1

Introduction

Imagine a scenario in which there is great concern for maintaining an adequate number of grocery stores to serve the food-purchasing public around the clock, so that consumers always have access to open grocery stores to buy food. In response to this concern, a new revenue stream is created for grocery stores. Previously grocery stores earned their revenues entirely from consumer purchases of food. Now consumers will also be required to pay in advance to ensure that grocery stores will be open for business twenty-four hours per day. Grocery stores will receive these payments from consumers merely for being open, even if no consumers buy any food. These new payments to grocery stores do not entitle consumers to purchase food at any particular price from the stores. The grocery stores' only obligation is to be open for business, ready to sell food.

A market to keep open grocery stores would be very peculiar and would raise several questions: Why are revenues from food purchases not sufficient to keep open grocery stores? Why should consumers have to pay to keep grocery stores open around the clock? How will the new revenue stream to grocery stores affect food prices and quantities sold in stores? If a store closes for reasons beyond its control, such as a snowstorm or tornado, does it still receive a payment? What about grocery stores that have their own reasons for not always being open, such as a kosher store that closes on Saturdays? Should grocery stores that do not always have fresh fruits and vegetables be rewarded less than stores that do have fresh produce? Finally, who is responsible for making all the rules necessary for this new market to operate?

As strange as this scenario seems, in many ways it describes the essence of electricity capacity markets. In a capacity market, suppliers of electric power, such as power plants, are paid in exchange for committing to have their capacity available for providing electricity to the electricity system – known as the *electricity grid* – so that adequate supplies of electricity are always available to meet demand. The supplier is not being paid for providing electricity; it is merely being paid to be available to provide electricity. The buyers of capacity, meanwhile, pay suppliers even though the benefits of capacity inure to the entire electricity system, rather than on a buyer-by-buyer basis.

There are reasons why we have electricity capacity markets but not markets to keep open grocery stores. Disappointed midnight grocery store customers experience inconvenience from having to wait until morning for an open store, but electricity can be a necessity that must be immediately available when needed. Grocery store customers can buy food in advance and save it for later, but electricity customers generally cannot store electricity for later use. Grocery stores can make their own individual decisions about when to be open, but the electricity system must be managed on a system-wide basis. These factors distinguish electricity markets from typical markets for goods and services. Whether these reasons justify creating electricity capacity markets is a more difficult question.

Each year in many parts of the United States, electricity distribution companies and other electricity retailers spend billions of dollars purchasing capacity in capacity markets. Although the business of these companies is to supply electric power to their customers, capacity is not power. Purchasing capacity in a capacity market thus does not entitle distribution companies to any electricity. These firms do not turn around and sell capacity to their customers. They purchase electricity itself from generators in separate transactions in what are called *energy markets*. Although capacity is forward-looking, the purchase of capacity is not an option or a futures contract. Purchasing capacity does not even entitle a utility to purchase electricity at a particular price at some future date, or to otherwise hedge against future increases in electricity prices. The generator that sells capacity is not being paid for generating electricity; instead, it is merely being paid to be available to generate electricity.

Capacity markets have become a significant part of the electric power industry, with many billions of dollars in payments per year. For example, capacity payments constitute approximately twenty percent of the charges from PJM, the United States' largest electricity system operator (PJM 2019, 1). In 2020, this amounted to over $7 billion in payments in PJM's capacity market (Monitoring Analytics 2021, 4, 17). Figure 1.1 illustrates capacity market payments since 2000 as compared with payments for the sale of electricity in PJM's wholesale energy market. Capacity markets have become an important source of revenue for generators and costs for consumers.

What these capacity payments are for, whether they are necessary, and how they are affecting the energy sector are therefore important questions of considerable importance. The goal of this book is to answer these and other related questions.

A THE CHALLENGE OF CAPACITY MARKETS

A modern industrial economy runs on massive amounts of energy, much of which is used in the form of electricity that is generated, transmitted, and distributed through the electricity grid. More than 8,000 power plants in the United States produce almost four trillion kilowatt hours of electricity annually, which five million miles of

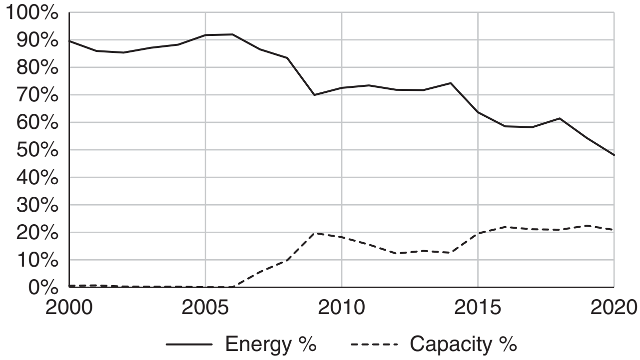

FIGURE 1.1 Share of total PJM payments, 2000–20
Source: Monitoring Analytics 2001–21.

transmission and distribution lines deliver to 150 million electricity customers (EIA 2019; EIA 2018; Weeks 2010). The operation and management of the modern electricity grid is so complex that it seems almost impossible that these systems are able to function as well and as reliably as they do. We take for granted that our electricity grid will provide power on demand at all times and in virtually any circumstances short of a catastrophic natural disaster.

