth were zero or negative, as happened over the past few years in leading developing countries with more mature market designs, then the efficiency objective may be more important. Usually, design options will be weighed based on effectiveness versus efficiency objectives, which run counter to each other in many cases. Therefore, defining and prioritizing a clear set of measurable objectives is critical in the choice of market design process.

Do Not Underestimate Structural Limitations and the Need for Fair Play

Structural limitations, such as vertical and horizontal integration or a high level of concentration in any of the segments such as generation and retail sales, make it difficult to achieve results using many market design models. Since no market design can solve structural concentration issues, structural changes, such as the separation of functions or legal limits to market share, should be priority activities. Likely, the necessary changes include full separation and independent governance of system operation functions, which have not been fully achieved in many developing countries.

In addition, where government-owned entities such as utilities are key actors, the importance of institutional governance is of utmost importance. It is necessary to ensure that state-owned entities have private-like orientation or at least the public sector goals are stated clearly and financed properly, without affecting the level playing field, proper efficient cost recovery, and price formation in wholesale markets.

Transmission Continues to Be a Bottleneck

In many developing countries, transmission is the least costly part of electricity infrastructure. Despite the fact that practices have been put in place to speed up transmission investment via, for example, better planning, the expansion of grids in a timely and nondiscriminatory fashion continues to be a recurrent challenge in many markets. Wholesale markets need robust transmission grids to function properly, especially with a large share of renewables. It is crucial to maintain a pace of investment aligned with the growth of generation and continuously improve access procedures. Transmission should be seen as a barrier that needs to be reduced as fast and cost-effectively as possible, with clearer and more effective expansion and execution instruments via multiple actors as well as simple, yet effective, cost-recovery mechanisms.

Stable Revenues via More Flexible Contracts Continue to Be a Key Design Feature

Capacity payments, such as those in Chile, provide a secured revenue stream for peaking power. Long-term contracts also provide such stability for energy and capacity projects. Markets may reduce costs by continuing to attract investments with a long-term amortization period. In developing markets that have large price-regulated retail segments, this will require continuing to provide long-term contracts, gradually developing flexible contracting mechanisms to cope with obsolescence risk, improving wholesale competition, and better managing supply security.

Some of the key features will still include a mechanism of competition for the market that continues to rely on planning to determine contracting levels and reliability requirements, coupled with better tools and governance to avoid over- or undercontracting. In addition, there should be contracts that are more flexible than standard PPAs, with different durations but certain standardization that makes them understandable and bankable for both regulated and unregulated consumer segments. System flexibility and variability, while not yet a generalized challenge for many developing countries, could be solved by allowing system operators to use new tools, such as storage, directly. In some cases, this may be preferable to a solution that relies on the perfect disaggregation of all ancillary services or requires storage to be a fully unregulated activity.

In countries where coal or other fossil-fuel generation still dominates, a clear and financed strategy to decommission assets may be needed to increase the penetration of renewables. Many countries, especially in Asia, relied heavily on long-term PPAs to attract needed investment in generation. Even at low levels of RE penetration, inflexibility in some of these contracts (similar to must-run provisions) are increasing the curtailment of renewables.

While renewable costs continue to quickly decrease, the risk of curtailment could derail efforts toward increasing renewable penetration. In addition to providing appropriate market design via ancillary services and other regulatory instruments to foster a flexible system, a specific transition and exit strategy may be needed to deal with premature plant retirements if RE goals are to be met. This should include appropriately financed transition plans, when appropriate, to deal with such retirements.

Accelerate Demand and Distributed Solutions, Including Electric Mobility

The industrial load in some developing regions represents a particularly large share of electricity demand, given low household demand. Such structural conditions can be helpful to foster demand-response programs to help improve system flexibility. Demand management will be key to manage large shares of renewable generation, especially in more extreme climates, such as hot countries close to the equator, where the demand for cooling is a large driver of demand growth. Distributed-generation and demand-management solutions may be more valuable in such contexts.

In addition, the rapid electrification of electric public transportation—already taking off in some cities in the developing world, especially in China, India, and some countries in Latin America—and its full two-way integration with the grid could become a real and sizable tool to manage variability. Thousands of electric buses already on the streets represent hundreds of megawatthours in electricity storage whose value to grid flexibility is waiting to be tapped. Wholesale market design also needs to be cognizant that a single centralized wholesale market will need to accept, rather than push away, other forms of energy trade that are less centralized and closer to the consumer.

Acknowledgments

The findings, interpretations, and conclusions expressed in this article are entirely those of the authors and should not be attributed in any manner to the institutions the authors work for or to their boards of directors and the countries they represent: IFC, the World Bank, IDB Invest, and IDB.

For Further Reading

F. Carvajal et al., *Brechas de Acceso a los Servicios de la Energía en América Latina y el Caribe*. Washington, D.C.: Banco Interamericano de Desarrollo, 2020.

V. Foster and A. Rana, “Rethinking power sector reform in the developing world,” World Bank, Washington, D.C., 2020. [Online]. Available: https://www.worldbank.org/en/topic/energy/publication/rethinking-power-sector-reform

J. Millan and M. von der Fehr N, *Keeping the Lights On: Power Sector Reform in Latin America,* N. Fehr, Ed. Washington, D.C.: Inter-American Development Bank, 2003.

S. Mocarquer, L. Barroso, H. Rudnick, B. Bezerra, and M. Pereira, “Energy policy in Latin America: The need for a balanced approach,” *IEEE Power Energy Mag.*, vol. 7, no. 5, pp. 26–35, 2009. [Online]. Available: https://www3.ing.uc.cl/publicaciones/balance-of-power-httpdx-doi-org10-1109mpe-2009-933417. doi: 10.1109/MPE.2009.933417.

A. Benn, P. Bodnar, and J. Mitchell, “Managing the coal capital transition,” Rocky Mountain Institute, Basalt, CO, 2018. [Online]. Available: https://rmi.org/managing-coal-capital-transition/

L. Viscidi and A. Yepez, “Clean energy auctions in Latin America,” Inter-American Development Bank, Washington, D.C., 2019. [Online]. Available: https://publications.iadb.org/publications/english/document/Clean_Energy_Auctions_in_Latin_America.pdf

“Development effectiveness overview (DEO) 2019,” Inter American Development Bank, Washington, D.C., 2019. [Online]. Available: https://publications.iadb.org/en/development-effectiveness-overview-deo-2019. doi: http://dx.doi.org/10.18235/0001797.

Biographies

Tonci Bakovic is with International Finance Corporation, Washington, D.C., USA.

Debabrata Chattopadhyay is with the World Bank, Washington, D.C., USA.

Fernando Cubillos is with IDB Invest, Washington, D.C., USA.

Marcelino Madrigal is with the IDB Invest., Washington, D.C., USA.

Adam Allerhand

charles f. brush

a pioneer of electric lights

This issue's "History" column explores Charles F. Brush's contributions to dynamos, arc lighting, and central stations. It addresses Brush's competition at the time and describes some of his activities after Thomson–Houston Electric took control of Brush Electric in 1889. It is a welcome treatment of an electric business pioneer who has not had much published about his life and work.

We welcome back Adam Allerhand for his third appearance on our pages. He is professor emeritus of chemistry at Indiana University Bloomington and listed among the 1,000 most-cited scientists in the 1965–1978 period. He is a fellow of the American Association for the Advancement of Science since 1987; the author of roughly 100 research papers, mostly in the field of nuclear magnetic resonance spectroscopy; and the author of *An Illustrated History of Electric Lighting* (2016).

John Paserba,
Associate Editor, "History"

THE AUTHOR OF AN ARTICLE OR a book about a single person is in danger of focusing heavily on that person's accomplishments and neglecting the broader picture so that the result is a hagiography. I will try not to fall into that trap. I start out by quoting, in a condensed version, from Harold C. Passer's superbly researched *The Electrical Manufacturers: 1875–1900*, published in 1953:

> The beginning of the arc-lighting industry can be divided into three stages. In the first stage, many inventors tried to develop a practical system, without succeeding. In the second stage, one of the many engineer-entrepreneurs carried out the act of innovation that founded the arc-lighting-equipment industry. That pioneer innovator was Charles Brush. He combined in one person the commercial ability to envision where arc lighting could be used successfully and the technical ability to invent an arc-lighting system that fitted his commercial vision. Most cities of any size were buying light for their streets from gas companies, a potential market for arc lamps. With this commercial vision before him, Brush solved the technical problems. The result was the spectacular success of the Brush Electric Company. In the third stage, engineer-entrepreneurs entered arc lighting after Brush had shown that it could be commercially successful. One of these entrepreneurs was Elihu Thomson. The Thomson–Houston system was probably the most reliable and lowest in operating cost when it was introduced.

Digital Object Identifier 10.1109/MPE.2020.3032804
Date of current version: 6 January 2021

After its introduction, in 1880, incandescent electric lighting gradually began to displace gas lighting in small indoor environments, but it failed to make inroads into arc lighting for streets and large indoor areas before high-intensity light bulbs were developed. In 1887, fully one-half of the generating capacity of the central stations in the United States came from arc light dynamos. In 1902, only 386,000 arc lamps were sold in the United States, compared to approximately 18 million incandescent bulbs. But that comparison does not present the whole picture from the standpoint of revenue, because roughly 170 million arc carbons were consumed that year, with a total price tag of some US$5 million, with each carbon burning for only about 8 h, compared to hundreds of hours per incandescent bulb. My estimate of the revenue from the sale of 18 million incandescent bulbs in 1902 is also approximately US$5 million. Both figures are based on prices in a 1904 General Electric catalog.