Exacerbating the difficulty of maintaining reliability, our electricity grid is becoming more complex as it faces new economic, technical, legal, and political challenges. Dramatic increases in natural gas production have reduced natural gas prices, leading to the construction of more natural gas power plants and the closing of coal and nuclear power plants. Increasingly dire warnings about the need to decarbonize rapidly to avoid catastrophic climate change have led to growth in the supply of solar and wind power generators that provide variable power to the grid. The costs of renewable power also are decreasing rapidly, further stimulating their growth. "Smart" grid technologies are being developed that have the capability of replacing the traditional electricity grid infrastructure with an interactive and multifunctional network. The innovations that are necessary to meet these diverse and pressing challenges have the potential to create great benefits that will make the electricity grid more responsive to the diverse needs of businesses and the public. Unfortunately, these innovations also will make the grid more complicated to operate.

Capacity markets have emerged as a key tool to manage the reliability of the electricity grid. They are increasingly important but also fiendishly difficult to understand, reflecting all the complexity of the grid itself and of electricity markets. Capacity markets are designed to ensure that the grid will have sufficient generation capacity to satisfy peak demand, so that the grid continuously provides a reliable supply of electric power. They do this by creating an additional revenue stream for resources that, in return for receiving capacity payments, incur an obligation to be available to provide power on demand. The expectation is that this additional revenue stream of capacity payments will encourage construction of new generation and allow some existing generation to remain in operation, to the extent necessary to achieve reliability.

In addition to their role in maintaining grid reliability, capacity markets also have important implications for the development of renewable energy sources and the environmental impacts of energy production. Capacity markets, which are intended to enhance reliability, must decide how to count generation capacity from variable renewable energy sources such as solar and wind that, due to intermittency, are not necessarily available to generate power on demand. This design challenge goes to the heart of what it means for a generation resource to satisfy its obligation to be available under capacity market rules. Questions also arise about whether and how capacity markets should treat generation resources that receive federal and state subsidies, many of which aim to create incentives for renewable and nuclear power. Capacity market designers must decide whether these subsidies are legitimate energy policies that capacity markets should accommodate or market distortions that capacity markets should seek to exclude or counteract. The stakes for the future of non-fossil energy sources are high, as revenues from capacity markets are an important determinant of the financial viability and ability to compete for both renewable and nuclear energy resources.

Although important to the economics and operation of the grid, capacity markets are complicated and hard to understand. Indeed, even many experts in electricity regulation concede their lack of comprehension of these markets. It is unclear what exactly is being transacted in capacity markets, precisely what market failures capacity markets are intended to correct, and how capacity markets operate. Some commenters have questioned whether capacity markets can properly be called markets at all, given how tightly controlled the transactions are and the fact that the sale of capacity in the market does not appear to transfer any identifiable benefit to the buyer (e.g., Farber 2018).

Capacity markets are a fundamentally pragmatic policy solution to a perceived problem of underinvestment in generation capacity – known as the "missing money" problem – that has been only thinly identified and explained. Evaluated at the practical level at which they were conceived and designed, capacity markets appear to be largely successful. Capacity markets seem to have alleviated any missing money problem and created sufficient incentives for investment in new generation. This can be construed, at least narrowly, as a success, although there is significant controversy about whether such incentives are excessive, to the detriment of electricity consumers.

Capacity markets, however, suffer from the inherent shortcomings of pragmatic policy solutions. Since there is no consensus about the root causes of the missing money problem, it is difficult to judge whether capacity markets are an optimal policy solution and, if they are not, how they can be improved. There is uncertainty and controversy about the level of missing money, and therefore whether capacity markets are sending the right market signals. Without a clear understanding of the market failures that capacity markets are attempting to redress, we cannot evaluate whether they are correcting or distorting electricity markets.

This book attempts to improve understanding of capacity markets and to advance the process of holding them accountable for their performance and consequences. The time is ripe for a searching reassessment of all aspects of capacity markets – of the reliability objectives that drive the markets, of the theories of market failure that are offered to justify capacity policies, of whether capacity markets are the best available option for achieving adequate reliability, and of how to improve capacity market design features.

We are now more than twenty years into the capacity market experiment in the United States. Two decades of experience have given us a more realistic and informed understanding of what capacity markets do well and not so well. Most important, experience has alleviated concerns that capacity markets will not procure adequate capacity. Indeed, capacity markets routinely procure excessive capacity, influenced by market rules that adopt conservative approaches to reliability. But capacity markets have struggled to employ competitive market forces. Demand remains entirely administrative, and supply has become increasingly so, with both offer caps and offer floors applied to constrain bidding. As regulation accretes, there is less and less of a "market" in capacity markets.

In addition, capacity market rules are becoming increasingly complicated, creating concerns that they are creating winners and losers through politics rather than neutral market forces. There is further controversy about how capacity markets designed to address the missing money problem should interact with other energy policies at the federal and state level that aim to address other concerns, such as subsidies designed to expand or preserve carbon-neutral power generation. Adding analytical depth to our understanding of capacity markets can help stimulate more insightful evaluation of their strengths and weaknesses and identify potential improvements.

Capacity markets are also facing a significantly changing context that will present both new options and new challenges. Technological innovations in electricity storage, demand response, and metering have the potential to reduce the incidence of market problems that provide the justification for capacity markets. Increasing development of intermittent renewable energy sources such as solar and wind will pose challenges for ensuring that electricity supply is always available to meet demand. Potential electrification of US. energy use, driven in part by efforts to transition away from fossil fuel use (Cleary 2019), may increase overall demand for electricity and shift electricity demand across periods of the day and seasons (Mai et al. 2018, xiv–xv). Capacity markets may need to evolve to ensure resource adequacy under these new conditions.

The electricity sector is transforming, and it needs policies that will channel change toward serving the public good. To date, the track record of capacity markets on this question is mixed at best. Even as they have succeeded relatively well in avoiding outages, they have become massively complicated conglomerations of rules and policies – too often the product of processes that favor incumbent stakeholders rather than market principles. Whether capacity markets can be reformed to stimulate rather than stifle innovation and adaptation remains very much an open question, and a question that is more important now than ever to answer.

B USING THIS BOOK

This book examines the economics, law, and politics of capacity markets. It is intended to appeal to energy experts and non-experts alike, across a range of disciplines, including energy, economics, business, political science, public policy, and law around the world. Indeed, the challenges facing capacity markets – harnessing market forces for the social good, creating networks that manage complexity, and achieving sustainability – are very much core challenges for a twenty-first century advanced industrial society.

The focus of the book is on capacity markets in the United States generally. That said, our description and analysis sometimes will focus on examples drawn from the capacity market in the largest electricity grid in North America – the PJM Interconnection in the Middle Atlantic region of the United States – as well as the capacity markets of the New York Independent System Operator (NYISO) and the New England Independent System Operator.

We have organized the fourteen chapters of this book into four general parts: Foundations, History, Market Elements, and Market Challenges.

Part I lays a foundation for understanding capacity markets. After this introduction, Chapter 2 offers a primer that introduces the essential elements of capacity markets that subsequent chapters will address in detail. Chapter 3 examines the process of restructuring by which many electricity markets in the United States in the 1990s moved away from traditional public utility regulation and toward competitive markets. Chapter 3 also explains regional transmission organizations (RTOs), which operate the electricity grid in areas with competitive markets. It is RTOs, under the oversight of the Federal Energy Regulatory Commission (FERC), that have created and operate capacity markets. Chapter 4 investigates the problem of reliability for electricity grids, and in particular the missing money theory which posits that competitive electricity markets operated by RTOs inadequately compensate electricity generation, justifying a regulatory intervention. Chapter 5 analyzes the various policy options that have been proposed to address the missing money problem in electricity markets, including capacity markets.

Part II conveys the twenty-plus-year history of capacity markets in the Northeast United States, separately operated by in PJM Interconnection, NYISO, and the New England ISO. Chapter 6 explains the first generation of capacity markets that RTOs created in the late 1990s immediately following the introduction of competitive electricity markets. First-generation capacity markets were originally seen as a means by which to add flexibility to traditional capacity requirements through a voluntary capacity auction, rather than as an altogether new capacity policy. Because of this, early capacity markets arose without much debate, theory, or detailed analysis. Soon after the turn of the century, however, the shortcomings of these first-generation capacity markets had become clear. As Chapter 7 explains, the three Northeast RTOs replaced their troubled first-generation capacity markets with more robust but also more complex second-generation capacity markets, which involved mandatory capacity auctions. Unlike the first-generation markets, second-generation capacity auctions treated capacity markets as a distinctive reliability policy. Neither FERC nor the RTOs, however, used the opportunity presented by this regulatory renovation to consider deeper questions about capacity markets, including a theoretical framework that could guide capacity market design.