Background

Alessandro Volta's invention of his electrolytic voltaic pile, announced in a letter read at the Royal Society

1540-7977/21©2021IEEE

in London, on 26 June 1800, revolutionized electrical research. The significance of Volta's invention—the production of a sustained electric current—was recognized immediately by many researchers. One of them was Cornish chemist and inventor Humphry Davy. In September 1800, he showed that charcoal acted as an electrical conductor and produced a "spark when made a medium of communication between the ends of the galvanic pile of Signor Volta." Davy did not suggest the use of this phenomenon for illumination. He would do that a decade later. In 1803, Russian physicist Vasily Petrov produced light by connecting closely spaced pieces of charcoal to the terminals of a very large battery bank. In 1810, Davy demonstrated, at the Royal Institution in London, a primitive arc lamp powered by a battery bank of 2,000 cells, which had a total surface of 83 m^2. More practical arc lamps would be invented after the introduction of steam-powered magnetoelectric generators during the 1840s. Early viable arc lamps used complicated clockwork mechanisms and other components to keep the gap between the tips of the two carbon rods constant as the carbons burned away gradually. When Brush generators, regulators, and arc lamps were being installed in 1878, dynamo electric generators had largely supplanted magnetoelectric ones.

Charles F. Brush

Charles F. Brush (Figure 1) was born on a farm near Cleveland, Ohio, on 17 March 1849. I have found not a single published book-size biography of this pioneer of electric lighting and central stations. An unpublished rough draft written around 1950 and an unpublished Ph.D. thesis dated 1967 can be accessed at digital.case.edu, a website of the Kelvin Smith Library at Case Western Reserve University, Cleveland. As a teenager, Brush experimented with batteries, magnets, telescopes, and microscopes. After high school, he enrolled at the University of Michigan, Ann Arbor, where he graduated, in 1869, with a degree in mining engineering. He moved to Cleveland, where he organized a chemical analysis business and, for a short time, engaged in the marketing of iron ore and pig iron. His start as an inventor of dynamos and arc lamps was facilitated by a childhood schoolmate, George W. Stockly, vice president and manager of the Telegraph Supply Company of Cleveland, a manufacturer of telegraph equipment and other electrical devices. I obtained the following timeline of contracts between Brush and Stockly's company from digital.case.edu.

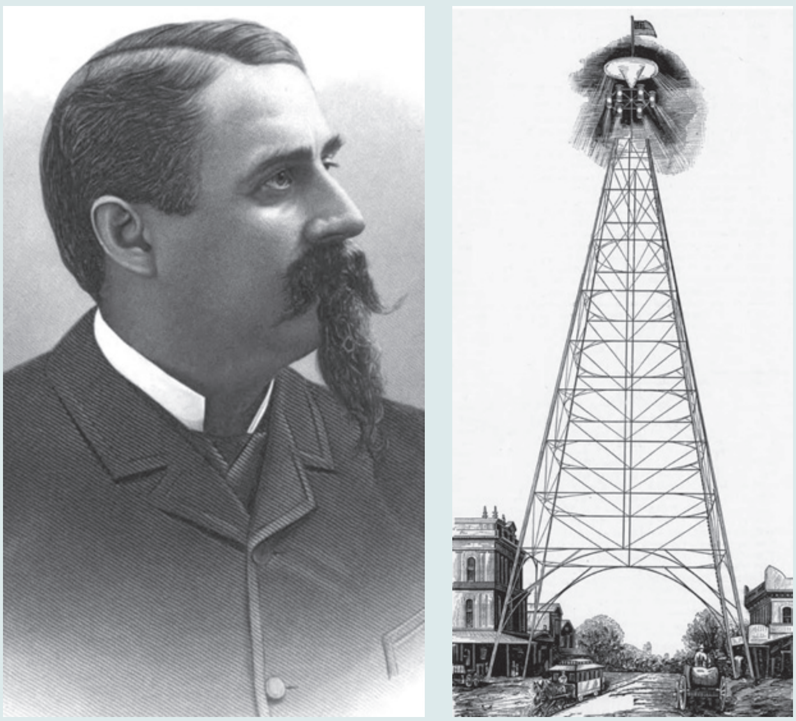

(a) (b)

figure 1. (a) Charles F. Brush. (Source: "Charles Francis Brush," *Scientific American Supplement*, vol. 18, p. 7,287, 1884.) (b) A Brush arc light tower in San Jose, California, in 1882. Brush Electric recommended placing its arc lights on tall towers, thus signaling that the lights were too bright for general use close to the ground. (Source: "Tower system of electric lighting," *Mechanics*, vol. 1, pp. 292–294, 1882.)

Stockly was impressed with Brush's ideas about the commercial possibilities of electric lighting, and he encouraged Brush to design and build a dynamo, which Brush did, in 1876. An agreement, dated 7 June 1876, granted the firm the exclusive right to manufacture and sell Brush dynamos and all the improved forms of that machine. The company assumed financial responsibility for development and marketing and for the expenses of patent applications. Brush would receive a royalty of 20% of the selling price for anything that was his invention. To preserve quality, Brush had to approve any deviation in the machine designs he supplied to the company.

The arrangement was modified in a memorandum of understanding, dated 1877, because production and sales were slow and the royalties too meager for Brush to live on. Brush would receive US$150 a month as an advance on his royalties and, in return, work full time on electric lighting development. A more detailed one-year agreement was signed on 24 March 1877 and extended for another year on 3 April 1878. A supplemental agreement between Brush and the Telegraph Supply was signed on 1 July 1880. On 19 August 1880, Telegraph Supply changed its name to Brush Electric because it had become apparent that its business would be mainly the manufacture and sale of Brush electric lighting equipment. An agreement between Brush and the company, signed 27 July 1886, gave Brush a one-time payment of US$46,666.67 and company stock with a value of US$500,000, in return for a full transfer of patent rights to the company.

Brush Electric

Telegraph Supply was incorporated, in 1875, with a capital of US$100,000, with Stockly as vice president and Mortimer D. Leggett, a former major general in the Union Army during the Civil War and a previous U.S. commissioner of Patents, as president. In May 1880, the company's factory was destroyed by fire, and nothing was saved but the contents of the iron safe. A larger building, but still a very small one, was soon found and equipped with manufacturing machinery. The plant churned out dynamos, arc lamps, and carbons at a rapid pace, and sales were brisk. Business increased so rapidly that, within a few months, it was clear that the 1880 facility would be too small. The company purchased a six-acre lot, on which a larger plant was completed in 1881 (Figure 2). The same year, the capital of the now Brush Electric Company was increased to US$3 million.

Revenue came from sales to private users and subsidiary organizations formed for furnishing lights from a central station and from sales of patents rights. The 1881 plant almost doubled in size by 1886. Stockly was president by then. A company report (in *American Electrical Dictionary for 1886*) that year claimed that the factory was "the most extensive electrical works in the world" and that "it enjoyed the lion's share of the arc light business." Figure 3 shows the very large Brush factory, circa 1888. In 1889, Thomson–Houston acquired control of Brush Electric. The circumstances that led to

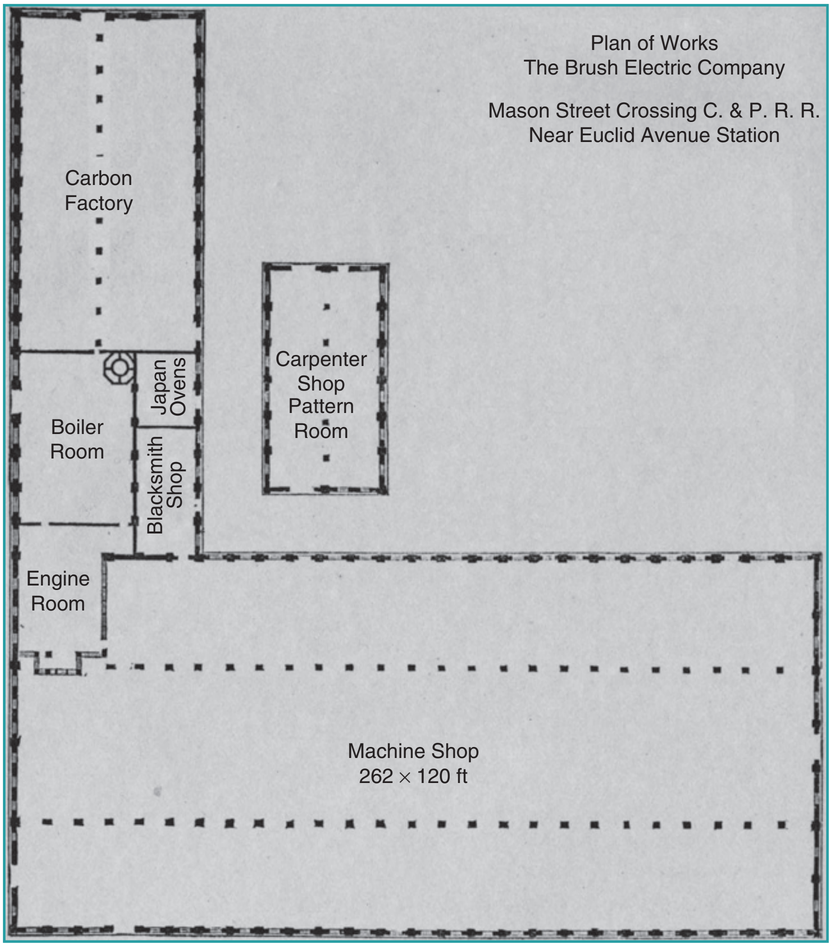


figure 2. The floorplan of the third Brush factory, which was completed in 1881. (Adapted from *The Brush Electric Light*. Cleveland, Ohio: Brush Electric Company, 1881.)