Part III dissects the fundamental elements of capacity markets: demand, supply, auction structure, performance incentives, and other aspects of market design. Part III starts with Chapter 8, focusing on demand in capacity markets. Unlike typical markets in which demand for a product arises organically from the preferences of buyers, RTOs administratively create demand curves for capacity markets. In doing so, RTOs have employed decision processes that reflect stakeholder interests rather than market principles, and the ensuing results tend to exhibit bias toward higher prices and quantities of capacity, resulting in higher revenues for generators. There thus appears to be a mismatch between capacity market demand and the actual value of capacity. Chapter 9 considers supply in capacity markets. The supply curve represents the aggregated bids of capacity suppliers. Different suppliers' bids indicate their differing opportunity costs of incurring capacity obligations – for example, existing suppliers tend to bid at prices close to zero, representing their variable costs, whereas new suppliers tend to bid at a non-zero price representing the fixed cost they need to incur to enter the market. Beyond issues of supply and demand, Chapter 10 examines important remaining questions of market design, such as auction format, opt-out programs, and performance incentives. The details of market design choices have important consequences for the outcomes and performance of capacity markets.

Part IV considers important challenges facing capacity markets. Chapter 11 addresses the problem of market power, which is a recurring difficulty in electricity markets. Market power is especially important in capacity markets, because of the inelasticity of both market supply and demand curves. RTOs have had limited success in curbing the exercise of market power in capacity markets through offer caps. The use of offer caps, however, undermines the ability of capacity markets to provide the benefits of a competitive market. Chapter 12 examines Minimum Offer Price Rules (MOPRs), a subject of intense controversy, which apply bid floors to capacity offers. Originally introduced to prevent vertically integrated firms from intentionally suppressing market prices for financial gain, MOPRs have broadened their targets to price suppression more generally. This regulatory expansion lacks justification, as there is no indication that capacity prices are too low. Indeed, the opposite problem appears more relevant. Chapter 13 turns away from the Northeast RTOs to focus on Texas, which addresses resource adequacy through its energy market rather than a capacity market. The Texas option is often touted as an alternative to capacity markets. An extended weather-related outage in February 2021 revealed some shortcomings in Texas's system, although it does not appear that capacity markets could have prevented the crisis.

The concluding chapter, Chapter 14, offers five lessons to be gleaned from the book's analysis, which collectively challenge some of the core assumptions on which capacity markets rest. First, no consensus exists regarding the missing money theory that is used to justify capacity markets. Second, the use of forward capacity markets, in which auctions occur several years in advance of the performance period, causes systematic errors in forecasting electricity demand, which in turn induces capacity markets to procure excess capacity and impose excess costs on consumers. Third, RTOs and FERC are preoccupied with concerns that capacity prices are too low, but in fact prices are higher than necessary. Fourth, capacity markets have become extremely complex, due in significant part to regulatory accretion that undermines the role of market forces. Fifth, capacity markets are designed to meet an engineering-based reliability standard that is based on a metric poorly matched to the goal of reliability.

Although this book was written with the objective of appealing in its entirety to a wide range of readers, different audiences may wish to focus on different parts of the book:

- Readers interested in learning more about capacity markets generally may want to focus on Part I (Foundations), Part II (History), and the Conclusion.
- Readers interested in understanding how capacity markets work may want to read Part I (Foundations), Part II (History), and Part III (Market Components).
- Readers interested in current policy controversies may want to read over Chapter 2 (Primer) and Part III (Market Components) and then closely read Part IV (Market Challenges).
- Readers interested in the implications of capacity markets for renewable energy development may want to read Part I (Foundations), Chapter 9 (Supply), Chapter 12 (Minimum Offer Price Rules), Chapter 13 (Texas), and the Chapter 14 (Conclusion).

C REFERENCES

Cleary K (2019) *Electrification 101*, Washington, DC: Resources for the Future.

Energy Information Administration (EIA) (2018) *Electric Power Annual*, Washington, DC.

Energy Information Administration (EIA) (2019) *Frequently Asked Questions*, Washington, DC.

Farber D (2018) The puzzle of capacity markets. *Legal Planet* (June 11).

Mai T, Jadun P, and Logan J et al. (2018) *Electrification Futures Study: Scenarios of Electric Technology Adoption and Power Consumption for the United States*, Golden, CO: National Renewable Energy Laboratory (Report TP-6A20–71500).

Monitoring Analytics, L.L.C. (2000–21) *State of the Market Report for PJM*, Eagleville, PA.

PJM Interconnection, L.L.C. (2019) *Understanding the Differences Between PJM's Markets* (March 6), Norristown, PA.

Weeks J (2010) US electrical grid undergoes massive transition to connect to renewables. *Scientific American* (April 28).