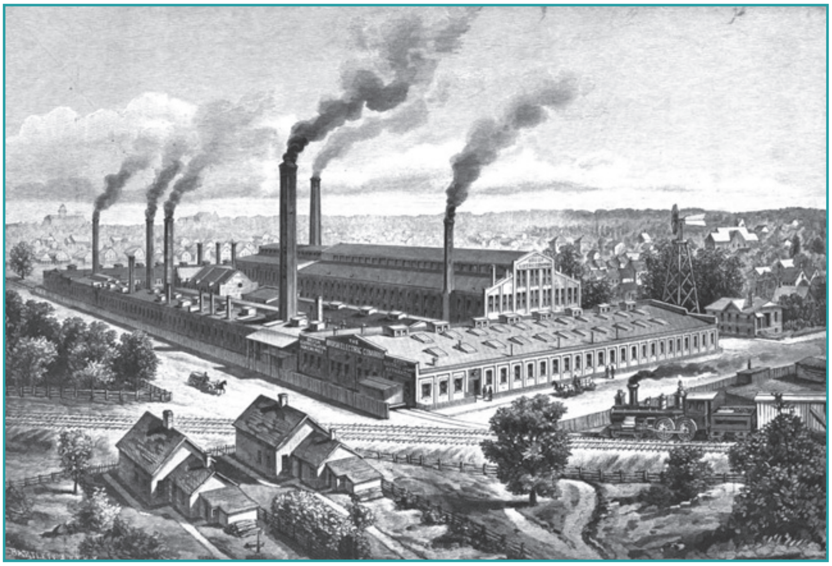


figure 3. The Brush Electric factory in 1888. (Source: *Dynamo-Electric Machines for Arc and Incandescent Lighting, Electric Motors, Electro-Plating Machines, Storage Batteries, Carbons, etc*. Cleveland, Ohio: Brush Electric Company, 1888.)

this takeover will be presented in the "Competition" section.

Brush Dynamos

Before 1877, manufacturers of arc lighting dynamos tested their machines mainly to ascertain that the devices would make arc lamps produce satisfactory light. Testing for efficiency was not a high priority, and little was known about the relative efficiencies of the different makes of dynamos. In the fall of 1877, the Franklin Institute, in Philadelphia, began systematic tests of Brush, Gramme, and Wallace–Farmer dynamos. Siemens dynamos were not tested because a machine was not made available. The Brush and Wallace–Farmer machines included arc lamps provided by those companies, but, as noted in *Journal of the Franklin Institute*, the Franklin Institute "quickly established the suitability of the Brush lamp as the source of light for all the machines." At the end of the trials, in the spring of 1878, a committee evaluated the results and concluded that "the small Brush machine, though somewhat less economical than the Gramme machine, or the large Brush machine, for the general production of light and of electrical currents is, of the various machines experimented with, the best adapted for the purposes of the Institute for the following reasons: It is admirably adapted to the production of currents of widely varying electromotive force and produces a good light.… It possesses great ease of repair."

Zénobe Theophile Gramme (1826–1901), born in Belgium, invented a dynamo more suitable for industrial uses than earlier ones. In partnership with French electrical engineer Hippolyte Fontaine (1833–1910), he founded Société des Machines Magnéto-Électriques Gramme, in France, in 1871, which began manufacturing the dynamos in 1872. Brush decided to design a dynamo more suitable for commercial arc lighting than the Gramme machine. He built his first dynamo in 1876 at his father's farm, with components made at Telegraph Supply and with other parts. In the absence of a steam engine, he attached the device to a horse-drawn treadmill used for sawing wood. It worked. He filed a patent application on 11 November 1876. Patent No. 189,997 was issued 24 April 1877.

Brush designed his dynamos (Figure 4) for optimally feeding a constant current to multiple arc lamps in series. The armature, commutator, and arrangement of the field magnets (Figure 5) all differed from Gramme's design. Gramme's circular armature rotated on a spindle between field magnet poles above and below, exerting a strong inducing action upon only the outer edge of the armature. Brush used two U-shaped field magnets facing each other across a gap in which a circular armature rotated, so both sides of the armature were exposed to the inductive influence of the field magnets.

Brush's "open-coil" armature was unprecedented. His bobbins of wire were not connected in a single circuit (closed coil). Instead, only each pair of diametrically opposite bobbins was connected, with the two free ends of the conductor thus formed attached to diametrically opposite segments of the commutator. Each pair of bobbins was independent of the other pairs. The popular No. 7 dynamo had four pairs of bobbins and a commutator with four separate rings of metal to accommodate them; each ring consisted of two nearly semicircular segments separated by gaps designed to enable the proper functioning of the current-collecting brushes (Figure 5). All dynamos prior to Brush's used closed-coil arrangements.

The sophistication of Brush's invention, when he was 27 years old and with a degree in mining engineering, was made apparent at a meeting of the American Institute of Electrical Engineers on 22 May 1891, at which the lecture "A Study of an Open-Coil Arc Dynamo" was presented. The resulting 14-page article, which contained many experimental data, failed to explain exactly why the open-coil dynamo was so successful for powering arc lights. The following is an excerpt:

> Of all the dynamo machines in use at the present day, perhaps the internal action of none is so little understood as that of the arc lighting machines of the open coil armature class. Much concerning the regulation and general behavior of these machines seems utterly at variance with what one would naturally expect

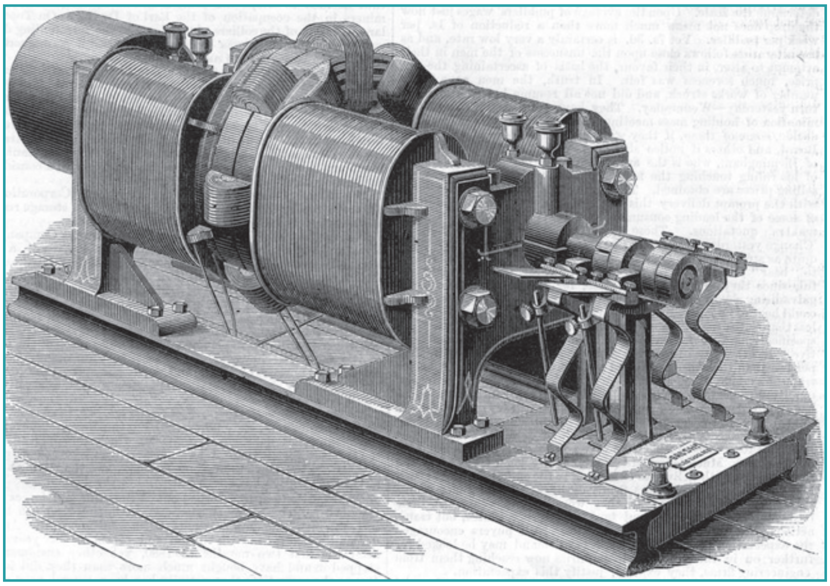


figure 4. The Brush 40-light No. 8 dynamo of 1880. It was 89 in long × 28 in wide × 36 in high, weighed 4,800 lb, and used with 2,000-candlepower lamps in series. (Source: *The Brush Electric Light*. Cleveland, Ohio: Brush Electric Company, 1881.)

from a superficial examination of the design and construction, and it is with the idea of throwing some light on this seeming mystery that the investigation to be described was undertaken.… Is the machine well suited for arc lighting? The vast numbers of these dynamos in daily use in all parts of the world is a practical answer which must carry more weight than any which might be suggested by a theoretical study of the machine.… It is to be regretted that time did not permit the measurement of the exact electromotive force in a separate coil of the armature …

Brush Arc Lamps

Brush filed his first patent application for an arc lamp on 28 September 1877; patent 203,411 was issued on 7 May 1878. Brush arc lamps were very successful because of their simple construction, which included no clockwork (Figure 6). Up to 16 or more Brush arc lamps could be operated in series, powered by the Brush No. 7 dynamo, with each carbon set lasting roughly 8 h. Lamps with two or three carbon sets for unattended 16- or 24-h operation were available, in which a Brush-invented mechanism automatically switched from one carbon set to another. It is noteworthy that the Brush No. 7 dynamo and larger ones had to be high-voltage generators to power arc lamps in series. The voltage drop across a Brush arc lamp was typically about 50 V, so a 16-lamp system would require an 800-V dynamo. The 40-light No. 8 dynamo (Figure 4) operated at 2,000 V.

Brush also developed carbons that were more suitable for the large-scale use of arc lighting than existing ones. Thin carbon points provide better illumination than thick ones, but they were not practical before Brush's invention because of their high resistance and rapid consumption. Brush solved the problem by copper-coating the carbons (U.S. patent 196,425, filed on 21 August 1877 and issued on 23 October

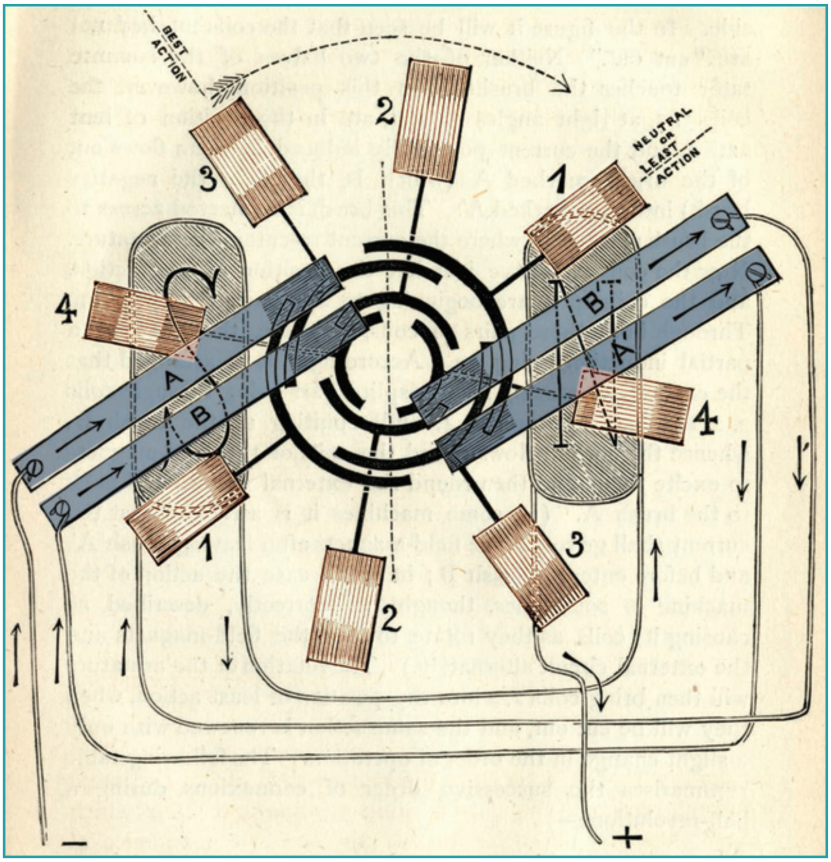


figure 5. The wiring diagram of the Brush No. 7 dynamo. (Adapted from Silvanus P. Thompson, *Dynamo-Electric Machinery*, 2nd ed. London: E. & F.N. Spon, 1886.)

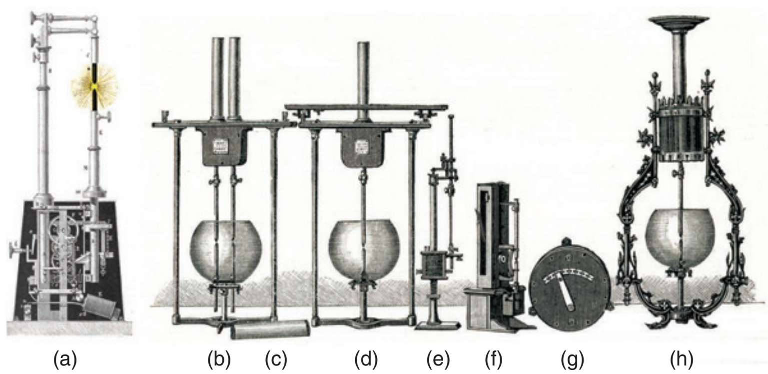


figure 6. (a) The Victor Serrin arc lamp, circa 1860. (Source: A. Guillemin, *The Application of Physical Forces*. London: Macmillan, 1877.) Complex arc regulators, such as Serrin's, became obsolete when Brush arc lamps entered the market. Brush arc lamps and accessories in 1880: (b) an arc lamp with two carbon sets, (c) a package of 25 carbons, (d) a lamp with a single carbon set, (e) a focusing lamp for projections by magic lanterns, (f) a headlight lamp for use in reflectors on ships and locomotives, (g) a dial attachment for controlling how many lamps in series are burning, and (h) an ornamental lamp. (Source: *The Brush Electric Light*. Cleveland, Ohio: Brush Electric Company, 1881.)

1877). An article about the history of arc light carbons published, in *Electrical World* on 5 January 1891, stated that "the present high position of the American carbon industry is due to two men, the celebrated inventor, Charles F. Brush, and Washington H. Lawrence, one of the founders with Mr. Brush of the Brush Electric Company, and now president of the National Carbon Company, Cleveland, Ohio." In 1877, Brush and Lawrence began a search for the best raw material for making arc carbons. They found it within a mile of the Brush factory, in the last product of the distillation of crude oil, which formed at the bottom of the stills of Standard Oil. This nonconducting refuse became a successful commercial arc carbon material after it was converted to a conducting state.

Brush Central Stations

The concept of a central station serving multiple customers would have been familiar where distributed gas service was available. Indeed, *Scientific American* published the following comment in 1857:

> From a grand reservoir of batteries, the electric fluid could be supplied through insulated wires to work engines in every part of a city, in the same manner that gas is furnished to support illumination in stores, houses, and workshops.

Table 1 presents the earliest documented electric central stations, all erected in the 1878–1880 period. First, some comments about the non-Brush stations. Electric streetlights powered from central stations made their debut in Europe, in 1878, in Paris and London. They used low-intensity arc lamps called *Jablochkoff candles*, which were invented by Russian electrical engineer Pavel Nikolayevich Yablochkov (Павел Николаевич Яблочков), who resided in Paris. Manufactured by Société Générale d'Électricité, the Jablochkoff candle was publicly introduced in 1877. It crossed the Atlantic but did not thrive in the United States. Jablochkoff candles lacked commercial staying power mainly because each one lasted only about 2 h, and no automatic mechanism for changing them had been developed. Table 1 provides clear evidence that Brush Electric was the pioneer of long-lasting commercial central station electric lighting.

Brush arc lights were initially used in isolated plants in 1878. More than 600 lamps for isolated plants were sold in 1879 for use in manufactories, mills, mines, hotels, stores, parks, steamers, seaside resorts, and places of similar character throughout the United States and other countries. I am omitting the story of the improved lead-acid battery Brush invented for use as an accessory in central stations when demand for power was low because Brush's batteries were not widely adopted.

April 1879: Cleveland

Cleveland was the logical target for getting approval from municipal authorities to replace gas streetlights, for the first time, with arc lights. Telegraph Supply installed the first electric streetlights in the United States in Cleveland in April 1879. A Brush No. 7 dynamo fed power to 12 arc lights that replaced 105 gas lights at Monumental Park (now Public Square) under a contract with the city, giving more illumination at a lower cost.

September 1879: San Francisco

The Brush series arc lamp system powered by the No. 7 dynamo was used by the central station installed in San Francisco in 1879, thought to be the first central station for multiple customers in the world. Canadian George H. Roe and partners incorporated the

table 1. Central stations for electric lighting systems in 1880.

				Lamps		
Year	**City**	**Location of Lights**	**Dynamos**	**Type**	**Number***	**Purpose**
1878	Paris	Streets	Gramme	Jablochkoff	142	Streetlights
1878	London	Thames embankment	Gramme	Jablochkoff	20	Streetlights
1879	Cleveland	Monumental Park	Brush	Brush arc	12	Streetlights
1879	San Francisco	Private buildings	Brush	Brush arc	50	Multiuse
1880	Montréal	Wharf	Brush	Brush arc	21	Wharf lights
1880	Grand Rapids, Michigan	Private buildings	Brush	Brush arc	18	Multiuse
1880	Detroit	Private buildings	Brush	Brush arc	16	Multiuse
1880	New York City	Broadway and an armory†	Brush	Brush arc	20	Multipurpose
1880	Menlo Park, New Jersey	Edison's laboratory and vicinity	Edison	Incandescent	~500	Experimental

*Initial number.
†Initial locations.

California Electric Light Company of San Francisco on 30 June 1879 and began offering Brush arc lighting to subscribers in September of that year. The initial station had a capacity of only 50 lights when a fire destroyed it on 24 April 1880. The company erected a new station at a different location almost immediately after the fire. In fewer than two years, demand exceeded capacity, and another station was added, which was enlarged in 1885. Yet another plant had to be added a few years later. The company had approximately 2,000 arc lamps operating in San Francisco in 1890 and by then had lighting plants in 17 other cities. After a series of mergers and acquisitions, the systems of the California Electric Light became part of Pacific Gas and Electric, formed in 1905.

Brush Electric was the pioneer of long-lasting commercial central station electric lighting.

June 1880: Montréal

The Board of Harbor Commissioners of Montréal evaluated various systems of electric lighting to ascertain if any of them might be advantageous for illuminating the Montréal wharves to enable cargo to be conveniently handled at night. After comparing the estimated cost with that of other systems, the board asked the Brush Electric Light Company of New York to furnish apparatus for lighting the central part of the harbor, on the condition "that the Board might purchase the apparatus or not as might seem fit after sufficient trial." Lighting began in June 1880, "and after seven weeks use, the apparatus was purchased, together with an increased number of lamps, so as to reach additional wharves, and the lighting was continued with satisfactory results till the close of navigation." One Brush dynamo powered a single 4.4-km circuit of 21 lamps in series, any 16 or fewer of which were switched on as needed. The length of the district covered was 1.7 km.

July 1880: Grand Rapids

Grand Rapids Electric Light and Power was organized on 22 March 1880 with William T. Powers as president. He had prospered in the sawmill and cabinetmaking businesses. During the 1860s, he built the West Side Water Power Canal at river frontage he had purchased. In July 1880, Grand Rapids Electric Light and Power began offering Brush arc light service to a few private customers, supplied from a Brush dynamo driven by a water turbine, thought to be the first commercial hydroelectric station in the world (see Allerhand in "For Further Reading.") Grand Rapids Electric Light and Power had 250 arc lamps in operation in 1891, 120 of them streetlights.

September 1880: Detroit

The Brush Electric Light Company of Detroit was incorporated in June 1880 by a group of businessmen led by Wells W. Leggett, a son of Mortimer D. Leggett of Cleveland, then president of Brush Electric. The first Brush installation in Detroit, with 18 arc lamps, began serving various business establishments on 13 September 1880. On 20 January 1881, the *Detroit Free Press* reported that 36 Brush electric lights were burning in the city, including outdoor lighting at a park "at private expense with no cost to the city," a comment possibly prompted by the fact that petitions to the city for electric street lighting had begun in September 1880. In April 1881, Brush Electric Light offered to provide the city with electric lights for US$50,000. The mayor and other city officials were not enthusiastic and dragged their feet. The Brush company had not inspired confidence by installing shoddy and unsightly poles for transmission to private customers and by doing so without city approval. Brush street lighting arrived in Detroit in 1884.

December 1880: New York City

Successfully replacing gas streetlights on Broadway with Brush arc lamps would provide a priceless seal of approval. The first step was to organize the Brush Electric Light Company of New York and obtain city approval. On 25 February 1879, a resolution was adopted by the Common Council, requesting the mayor, comptroller, and commissioner of public works (the Gas Commission) "to have experiments made to test the practicability of lighting the Central Park and the other public parks or places, streets and avenues of this city with electric light, and with a view also of determining the relative cost of the two systems, viz.: gas or electric light; such experiment to be without expense to the city." On 6 November 1880, a communication was received from the Brush Electric Light Company of New York, requesting permission "to erect ornamental lampposts, of iron, on Broadway, from Fourteenth street to Thirty fourth street, for the purpose of lighting said thoroughfare by the Brush electric light; also, the necessary lamp-posts or poles in Twenty-fifth street, from its station to Broadway, for the wires for said purpose, all the work to be done, and lamps, wires, etc., maintained without expense to the city." The length of said portion of Broadway is 1 mi.

Permission was granted on 29 November 1880, provided "that all the work be done and the lamps, wires, etc., maintained by the said Brush Electric Light Company at its own expense, and the Corporation of the City of New York be not committed to any expense for such lighting, and also provided that such permit shall be revocable by the Commissioner of Public Works at any time." The company constructed a central station and installed cast iron lampposts, each 20 ft high, one at the intersection of each street crossing Broadway at an average distance between lamps of 260 ft. For the first time, Brush used his double-arc carbon lamp, which would burn for 16 h without carbon replacement. On 20 December 1880, the section of Broadway between 14th Street and 26th Street was lighted by 15 double-arc lamps (Figure 7), powered by a five-dynamo plant, the largest yet.

The installation was extended to 34th Street in January 1881, with 22 lamps in operation, without cost to the city for five months. Then, the city decided that “the light having proved satisfactory, and an advantageous offer being made by the Company for lighting certain streets and public places, the officers designated in the charter to make contracts for lighting public streets made a contract with this Company to furnish electric lights for eleven months, commencing June 1, 1881,” at a cost estimated to be slightly lower than that of the displaced gas lamps. The contract included Broadway and Fifth Avenue between 14th Street and 34th Street, 14th Street between Fourth Avenue and Fifth Avenue, 34th Street between Fifth Avenue and Broadway, Union Square, and Madison Square. There were 55 Brush arc lamps illuminating streets and other public areas in New York City by the end of 1881.

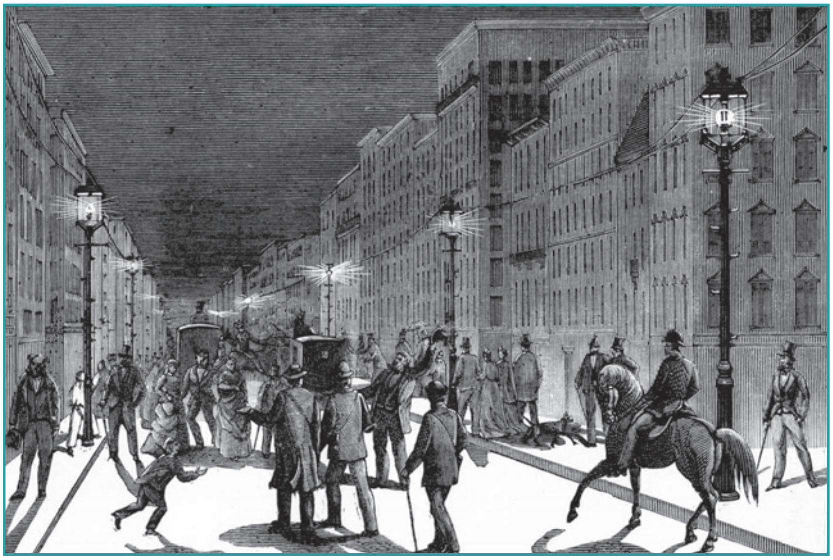


figure 7. Brush streetlights on Broadway, in New York City, at the start of 1881. (Source: É. Alglave and J. Boulard, *La Lumière Électrique*. Paris, France: Librairie de Firmin-Didot et Cie, 1882, p. 95.)

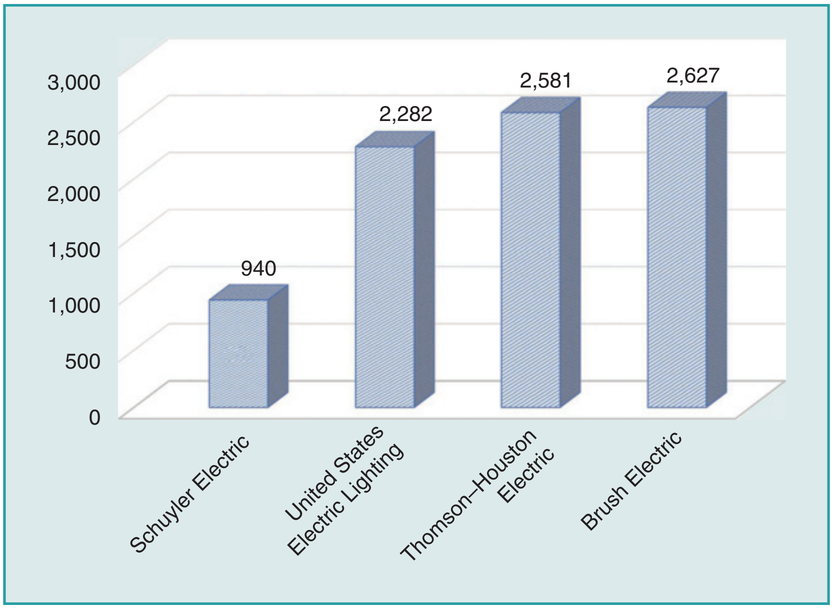


figure 8. The number of arc lamps of major manufacturers in New York City in 1891.

A report by the Department of Public Works, dated 31 December 1881, stated that “the electric lamps now in use have been kept lighted with great regularity and have shown remarkable steadiness or freedom from wavering or blinking.” The report also mentioned that there were 23,466 gas lamps in use in streets and other public areas, without commenting about them. The success of the New York system gave Brush arc lighting a competitive advantage. In the period of 1881–1885, Brush Electric annually doubled its production of arc lamps until, in 1885, the company was producing 16 times as many arc lights as it had in 1881. There were arc lighting systems of 15 manufacturers in use in New York City in 1891, but there were more installed Brush arc lamps than any other brand (Figure 8). However, Brush Electric had ceased to exist as an independent corporation two years earlier.

Competition

The Brush arc lighting system had very little competition in the United States when it was introduced commercially in 1878. In 1881, there were roughly 5,000 Brush arc lights burning in the United States, representing 75–80% of the total. In London in 1881, the rival arc lighting systems of Anglo–American Brush and Siemens Brothers (London) were put on a year’s trial in the streets of two separate districts. At the end of the period, the Streets Committee decided to continue lighting the two districts for another year. The Brush company agreed to light its district for £800. The Siemens offer to light its district for £3,600 was declined.

In 1885, there were approximately 96,000 arc lights in use in the United States, with Brush lamps constituting the largest share. In 1888, there were 21 manufacturers offering systems of electric lighting in the United States. Such a crowded field offered ample opportunities for the strong to swallow the weak. By then, Brush Electric was vulnerable. It suffered from several weaknesses. First, there was no research and development team. Innovation stopped when Brush stopped inventing. Second, Brush's patents did not prevent competitors from marketing inferior and superior arc lighting systems after Brush Electric led the way. Third, Stockly perceived the potential threat of competition as early as 1882, when there was no serious opposition, but omitted due diligence in trying to address it.

Outside of electricity, Brush's contributions as a scientist did not stand the test of time.

A potential rival was American Electric, with the main asset of Elihu Thomson, an inventor of his own arc lighting system. Thomson participated in the Franklin Institute dynamo tests in 1877–1878 and saw the commercial possibilities of arc lighting. Although he had not yet perfected his system in 1882, Thomson had great potential. Stockly bought a controlling interest in American Electric, thinking the deal included Thomson, but Thomson had already decided to leave with his patents because of the company's marketing shortcomings. Thomson's contract with American Electric stipulated he could terminate the agreement and have the patents revert to the patentees (himself and Edwin J. Houston) if the company did not handle the arc lighting business with "reasonable diligence." Stockly had purchased nothing more than machinery and factory space.

Stockly sold his holdings in American Electric to a group of Massachusetts entrepreneurs who, together with Thomson, reconstituted the company as Thomson–Houston in 1883. After that, Brush Electric used patent infringement litigation to attack competition. By 1884, the company was involved in suits against users of arc lighting systems of five manufacturers and against users of arc carbons of two manufacturers. Brush Electric, having enjoyed spectacular growth, was absorbed by Thomson–Houston in 1889.

Thomson–Houston was incorporated in Connecticut on 17 April 1883 but established headquarters in Lynn, Massachusetts. In 1886, Thomson–Houston was capitalized at a meager US$125,000, compared to US$3 million for Brush Electric. But with two key individuals on board, marketing expert Charles A. Coffin as vice president and Thomson as "electrician," the company enjoyed impressive growth. Thomson–Houston was, by far, the dominant supplier of central station machinery in the United States in 1891, with more than 600 central stations (Figure 9). Only 199 central stations were equipped with Brush machinery at the time. Thomson–Houston merged with Edison General Electric in 1892 to form General Electric, with Coffin as president. Thomson's main contribution at General Electric was in the development of commercial X-ray equipment, soon after Wilhelm Konrad Röntgen announced his discovery of X-rays in 1895. Thomson is thought to be the first to demonstrate that X-rays can be harmful. In 1895, General Electric closed the Brush factory in Cleveland and moved the facility's contents to Lynn, the location of the Thomson–Houston plant.

Brush After 1889

After Thomson–Houston took control of Brush Electric in 1889, Brush moved on to various activities. He designed and built on his property what is thought to be the first wind-powered electric generator in the United States, but not in the world (Figure 10). Outside of electricity, Brush's contributions as a scientist did not stand the test of time. Beginning in

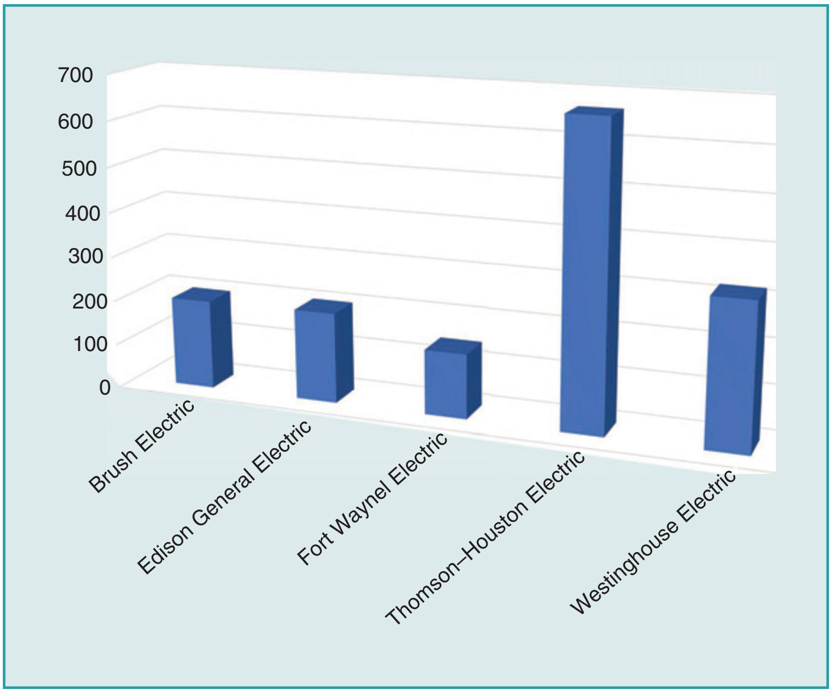


figure 9. The number of central stations using equipment of major manufacturers in the United States in 1891.

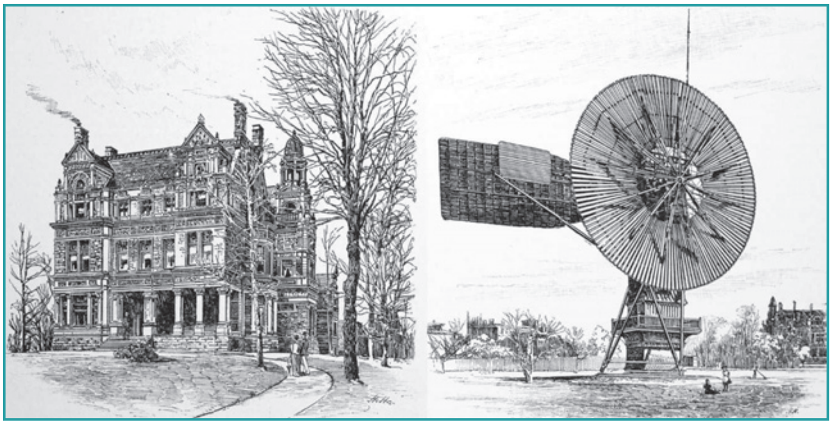


figure 10. Brush's mansion and a windmill-driven electric power generator on his property in 1890. (Source: "Charles Francis Brush," *Harper's Weekly*, vol. 34, pp. 583–584, 1890.)

the 1890s and for much of the rest of his life, Bush was preoccupied with the "ether of space" and gravitation. In 1898, he published experiments that purported to show the existence of a new gas "with enormous heat conductivity at low pressures." He called this new gas *etheron* because he inferred that it "fills all celestial space." In 1911 and then again in 1921 and 1922, he published papers about a "kinetic theory of gravitation." He believed that "the kinetic energy of the ether is the fundamental cause of gravitation."

As a businessman, Brush successfully and profitably shepherded the introduction of German air liquefaction manufacturing into the United States. German scientist and industrialist Carl von Linde had invented a process for industrial-scale refrigeration, and in 1879 he and his partners founded Gesellschaft für Linde's Eismaschinen, a maker of refrigeration machinery. He then invented a process for industrial-scale air liquefaction and entered that business, too. Air liquefaction led to the large-scale manufacture of high-purity oxygen for oxy-acetylene welding and other uses. Linde decided to establish an American subsidiary. He filed a patent application in the United States, in 1895, for a "process of producing low temperatures, the liquefaction of gases, and the separation of the constituents of a gaseous mixture," but his efforts to acquire an American patent stalled because of a pending application for air liquefaction by Charles E. Tripler.

Brush offered to help in exchange for a share in Linde's American venture. Brush filed a second patent on Linde's behalf in 1900 for the apparatus to carry out Linde's process, and he succeeded in having the conflicting application rejected in 1902. Linde's two patent applications were approved in 1903, both one-third assigned to Brush (patents 727,650 and 728,173). Negotiations between Linde and Brush followed. Linde Air Products was born in 1907. It was absorbed by newly incorporated Union Carbide & Carbon in 1917, which changed its name to Union Carbide in 1957.

As a philanthropist, in 1928, Brush created a foundation for "furtherance of research in the field of eugenics and in the regulation of the increase of population," in honor of his son Charles Francis Brush Jr., with an initial bequest of US$500,000. The Brush Foundation's first grant of US$5,000 was awarded in 1929 to establish a birth control clinic in Cleveland. With the discrediting of the eugenics movement after Nazi Germany used it to justify genocide, the foundation's focus since World War II has been to advance reproductive health and rights through targeted philanthropy.

In 1927, Brush lost his son and a granddaughter. Charles Jr. gave a blood transfusion to a very ill daughter. She did not survive. Charles Jr. died on 29 May 1927 at age 33, from complications caused by the transfusion. He had studied chemistry and physics at Harvard University and with a partner founded Brush Laboratories with his father's help in 1921. There, he studied the properties of the very light metal beryllium for the purpose of developing industrial uses for it. In 1931, Brush Laboratories developed a method for extracting beryllium metal from ore, prompting a name change to Brush Beryllium. Additional name changes followed through the years. It is now Materion of Mayfield Heights, Ohio, a suburb of Cleveland. Materion manufactures advanced-performance alloys and composite materials. Charles F. Brush died on 15 June 1929 in Cleveland at age 80.

For Further Reading

C. F. Brush, *The Brush System of Electric Lighting*. Cleveland, OH: Wiseman & Harvey, 1879.

The Brush Electric Light. Cleveland, Ohio: Brush Electric Company, 1881.

M. E. Thompson, "A study of an open coil arc dynamo," *Trans. AIEE*, vol. 8, no. 1, pp. 375–395, 1891. doi: 10.1109/T-AIEE.1891.5570153.

H. C. Passer, *The Electrical Manufacturers—1875–1900*. Cambridge, MA: Harvard Univ. Press, 1953.

H. J. Eisenman, *Charles F. Brush: Pioneer Innovator in Electrical Technology*. Cleveland, OH: Case Western Reserve Univ., 1967.

A. Allerhand, "Hydroelectric power: The first 30 years," *IEEE Power Energy Mag.*, vol. 18, no. 5, pp. 76–87, Sept./Oct. 2020. doi: 10.1109/MPE.2020.2999791.

The Brush Electric Company, *American Electrical Directory for 1886*, Fort Wayne, IN: Star Iron Tower Co., p. 331, 1886.

"Report of the Committee on Dynamo-Electric Machines," *Journal of the Franklin Institute*, vol. 105, p. 296, 1878.

p&e

educating engineers

a time-honored text

THIS ISSUE'S "BOOK REVIEW" column discusses *Electric Power Principles: Sources, Conversion, Distribution and Use,* second edition, written by James L. Kirtley. The reviewer writes, "As a power engineering teacher and researcher, I recommend this book as one of the best electric power systems books."

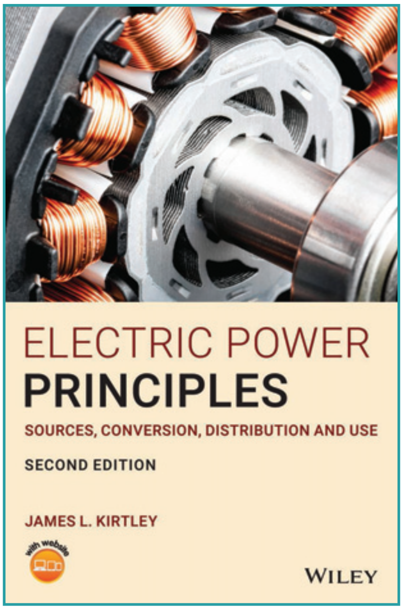
Electric Power Principle: Sources, Conversion, Distribution and Use

By James L. Kirtley

Electric Power Principle: Sources, Conversion, Distribution and Use, the 2020 edition published by Wiley, is aimed at educating engineers and researchers in designing, developing, and operating new types of electric power systems. The initial use of this text was two courses at the Massachusetts Institute of Technology (MIT): 6.061 Introduction to Electric Power Systems and 6.685 Electric Machines. This text has been used and revised for more than three decades. It has educated numerous MIT undergraduate and graduate students, including myself as a postdoctoral fellow when I was at MIT.

Kirtley is an engineer, researcher, and educator in the field of electrical engineering, with an emphasis on power systems and electric machinery. He has close to 50 years of experience working in these areas while at MIT, nurturing students over several decades. He worked for General Electric as an electrical engineer in the large steam turbine generator department. He joined Satcon Technology Corp. as the vice president and general manager of the tech center and a chief scientist. Based on his contributions, he was elected to the U.S. National Academy of Engineering, and he is a Fellow of IEEE.

Digital Object Identifier 10.1109/MPE.2020.3033370
Date of current version: 6 January 2021



All of the key concepts are discussed from fundamental physics, with progress step by step.

This book has 16 chapters. The first covers modern electric power systems' basic structures. Chapter 2 covers electrical engineering fundamentals, while Chapter 3 expands on core theories and transmission lines. Chapter 4 examines common polyphase systems, and Chapter 5 covers electric circuit theory with mathematic derivations. Chapter 6 focuses on transformers for both single- and three-phase structures, while Chapter 7 examines polyphase transmission and distribution lines and introduces the per-unit system. Chapter 8 investigates the fundamentals of electromagnetic forces and loss mechanisms based on the energy-conversion process. Chapter 9 dives into synchronous machines, ranging from basic modeling to application examples. Chapter 10 is about system analysis and protection, such as fault handling, and Chapter 11 presents load flow in power systems with Gauss–Seidel and Newtown–Raphson iterative techniques. Chapter 12 presents the most common power electronic circuits. Chapter 13, a new chapter in this edition, reveals energy storage, including battery modeling and its associated power electronics. Chapter 14 elaborates on classical induction machine developments, and Chapter 15 examines dc machines. The final chapter examines permanent magnet materials.

The author's academic and industry experiences are clearly evident. All of the key concepts are discussed from fundamental physics, with progress step by step. Real-world examples are given to visualize the discussed items. Each chapter's problem set strengthens the presented concepts.

A comprehensive companion website is hosted by Wiley at www.wiley.com/go/Kirtley/electricpowerprinciples. The site supports both instructors and students. Exercises for each chapter are available along with code to implement complicated concepts and generate figures and references for further reading.

The author asserts that this book is suitable for third-year undergraduate electrical engineering students. The material requires a solid background in electrical engineering, multivariable calculus, basic differential equations, electric circuit theory, and Maxwell's equations. As a power engineering teacher and researcher, I recommend this book as one of the best electric power systems books available.

Editor's note: The author of this review is a postdoctoral fellow at the Massachusetts Institute of Technology, under the supervision from Prof. James Kirtley.

—Christopher H.T. Lee

Digital Object Identifier 10.1109/MPE.2020.3043896

PES meetings

for more information, www.ieee-pes.org

THE IEEE POWER & ENERGY Society's (PES's) website (http://www.ieee-pes.org) features a meetings section, which includes calls for papers and additional information about each of the PES-sponsored meetings. Please check the conference website for the most current information and any venue changes due to the pandemic.

January 2021

IEEE PES 2021 Joint Technical Committee Meeting (JTCM 2021), 10–15 January, virtual event, United States, contact Solveig Ward, sward@quanta-technology.com, www.pestechnical.org

February 2021

IEEE PES Innovative Smart Grid Technologies (ISGT 2020), 15–18 February, virtual event, contact Kathy Heilman, kathy.heilman@ieee.org, https://ieee-isgt.org/

May 2021

IEEE International Conference on Electrical Machines and Drives (IEMDC 2021), 16–19 May, Hartford, Connecticut, United States, contact Ali Bazzi, alibassi@ieee.org, http://iemdc-conference.org/

Digital Object Identifier 10.1109/MPE.2020.3033371
Date of current version: 6 January 2021

June 2021

IEEE PowerTech Madrid (PowerTech 2021), 28 June–2 July, Madrid, Spain, contact Tomas Gomez San-Roman, tomas.gomez@comillas.edu, https://www.powertech2021.com/

July 2021

IEEE PES General Meeting (GM 2021), 18–22 July, Washington, D.C., United States, contact Roseanne Jones, roseanne.jones@ieee.org

September 2021

IEEE PES GT&D International Conference and Exposition, Istanbul (GTD), 14–17 September, Istanbul, Turkey, contact Omer Usta, usta@ieee.org, https://ieee-gtd.org/

November 2021

IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC 2021), 21–23 November, Thiruvananthapuram, India, contact Boby Philip, boby.philip@ieee.org

December 2021

IEEE PES Innovative Smart Grid Technologies Asia (ISGT Asia 2021), 5–8 December, Brisbane, Australia, contact Tapan Saha, saha@itee.uq.edu.au

For more information on additional technical committee meetings, webinars, and events, please visit our IEEE PES calendar: https://www.ieee-pes.org/meetings-and-conferences/conference-calendar.

p&e

of negative prices might become less frequent. In fact, as indicated in the previous section, more flexible demand from energy storage and power-to-gas facilities will find it particularly convenient to absorb excess generation when prices reach very low levels. Thus, flexible demand might prevent prices from falling into negative territory. The same applies to electric vehicle charging.

Even without negative prices and with a less binomial price distribution, it is likely that the overall level of revenues accruing to (backup) conventional generation might be insufficient to cover fixed costs. In the medium term, this could lead to a reduction in conventional capacity (in the short term, fixed costs are mostly sunk), negatively affecting system adequacy. The exact policy implications of this possible trend have been the source of much debate over the years.

One viewpoint in this debate is that energy-only markets will continue to provide sufficient price signals and revenues to promote investments in generation capacity. There is a lot of merit in this position, and, in fact, this is the setup in most other markets (for commodities and services), even though the specificities of electricity system operation, with its need for instantaneous balance and its still-limited storage possibilities, might exacerbate its implications in the electricity sector. Capacity shortages will be signaled by frequent electricity price spikes, possibly to the VOLL. This will, in the short run, promote demand response and other forms of demand flexibility; at prices equal to the VOLL, consumers should be, by definition, indifferent toward consuming or not consuming. In the longer run, more frequent high prices would attract needed additional capacity in the market. The European Commission has leaned toward this view of the electricity market. However, as just indicated, this view implies that prices occasionally, and possibly more often than that, reach the VOLL (in the order of several thousands of euros per megawatt hour, or several euros per kilowatt hour). In principle, there is nothing wrong with this, except that this sort of price dynamic might be socially and politically unacceptable.

Capacity Remuneration Mechanisms

As a result, even the European Commission recognizes the need, in some specific circumstances and mostly as a transitory measure, for the deployment of capacity remuneration mechanisms (CRMs). The recast of the Electricity Regulation [Regulation (EU) 2019/943], as part of the Clean Energy for All Europeans package, has set strict rules in this regard. In particular, CRMs can be introduced only to address residual adequacy concerns that cannot be dealt with by measures that member states have to introduce to eliminate any regulatory distortions. Among the different CRMs available, legislation favors the use of strategic reserve.

There is an expectation that residual adequacy concerns will be temporary. Therefore, a good feature of CRMs is that of self-regulation, in the sense of being able to provide the necessary additional stimulus when adequacy concerns emerge but automatically "retreat" when these concerns recede. Among the CRMs proposed or being implemented, the one based on reliability options, among others, appears to have such characteristics. Reliability options, used as a way of addressing long-term resource adequacy, were first proposed by Pérez-Arriaga in 1999. As the name *reliability options* suggests, they are option contracts that require generators and other adequacy providers, in exchange for a fixed fee, to pay the holder of the contract, in any market time unit, any positive difference between the equilibrium market price (p) and a predefined strike price (s). In so doing, these contracts aim to incentivize generators and other adequacy providers to be available when the market price is high (above the strike price), signaling the tightening of the demand-supply relationship, because in this way, they will be able to honor payment under the contract with the higher revenues from the market.

The main advantage of the reliability option mechanism vis-à-vis other possible CRMs is twofold. On the one hand, it does not require the predefinition of scarcity periods (when the contracted resources are expected to be available), as it uses the occurrence of high market prices—above the strike price—for this purpose. On the other hand, the reliability option mechanism is, admittedly, more market oriented because if the strike price is set at an appropriately high level (as discussed later in this article), it activates only when the system is close to rationing and does not affect the spot market under normal or even tight conditions. In this sense, reliability options can coexist well with commercial long-term financial contracts (e.g., futures or contracts for difference), which provide hedges against price volatility for both consumers and generators.

Among EU member states, Ireland and Italy have decided to implement reliability option-based schemes. Outside the EU, reliability options have been implemented in Colombia and by ISO New England, a regional system operator in the United States. However, in the Irish and Italian implementations, the distinction between reliability options and price risk-hedging instruments has been somewhat lost due to the fact that the strike price is closely linked to the costs or the expected offer levels of a peaking generating unit (capped at €500/MWh in the Irish implementation).

Setting the strike price in this way might risk replacing market dynamics with an administratively set, two-part wholesale price, which was typical of the regulated generation sector of the 1990s. The reason usually provided for this approach is to control the potential abuse of market power, which, in a competitive market, is better controlled through competition policy rather than ex-ante regulation.

Over time, other resources, such as demand response and storage, have entered and will continue to enter the market, which might be promoted by prices higher than the costs of peaking units. The relatively low strike prices of reliability options can act as de facto price caps that may discourage the participation of such new resources. While reliability options typically operate only in the day-ahead market (although in Ireland, the market reference price is a weighted average of the day-ahead, within-day, and balancing markets prices) and demand response and storage as well as other decentralized resources can also operate in the intraday and ancillary services markets, there does not seem to be any good reason to effectively exclude such decentralized resources from the day-ahead market.

In my view, the strike price of reliability options, because of their purpose to ensure adequacy and not to regulate the wholesale price, should be set at a level that represents the conceptual discriminant between market functioning, including under tight conditions, and the situation in which nothing could prevent prices to increase up to the VOLL. In their 2002 seminal work on reliability options, Vázquez et al. suggested that the strike price could be considered "as a frontier between the normal energy prices ($p < s$) and the near-rationing or emergency prices ($p > s$)." At that time, a markup of 25% above the variable cost of the most expensive generator expected to produce was considered sufficient. Today, with additional resources available to ensure adequacy, such a reference to generation capacity appears outdated, but the general consideration of the strike price as the frontier to near-rationing levels remains valid. This frontier level for the strike price is clearly not the cost of a peaking unit in the market because some other resources, such as demand response, storage, and other decentralized resources, would be attracted to the market only by much higher prices.

Such higher prices have occurred in the EU electricity market in recent years without creating any particular disruption. According to the ACER/CEER 2019 Market Monitoring Report, hourly day-ahead prices higher than three times the (theoretical) variable cost of a gas-fired power plant have occurred close to 3,500 times from 2015 to 2018, with nearly 1,500 occurrences in 2016 alone. In fact, in 2017, when defining the harmonized maximum clearing price for the day-ahead market coupling, ACER set an initial value of €3,000/MWh, with the possibility of an automatic upward dynamic adjustment if the clearing price in the market exceeded 60% of the applicable maximum value. ACER's decision shows that even prices in the thousands of euros per megawatt hour, and not just those in the hundreds of euros, can be expected and should not be considered as anomalous—to the extent that they lead to an upward revision of the maximum clearing price.

Let me conclude with one final observation regarding the design of reliability option mechanisms, which is often overlooked. Reliability option schemes clearly provide a more stable revenue stream for contracted reliability resources. However, unless they include a penalty for nondelivery when the option is called, they may be seen as not providing additional incentives for adequacy providers to be available during times of scarcity.

This is clear if we consider, as an example, the difference in the payouts (including variable costs, as fixed costs do not depend on production levels) for generators between producing and not producing at times when the market price exceeds the strike price, both with and without a reliability option. Without a reliability option contract, the incentive for a generator to produce when prices are high would be the (gross) margin that it could obtain from selling its electricity in the market, i.e., the difference between the price it receives and its variable costs. If the generator has entered into a reliability option contract, it would receive the contract fee irrespective of the price level in the market and whether it produces or not. Moreover, if prices rise above the strike price, it would have to make the contract payment of the difference between the market price and the strike price, again, irrespective of whether it produces or not. Therefore, the incentive to produce would be the margin that it can obtain from producing and selling its electricity in the market, i.e., the difference between the price it receives from the market and its variable costs, the same as in the case without the reliability option contract.

Only when a nondelivery penalty—imposed on the contracted generator that fails to produce at high prices—is included in the reliability option scheme does the incentive to produce when prices are higher than the strike price become greater than without such an option. The Irish implementation does not envisage nondelivery penalties while the Italian one has two sets of penalties. Note that the inclusion of a penalty neither presupposes nor is equivalent to physical prequalification, as required by some schemes. The penalty is still a financial charge, increasing the incentive to produce when the market price is higher than the strike price, but the writer of the option still needs to secure the necessary physical generation capacity during times of high prices.

Concluding Remarks

As more countries are tempted to introduce CRMs to address residual adequacy concerns raised by the greater penetration of renewable-based generation, prompting them to consider reliability option schemes as the preferred approach due to their market-oriented characteristics, these schemes must be appropriately designed. In particular, the strike price should not interfere with the functioning of the energy market under normal or even tight conditions and penalties should be envisaged that reinforce incentives for adequacy resources to be available during times of scarcity.

For Further Reading

I. J. Pérez-Arriaga, "Reliability in the new market structure (part 1)," *IEEE Power Eng. Rev.*, vol. 19, no. 12, pp. 4–5, Dec. 1999. doi: 10.1109/MPER.1999.808165.

C. Vázquez, M. Rivier, and I. J. Pérez-Arriaga, "A market approach to long-term security of supply," *IEEE Trans. Power Syst.*, vol. 17, no. 2, pp. 349–357, May 2002. doi: 10.1109/TPWRS.2002.1007903.

P. C. Bhagwat and L. Meeus, "Reliability options: Can they deliver on their promises?" *Electricity J.*, vol. 32, no. 10, p. 106,667, 2019. doi: 10.1016/j.tej.2019.106667.

Alberto Pototschnig

capacity remuneration

advancing the spread of renewables

THE EUROPEAN GREEN DEAL, setting the decarbonization strategy of the European Union (EU) for the period to 2050, envisages a 50–55% reduction in greenhouse gas emissions with respect to 1990 levels by 2030 and carbon neutrality by 2050. This ambitious pathway requires a much greater penetration of renewables in the energy system. The current 32% renewable penetration target for final energy consumption set for 2030 only two years ago is, therefore, likely to be increased to 38–40%. The electricity sector typically contributes more than proportionately to the achievement of the overall target. This means that we can expect renewable-based generation to represent approximately two-thirds of final electricity consumption by 2030, up from roughly 30% in 2019.

A large share of the increased renewable-based generation capacity would be wind and solar. These technologies are characterized by zero or near-zero marginal costs, but when compared to conventional generation, they are also more variable and less dependable in providing electricity with a profile that follows the pattern of demand. Therefore, the security of supply requires the availability of backup dispatchable capacity to follow and serve the residual load—the difference between demand- and renewable-based generation.

The merit-order effect of the greater penetration of renewable-based generation implies that this flexible backup capacity, which, in the shorter term, will be mostly provided by conventional (gas-fired) plants, will be called to produce for a decreasing and highly variable number of hours and thus require higher revenues when it generates.

Therefore, a simplistic assessment of the implications of this change in the generation mix and cost structure on the electricity price profile suggests a larger number of hours when the electricity price in the market will be set by renewable-based generation at zero or very low levels and a few hours in which prices might reach very high levels. This could possibly be up to the value of lost load (VOLL) when available generation capacity is unable to meet demand, to allow nonrenewable-based generation capacity to recover its fixed costs.

However, a more elaborate assessment could recognize that the distribution of prices might not necessarily be as binomial as it might appear at first, especially considering technological development and the emergence of new technologies and facilities in the market. For example, more variable prices may enable energy storage facilities to take advantage of arbitrage opportunities by charging when prices are low and discharging at higher prices. The future will also see power-to-gas technologies develop, which will allow for the storage of energy in molecular form over longer periods of time. The further development of demand-side response and other decentralized resources may also contribute to curbing price spikes. Therefore, low prices might not always be that low, and high prices might not always be that high.

Negative Market Prices and Renewable Support Schemes

In this context, specific attention should be paid to negative market prices. According to the 2019 Market Monitoring Report of the European Union Agency for the Cooperation of Energy Regulators (ACER) and the Council of European Energy Regulators (CEER), there were more than 500 occurrences of negative hourly prices in the EU day-ahead market in both 2017 and 2018. Different circumstances can lead to negative prices occurring. In the past, these occurrences were fairly rare and mostly caused by the limited flexibility, in technical and economic terms, of some conventional generators. It was cheaper to produce at negative prices than to switch generating units off for one or a few hours. More recently, negative prices have also been the result of ill-designed renewable support schemes. Some of these schemes have been providing supplemental revenues—through feed-in premiums—to renewable-based generation plants, even during times of negative prices. This means that such plants are incentivized to generate even when there is excess production in the system, at least as long as the supplemental revenue component compensates for the negative price.

Therefore, it seems urgent that the design of feed-in premiums and similar schemes excludes support when the market signals excess supply through negative prices. This is a no-regret move, even though, in the future, the occurrence

(continued on p. 98)

Digital Object Identifier 10.1109/MPE.2020.3033372
Date of current version: 6 January 2021

advertisers index

The Advertisers Index contained in this issue is compiled as a service to our readers and advertisers: the publisher is not liable for errors or omissions although every effort is made to ensure its accuracy. Be sure to let our advertisers know you found them through *IEEE Power & Energy Magazine*.

Company	page#	URL	Phone
ASPEN, Inc.	CVR 4	www.aspeninc.com	+1 650 347 3997
Bigwood Systems, Inc.	6	www.bigwood-systems.com	+1 607 257 0915
Craig Wire Products	15	www.craigwire.com	+1 770 920 2222
CYME	CVR 2	www.eaton.com/cyme	+1 800 361 3627
DIgSILENT GmbH	3	www.digsilent.de/powerfactory	
EasyPower LLC	17	www.EasyPower.com/demo	
ETAP	13	etap.com/20.5	+1 949 900 1000
Mitsubishi Power	11	ChangeInPower.com/TOMONI	
P & R Technologies	18	www.pr-tech.com	+1 800 722 8078
Powertech Labs, Inc.	14	www.dsatools.com	
PowerWorld Corporation	8	www.powerworld.com	+1 217 384 6330
PSCAD a product of MHI	7	pscad.com	
RTDS Technologies, Inc.	9	www.rtds.com	
SKM Systems Analysis, Inc.	5	www.skm.com	+1 800 500 4SKM
Tana Wire Markers	97	www.tanawiremarker.com	+1 573 796 3812

445 Hoes Lane, Piscataway, NJ 08854

IEEE power & energy magazine representative

Erik Henson
Naylor Association Solutions
+1 352 333 3443, fax: +1 352 331 3525
ehenson@naylor.com

Digital Object Identifier 10.1109/MPE.2019.3042575